

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

³ Reed Clark Benkendorf^{1,2}, Emily J. Woodworth^{1,2}, Jane E. Ogilvie^{3,4},
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

Reed Clark Benkendorf^{*1,2}, Jane E. Ogilvie^{3,4} , Emily J. Woodworth^{1,2}

⁴ **Abstract**

5) 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 syncology in some instances may begin to lag behind other
9 kingdoms.

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

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

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

*Correspondence: reedbenkendorf2021@u.northwestern.edu

20 real world scenarios.

21

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

24 **1 | INTRODUCTION**

25 The inability to reliably identify plants down to terminal taxon can limit our understanding of ecosystem
26 function and interactions (Bortolus (2008)). This is especially true for genera where many species are defined
27 upon ecological and behavioral rather than morphological properties, and hence often serve as bioindicators
28 of habitat (e.g. different species of Sagebrush- *Artemisia* L., Willows - *Salix* L., and Sedges - *Carex* L.)
29 (Gage & Cooper (2013)). The lack of species level data can hinder our understanding of the breadth of
30 habitat which some species occupy, and the interactions they have with other species. Current methods
31 to ameliorate this situation include: ignoring these ecologically relevant levels of detail, revisiting plots
32 as diagnostic material becomes temporally available, assistance from taxonomic specialists, or the use of
33 barcoding or other molecular techniques (CITE). The identification of organisms to terminal taxon is often
34 mired by lack of diagnostic characters (e.g. flowers, fruits, roots or combinations thereof), an increasing lack
35 of taxonomic experts (Hebert *et al.* (2003)) and increasingly the description of cryptic species (Janzen *et al.*
36 (2017), Oliver *et al.* (2009)). And revisiting field sites to identify material using morphological or chemical
37 approaches, can be resource intensive and often does not work.

38 Recently barcoding (the identification of a sample from a single organism *e.g.* a piece of leaf), and metabar-
39 coding (the identification of a sample containing a mix of organisms *e.g.* soil), have shown considerable
40 promise in all Kingdoms of Life (Ruppert *et al.* (2019)). With plants the identification of members of cer-
41 tain clades using barcoding has been quite successful (Kress (2017)), whereas with many other clades results
42 have been elusive (Liu *et al.* (2014), Group *et al.* (2011), Coissac *et al.* (2012)), however metabarcoding
43 incurs additional challenges to those which exist for the currently implemented barcodes (Li *et al.* (2015),
44 Kress & Erickson (2007), Group *et al.* (2009), Coissac *et al.* (2012)). Particular challenges with the high
45 copy number barcodes (e.g. ITS2, *rbcL*, *matK*, *trnH-psbA*) include the utilization their rates of divergence,
46 gene tree conflict, and hybridization (Coissac *et al.* (2016), Fazekas *et al.* (2009)).

47 Currently the largest plant systematic endeavor ever undertaken, the Plant and Fungal Tree of Life
48 (PAFTOL) undertaken by the Royal Botanic Gardens Kew, is approaching completion (Baker *et al.*
49 (2021a)). This data set will contain hybridization capture (Hyb-Seq) data from at least one species in each

50 genus of the plant kingdom using the popular Angiosperms353 (A353) probes, at 353 single-copy orthologous
51 loci, (Baker *et al.* (2021a), Johnson *et al.* (2019)), resulting in over 14,000 represented species. These
52 publicly available data serve to provide a taxonomically comprehensive backbone for plant metabarcoding,
53 and the A353 probes are currently being used in many other plant phylogenetic issues increasing the
54 sampling depth of many clades (Baker *et al.* (2021b)). Data from the 10kP project, which seeks to develop
55 reference genomes from a phylogenetically diverse suite of plants will contribute many more species upon
56 its intended completion, now slated to be by 2030, similar projects which seek to sequence high amounts
57 of genomes in geographic regions e.g. the ‘Darwin Tree of Life’ which will sequence all described taxa in
58 Britain and Ireland, are being undertaken which will contribute data sets applicable to enormous spatial
59 domains (Cheng *et al.* (2018), Life Project Consortium *et al.* (2022), Lewin *et al.* (2022)). These data will
60 promote the ability to apply metabarcoding to resolve a diversity of questions relevant to theoretical and
61 applied ecology (Kress (2017), Hollingsworth *et al.* (2016)). However, the application of metabarcoding
62 still face challenges relating to the enormity of the genomic data sets and the computational power required
63 to process sequence data. Herein we have resolved major components of the problems of identifying plant
64 material without diagnostic morphological character states using the A353 Hyb-Seq probes (Johnson *et al.*
65 (2019)), and custom species sequence databases derived via species distribution modelling, and temporal
66 filtering.

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

81 Considerable amounts of species interactions vary both in space and time (CaraDonna *et al.* (2021)). In many

82 regions of the world, contrasts in the flowering periods of many plant species, can provide an additional filter
83 for identifying material in certain types of metagenomic samples (Janzen (1967), Newstrom *et al.* (1994)).
84 In high elevation temperate regions, pollination interactions vary temporally and are characterized by high
85 turnover in active periods of species (CaraDonna *et al.* (2017)), however the overall shorter extent of the
86 active growing season in these systems results in the presence of few to any natural breaks, which reduces
87 the utility of these to operate as filters in the post-processing of sequence matches. Nonetheless, we work
88 develop a general approach which seems applicable to many areas which utilize the temporal dimension for
89 classifying sequences in metagenomic samples (but see Davis *et al.* (2022)).

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

¹⁰⁸ **2 | METHODS**

¹⁰⁹ **Study System & Field Work**

¹¹⁰ Observations and sample collection was conducted at The Rocky Mountain Biological Laboratory (RMBL;
¹¹¹ 38°57.5" N, 106°59.3" W (WGS 84), 2900 m.a.s.l.), Colorado, USA (APPENDIX 1 for site information). Pol-
¹¹² linator observations of *Bombus* Latreille spp. (Apidae Latreille) were conducted from June - August of 2015
¹¹³ in six study sites characterized by high-montane/subalpine Parkland vegetation communities. Observations
¹¹⁴ of *Bombus* foraging took place for one hour at each field site in three 100m transects, where all flowers were
¹¹⁵ also counted and placed into abundance bins. Corbiculae loads were, non-lethally, collected once from all
¹¹⁶ Queen individuals encountered.

¹¹⁷ **2.1 | Spatial Analyses**

¹¹⁸ **2.1.1 Candidate Species** We downloaded from the Botanical Information and Ecology Network ‘BIEN’
¹¹⁹ (Maitner (2022)) all records adjacent to the field sites to develop an ecologically relevant list of vascular
¹²⁰ plant species, with expected biotic pollination, which may be present at the study area . We then generated
¹²¹ Species Distribution Models (SDMs) to predict their distribution throughout the study area. These maps of
¹²² potential distribution served as a reference to reduce the list of species to include in the genomic sequence
¹²³ databases.

¹²⁴ In order to minimise the number of species for which SDM’s were to be generated, BIEN was queried at a
¹²⁵ distance of up to 100km from our study area and all plant species records were downloaded. In order to
¹²⁶ emulate the stochasticity of botanical collecting and offset the number of records associated with the research
¹²⁷ station, this data set was bootstrap re-sampled 250 times, with 90% of samples selected, to create a testing
¹²⁸ data set. The median of the logistic regression assessing the probability of occurrence of a species record as
¹²⁹ a function of distance from the study area was used as a threshold distance, under which, to include species
¹³⁰ as candidates for distribution modelling.

¹³¹ **2.1.2 Distribution Modelling** We used all occurrence records from BIEN ($n = 23,919$) within a 50km
¹³² border of the Omernik level 3 ecoregion, which includes the study area (*No. 21 “Southern Rockies”*) to
¹³³ construct the species distribution model (Omernik (1987)). These records were copied into two, initially
¹³⁴ identical, sets, one for generating machine learning models (ML; Random Forest, and Boosted Regression
¹³⁵ Tree’s), and the other for Generalised Linear (GLM) and Generalized Additive Models (GAM) (Barbet-

136 Massin *et al.* (2012)). Ensembled predictions have been shown to outperform their constituent
137 models, on average, and to reduce the ecological signal to the analytical noise of individual runs
138 (Araujo & New (2007)). No single method of producing SDMs has been shown to universally
139 outperform others when faced with a large and diverse number of applications, in our case a
140 great number of species with differing biologies and ecologies (Elith* *et al.* (2006), Qiao *et al.*
141 (2015)). In the spirit of these findings, multiple families of models, which can be generated
142 together as they have similar requirements regarding the number and ratios of Presence to
143 Absence records were ensembled together (Barbet-Massin *et al.* (2012)).

144 We then generated 4,029 absence points, locations where the focal taxon is anticipated missing, through a
145 random stratification of 19% of the land cover in the area and included them in (Land Management (2019)).
146 To achieve a larger absence data set, we generated 1,000 pseudo-absence records for each taxon by randomly
147 selecting coordinates located at least 10km away from an occurrence record. For ML models, these pseudo-
148 absences were reduced so that the ratio of presence to absence records were balanced (Barbet-Massin *et al.*
149 (2012)). To achieve this, we removed absence records inside of 10% of the mean sample value of any predictor
150 variable the presence records; the required number of absence records were then randomly sampled.

151 To predict the potential distribution of each species we used 26 environmental variables at 30m resolution,
152 six related to climate, five soil, four topographic, four related to cloud cover, with the remaining reflecting
153 assorted abiotic parameters (Wilson & Jetz (2016), Wang *et al.* (2016), Hengl *et al.* (2017), Robinson *et al.*
154 (2014)) (**APPENDIX 6**). These publicly available data sets, were selected as they . . . For linear
155 regression models these predictors underwent both *vifstep* (theta = 10, max observations = 12,500) and
156 *vifcor* (theta = 0.7, max observations = 12,500) to detect highly correlated variables, and collinear features
157 were removed leaving 16 variables (Naimi *et al.* (2014)).

158 Modelling: Random Forest and Boosted Regression Trees, were sub sampled with 30% test and two replicates
159 each before weighted ensemble based on True Skill Statistics (tss) (Naimi & Araujo (2016)). Generalised
160 linear models (GLM) and Generalised additive models (GAM) with 30% sub sampling and three replicates
161 each were also ensembled using the tss (Naimi & Araujo (2016), @). TSS was chosen as the ensemble
162 criterion as it has been shown to work across a wide range of species occurrences prevalence (Allouche *et*
163 *al.* (2006)). The results of these models were extracted on a cell-by-cell basis to a polygon feature derived
164 from a minimum-spanning tree which encompasses the study sites, and species from either ensemble with
165 greater than 50% mean habitat suitability across all cells were considered present for further purposes (Prim
166 (1957)).

167 535 species were modelled using Generalized Linear Models and Generalized Additive Models. 534 species

were modelled using Random Forest and Boosted Regression Trees. To evaluate the accuracy of the species distribution models, additional presence records from GBIF ($n = 61,789$), and AIM ($n = 12,730$) were used as test and training sets ($n = 74,519$) for logistic regression (Occdownload Gbif.Org (2021), Land Management (2019)). Additional novel absence records were generated from the AIM data set to create a data set where each species has balanced presence and absences. Eleven or more paired presence and absence records were required for this testing, resulting in 334 species being included in the logistic regression ($Mdn = 110.0$, $\bar{x} = 223.1$, max = 1568 record pairs used) with a 70% test split (Kuhn (2022)).

2.2 | Molecular Lab Work

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

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

2.2.2 | Plant Genomic DNA Extraction Plant genomic DNA was isolated from $\sim 1 \text{ cm}^2$ of leaf tissue from silica-gel dried or herbarium material using a modified cetyltrimethylammonium (CTAB) protocol (Doyle & Doyle (1987)) that included two chloroform washes. DNA was quantified using a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and Qubit fluorometer (Thermo Fisher Scientific).

2.2.3 | Pollen Genomic DNA Extraction Pollen genomic DNA was extracted from corbiculae using a CTAB based protocol modified from Lahlamgiah et al. and Guertler et al. (2014, 2014). A SDS extraction buffer (350 μL , 100mM Tris-HCl, 50 mM EDTA, 50 mM NaCl, 10% SDS v/v., pH 7.5) was added followed by

196 vortexing to allow dissolution of corbiculae. Pollen grains were then macerated with Kontes Pellet Pestles,
197 and the tip of these washed with 130 μ L of the SDS extraction buffer, samples were then incubated for
198 1 hour at 30°C. This was followed by the addition of 10% CTAB solution (450ul, of 20 mM Tris-Cl pH.
199 8.0, 1.4 M NaCl, 10 mM EDTA pH 7.5, 10% CTAB, 5% PVP, ~85% Deionized water) and RNase (10
200 μ L of 10 mg/mL) and samples were incubated for 40 minutes at 37°C, on heat block (Multi-Blok, Thermo
201 Fisher Scientific, Waltham Massachusetts) set to 40°C. After 20 minutes incubation, Proteinase K (15 μ L of
202 20mg/ml) and DTT (12.5 μ L of 1M in water) were added, and the samples were further incubated at 60°C
203 for 1 hour. Samples were then incubated overnight at 40°C. 500 μ L of Phenol-Chloroform-Isoamyl alcohol
204 (25:24:1) were added, vortexed, and centrifuged at 10,000 rpm for 10 minutes and the aqueous phase was
205 pipetted to a 1.5 ml centrifuge tube.

206 To precipitate the DNA, chilled Isopropyl alcohol & 3 mM Sodium acetate (5:1) equivalent to $\frac{2}{3}$ of the
207 volume of sample were added, with 1 hour of chilling at -20°C, followed by 10 minutes of centrifuging at
208 13,000 rpm. The supernatant was pipetted to a new 1.5 ml centrifuge tube, and 70% EtOH (400 μ L) were
209 added before chilling at -20°C for 20 minutes followed by centrifugation at 13,000 rpm for 10 minutes. Both
210 tubes were then washed with 75% EtOH (400 μ L), inverted, centrifuged at 13,000 rpm for 4 minutes, and
211 the solution discarded, then washed with 95% EtOH (400 μ L), inverted, centrifuged at 13,000 rpm for 4
212 minutes, and the solution discarded. Pellets were dried at room temperature overnight before resuspension
213 in nuclease free H₂O. Extractions were assessed using a Nanodrop 2000 (Thermo Fisher Scientific) and Qubit
214 fluorometer (Thermo Fisher Scientific). DNA extracts were then cleaned using 2:1 v./v. Sera-Mag beads
215 (Cytiva, Little Chalfont, UK) to solute ratio following the manufacturer's protocol, eluted in 0.5x TE, and
216 the eluent allowed to reduce by half volume in ambient conditions. DNA was quantified using a Qubit
217 fluorometer.

218 **2.2.4 | Fragmentation, Library Preparation & Target Enrichment** Library preparation was per-
219 formed using the NEBNext Ultra II FS-DNA Library Prep Kit for Illumina (New England BioLabs, Ipswich,
220 Massachusetts, USA) using slightly modified manufacturers recommendation. Fragmentation was performed
221 at $\frac{1}{2}$ volume of reagents and $\frac{1}{4}$ enzyme mix for 40 minutes at 37°C, with an input of 500 ng cleaned DNA.
222 Adapter Ligation and PCR enrichment were performed with $\frac{1}{2}$ volumes, while cleanup of products was
223 performed with $\frac{1}{2}$ volume of SPRI beads (Beckman Coulter, Indianapolis, Indiana, USA) and recommended
224 volumes of 80% v./v. ethanol washes. The exception was the herbarium specimens which were not frag-
225 mented and only end repaired, with similar library preparation of all samples. Products were analysed on
226 4% agarose gels, and a Qubit fluorometer. Libraries were pooled and enriched with the Angiosperms 353

227 probe kit V.4 (Arbor Biosciences myBaits Target Sequence Capture Kit) by following the manufacturer's
228 protocol and Brewer et al. 2019. Sequencing was performed using an Illumina mi-Seq with 150-bp end reads,
229 (NUSeq Core, Chicago, Illinois).

230 **2.2.5 | Computational Processes and Analyses.**

231 **2.2.5.1 | Reference Library Data Processing** Sequences were processed using Trimmomatic, which
232 removed sequence adapters, clipped the first 3 bp, discarding reads less than 36 bp, and removing reads
233 if their average PHRED score dropped beneath 20 over a window of 5 bp (Bolger & Giorgi (2014), Tange
234 (2021)). Contigs were generated using HybPiper using target files created by M353 (Johnson *et al.* (2016),
235 McLay *et al.* (2021)).

236 **2.2.5.2 | Sequence Identification** A custom Kraken2 database was created by downloading represen-
237 tative species of each genus indicated as being present in the study area by the spatial analyses from the
238 Sequence Read Archive (SRA) NCBI (Wood *et al.* (2019)). These sequences were processed in the same
239 manner as our novel sequences. The Kraken2 database was built using default parameters. Kraken2 was run
240 on sequences using default parameters (*APPENDIX 5*). Following Kraken2, Bracken was used to classify
241 sequences to terminal taxa (Lu *et al.* (2017)). Results from both Kraken2 and Bracken, results were reclas-
242 sified manually to identify terminal taxa. For example, when only a single species of a genus was known
243 in the study area, but our database used a representative of another taxon in the genus, this species was
244 coded as the result. The re-coding of sequences from another representative species for the genus to the sole
245 RMBL representative allowed the identification of *XX & %* more species.

246 **2.2.5.3 | Identification of Sequence Matching Loci** A local NCBI database was built using the same
247 processed novel and downloaded sequences as the previous database (Camacho *et al.* (2009)).

248 **2.2.5.4 | Morphological Pollen identification**

249 To develop a reference library of pollen grains which may be present in corbiculae loads, an image reference
250 collection of fuchsin-jelly stained (Beattie (1971)) slides was assembled from slides previously prepared by the
251 authors (n = 21), and other researchers (n = 38) (Brosi & Briggs (2013)). Using five years of observational
252 data on *Bombus* Queen Bee foraging at these studies sites (Ogilvie unpublished), as well as the Vascular
253 Plant Checklist (Frase & Buck (2007)), an additional 62 voucher slides for species were prepared and imaged

254 at 400x (Leica DMLB, Leica MC170 HD Camera, Leica Application Suite V. 4.13.0) from non accessioned
255 herbarium collections to supplement the number of species and clades covered (Appendix 3).
256 We used Divisive Hierarchical Clustering techniques to determine which plant taxa were distinguishable via
257 light microscopy, and to develop a dichotomous key to pollen morphotypes. Ten readily discernible categorical
258 traits were collected from each specimen in the image collection. These traits were transformed using Gower
259 distances, and clustered using Divisive Hierarchical clustering techniques (Maechler *et al.* (2022)). Using
260 the cluster dendrogram, elbow plot, and heatmaps (Hennig (2020)), of these results morphological groups
261 of pollen which could not be resolved via microscopy were delineated, and a dichotomous key was prepared
262 (APPENDIX NO.). This key was then used to identify the pollen grains sampled from corbiculae loads to
263 morphotypes in a consistent manner. To prepare the pollen slides from corbiculae, all corbiculae loads were
264 broken apart and rolled using dissection needlepoints to increase heterogeneity of samples. *Cerca* 0.5mm²
265 of pollen was placed onto a ~4mm² fuchsin jelly cube (Beattie (1971)) atop a graticulated microscope slide,
266 with 20 transects and 20 rows (400 quadrants) (EMS, Hartfield, PA). The jelly was melted, with stirring,
267 until pollen grains were homogeneously spread across the microscope slide. Slides were sealed with Canada
268 Balsam (Ruble Colours, Willits, CA) followed by sealing with nail polish to prevent oxidation; all samples
269 are noted in APPENDIX 3. To identify the pollen present in corbiculae loads, light microscopy at 400x
270 (Zeiss Axioscope A1) was used. In initial sampling in three transects, each pollen grain was identified to
271 morphotype and counted; an additional two transects were scanned for morphotypes unique to that slide,
272 if either transect contained an unique morphotype than all grains in that transect were also identified and
273 counted. Subsequent to the first round of sampling, non-parametric species richness rarefaction curves
274 (Oksanen *et al.* (2022)), and non-parametric species diversity rarefaction curves were used to assess the
275 completeness of sampling (Chao *et al.* (2014), Hsieh *et al.* (2020)). Slides not approaching the asymptote
276 of the rarefaction curve were then re-sampled, and analysed iteratively for up to a total of seven transects
277 APPENDIX 2.

278 2.3 | Temporal Analyses

279 To estimate the duration of dates in which plant species were flowering weibull estimates of several pheno-
280 logical parameters all spatially modelled taxa were developed (Belitz *et al.* (2020), Pearse *et al.* (2017)).
281 Only BIEN records which occurred in the Omernik Level 4 Ecoregions within 15km of the study area (n =
282 5 Level 4 Ecoregions, or conditionally 6 ecoregions if enough records were not found in the nearest 5),
283 and which were from herbarium records were included. To remove temporally irrelevant herbarium records,
284 i.e. material collected during times which flowering is impossible at the study area due to snow cover, we

used the SnowUS data set (Iler *et al.* (2021), Tran *et al.* (2019)) from 2000-2017 were analyzed for the first three days of contiguous snow absence, and the first three days of contiguous snow cover in Fall. Herbarium records after the 3rd quantile for melt, and the 1st quantile for snow cover of these metrics were removed. Species with > 10 records had their weibull distributions generated for the date when 10% of individuals had begun flowering, when 50% were flowering, and when 90% of individuals had flowered, we used the initiation and cessation dates, respectively, as effective start and ends of flowering.

291 2.4 | Floral Observations

292 3 | RESULTS

293 3.1 | Spatial Analyses

294 [Table 1 about here.]

295 [Table 2 about here.]

296 The median (25.009 km) of the logistic regression assessing the probability of occurrence of a species record as
297 a function of distance from the study area was used as a threshold distance to include species for distribution
298 modelling. A 2-sample test for equality of proportions with continuity correction (X-squared = 13.254, df
299 = 1, p-value = 0.000136, 95% CI 0.04-1.00) was used to test whether more of the records located in the
300 broad ecological sites present at the field station, between the distance of the median (25.009 km) to the
301 third quantile (ca 43.830 km) of the regression distance, were true presences at the field station. Including
302 these records would have resulted in modelling an additional 222 species distributions of which 30 are true
303 presences, these taxa were not modelled.

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

306 In the area of the minimum-spanning tree encompassing the field sites, of the 554 vascular plants with biotic
307 pollination syndromes, the 493 ML ensembles accurately predicted the presence of 362 (65.3%), incorrectly
308 predicted the presence of 64 (11.6%), incorrectly predicted 34 true presences (6.1%) as being absent, and
309 correctly predicted the true absence of 33 (6.0%). The balanced accuracy of the ensembled models is 0.627
310 (Sensitivity = 0.340, Specificity 0.914). Of the 554 vascular plants with biotic pollination syndromes, the
311 475 LM ensembles accurately predicted the presence of 286 (51.6%), incorrectly predicted the presence of

312 41 (14.3%), incorrectly predicted 93 true presences (16.8%) as being absent, and correctly predicted the
313 true absence of 55 (9.9%). The balanced accuracy of the ensembled models is 0.664 (Sensitivity = 0.573,
314 Specificity 0.754). Of the 554 vascular plants with biotic pollination syndromes in the flora 13 (2.3%) were
315 in the Orchid family and 41 (7.4%) are non-natives, both of which are restricted from the database, and can
316 only reduce the number of true predicted presences by roughly 10%.

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

322 3.2 | Microscopic Pollen identification

323 Using the fuchsin jelly preparation and light microscopic analyses of grains and scoring of 10 character
324 states resulted in the establishment of 28 morphotypes which grains could be classified into. APPENDIX
325 7. 60 samples were counted and based on rarefaction **had over % of expected morphotypes found**
326 (morphotype richness, $\bar{x} = 4.5$, Mdn = 4, min = 1, max = 9), all samples had expected morphotype diversity
327 reach the asymptote APPENDIX 8. The number of counted pollen grains in each sample range from (MIN
328 - 16,293, $\bar{x} = 2788.685$, Mdn = 1453).

329 [Figure 1 about here.]

330 3.3 | Metabarcoding Pollen identification

331 54 corbiculae loads had DNA extracted and underwent various steps towards hyb-seq, in the end a total of
332 44 corbiculae samples were sequenced, 7,752,353 reads were recovered from sequencing. The number of reads
333 per sequence varied widely (range = 76 - 508,795, $\bar{x} = 176,189.8$, Mdn = 138,395). Of the possible 353 loci,
334 the number which were recovered from each sample, and informative to BLAST were range = 24 - 353, $\bar{x} =$
335 305.5, Mdn = 331. The number of reads per loci from across all samples had a range of 178 - 506,653, $\bar{x} =$
336 20,688, Mdn = 12,616.

337 APPENDIX X Reads Per Loci.

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

340 reads were matched using BLAST. Based upon subjective review of the three classifiers **APPENDIX X**
341 **MOLECULAR NETWORKS - 3 DIFFERENT ONES**, BLAST was chosen as the classification
342 method which yielded the most probable results by the field ecologist, and it's values were used for all
343 subsequent analyses.

344 [Table 3 about here.]

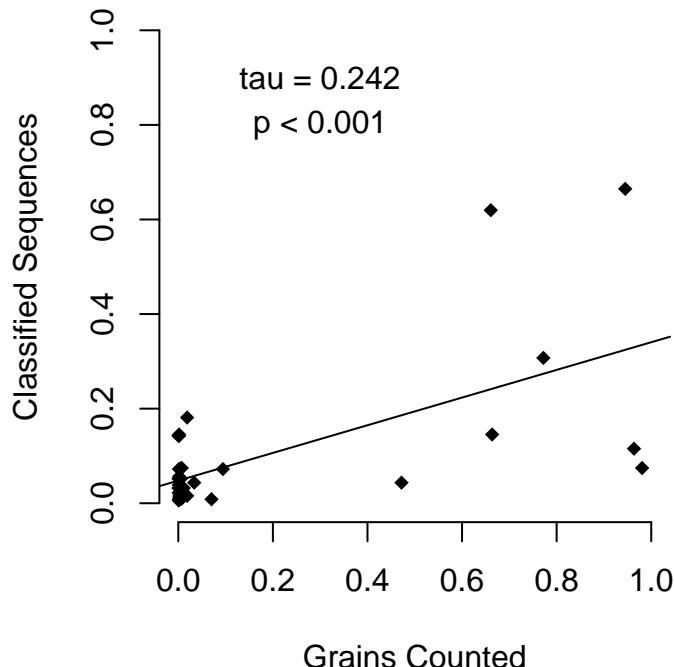
345 The initial classification of sequences which were made by BLAST were reviewed programmatically, using
346 predicted presence of the species (from spatial modelling), modelled flowering time (from temporal mod-
347 elling), and taxonomy (from existing sources). A sequential process was utilized which reassigned sequences
348 based on binary combinations of the factors above (Appendix XX). Given the relative sparsity of the num-
349 ber, and relatedness, of species represented in the sequence database this was performed to: 1) Identify
350 locally present species represented by surrogates in the DB 2) Reduce false classifications of focal species 3)
351 Identify high confidence sequence matches. Each of the top ten taxa which were identified by BLAST from
352 the aforementioned process composed 449 distinct records (CONFUSING WORDS), of these 56.35% of the
353 reads were classified to species 47.95% which represent of the total classified sequences, 35.19% of the reads
354 were classified to genus representing 47.95% of the classified reads, and 8.46% of the records were classified
355 to genus 4.61% of total sequences.

356 *Sequences which were assigned to genera without any species predicted by spatial analyses, were investigated*
357 *by hand after post-processing steps. These were manually assigned to a variety of ranks, occasionally to*
358 *genus #COUNTS# and family #COUNTS#, by consultation of the alpha-taxonomic literature (Sadeghian*
359 *et al. (2015), Sennikov & Kurtto (2017), Rabeler & Wagner (2016), Pusalkar & Singh (2015), Moore &*
360 *Bohs (2003), Weber (1998)). <- went back on this as new function captures most of it, but don't throw*
361 *out.*

362 To determine at which level species in pollen loads could be detected the results of light microscopy were
363 compared to the molecular results. The pollen samples contained three morphotypes which could readily
364 be identified via microscopy. Two of these mapped to the clades (Boraginaceae & Heliantheae Alliance),
365 and one to a Asteraceae less Heliantheae. Boraginaceae grains were detected in 64.3% of samples where the
366 proportion of target grains were between 0.01-1 ($n = 14$ Mdn = 0.572). Asteraceae type 1, non-helianthoids,
367 were detected in 20% of samples where the proportion of target grains were between 0.001-0.01 ($n = 5$ Mdn
368 = 0.002) Asteraceae type 2, Helianthoids, were detected in 75% of samples where the proportion of target
369 grains were between 0.001-0.01 ($n = 8$ Mdn = 0.003). Both morphotypes of Asteraceae pollen were detected
370 in 100% of samples where the proportion of target grains were between 0.01-1 ($n = 3$ Mdn = 0.011), and

³⁷¹ Ericaceae were detected in 50% of samples where the proportion of target grains were between 0.001-0.1 (n
³⁷² = 2 Mdn = 0.01).

Correlation of Proportion Counted Grains and Sequence Reads



³⁷³ To detect whether the sequencing reads were semi-quantitative the subset of all pollen morphotypes distinguishable by microscopy were compared to the sequence reads. In all instances sequence reads were pooled to the highest taxonomic rank associated with the morphotype, e.g. if both species of *Mertensia*, or one species and read only classified to genus were present in a sample, the reads were summed. The total percentage of the ten most abundant grains per sample were then relativized to constitute the entire sample.

³⁷⁴ The relationship between the number of pollen grains in a sample and the number of sequence reads is roughly ³⁸⁰ curvilinear, where grains which are present in trace amounts are overestimated by sequence counts, while ³⁸¹ grains present in high amounts are underestimated. This is likely due to the proportion of high false positives ³⁸² which occur in the classification process with NGS (BELL NOVEMBER 2021). There was strong evidence ³⁸³ of a strong correlation between the proportion of grains per morphotype and the number of sequences per group (0.242, p < 0.0001, n = 37).

385 **3.4 | Temporal Analyses**

386 The first date of modeled snow melt in the Gothic area ($n = 17$, $\bar{x} = 137.9$, Mdn = 135, 3rd quantile =
387 151), and the first date of a consistent winter snow base ($n = 17$, $\bar{x} = 299.9$, Mdn = 300, 1st quantile =
388 291) from 2000-2017, were used as delimiters for the inclusions of herbarium records in modelling. Of the
389 439 species predicted likely present in the area via logistic regression, 332 species (64.4%) with more than 10
390 records in the focal level 4 ecoregions ($\bar{x} = 35.016$, Mdn = 35, max = 96) had weibull estimates calculated,
391 an additional 56 species (11.2%) with enough contributing records from the “Sedimentary Mid-Elevation
392 Forests”, a large ecoregion generally just beneath the elevation bands occupied by the five ecoregions around
393 the study area had weibull estimates also calculated ($\bar{x} = 13.868$, Mdn = 13, max = 24).

394 Only 58 of these 388 species ($n = 34.568$, Mdn = 31) were able to be compared to plot based observational
395 data from the long term (1974–2012) data set (CITE). Of these species relatively high accord was observed
396 between the long-term ground truthed data set, and the modelled species. There was very strong evidence
397 that the weibull estimates were positively associated with the observed onset ($r^2 = 0.72$, $p < 0.0001$, tau =
398 0.61) and peak ($r^2 = 0.70$, $p < 0.0001$, tau = 0.65) of flowering, and that the number of herbarium samples
399 had a moderate effect on the estimates ($p = 0.004$ and $p = 0.034$ respectively). There was very strong
400 evidence that the weibull estimates had a positive association with the observed cessation of flowering ($r^2 =$
401 0.4339, $p < 0.0001$, tau = 0.489), however there was no evidence that sample size had an effect ($p = 0.349$).
402 There was moderate evidence that the weibull estimates, with an effect of sample size, had a weak positive
403 association with the observed duration of flowering ($p = 0.0401$, $r^2 = 0.07$, tau = 0.17).

404 [Figure 2 about here.]

405 **3.5 | Floral Observations**

406 The six sites were surveyed once weekly from May 27-July 27 for a total of 52 hours from . A total of 723
407 queen-pollen foraging interactions were observed (range per bee species by week range = 1-18, $\bar{x} = 3.46$,
408 Mdn = 2), with a range of total observed interactions per bee species across this time period (min = 1, \bar{x}
409 = 59.08, Mdn = 19, max = 184). Plants varied widely in the number of interactions which they partook in
410 with each species of bee (range per plant species by week min = 1 - 20, $\bar{x} = 3.51$, Mdn = 2), with a range
411 of total observed interactions per plant species over this time period (min = 1, $\bar{x} = 20.26$, Mdn = 4, max =
412 141). The number of plant species which bees were observed interacting with varied more narrowly (range
413 = 1 - 18, $\bar{x} = 8$, Mdn = 6).

⁴¹⁴ A total of 66 corbiculae loads were collected from Bees, 64 of them from Queens.

⁴¹⁵ [Figure 3 about here.]

⁴¹⁶ 3.6 | Integrated Observational, Molecular, and Palynological Network

⁴¹⁷ While the spatial results were used to declare the taxonomic composition of the sequence database, temporal
⁴¹⁸ results were used in consideration with plant phylogeny to retroactively, reassign the assignment of sequences
⁴¹⁹ to taxa. Essentially, if a sequence was identified to a taxon which was not known from the field site

⁴²⁰ For example a common UNKNOWN sequence mapped to the Asteraceae family, but which was flagged
⁴²¹ by temporal filters and is present in both *B. nevadensis* Cresson and *B. rufocinctus* Cresson pollen is
⁴²² most likely *Frasera* Walter, failed extraction. A similar likely mismatch could be between what was fide
⁴²³ molecular evidence as *Agastache pallidiflora* (A. Heller) Rydb. but where feeding was infrequently observed
⁴²⁴ on *Pedicularis* L., likely due to this entire order being represented by only a single molecular reference species.

⁴²⁵ It is not unlikely that much of the difference in the results between the observational and molecular work
⁴²⁶ are attributable to the challenges in detecting rare events in these smaller sizes. For example, no more than
⁴²⁷ 10 bee corbiculae loads per species were sequenced with the Mdn = 7 . . . , and the median of interactions
⁴²⁸ with the top 5 plant sizes constituted 0.8142857 of the top.

⁴²⁹ . . . many of our results indicate foraging on *Viola* L. spp, zygomorphic flowers with architecture which would
⁴³⁰ require subtle handling and strength to reach the pollen and nectar loads. . . . Or the *Epilobium* L. spp.
⁴³¹ results likely indicating that a species of *Chamerion* Seg. such as *C. angustifolium* (L.) Scop. or *latifolium*
⁴³² (L.) Sweet is occasionally utilized, as it supported by limited palynology data.

⁴³³ 4 | DISCUSSION

⁴³⁴ ~ What we DEMONSTRATED ~ We have demonstrated how the Angiosperms533 hyb-seq probes may
⁴³⁵ be used for plant barcoding in a metagenomic context (Johnson *et al.* (2019), Hollingsworth *et al.* (2016)).
⁴³⁶ This was exemplified in an ecologically relevant scenario, where the results have immediate implications
⁴³⁷ for natural history driven fundamental science and land management. The test pollen loads contained
⁴³⁸ a number of closely related taxa, some in notoriously morphologically difficult clades with rapid rates of
⁴³⁹ diversification (e.g. *Mertensia*, *Lupinus* L.), at naturally occurring proportions (Nevado *et al.* (2016), Nazaire
⁴⁴⁰ & Hufford (2014)). We incorporated spatial and temporal approaches for creating custom sequence databases

441 an approach which is readily applicable to any lab group with the capacity to perform next-generation
442 sequencing across the entirety of multiple continents, and which we expect to be highly beneficial in many
443 study areas. By combining insights from these novel approaches with an extensive observational field based
444 study we show how these methods may be applied to test a variety of hypotheses related to ecological
445 interactions.

446 ~ **What CHALLENGES we FACED** ~ We anticipate that many of the complications which we faced,
447 using opportunistically collected pollen loads and the first implementation of this method may readily be
448 overcome. It seems apparent that we had issues detecting pollen from several genera of plants, based upon
449 these and other observational studies most likely *Vicia* L., *Lathyrus* L., and *Frasera* (Inouye (1980), Pleasants
450 (1980)); this is most likely related to user error in obtaining high quality DNA during the plant reference
451 library generation period. (**REED SHOULD HUNT FOR LOCI RETURNS FROM BLAST LIKE**
452 **HE DID POLLEN**). Additional complications seem to relate to the presence of closely related false
453 positives, e.g. frequent classifications of sequences as *Trollius* L., *Caltha* L., and *Thalictrum* L. alongside a
454 more common species in the family, e.g. *Delphinium*. L. Many of our errors are known to us and multiple
455 mnemonics are in APPENDIX XX to assist others in future attempts to achieve better results. However, the
456 line between false and negative positives may be blurred in some of these instances and warrant further work,
457 for example **Ericaceae pollen grains were observed in a number of samples in trace quantities....**
458 We used only a relatively few amount of species records for both the spatial and temporal modelling, we
459 feel this was an appropriate measure to show that these methods would work with data sets which were
460 readily available without considerable data ‘scraping’. We anticipate that the number of records available
461 to all researchers have grown rapidly since the inception of our modelling (2019) to date, largely as a result
462 of herbarium digitization efforts, and that speak of a massive ‘one stop shop’ of an integrated database is
463 just over the horizon (Feng *et al.* (2022), Soltis (2017)). The SDM’s which we generated, with relatively few
464 occurrence records and few modelling iterations, performed beyond expectations, likely due to the utility
465 of the predictor variables and strong alignment of vegetation by orographic precipitation in the study area.
466 However, we had difficulties in evaluating our predictions in an operational context. We utilized the database
467 query approach, to only model species with a high probability of not being dispersal limited to the focal area,
468 and focused on a relevant subset of many of these species ranges to reduce the contributions of range wide
469 adaptions on habitat (Sork (2018), Joshi *et al.* (2001)). While the models worked well compared to both
470 test, and validation with external point data, moving from points to polygon features was more difficult. We
471 were able to compare our results to 1) a Flora, 2) lists of plants used by Bumble Bees at plots; the former
472 inappropriate in that it contained a great number of species which we sought to use modelling to reduce *e.g.*

473 all strictly alpine species, and the latter inappropriate in that it contained only species relevant to *Bombus*
474 but had no official ‘absence’ data. Further given the size of the minimum spanning tree (AREA???) which
475 we extracted points to, a formal floristic inventory would still be a time intensive process. Accordingly,
476 we expect the real results of our data lay somewhere in between these two evaluations; with an excess of
477 species predicted present (Dubuis *et al.* (2011), Calabrese *et al.* (2014), Pinto-Ledezma & Cavender-Bares
478 (2021)), but few enough that they lend themselves to metabarcoding. We observe that our models seemed
479 very capable of effectively identifying alpine species and removing them in binomial contexts. Difficulties in
480 temporal models related to variability in drivers of flowering phenology.

481 ~ **WHAT we learned about Bee foraging (BASIC)** ~ These results show that the overall results
482 between **Bumble Bee ecology** observational and barcoding are largely congruent. But that ... We
483 analyzed pollen loads from all of the most common bumble bee species in the area (Pyke (1982)) Future
484 analyses of the long term data set...

485 ~ **WHAT this tells us about Bee foraging (APPLIED)** ~ Some foraging preferences of *Bombus*,
486 both at this field site and across a great many localities globally emerge from this work, which reiterates the
487 needs for land managers to maintain relatively high amounts of members of the Fabaceae, Boraginaceae, and
488 Ranunculaceae, in Western North American montane landscapes (Goulson *et al.* (2005), Goulson (2010),
489 Liang *et al.* (2021), Bontsutsnaja *et al.* (2021)). Numerous historic, and some ongoing, land management
490 practices reduce the ability of many landscapes to support stable populations of *Bombus*. Historic livestock
491 grazing was often associated with the targeted removal of many species of plants which are known to have
492 compounds toxic to cattle. In particular, the removal of locoweeds (Fabaceae: *Astragalus* L. & *Oxytropis*
493 DC.) and larkspurs (Ranunculaceae: *Delphinium*) were common across public lands administered by the
494 United States Forest Service (Ralphs & Ueckert (1988), Aldous (1919), Ralphs *et al.* (2003)). Further
495 actions, generally initiated by early settlers, involved the channelization and incising of streams, culling of
496 beavers, and leaving cattle concentrated on higher order stream banks for significant periods of time, all
497 processes which lower the water tables and reduced the extent of stream-associated [riverine] wetlands and
498 the mesic meadows fringes which provide habitat for many species of tall *Mertensia* (Boraginaceae, e.g. *M.*
499 *ciliata* Torr. G. Don.) widely distributed across Western North America, and to an extent *Delphinium*
500 *barbeyi* (Huth) Huth and many species of native *Trifolium* L. (Dahl (1990), Naiman *et al.* (1988), Belsky
501 *et al.* (1999), Cooke & Reeves (1976)). Fire suppression further resulted in the succession of many Aspen
502 (*Populus tremuloides* Michx.) groves to Conifer stands, decreasing the mosaic of age structured habitats
503 in many landscapes, adversely effects habitat for tall *Mertensia* species and several species of *Delphinium*

504 (Brewen *et al.* (2021), Keane (2002)). Finally the effects of Nitrogen deposition, especially given the West's
505 rapidly growing population still pose adverse effects on the abundance of a variety of species of Fabaceae
506 at Urban-Rural interfaces (see Stevens *et al.* (2018), Fenn *et al.* (2003)). Current solutions to these issues,
507 involve targeted burns, reintroduction of beavers and beaver habitat analogs, and the possibility of re-seeding
508 a variety of 'locoweeds' and 'larkspurs' in areas now seldom used, or only used for early, grazing. The highly
509 enthusiastic response of land managers, and homeowners, to plant *Asclepias* L., using genetically appropriate
510 materials, to improve Monarch Butterfly (*Danaus plexippus* L.) habitat provides an effective framework for
511 the latter (Oberhauser *et al.* (2015), Basey *et al.* (2015)).

512 ~ **WHERE we see spatial/temporal going** We have concerns regarding the number of persons training
513 to become and practice botany, and grave concerns regarding the funding mechanisms for floristic and field
514 based botanical research and for centralized authorities to produce consensus opinions on alpha taxonomy
515 (Prather *et al.* (2004b), Kramer & Havens (2015), Prather *et al.* (2004a), Crisci *et al.* (2020), Manzano
516 (2021), Stroud *et al.* (2022)). To reduce the effects of a low population density of botanists on the mainte-
517 nance of and production of flora's and to foster meta-genomics across landscapes without field stations we
518 utilized Species Distribution Modelling to generate predictive species lists. In this proof of concept example
519 we performed several iterations of modelling runs, and several approaches (i.e. the 'linear models', and the
520 'machine learning'), which took notable amounts of compute power. We suspect the possible deleterious
521 nature of this endeavor may be reduced by: 1) more field surveying by crews will reduce the need to generate
522 as many species 2) fewer runs of models, 3) only running machine learning models which do not require an
523 explicitly process to reduce spatial autocorrelation. However, given the time required to perform all aspects
524 of a study, even our amount of computation was negligible. Further, we are very optimistic about the pos-
525 sibility for persons to perform these tasks, as mentioned we utilized roughly only one quarter of the records
526 which were digitally available for presence, and we suspect others will have enough records to perform this
527 process nearly anywhere else in the temperate. In certain scenarios modelling of predicted species via more
528 formally tailored S(tacked)-SDM or J(oint)-SDM approaches may be beneficial (Wilkinson *et al.* (2021),
529 Pinto-Ledezma & Cavender-Bares (2021), Schmitt *et al.* (2017)).

530 Tandem to the lack of continued expertise required to generate and maintain species lists, is the expertise
531 required to continue tracking when major phenological events occur in many plant species at relatively fine
532 scales or under novel climates. Knowledge of these events is currently limited to general time periods of only
533 a handful of phenological events and groups of organisms (e.g. flowering initiation, or trees) (Prather *et al.*
534 (2004a), Li *et al.* (2016)). While many programs and initiatives exist to collect phenological information on

535 subsets of easily identifiable charismatic species to detect major trends in phenology, these capture only a
536 subset of the extent diversity (Betancourt *et al.* (2005), Havens *et al.* (2007)). In many instances it appears
537 that while landscapes respond similarly to environmental variables which predict phenological responses,
538 that individual species vary widely in their responses to similar environmental cues, or respond to different
539 cues (Augspurger & Zaya (2020), Xie *et al.* (2015), Xie *et al.* (2018), CaraDonna *et al.* (2014)). **As can**
540 **be seen here, predictions of when a single, major phenological event occurs is already data**
541 **limited, with sample size having an effect on the subset of species which we could even generate**
542 **weibull estimates for. ; check assumptions of model again and just do a quasi something A**
543 more promising approach for the tropics may lay in circular statistics (Park *et al.* (2022)).

544 ~ **WHERE we see MOLECULAR going** The nearly complete Plant and Fungal Tree of Life (PAFTOL)
545 will provide a comprehensive phylogenetic backbone of the entire plant kingdom, and the inclusion of
546 A353 probes with lineage specific probe sets is common in producing massive genetic datasets (Baker *et*
547 *al.* (2021b)). We predict that the A353 probes which it is utilizing to work nearly immediately for DNA
548 barcoding of whole plant material, and that more elaborate validation studies in controlled metabarcoding
549 settings, utilizing existing experimental designs, will have favorable results (Bell *et al.* (2017), Bell *et al.*
550 (2019), Bell *et al.* (2021), Lamb *et al.* (2019)). In particular the harvesting of loci with more variation
551 in certain lineages, and or with more variable flanking regions, will prove promising for identifying closely
552 related plant material (CITE). We suspect that conserved reaches of genes resulted in the high amounts
553 of reads in somewhat obscure species. Given that the A353 loci are nuclear, single copy, and a variety are
554 present the possibility of identifying target loci for quantitative purposes is high, without continual PCR
555 enrichment is possible; this would align with relatively high efficacy of WGS (Lang *et al.* (2019), Peel *et al.*
556 (2019), Bell *et al.* (2021)). Recent evidence indicates that the potential for identifying nearly cryptic taxa
557 and even infra-specific inference, of either whole plant material, and perhaps in metagenomic context are
558 possible (Ottenlips *et al.* (2021), Wenzell *et al.* (2021), Loke *et al.* in prep, Slimp *et al.* (2021), Beck *et al.*
559 (2021)). We further believe that in synthetic phylogenetic trees - with incorporation of NGS backbones - will
560 allow in automatic reassignment of reads as a function of phylogenetic distance with measures of uncertainty
561 (Hinchliff *et al.* (2015), Smith & Brown (2018), Baker *et al.* (2021a)).

562 **5 | CONCLUSION**

563 We believe that the combination of spatial and temporal models, united and guided by localized natural
564 history knowledge, provides the essential components of a bayesian framework for approaching the coarse
565 elucidation of ecological interactions using DNA Barcoding. Herein we crudely utilized this thinking via
566 binary outcomes, should a species predicted be predicted present or not? Is it unequivocally flowering
567 or not? Myriad data show biological systems and ecological interactions have more variance than can be
568 reasonably discretely parsed. We expect that within a bayesian framework studies of pollinator behavior
569 may be enacted via this approach at a landscape level, e.g. the scale of an entire drainage basin such as the
570 Gunnison which is quickly becoming one of the worlds few model ecosystems. We hope that the promise of
571 A353 probes as tools for metabarcoding play a role in these endeavors.

572 **AUTHOR CONTRIBUTIONS:** R.C.B conducted botanical collections, conducted all molecular lab
573 work, lead all analyses, and writing. J.E.O conceived, designed, and conducted all ecological fieldwork,
574 assisted with analyses, and writing. E.J.W. prepared, imaged, and collected trait data on pollen reference
575 slides, and assisted with analysis of trait data and writing a dichotomous key. S.T. assisted with spatial
576 analyses and writing. P.J.C assisted with ecological analyses and writing. J.B.F. conceived, and designed all
577 lab work, analyses, and integration of approaches, assisted with writing, and secured funding for molecular
578 work.

579 **ACKNOWLEDGMENTS:** Nyree Zerega for assistance obtaining herbaria loans and accessioning our
580 collections at CHIC. Pat Herendeen for assistance with virtually all aspects of preparing pollen vouchers
581 and the identification process. Ian Breckheimer for sharing the SDM predictor variables. Hilary Noble, Zoe
582 Diaz-Martinez, Angela McDonnell, & Elena Loke for assistance with genomic library preparation. We thank
583 the curators at the following herbaria for supplying tissue: Ben Legler at Stillinger (ID), Charles (Rick)
584 Williams at Ray J. Davis (IDS), (B)Ernie Nelson at Rocky Mountain (RM); and the collectors: D. Knoke,
585 L. Brummer, J. Boyd, C. Davidson, I. Gilman, M. Kirkpatrick, S. McCauley, J. Smith, K. Taylor, & C.
586 Williams. David Giblin & Mare Nazaire for sharing relevant sections of an advanced draft of FNA V. 15.
587 The Bureau of Land Management is thanked as many plant specimens were collected by R.C.B as a partner
588 or contractor to the agency; Sarah Burnett and Lauren Price are thanked for sharing AIM data. Sanda and
589 New England Biotech are gratefully acknowledged for technical support and generously sharing samples.
590 T.C.H. Cole for sharing the Angiosperm Phylogeny 4 colour palette. The Program in Plant Biology and
591 Conservation is thanked for funding.

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

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

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

596 **ORCID**

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

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

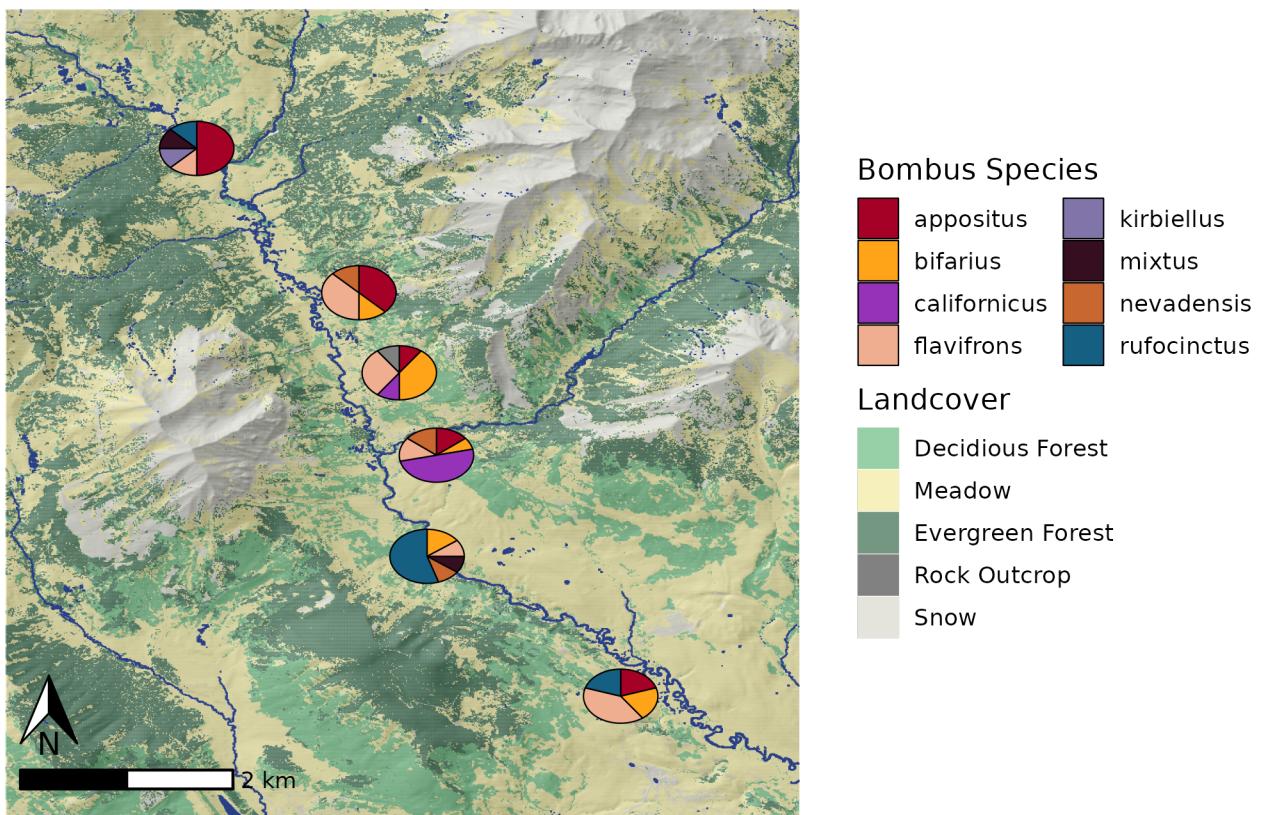
599 Jane Ogilvie <https://orcid.org/0000-0001-8546-0417>

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

601 **References**

602 **Supporting**

Origins of Corbiculae Loads



Upper East River Valley, Colorado

605 Appendix 2 - Species Distribution Models Predictors

Layer	LM	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

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 foliacium</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>Lupinus bakeri</i> Greene	Fabaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Vicia americana</i> Muhl. ex Willd.	Fabaceae	CHIC tba	S	Montana, Carbon	4.VII.2019	tba	10020.8
<i>Vicia americana</i> Muhl. ex Willd. var. minor Hook.	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>Frasera speciosa</i> Douglas ex Griseb	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 capitatum</i> Douglas ex. Benth	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>Hydrophyllum fendleri</i> (Gray) Heller	Hydrophyllaceae	CHIC tba	S	Arizona, Coconino	17.VII.2020	tba	617.7
<i>Agastache pallidiflora</i> (Heller) Rydberg	Lamiaceae	CHIC tba	S	Arizona, Coconino	17.VII.2020	tba	617.7
<i>Chamerion angustifolium</i> (L.) Holub	Onagraceae	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	982.5
<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

609 Appendix 4 - 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 lemnmonii</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 multiflora</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 stenomeria</i> (A. Gray) Greene	Asteraceae	ID, Idaho, Marshall Mountains	tba	Novo	E.J.W.	2021
<i>Senecio hydrophilus</i> Nutt.	Asteraceae	ID, Custer, E. fl. Salmon River	tba	Novo	E.J.W.	2021
<i>Senecio integrerrimus</i> Nutt.	Asteraceae	CO, Gunnison, Gothic	None	Physical	P.J.C.	2014
<i>Senecio serrula</i> Hook.	Asteraceae	CO, Gunnison, Washington Gulch	tba	Novo	E.J.W.	2021
<i>Senecio wootonii</i> Greene	Asteraceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Solidago lepida</i> DC.	Asteraceae	ID, Idaho, American River	tba	Novo	E.J.W.	2021
<i>Symphotrichum foliacum</i> (Lindl. ex DC.) G.L. Nesom	Asteraceae	CO, Gunnison, Gothic	None	Image	H.M.B.	2010
<i>Symphotrichum 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.

613 POLLEN CLUSTER RESULTS SHOULD BE HERE

- 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)

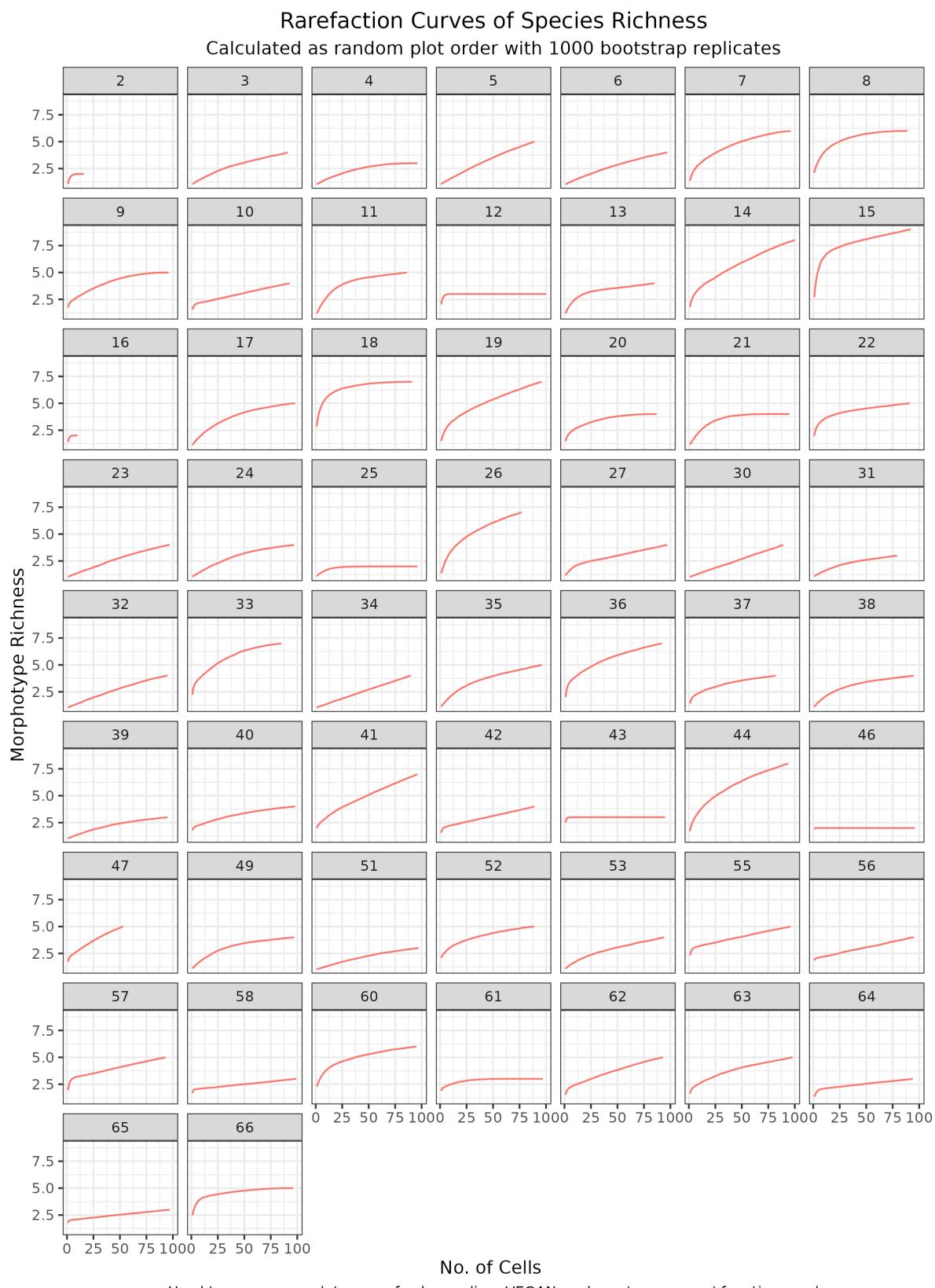
34b: Apertures not annulate – grains appear more or less circular (~ERICACEAE)

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



Used to assess completeness of subsampling. VEGAN package 'specaccum' function used

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

Confidence Interval of 99% with 1000 Bootstrap replicates

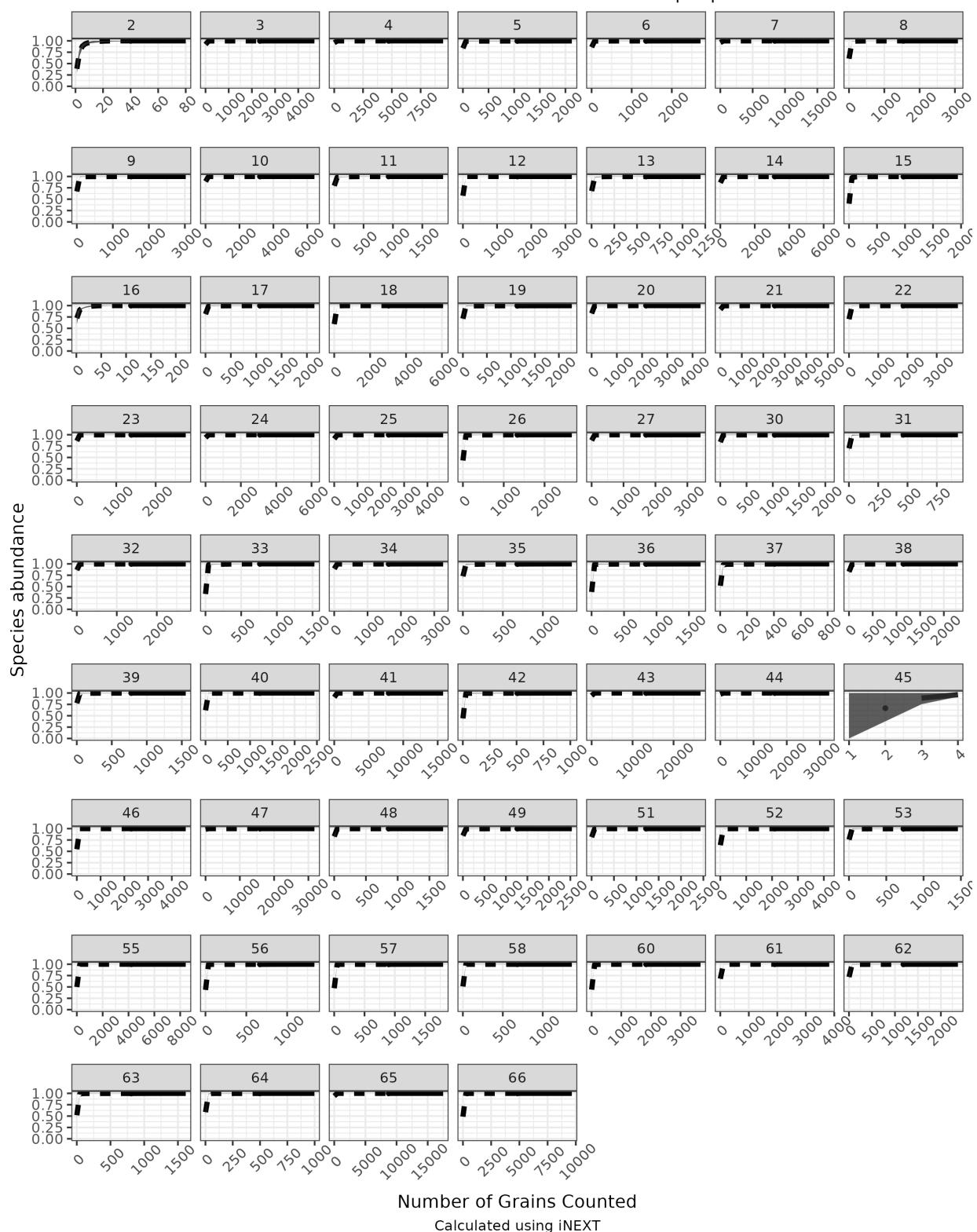


Table 1: All species present in the Reference Sequence Databases
(Kraken and BLAST)

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
		<i>Solidago chilensis</i>
		<i>Stilpnolepis intricata</i>
		<i>Symphyotrichum foliaceum</i>
		<i>Taraxacum cucullatum</i>
		<i>Taraxacum officinale</i>
		<i>Tonestus lyallii</i>
		<i>Townsendia formosa</i>
	Campanulaceae	<i>Campanula argaea</i>
		<i>Campanula rotundifolia</i>
Boraginales	Boraginaceae	<i>Cynoglossum amplifolium</i>
		<i>Cynoglossum anchusoides</i>
		<i>Cynoglossum pringlei</i>
		<i>Mertensia ciliata</i>
		<i>Mertensia fusiformis</i>
	Hydrophyllaceae	<i>Hydrophyllum canadense</i>
		<i>Hydrophyllum capitatum</i>
		<i>Hydrophyllum fendleri</i>
		<i>Nemophila menziesii</i>
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>
	Polygonaceae	<i>Rumex induratus</i>
		<i>Rumex spinosus</i>
Celastrales	Celastraceae	<i>Parnassia faberi</i>
		<i>Parnassia palustris</i>
		<i>Paxistima canbyi</i>
Ericales	Ericaceae	<i>Gaultheria prostrata</i>
		<i>Moneses uniflora</i>
		<i>Orthilia secunda</i>
		<i>Vaccinium vitis-idaea</i>
	Polemoniaceae	<i>Collomia grandiflora</i>
		<i>Ipomopsis aggregata</i>
		<i>Phlox douglasii</i>
	Primulaceae	<i>Androsace studiosorum</i>
		<i>Androsace vitaliana</i>
Fabales	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)

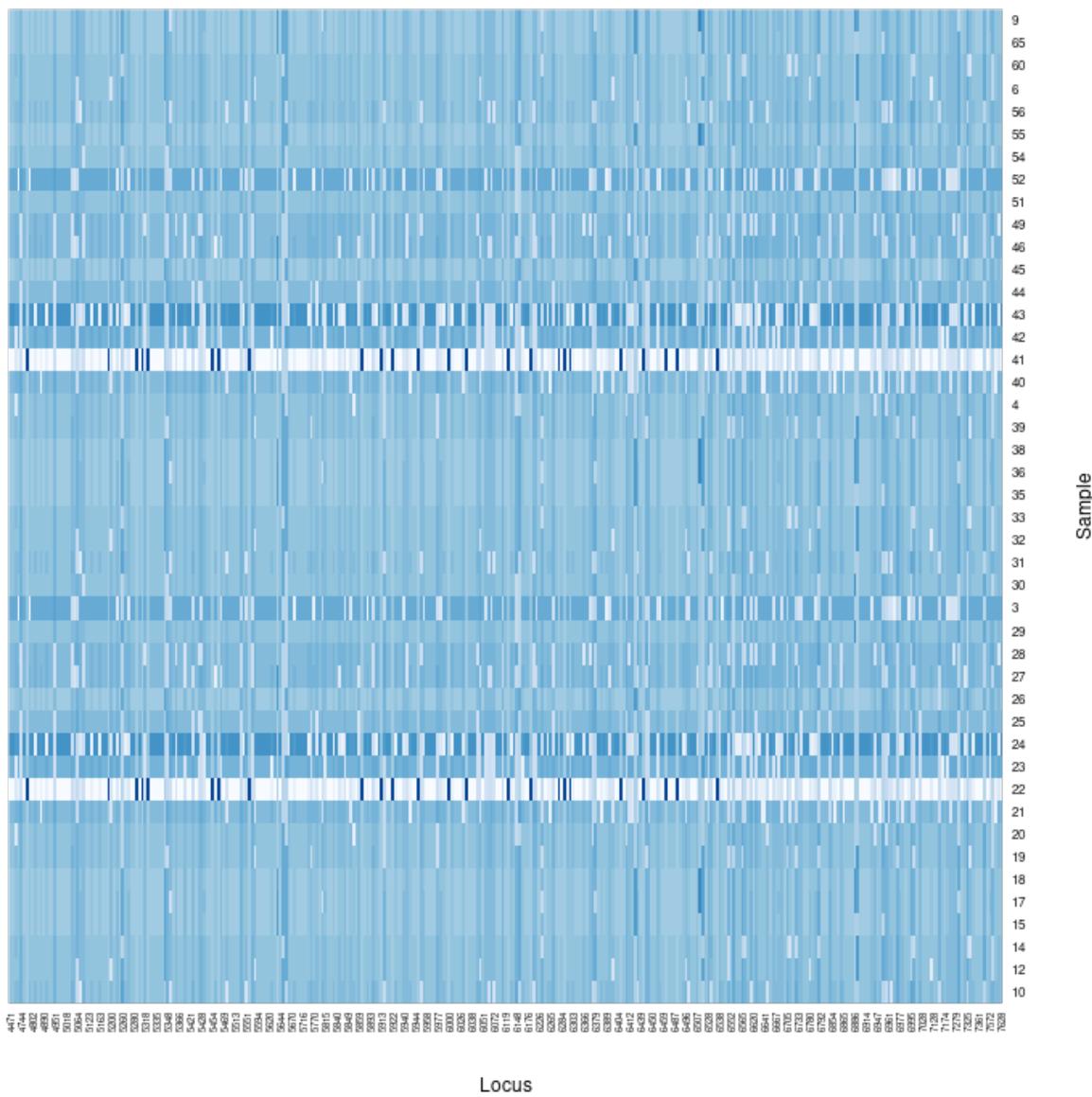
Table 1: All species present in the Reference Sequence Databases
(Kraken and BLAST) (*continued*)

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

x geum* \end{longtable}

628 Appendix XX - Reads Per Loci

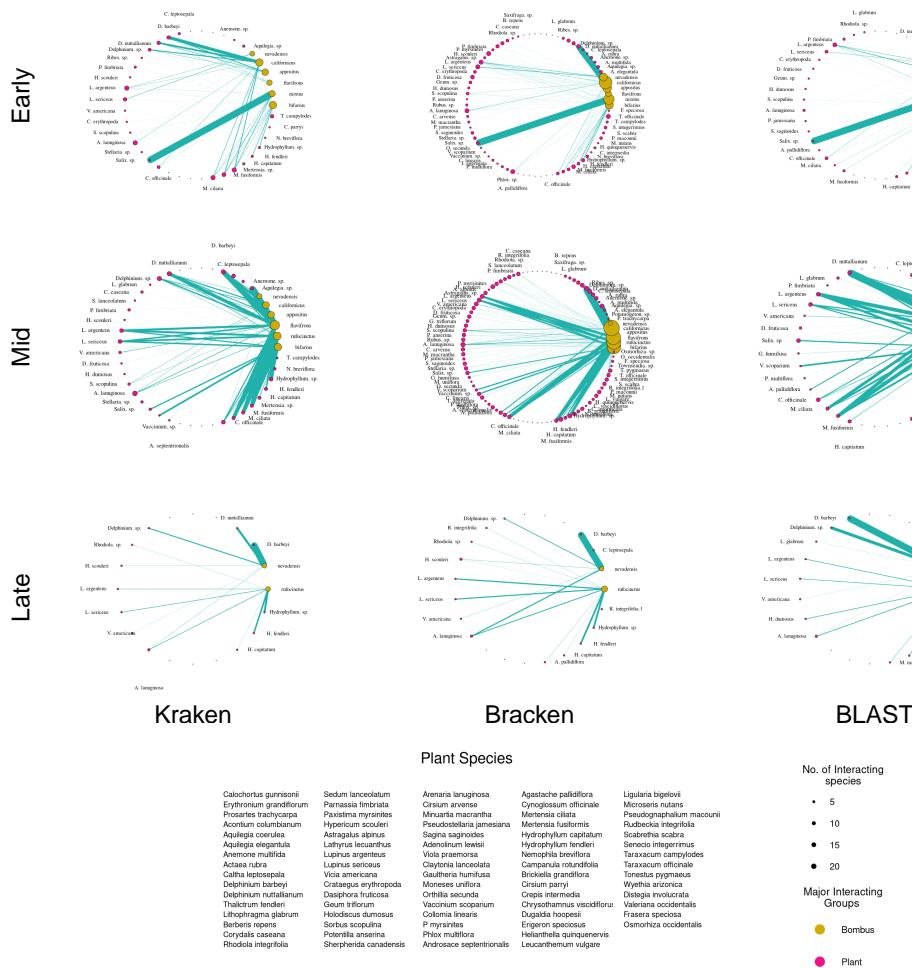
Percent matched reads per locus by sample



629

630 Appendix XX - Comparison of Kraken2, Bracken, and BLAST

Comparision of Foraging Patterns from Three Sequence Alignment Algorithms



632 Appendix XX - Models used for Species Distribution Model Ensembles

633 *Generalised Linear Models (GLM)*

634 *Generalised Additive Models (GAM)*

635 The two machine learning models utilize Ensemble learning.

636 Decision trees, ...

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

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

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

652 Random Forest have high bias and low variance, where boosted regressions trees have low bias and high
653 variances.

Table 1: Subset of Possible Combinations for re-classifying Sequences by Incorporating Ecological Factors

Spatial	Temporal	Congener	Confamilial	Congeners	Confamilials	Condition	Return	Rank
1	1	1	1	0	0	A.1	Input	Species
1	1	1	1	1	0	A.2	Input	Species
1	1	1	1	0	1	A.3	Input	Species
1	1	1	1	1	1	A.4	Input	Species
1	1	1	0	0	0	A.5	Input	Species
1	1	1	0	1	0	A.6	Input	Species
1	1	0	1	0	0	A.7	Input	Species
1	1	0	1	0	1	A.8	Input	Species
1	1	0	0	0	0	A.9	Input	Species
1	0	1	1	0	0	B.1	Congener	Species
1	0	1	1	0	1	B.2	Congener	Species
1	0	1	0	0	0	B.3	Congener	Species
1	0	1	1	1	0	C.1	Congener	Genus
1	0	1	1	1	1	C.2	Congener	Genus
1	0	1	0	1	0	C.3	Congener	Genus
1	0	0	1	0	0	D.1	Confamilial	Species
1	0	0	1	0	1	E.1	Confamilial	Family
1	0	0	0	0	0	F.1	Input	Species
0	0	1	1	0	0	G.1	Congener	Species
0	0	1	1	0	1	G.2	Congener	Species
0	0	1	0	0	0	G.3	Congener	Species
0	0	1	1	1	0	H.1	Congener	Genus
0	0	1	1	1	1	H.2	Congener	Genus
0	0	1	0	1	0	H.3	Congener	Genus
0	0	0	1	0	0	I.1	Confamilial	Species
0	0	0	1	0	1	J.1	Confamilial	Family

Note, for both ‘Congener’ and ‘Confamilial’ (*in the singular*) ‘1’ denotes that a species is present; in a sense the genus is monotypic in space and time. For both ‘Congeners’ and ‘Confamilials’ (*in the plural*), ‘1’ denotes that two or more species are present; ‘Confamilial’ again representing a monotypic entity in space and time.

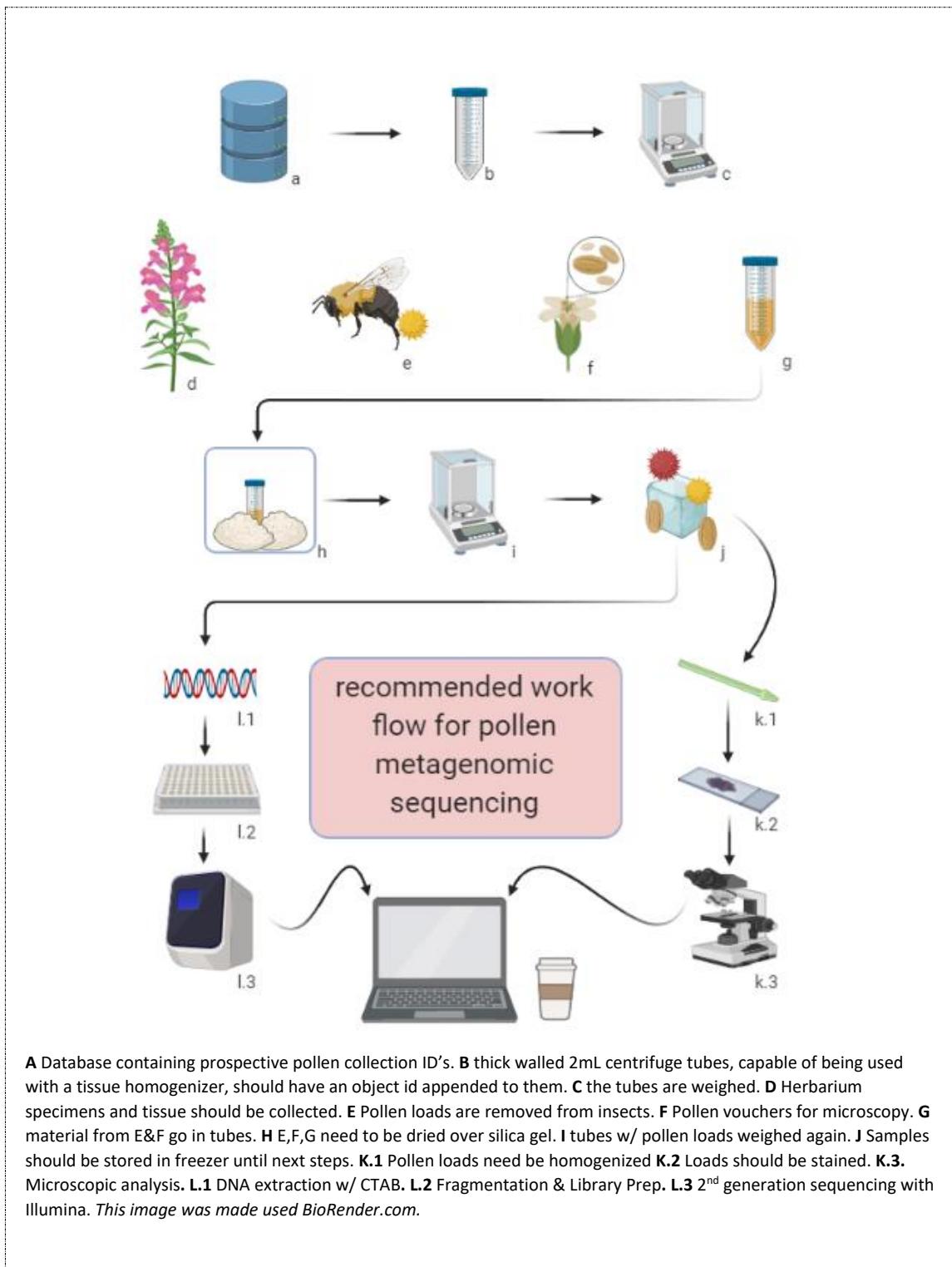
$$\begin{aligned} \text{Spatial} == 1 \& \text{ Temporal} == 1 \sim \mathbf{A} \\ \text{Spatial} == 1 \& \text{ Temporal} == 0 \& \text{ Congener} = 1 \sim \mathbf{B} \end{aligned}$$

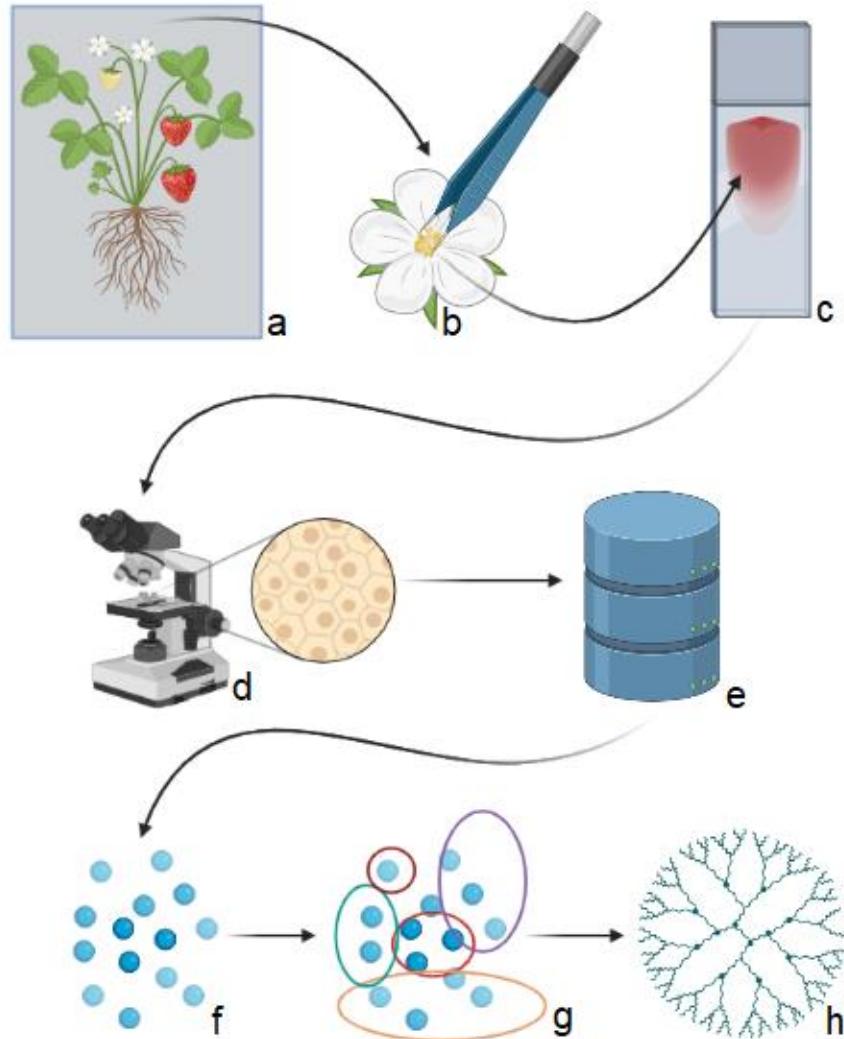
The temporal dimension is now buffered and a form of \mathbf{A} is employed
 $\text{Spatial} == 1 \& \text{ Temporal} +/- \text{ Buffer} == 1 \sim \mathbf{X}$

$$\begin{aligned} \text{Spatial} == 1 \& \text{ Temporal} == 0 \& \text{ Congeners} >= 2 \sim \mathbf{C} \\ \text{Spatial} == 1 \& \text{ Temporal} == 0 \& \text{ Congeners} == 0 \& \text{ Confamilial} == 1 \sim \mathbf{D} \\ \text{Spatial} == 1 \& \text{ Temporal} == 0 \& \text{ Congeners} == 0 \& \text{ Confamilial} >= 2 \sim \mathbf{E} \\ \text{Spatial} == 1 \& \text{ Temporal} == 0 \& \text{ Congener|s} == 0 \& \text{ Confamilial|s} == 0 \sim \mathbf{F} \end{aligned}$$

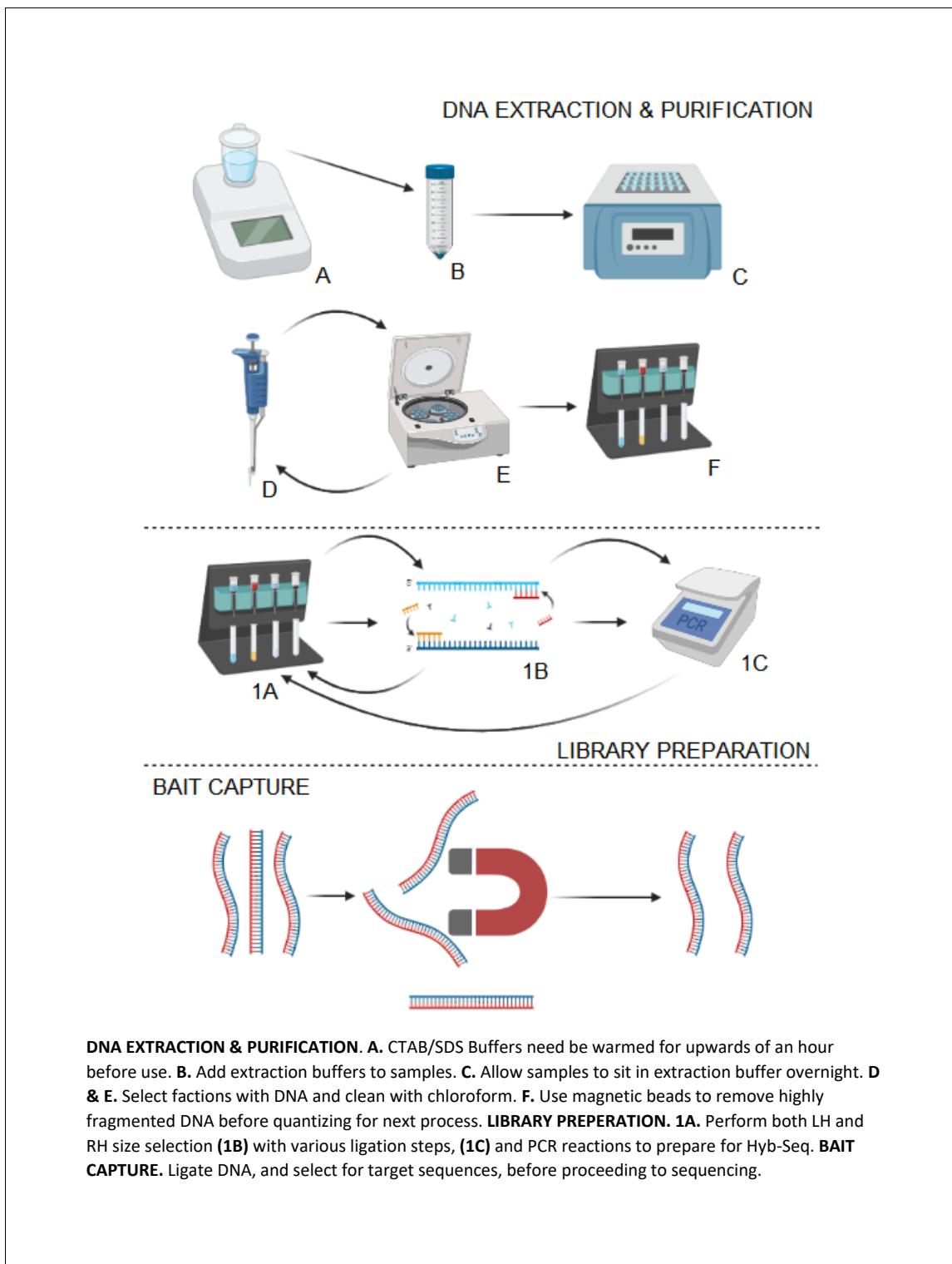
$$\begin{aligned} \text{Spatial} == 0 \& \text{ Temporal} == 0 \& \text{ Congener} == 1 \sim \mathbf{G} \\ \text{Spatial} == 0 \& \text{ Temporal} == 0 \& \text{ Congeners} == 1 \sim \mathbf{H} \\ \text{Spatial} == 0 \& \text{ Temporal} == 0 \& \text{ Confamilial} == 1 \sim \mathbf{I} \\ \text{Spatial} == 0 \& \text{ Temporal} == 0 \& \text{ Confamilials} == 1 \sim \mathbf{J} \end{aligned}$$

While the overall order matters, \mathbf{X} in particular may significantly alter conclusions.





A. An herbarium collection in flower from which pollen may be removed. **B.** The careful removal of pollen from dehiscent anthers using dissection tools. **C.** Place the pollen on a fuchsin jelly cube and melt it with stirring on a hotplate, for ca. 30 seconds. **D.** Microscopic imaging, of specimens and collection of character trait data. **E.** Immediate input and accessioning of files to a database. **F.** Ordination of traits into 2-dimensional space. **G.** Agglomerative clustering of data points into similar groups. **H.** Recovery of bifurcating decisions in development of clusters, or handwritten keys to visually diagnosable groups.



DNA EXTRACTION & PURIFICATION. A. CTAB/SDS Buffers need to be warmed for upwards of an hour before use. B. Add extraction buffers to samples. C. Allow samples to sit in extraction buffer overnight. D & E. Select fractions with DNA and clean with chloroform. F. Use magnetic beads to remove highly fragmented DNA before quantizing for next process. **LIBRARY PREPARATION.** 1A. Perform both LH and RH size selection (1B) with various ligation steps, (1C) and PCR reactions to prepare for Hyb-Seq. **BAIT CAPTURE.** Ligate DNA, and select for target sequences, before proceeding to sequencing.

CTAB-DNA POLLEN EXTRACTIONS

Adapted from Lalhmangiahi et. al & Guertler et al. by Benkendorf, Fant, & Noble.

SAMPLE PREPARATION AND GRINDING

- a1) Add 380 μ L extraction buffer (100 mM Tris-HCl, 50 mM EDTA, 50 mM NaCl, 10% SDS, pH 7.5). This solution will need to be warm enough for the SDS to be in solution, requires heat and stirring.
- a2) Vortex samples at speed > 2000, until pellet breaks apart, ca. 20-30 seconds.
- a3) Use the Pellet Pestle Motor (Kontes) for ca. 15 seconds to macerate samples.
- a4) Add 100 μ L extraction buffer to wash the tip of the pestle into the centrifuge tube, and burst bubbles.
- a5) Allow to sit at 35°C for 1 hour, use vortex occasionally if sedimentation of pollen occurs.

EXTRACTION AND ISOLATION OF DNA

- b1) Warm CTAB buffer to remove any precipitants if present.
- b2) Add 480 μ L 10% CTAB buffer.
- b3) Add 10 μ L RNase (10mg/mL); invert by hand, incubate for 40 minutes at 37°C, increase heat to 60°C wait 20 minutes before continuing to b4.
- b4) Add 15 μ L proteinase K (20mg/mL) & 12.5 μ L DTT (1 molar in H₂O); invert by hand, incubate for 1 hr. at 60°C.
- b5) Incubate overnight at 40°C (*note: this is a hard stopping point*)
- b6) Add 500 μ L of Phenol-chloroform-isoamyl alcohol vortex samples, centrifuge at 10,000 rpm (10 min.)
- b7) Transfer the uppermost aqueous layer to a new 2 mL centrifuge tube.

DNA PRECIPITATION

- c1) Add slightly chilled Isopropyl alcohol & Sodium Acetate 3mM 5:1, equivalent to ca. 2/3 of the removed layer. Store at -20°C, 1 hour to allow precipitation.
(Note: potential stopping point for a day or more, samples can stay at -20°C for days)
- c2) Centrifuge at 13,000 rpm for 10 minutes.
- c3) Pour supernatant into new 2mL centrifuge tube, add 400 μ L 70% EtOH. Store at -20°C for 20 minutes.
- c4) Spin at 13,000 rpm for 10 minutes, discard supernatant.
for both tubes the following steps apply
- c3) Add 400 μ L of 75% EtOH, invert tube x3, centrifuge at 13,000 rpm for 4 minutes; discard supernatant
- c4) Add 400 μ L of 95% EtOH, invert tube x3, centrifuge at 13,000 rpm for 4 minutes, discard supernatant
- c5) Dry tubes in vacuum centrifuge for 30 minutes on medium heat at 15 mmHG.

RESUSPENSION OF DNA

- d1) Add 40 μ L of dna free H₂O to sample.
- d2) place on heat block at 37°C until pellet resuspends with occasional use of vortexes.

NOTES: a 10% CTAB preparation will not readily stay in solution, maintain it on heatblock until you are ready to use it. After adding it to extraction tubes move them to heat block immediately (*i.e.* in batches of 5-10).

Solutions

Extraction buffer (100 mM Tris-HCl, 50 mM EDTA, 50 mM NaCl, 10% SDS - pH 7.5, ca. 32 mL H₂O)

For 100 samples (50 mL solution)

10 grams SDS (Sodium Dodecyl Sulfate , d = 1.01 g/cm3)

146.1 mg Sodium Chloride (NaCl, mw = 58.4 g/mol)

930.6 mg EDTA (EthyleneDiamineTetraacetic Acid Disodium Salt dihydrate, mw = 372.24 g/mol)

Add 20 mL deH₂O

5 mL Tris-HCl pH 8.0 (1 molar- kept in fridge)

Fill to 50 mL with deH₂O

Auto clave on 'Liquid' setting for 15 minutes.

Dissolution may require heat and stirring (3 & 4 settings respectively, ca. 15 min.)

10% CTAB solution (20 mM Tris-Cl pH 8.0, 1.4 M NaCl, 10 mM EDTA pH 7.5, 10% CTAB, 5% PVP, 40 mL DiH₂O)

For 100 samples (50 mL solution)

add ~30 mL deH₂O,

1 ml Tris-HCl pH 8.0 (1 molar- kept in fridge; 2-Amino-2-(hydroxymethyl)propane-1,3-diol)

4.08 g Sodium Chloride (NaCl, mw = 58.4 g/mol)

4 mL EDTA pH 7.5 (0.125 molar – kept in fridge; 2,2',2'',2'''-(Ethane-1,2-diyl)dinitrilo)tetraacetic acid)

5 g CTAB (hexadecyl(trimethyl)ammonium bromide, mw = 364.45, FYI this is 274 mM)

Auto clave on 'Liquid' setting for 15 minutes.

2.5 g PVP-40 (1-ethenylpyrrolidin-2-one) – add after autoclave

Fill to 50 mL with deH₂O

Dissolution of PVP will require 2-3 hrs, at 65°C with stirring. Before use allow one hour of stirring and heat to resuspend all salts in the solution.

Sodium acetate solution (3mM)

For 100 samples (10 mL solution)

20.4 mg Sodium Acetate trihydrate (mw = 136.08 g/M)

to 50 mL deH₂O

Auto clave on 'Liquid' setting for 15 minutes.

Phenol-chloroform Isoamyl alcohol (25:24:1) Saturated with 10 mM Tris pH 8.0, EDTA

For 100 samples (50 mL solution) (no need to make, is bought)

25 mL Phenol

24 mL Chloroform (Trichloromethane)

1 mL Isoamyl alcohol

Literature cited

Lalhmangaibi, R., Ghatak, S., Laba, R., Gurusubramian, G., Jumar, N.S. *Protocol for Optimal Quality and Quantity Pollen DNA Isolation from Honey Samples*. 2014. Journal of Biomolecular Techniques 25:92-95

Guertler, P., Eicheldinger, A., Muschler, P., Goerlich, O., Bursch, U. *Automated DNA extraction from pollen in honey* 2014. Food Chemistry 149:302-306

662 THIS SHOULD BE TURNED INTO A SMALLER PNG, AND HAVE THE NUMBER OF SEQUENCED
663 METAGENOMIC SAMPLES PLACED INTO THAT COLUMN AND INCLUDED IN TEXT ~~~ NEED
664 THIS !!!

Table 1: Queen Bee Pollen Loads examined

Subgenus	Species	Author	Tongue Length	Microscope Slides	Metagenome Samples
Subterraneobombus Vogt	<i>B. appositus</i>	Cresson 1879	Long	11	NA
Pyrobombus Dalla Torre	<i>B. bifarius</i>	Cresson 1879	Short	11	NA
Thoracobombus Dalla Torre	<i>B. californicus</i>	Smith 1854	Long	8	NA
Pyrobombus Dalla Torre	<i>B. flavifrons</i>	Cresson 1864	Medium	13	NA
Pyrobombus Dalla Torre	<i>B. mixtus</i>	Cresson 1879	Short	3	NA
Pyrobombus Dalla Torre	<i>B. nevadensis</i>	Cresson 1874	Long	5	NA
Bombias Robertson	<i>B. nevadensis</i>	Cresson 1864	Short	13	NA
Cullumanobombus Vogt	<i>B. rufocinctus</i>	Kirby 1837	Short	1	NA
Pyrobombus Dalla Torre	<i>B. sylvicola</i>				

^a All subgenera follow the system of Williams et al. 2008, and placements were found from the NMH website.

^b Tongue Lengths collected from Pyke et al. 2012

666 Ackerfield, J. (2015). *Flora of colorado*. BRIT Press Fort Worth.

667 Alarcón, R. (2010). Congruence between visitation and pollen-transport networks in a california plant–

- 668 pollinator community. *Oikos*, **119**, 35–44. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2009.17694.x>
- 669
- 670 Aldous, A.E. (1919). *Eradicating tall larkspur on cattle ranges in the national forest*. US Department of
671 Agriculture.
- 672 Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models:
673 Prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, **43**, 1223–1232.
- 674 Allred, K.W. & Ivey, R. (2012). Flora neomexicana III: An illustrated identification manual. *Lulu. com*.
- 675 Araujo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *Trends in ecology & evolution*,
676 **22**, 42–47.
- 677 Augspurger, C.K. & Zaya, D.N. (2020). Concordance of long-term shifts with climate warming varies among
678 phenological events and herbaceous species. *Ecological Monographs*, **90**, e01421.
- 679 Baker, W.J., Bailey, P., Barber, V., Barker, A., Bellot, S., Bishop, D., Botigué, L.R., Brewer, G., Carruthers,
680 T., Clarkson, J.J., Cook, J., Cowan, R.S., Dodsworth, S., Epitawalage, N., Françoso, E., Gallego, B.,
681 Johnson, M.G., Kim, J.T., Leempoel, K., Maurin, O., Mcginnie, C., Pokorny, L., Roy, S., Stone, M.,
682 Toledo, E., Wickett, N.J., Zuntini, A.R., Eiserhardt, W.L., Kersey, P.J., Leitch, I.J. & Forest, F. (2021a).
683 A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Systematic Biology*,
684 **71**, 301–319. Retrieved from <https://doi.org/10.1093/sysbio/syab035>
- 685 Baker, W., Dodsworth, S., Forest, F., Graham, S., Johnson, M., McDonnell, A., Pokorny, L., Tate, J., Wicke,
686 S. & Wickett, N. (2021b). Exploring Angiosperms353: An open, community toolkit for collaborative
687 phylogenomic research on flowering plants. *American Journal of Botany*, **108**.
- 688 Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012). Selecting pseudo-absences for species
689 distribution models: How, where and how many? *Methods in ecology and evolution*, **3**, 327–338.
- 690 Barker, D.A. & Arceo-Gomez, G. (2021). Pollen transport networks reveal highly diverse and temporally
691 stable plant–pollinator interactions in an Appalachian floral community. *AoB PLANTS*, **13**. Retrieved
692 from <https://doi.org/10.1093/aobpla/plab062>
- 693 Basey, A.C., Fant, J.B. & Kramer, A.T. (2015). Producing native plant materials for restoration: 10 rules
694 to collect and maintain genetic diversity. *Native Plants Journal*, **16**, 37–53.
- 695 Beattie, A. (1971). A technique for the study of insect-borne pollen. *The Pan-Pacific Entomologist*, **47**, 82.
- 696 Beck, J.B., Markley, M.L., Zielke, M.G., Thomas, J.R., Hale, H.J., Williams, L.D. & Johnson, M.G. (2021).
697 Are palmer's elm-leaf goldenrod and the smooth elm-leaf goldenrod real? The Angiosperms353 kit
698 provides within-species signal in solidago ulmifolia sl. *Systematic Botany*, **46**, 1107–1113.
- 699 Belitz, M.W., Larsen, E.A., Ries, L. & Guralnick, R.P. (2020). The accuracy of phenology estimators for use
700 with sparsely sampled presence-only observations. *Methods in Ecology and Evolution*, **11**, 1273–1285.

- 701 Bell, K.L., Burgess, K.S., Botsch, J.C., Dobbs, E.K., Read, T.D. & Brosi, B.J. (2019). Quantitative and
702 qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Molecular*
703 *Ecology*, **28**, 431–455.
- 704 Bell, K.L., Fowler, J., Burgess, K.S., Dobbs, E.K., Gruenewald, D., Lawley, B., Morozumi, C. & Brosi, B.J.
705 (2017). Applying pollen DNA metabarcoding to the study of plant–pollinator interactions. *Applications*
706 *in plant sciences*, **5**, 1600124.
- 707 Bell, K.L., Petit III, R.A., Cutler, A., Dobbs, E.K., Macpherson, J.M., Read, T.D., Burgess, K.S. & Brosi,
708 B.J. (2021). Comparing whole-genome shotgun sequencing and DNA metabarcoding approaches for
709 species identification and quantification of pollen species mixtures. *Ecology and evolution*, **11**, 16082–
710 16098.
- 711 Belsky, A.J., Matzke, A. & Uselman, S. (1999). Survey of livestock influences on stream and riparian
712 ecosystems in the western united states. *Journal of Soil and water Conservation*, **54**, 419–431.
- 713 Bergman, P., Molau, U. & Holmgren, B. (1996). Micrometeorological impacts on insect activity and plant
714 reproductive success in an alpine environment, swedish lapland. *Arctic and alpine research*, **28**, 196–202.
- 715 Betancourt, J.L., Schwartz, M.D., Breshears, D.D., Cayan, D.R., Dettinger, M.D., Inouye, D.W., Post, E.
716 & Reed, B.C. (2005). Implementing a US national phenology network.
- 717 Bingham, R.A. & Orthner, A.R. (1998). Efficient pollination of alpine plants. *Nature*, **391**, 238–239.
- 718 Bolger, A. & Giorgi, F. (2014). Trimmomatic: A flexible read trimming tool for illumina NGS data. *Bioin-*
719 *formatics*, **30**, 2114–2120.
- 720 Bontsutsnaja, A., Karise, R., Mand, M. & Smagghe, G. (2021). Bumble bee foraged pollen analyses in spring
721 time in southern estonia shows abundant food sources. *Insects*, **12**, 922.
- 722 Bortolus, A. (2008). Error cascades in the biological sciences: The unwanted consequences of using bad
723 taxonomy in ecology. *AMBIO: A journal of the human environment*, **37**, 114–118.
- 724 Brewen, C.J., Berrill, J.-P., Ritchie, M.W., Boston, K., Dagley, C.M., Jones, B., Coppoletta, M. & Burnett,
725 C.L. (2021). 76-year decline and recovery of aspen mediated by contrasting fire regimes: Long-unburned,
726 infrequent and frequent mixed-severity wildfire. *Plos one*, **16**, e0232995.
- 727 Brito-Morales, I., Molinos, J.G., Schoeman, D.S., Burrows, M.T., Poloczanska, E.S., Brown, C.J., Ferrier,
728 S., Harwood, T.D., Klein, C.J., McDonald-Madden, E. & others. (2018). Climate velocity can inform
729 conservation in a warming world. *Trends in ecology & evolution*, **33**, 441–457.
- 730 Brosi, B.J. & Briggs, H.M. (2013). Single pollinator species losses reduce floral fidelity and plant reproductive
731 function. *Proceedings of the National Academy of Sciences*, **110**, 13044–13048.
- 732 Calabrese, J.M., Certain, G., Kraan, C. & Dormann, C.F. (2014). Stacking species distribution models and
733 adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, **23**, 99–112.

- 734 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T.L. (2009).
735 BLAST+: Architecture and applications. *BMC bioinformatics*, **10**, 1–9.
- 736 Cameron, S.A. & Sadd, B.M. (2020). Global trends in bumble bee health. *Annual review of entomology*, **65**,
737 209–232.
- 738 CaraDonna, P.J., Burkle, L.A., Schwarz, B., Resasco, J., Knight, T.M., Benadi, G., Blüthgen, N., Dormann,
739 C.F., Fang, Q., Fründ, J. & others. (2021). Seeing through the static: The temporal dimension of
740 plant–animal mutualistic interactions. *Ecology Letters*, **24**, 149–161.
- 741 CaraDonna, P.J., Iler, A.M. & Inouye, D.W. (2014). Shifts in flowering phenology reshape a subalpine plant
742 community. *Proceedings of the National Academy of Sciences*, **111**, 4916–4921.
- 743 CaraDonna, P.J., Petry, W.K., Brennan, R.M., Cunningham, J.L., Bronstein, J.L., Waser, N.M. & Sanders,
744 N.J. (2017). Interaction rewiring and the rapid turnover of plant–pollinator networks. *Ecology letters*,
745 **20**, 385–394.
- 746 Chao, A., Gotelli, N.J., Hsieh, T.C., Sande, E.L., Ma, K.H., Colwell, R.K. & Ellison, A.M. (2014). Rarefac-
747 tion and extrapolation with hill numbers: A framework for sampling and estimation in species diversity
748 studies. *Ecological Monographs*, **84**, 45–67.
- 749 Cheng, S., Melkonian, M., Smith, S.A., Brockington, S., Archibald, J.M., Delaux, P.-M., Li, F.-W., Melko-
750 nian, B., Mavrodiev, E.V., Sun, W., Fu, Y., Yang, H., Soltis, D.E., Graham, S.W., Soltis, P.S., Liu,
751 X., Xu, X. & Wong, G.K.-S. (2018). 10KP: A phylogenetic genome sequencing plan. *GigaScience*, **7**.
752 Retrieved from <https://doi.org/10.1093/gigascience/giy013>
- 753 Coissac, E., Hollingsworth, P.M., Lavergne, S. & Taberlet, P. (2016). From barcodes to genomes: Extending
754 the concept of DNA barcoding.
- 755 Coissac, E., Riaz, T. & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants
756 and animals. *Molecular ecology*, **21**, 1834–1847.
- 757 Colla, S.R., Gadallah, F., Richardson, L., Wagner, D. & Gall, L. (2012). Assessing declines of north american
758 bumble bees (*bombus* spp.) Using museum specimens. *Biodiversity and Conservation*, **21**, 3585–3595.
- 759 Cooke, R.U. & Reeves, R.W. (1976). *Arroyos and environmental change in the american south-west*. Claren-
760 don Press.
- 761 Crisci, J.V., Katinas, L., Apodaca, M.J. & Hoch, P.C. (2020). The end of botany. *Trends in Plant Science*,
762 **25**, 1173–1176.
- 763 Cronquist, A., Holmgren, A.H., Holmgren, N.H., Reveal, J.L., Holmgren, P.K., Barneby, R & others.
764 (1977+). *Intermountain flora. Vascular plants of the intermountain west, USA volume six. The mono-*
765 *cotyledons*. Columbia University.
- 766 Dahl, T.E. (1990). *Wetlands losses in the united states, 1780's to 1980's*. US Department of the Interior,

- 767 Fish; Wildlife Service.
- 768 Davis, C.C., Lyra, G.M., Park, D.S., Asprino, R., Maruyama, R., Torquato, D., Cook, B.I. & Ellison, A.M.
769 (2022). New directions in tropical phenology. *Trends in Ecology & Evolution*.
- 770 Dobrowski, S.Z. & Parks, S.A. (2016). Climate change velocity underestimates climate change exposure in
771 mountainous regions. *Nature Communications*, **7**, 1–8.
- 772 Doyle, J.J. & Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue.
773 *Phytochemical Bulletin*, **19**, 11–15.
- 774 Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J.-P. & Guisan, A. (2011). Predicting spatial
775 patterns of plant species richness: A comparison of direct macroecological and species stacking modelling
776 approaches. *Diversity and Distributions*, **17**, 1122–1131.
- 777 Elith*, J., H. Graham*, C., P. Anderson, R., Dudik, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann,
778 F., R. Leathwick, J., Lehmann, A. & others. (2006). Novel methods improve prediction of species'
779 distributions from occurrence data. *Ecography*, **29**, 129–151.
- 780 Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Percy, D.M., Graham, S.W., Barrett, S.C., Newmaster,
781 S.G., Hajibabaei, M. & Husband, B.C. (2009). Are plant species inherently harder to discriminate than
782 animal species using DNA barcoding markers? *Molecular Ecology Resources*, **9**, 130–139.
- 783 Feng, X., Enquist, B.J., Park, D.S., Boyle, B., Breshears, D.D., Gallagher, R.V., Lien, A., Newman, E.A.,
784 Burger, J.R., Maitner, B.S. & others. (2022). A review of the heterogeneous landscape of biodiversity
785 databases: Opportunities and challenges for a synthesized biodiversity knowledge base. *Global Ecology
786 and Biogeography*.
- 787 Fenn, M.E., Baron, J.S., Allen, E.B., Rueth, H.M., Nydick, K.R., Geiser, L., Bowman, W.D., Sickman, J.O.,
788 Meixner, T., Johnson, D.W. & others. (2003). Ecological effects of nitrogen deposition in the western
789 united states. *BioScience*, **53**, 404–420.
- 790 Flora of North America Editorial Committee, eds. (1993+). *Flora of north america north of mexico [online]*.
791 Oxford University Press on Demand.
- 792 Frase, 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
- 793 Gage, E. & Cooper, D.J. (2013). Historical range of variation assessment for wetland and riparian ecosystems,
794 u.s. Forest service rocky mountain region
- 795 Goulson, D. (2010). *Bumblebees: Behaviour, ecology, and conservation*. Oxford University Press on Demand.
- 796 Goulson, D., Hanley, M.E., Darvill, B., Ellis, J. & Knight, M.E. (2005). Causes of rarity in bumblebees.
797 *Biological conservation*, **122**, 1–8.
- 798 Goulson, D., Lye, G. & Darvill, B. (2008). The decline and conservation of bumblebees. *Annual review of*

- 800 entomology, **53**, 191–208.
- 801 Grixti, J.C., Wong, L.T., Cameron, S.A. & Favret, C. (2009). Decline of bumble bees (*bombus*) in the north
802 american midwest. *Biological conservation*, **142**, 75–84.
- 803 Group, C.P.W., Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., Bank,
804 M. van der, Chase, M.W., Cowan, R.S., Erickson, D.L. & others. (2009). A DNA barcode for land
805 plants. *Proceedings of the National Academy of Sciences*, **106**, 12794–12797.
- 806 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.,
807 Chen, S.-L. & others. (2011). Comparative analysis of a large dataset indicates that internal transcribed
808 spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National
Academy of Sciences*, **108**, 19641–19646.
- 809 Havens, K., Vitt, P., Schwarz, J., Orr, B. & Crimmins, T. (2007). Chicago botanic garden's conservation
810 and outreach efforts on climate change. *BGjournal*, **4**, 13–16.
- 811 Hebert, P.D., Cywinska, A., Ball, S.L. & DeWaard, J.R. (2003). Biological identifications through DNA
812 barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, 313–321.
- 813 Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A.,
814 Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B. & others. (2017). SoilGrids250m:
815 Global gridded soil information based on machine learning. *PLoS one*, **12**, e0169748.
- 816 Hennig, C. (2020). *Fpc: Flexible procedures for clustering*. Retrieved from [https://CRAN.R-project.org/
817 package=fpc](https://CRAN.R-project.org/package=fpc)
- 818 Hinchliff, C.E., Smith, S.A., Allman, J.F., Burleigh, J.G., Chaudhary, R., Coghill, L.M., Crandall, K.A.,
819 Deng, J., Drew, B.T., Gazis, R. & others. (2015). Synthesis of phylogeny and taxonomy into a compre-
820 hensive tree of life. *Proceedings of the National Academy of Sciences*, **112**, 12764–12769.
- 821 Hitchcock, C.L. & Cronquist, A. (2018). *Flora of the pacific northwest: An illustrated manual*. University
822 of Washington Press.
- 823 Hollingsworth, P.M., Li, D.-Z., Bank, M. van der & Twyford, A.D. (2016). Telling plant species apart with
824 DNA: From barcodes to genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*,
825 **371**, 20150338.
- 826 Hsieh, T.C., Ma, K.H. & Chao, A. (2020). *iNEXT: Interpolation and extrapolation for species diversity*.
827 Retrieved from http://chao.stat.nthu.edu.tw/wordpress/software_download/
- 828 Iler, A.M., Humphrey, P.T., Ogilvie, J.E. & CaraDonna, P.J. (2021). Conceptual and practical issues limit
829 the utility of statistical estimators of phenological events. *Ecosphere*, **12**, e03828.
- 830 Inouye, D.W. (1980). The effect of proboscis and corolla tube lengths on patterns and rates of flower
831 visitation by bumblebees. *Oecologia*, **45**, 197–201.

- 833 Janzen, D.H. (1967). Synchronization of sexual reproduction of trees within the dry season in central america.
834 *Evolution*, **21**, 620–637.
- 835 Janzen, D.H., Burns, J.M., Cong, Q., Hallwachs, W., Dapkey, T., Manjunath, R., Hajibabaei, M., Hebert,
836 P.D. & Grishin, N.V. (2017). Nuclear genomes distinguish cryptic species suggested by their DNA
837 barcodes and ecology. *Proceedings of the National Academy of Sciences*, **114**, 8313–8318.
- 838 *Jepson flora project*. (2020).
- 839 Johnson, M.D., Fokar, M., Cox, R.D. & Barnes, M.A. (2021). Airborne environmental DNA metabarcoding
840 detects more diversity, with less sampling effort, than a traditional plant community survey. *BMC Ecology*
841 and *Evolution*, **21**, 1–15.
- 842 Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J. & Wickett,
843 N.J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput
844 sequencing reads using target enrichment. *Applications in plant sciences*, **4**, 1600016.
- 845 Johnson, M.G., Pokorny, L., Dodsworth, S., Botigue, L.R., Cowan, R.S., Devault, A., Eiserhardt, W.L.,
846 Epitawalage, N., Forest, F., Kim, J.T. & others. (2019). A universal probe set for targeted sequencing
847 of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic biology*,
848 **68**, 594–606.
- 849 Joshi, J., Schmid, B., Caldeira, M., Dimitrakopoulos, P., Good, J., Harris, R., Hector, A., Huss-Danell, K.,
850 Jumpponen, A., Minns, A. & others. (2001). Local adaptation enhances performance of common plant
851 species. *Ecology Letters*, **4**, 536–544.
- 852 Keane, R.E. (2002). Cascading effects of fire exclusion in rocky mountain ecosystems: A literature review.
- 853 Kramer, A.T. & Havens, K. (2015). Report in brief: Assessing botanical capacity to address grand challenges
854 in the united states. *Natural Areas Journal*, **35**, 83–89.
- 855 Kress, W.J. (2017). Plant DNA barcodes: Applications today and in the future. *Journal of systematics and*
856 *evolution*, **55**, 291–307.
- 857 Kress, W.J. & Erickson, D.L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL
858 gene complements the non-coding trnH-psbA spacer region. *PLoS one*, **2**, e508.
- 859 Kuhn, M. (2022). *Caret: Classification and regression training*. Retrieved from <https://CRAN.R-project.org/package=caret>
- 860 Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G. & Taylor, M.I. (2019). How quantitative is
862 metabarcoding: A meta-analytical approach. *Molecular ecology*, **28**, 420–430.
- 863 Land Management, B. of. (2019). U.s. Department of interior bureau of land management, BLM - assess-
864 ment, inventory, and monitoring (AIM) terrestrial indicators raw dataset. Retrieved from <https://gbp-blm-egis.hub.arcgis.com/pages/aim>

- 866 Lang, D., Tang, M., Hu, J. & Zhou, X. (2019). Genome-skimming provides accurate quantification for pollen
867 mixtures. *Molecular Ecology Resources*, **19**, 1433–1446.
- 868 Lesica, P., Lavin, M. & Stickney, P.F. (2012). *Manual of montana vascular plants*. Brit Press.
- 869 Lewin, H.A., Richards, S., Aiden, E.L., Allende, M.L., Archibald, J.M., Bálint, M., Barker, K.B., Baumgart-
870 ner, B., Belov, K., Bertorelle, G., Blaxter, M.L., Cai, J., Caperello, N.D., Carlson, K., Castilla-Rubio,
871 J.C., Chaw, S.-M., Chen, L., Childers, A.K., Coddington, J.A., Conde, D.A., Corominas, M., Crandall,
872 K.A., Crawford, A.J., DiPalma, F., Durbin, R., Ebenezer, T.E., Edwards, S.V., Fedrigo, O., Flieck, P.,
873 Formenti, G., Gibbs, R.A., Gilbert, M.T.P., Goldstein, M.M., Graves, J.M., Greely, H.T., Grigoriev,
874 I.V., Hackett, K.J., Hall, N., Haussler, D., Helgen, K.M., Hogg, C.J., Isobe, S., Jakobsen, K.S., Janke,
875 A., Jarvis, E.D., Johnson, W.E., Jones, S.J.M., Karlsson, E.K., Kersey, P.J., Kim, J.-H., Kress, W.J.,
876 Kuraku, S., Lawniczak, M.K.N., Leebens-Mack, J.H., Li, X., Lindblad-Toh, K., Liu, X., Lopez, J.V.,
877 Marques-Bonet, T., Mazard, S., Mazet, J.A.K., Mazzoni, C.J., Myers, E.W., O'Neill, R.J., Paez, S.,
878 Park, H., Robinson, G.E., Roquet, C., Ryder, O.A., Sabir, J.S.M., Shaffer, H.B., Shank, T.M., Sherkow,
879 J.S., Soltis, P.S., Tang, B., Tedersoo, L., Uliano-Silva, M., Wang, K., Wei, X., Wetzer, R., Wilson,
880 J.L., Xu, X., Yang, H., Yoder, A.D. & Zhang, G. (2022). The earth BioGenome project 2020: Starting
881 the clock. *Proceedings of the National Academy of Sciences*, **119**, e2115635118. Retrieved from
882 <https://www.pnas.org/doi/abs/10.1073/pnas.2115635118>
- 883 Liang, H., Zhao, Y.-H., Rafferty, N.E., Ren, Z.-X., Zhong, L., Li, H.-D., Li, D.-Z. & Wang, H. (2021).
884 Evolutionary and ecological factors structure a plant–bumblebee network in a biodiversity hotspot, the
885 himalaya–hengduan mountains. *Functional Ecology*, **35**, 2523–2535.
- 886 Life Project Consortium, D.T. of, Blaxter, M., Mieszkowska, N., Palma, F.D., Holland, P., Durbin, R.,
887 Richards, T., Berriman, M., Kersey, P., Hollingsworth, P., Wilson, W., Twyford, A., Gaya, E., Lawniczak,
888 M., Lewis, O., Broad, G., Howe, K., Hart, M., Flieck, P. & Barnes, I. (2022). Sequence locally, think globally:
889 The darwin tree of life project. *Proceedings of the National Academy of Sciences*, **119**, e2115642118.
890 Retrieved from <https://www.pnas.org/doi/abs/10.1073/pnas.2115642118>
- 891 Li, X., Jiang, L., Meng, F., Wang, S., Niu, H., Iler, A.M., Duan, J., Zhang, Z., Luo, C., Cui, S. & others.
892 (2016). Responses of sequential and hierarchical phenological events to warming and cooling in alpine
893 meadows. *Nature Communications*, **7**, 1–8.
- 894 Liu, J., Shi, L., Han, J., Li, G., Lu, H., Hou, J., Zhou, X., Meng, F. & Downie, S.R. (2014). Identification
895 of species in the angiosperm family apiaceae using DNA barcodes. *Molecular ecology resources*, **14**,
896 1231–1238.
- 897 Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y. & Chen, S. (2015). Plant DNA barcoding: From gene
898 to genome. *Biological Reviews*, **90**, 157–166.

- 899 Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009). The velocity of
900 climate change. *Nature*, **462**, 1052–1055.
- 901 Lu, J., Breitwieser, F.P., Thielen, P. & Salzberg, S.L. (2017). Bracken: Estimating species abundance in
902 metagenomics data. *PeerJ Computer Science*, **3**, e104.
- 903 Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2022). *Cluster: Cluster analysis basics*
904 and extensions. Retrieved from <https://CRAN.R-project.org/package=cluster>
- 905 Maitner, B. (2022). *BIEN: Tools for accessing the botanical information and ecology network database*.
906 Retrieved from <https://CRAN.R-project.org/package=BIEN>
- 907 Manzano, S. (2021). Flippant attitudes towards plant identification jeopardize early career botanists. *Trends*
908 in *Plant Science*, **26**, 987–988.
- 909 McLay, T.G., Birch, J.L., Gunn, B.F., Ning, W., Tate, J.A., Nauheimer, L., Joyce, E.M., Simpson, L.,
910 Schmidt-Lebuhn, A.N., William J & others. (2021). New targets acquired: Improving locus recovery
911 from the Angiosperms353 probe set. *Applications in plant sciences*, **9**.
- 912 Mohlenbrock, R.H. (2002). *Vascular flora of illinois*. SIU Press.
- 913 Moore, A.J. & Bohs, L. (2003). An ITS phylogeny of balsamorhiza and wyethia (asteraceae: heliantheae).
914 *American Journal of Botany*, **90**, 1653–1660.
- 915 Naiman, R.J., Johnston, C.A. & Kelley, J.C. (1988). Alteration of north american streams by beaver.
916 *BioScience*, **38**, 753–762.
- 917 Naimi, B. & Araujo, M.B. (2016). Sdm: A reproducible and extensible r platform for species distribution
918 modelling. *Ecography*, **39**, 368–375.
- 919 Naimi, B., Hamm, N. a.s., Groen, T.A., Skidmore, A.K. & Toxopeus, A.G. (2014). Where is positional
920 uncertainty a problem for species distribution modelling. *Ecography*, **37**, 191–203.
- 921 Nazaire, M. & Hufford, L. (2014). Phylogenetic systematics of the genus mertensia (boraginaceae). *Systematic*
922 *Botany*, **39**, 268–303.
- 923 Nevado, B., Atchison, G.W., Hughes, C.E. & Filatov, D.A. (2016). Widespread adaptive evolution during
924 repeated evolutionary radiations in new world lupins. *Nature communications*, **7**, 1–9.
- 925 Newstrom, L.E., Frankie, G.W. & Baker, H.G. (1994). A new classification for plant phenology based on
926 flowering patterns in lowland tropical rain forest trees at la selva, costa rica. *Biotropica*, **26**, 141–159.
- 927 Oberhauser, K.S., Nail, K.R. & Altizer, S. (2015). *Monarchs in a changing world: Biology and conservation*
928 of an iconic butterfly. Cornell University Press.
- 929 Occdownload Gbif.Org. (2021). Occurrence download. Retrieved from <https://www.gbif.org/occurrence/>
930 download/0206948-200613084148143
- 931 Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos,

- 932 P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D.,
933 Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly,
934 M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M.-H., Ribeiro Cunha,
935 E., Smith, T., Stier, A., Ter Braak, C.J.F. & Weedon, J. (2022). *Vegan: Community ecology package*.
936 Retrieved from <https://CRAN.R-project.org/package=vegan>
- 937 Oliver, P.M., Adams, M., Lee, M.S., Hutchinson, M.N. & Doughty, P. (2009). Cryptic diversity in vertebrates:
938 Molecular data double estimates of species diversity in a radiation of australian lizards (diplodactylus,
939 gekkota). *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2001–2007.
- 940 Omernik, J.M. (1987). Ecoregions of the conterminous united states. *Annals of the Association of American
941 geographers*, **77**, 118–125.
- 942 Ottenlips, M.V., Mansfield, D.H., Buerki, S., Feist, M.A.E., Downie, S.R., Dodsworth, S., Forest, F., Plun-
943 kett, G.M. & Smith, J.F. (2021). Resolving species boundaries in a recent radiation with the An-
944 giosperms353 probe set: The lomatium packardiae/l. anomalam clade of the l. triternatum (apiaceae)
945 complex. *American journal of botany*, **108**, 1217–1233.
- 946 Park, D.S., Lyra, G.M., Ellison, A.M., Maruyama, R.K.B., Torquato, D. dos R., Asprino, R.C., Cook, B.I. &
947 Davis, C.C. (2022). Herbarium records provide reliable phenology estimates in the understudied tropics.
948 *Journal of Ecology*.
- 949 Pearse, W.D., Davis, C.C., Inouye, D.W., Primack, R.B. & Davies, T.J. (2017). A statistical estimator
950 for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution*, **1**,
951 1876–1882.
- 952 Peel, N., Dicks, L.V., Clark, M.D., Heavens, D., Percival-Alwyn, L., Cooper, C., Davies, R.G., Leggett,
953 R.M. & Yu, D.W. (2019). Semi-quantitative characterisation of mixed pollen samples using MinION
954 sequencing and reverse metagenomics (RevMet). *Methods in Ecology and Evolution*, **10**, 1690–1701.
- 955 Pepin, N., Arnone, E., Gobiet, A., Haslinger, K., Kotlarski, S., Notarnicola, C., Palazzi, E., Seibert, P.,
956 Serafin, S., Schöner, W. & others. (2022). Climate changes and their elevational patterns in the mountains
957 of the world. *Reviews of geophysics*, **60**, e2020RG000730.
- 958 Pinto-Ledezma, J.N. & Cavender-Bares, J. (2021). Predicting species distributions and community compo-
959 sition using satellite remote sensing predictors. *Scientific Reports*, **11**, 1–12.
- 960 Pleasants, J.M. (1980). Competition for bumblebee pollinators in rocky mountain plant communities. *Ecol-
961 ogy*, **61**, 1446–1459.
- 962 Pornon, A., Andalo, C., Burrus, M. & Escaravage, N. (2017). DNA metabarcoding data unveils invisible
963 pollination networks. *Scientific Reports*, **7**, 1–11.
- 964 Prather, L.A., Alvarez-Fuentes, O., Mayfield, M.H. & Ferguson, C.J. (2004a). Implications of the decline in

- 965 plant collecting for systematic and floristic research. *Systematic Botany*, **29**, 216–220.
- 966 Prather, L.A., Alvarez-Fuentes, O., Mayfield, M.H. & Ferguson, C.J. (2004b). The decline of plant collecting
967 in the united states: A threat to the infrastructure of biodiversity studies. *Systematic Botany*, **29**, 15–28.
- 968 Prim, R.C. (1957). Shortest connection networks and some generalisations. *Bell System Technical Journal*,
969 **36**, 1389–1401.
- 970 Pusalkar, P.K. & Singh, D.K. (2015). Taxonomic rearrangement of arenaria (caryophyllaceae) in indian
971 western himalaya. *Journal of Japanese Botany*, **90**, 77–91.
- 972 Pyke, G.H. (1982). Local geographic distributions of bumblebees near crested butte, colorado: Competition
973 and community structure. *Ecology*, **63**, 555–573.
- 974 Qiao, H., Soberon, J. & Peterson, A.T. (2015). No silver bullets in correlative ecological niche modelling:
975 Insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and
976 Evolution*, **6**, 1126–1136.
- 977 Rabeler, R.K. & Wagner, W.L. (2016). New combinations in odontostemma (caryophyllaceae). *PhytoKeys*,
978 **77**.
- 979 Ralphs, M.H. & Ueckert, D.N. (1988). Herbicide control of locoweeds: A review. *Weed Technology*, **2**,
980 460–465.
- 981 Ralphs, M., Woolsey, L. & Bowns, J. (2003). Mechanism by which ammonium fertilizers kill tall larkspur.
982 *Rangeland Ecology & Management/Journal of Range Management Archives*, **56**, 524–528.
- 983 Robinson, N., Regetz, J. & Guralnick, R.P. (2014). EarthEnv-DEM90: A nearly-global, void-free, multi-
984 scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. *ISPRS Journal of
985 Photogrammetry and Remote Sensing*, **87**, 57–67.
- 986 Ruppert, K.M., Kline, R.J. & Rahman, M.S. (2019). Past, present, and future perspectives of environmental
987 DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global
988 eDNA. *Global Ecology and Conservation*, **17**, e00547.
- 989 Sadeghian, S., Zarre, S., Rabeler, R.K. & Heubl, G. (2015). Molecular phylogenetic analysis of arenaria
990 (caryophyllaceae: Tribe arenarieae) and its allies inferred from nuclear DNA internal transcribed spacer
991 and plastid DNA rps16 sequences. *Botanical Journal of the Linnean Society*, **178**, 648–669.
- 992 Sarro, E., Tripodi, A. & Woodard, S.H. (2022). Bumble bee (*bombus vosnesenskii*) queen nest searching
993 occurs independent of ovary developmental status. *Integrative Organismal Biology*, **4**, obac007.
- 994 Schmitt, S., Pouteau, R., Justeau, D., De Boissieu, F. & Birnbaum, P. (2017). Ssdm: An r package to
995 predict distribution of species richness and composition based on stacked species distribution models.
996 *Methods in Ecology and Evolution*, **8**, 1795–1803.
- 997 Sennikov, A.N. & Kurtto, A. (2017). A phylogenetic checklist of sorbus sl (rosaceae) in europe.

- 998 Sickel, W., Ankenbrand, M.J., Grimmer, G., Holzschuh, A., Hartel, S., Lanzen, J., Steffan-Dewenter, I. &
999 Keller, A. (2015). Increased efficiency in identifying mixed pollen samples by meta-barcoding with a
1000 dual-indexing approach. *BMC ecology*, **15**, 1–9.
- 1001 Slimp, M., Williams, L.D., Hale, H. & Johnson, M.G. (2021). On the potential of Angiosperms353 for
1002 population genomic studies. *Applications in Plant Sciences*, **9**.
- 1003 Smith, S.A. & Brown, J.W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American journal*
1004 *of botany*, **105**, 302–314.
- 1005 Soltis, P.S. (2017). Digitization of herbaria enables novel research.
- 1006 Sork, V.L. (2018). Genomic studies of local adaptation in natural plant populations. *Journal of Heredity*,
1007 **109**, 3–15.
- 1008 Stevens, C.J., David, T.I. & Storkey, J. (2018). Atmospheric nitrogen deposition in terrestrial ecosystems: Its
1009 impact on plant communities and consequences across trophic levels. *Functional Ecology*, **32**, 1757–1769.
- 1010 Stroud, S., Fennell, M., Mitchley, J., Lydon, S., Peacock, J. & Bacon, K.L. (2022). The botanical education
1011 extinction and the fall of plant awareness. *Ecology and Evolution*, **12**, e9019.
- 1012 Suchan, T., Talavera, G., Saez, L., Ronikier, M. & Vila, R. (2019). Pollen metabarcoding as a tool for
1013 tracking long-distance insect migrations. *Molecular Ecology Resources*, **19**, 149–162.
- 1014 Tange, O. (2021). GNU parallel 20220322 (savannah). Retrieved from <https://doi.org/10.5281/zenodo.6377950>
- 1016 Tran, H., Nguyen, P., Ombadi, M., Hsu, K., Sorooshian, S. & Qing, X. (2019). A cloud-free MODIS snow
1017 cover dataset for the contiguous united states from 2000 to 2017. *Scientific data*, **6**, 1–13.
- 1018 Wang, T., Hamann, A., Spittlehouse, D. & Carroll, C. (2016). Locally downscaled and spatially customizable
1019 climate data for historical and future periods for north america. *PloS one*, **11**, e0156720.
- 1020 Weber, W. (1998). New names and combinations in asteraceae: Heliantheae-ecliptinae. *Phytologia*, **85**,
1021 19–21.
- 1022 Wenzell, K.E., McDonnell, A.J., Wickett, N.J., Fant, J.B. & Skogen, K.A. (2021). Incomplete reproductive
1023 isolation and low genetic differentiation despite floral divergence across varying geographic scales in
1024 castilleja. *American Journal of Botany*, **108**, 1270–1288.
- 1025 Wilkinson, D.P., Golding, N., Guillera-Arroita, G., Tingley, R. & McCarthy, M.A. (2021). Defining and
1026 evaluating predictions of joint species distribution models. *Methods in Ecology and Evolution*, **12**, 394–
1027 404.
- 1028 Williams, P.H. (1982). The distribution and decline of british bumble bees (*bombus latr.*). *Journal of*
1029 *Apicultural Research*, **21**, 236–245. Retrieved from <https://doi.org/10.1080/00218839.1982.11100549>
- 1030 Wilson, A.M. & Jetz, W. (2016). Remotely sensed high-resolution global cloud dynamics for predicting

- 1031 ecosystem and biodiversity distributions. *PLoS biology*, **14**, e1002415.
- 1032 Wood, D.E., Lu, J. & Langmead, B. (2019). Improved metagenomic analysis with kraken 2. *Genome biology*,
1033 **20**, 1–13.
- 1034 Xie, Y., Wang, X. & Silander Jr, J.A. (2015). Deciduous forest responses to temperature, precipitation, and
1035 drought imply complex climate change impacts. *Proceedings of the National Academy of Sciences*, **112**,
1036 13585–13590.
- 1037 Xie, Y., Wang, X., Wilson, A.M. & Silander Jr, J.A. (2018). Predicting autumn phenology: How deciduous
1038 tree species respond to weather stressors. *Agricultural and Forest Meteorology*, **250**, 127–137.
- 1039 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,
1040 D.-Z. & Wang, H. (2019). The topological differences between visitation and pollen transport networks:
1041 A comparison in species rich communities of the himalaya–hengduan mountains. *Oikos*, **128**, 551–562.
1042 Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/oik.05262>

1043 List of Figures

1044	1	Phylogenetic tree of all biotically pollinated plant genera in the study area. The innermost	63
1045		ring indicates every genus which Queen Bee's were observed to visit. The intermediate ring	
1046		indicates that at least a single morphological pollen voucher slide was prepared for a member	
1047		of the genus. The outermost ring indicates that sequence data were available for at least a	
1048		member of that genus. Branch colors follow APG 4.	
1049	2	Modelled dates of when major flowering events occurred	64
1050	3	Number of the ten most commonly visited plants which are also in the top ten most common	
1051		sequences	65

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.

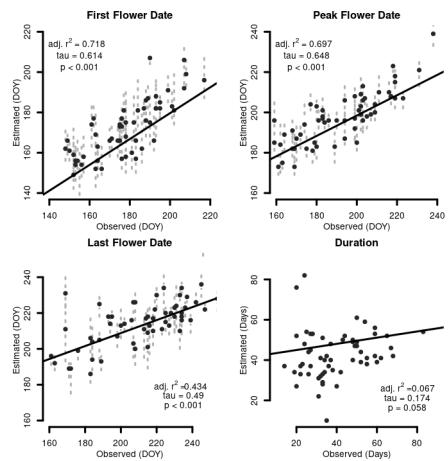


Figure 2: Modelled dates of when major flowering events occurred

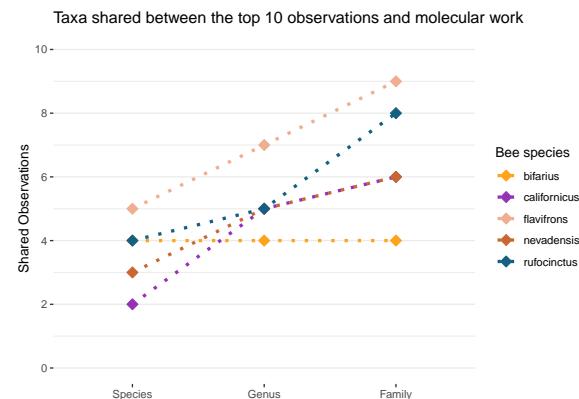


Figure 3: Number of the ten most commonly visited plants which are also in the top ten most common sequences

₁₀₅₂ **List of Tables**

₁₀₅₃ 2	Logistic regression assessing accuracy of SDMs	67
₁₀₅₄ 3	Species Distribution Modeling evaluation contingency table	68
₁₀₅₅ 4	Post classification of Sequences via Taxonomy and Ecology	69

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	Absence	25620	3838	11130	1653
	Presence	6614	28248	2758	12024

Table 4: Post classification of Sequences via Taxonomy and Ecology

Condition	No. Class.	Prcnt. Class.	Total Seqs	Rank
A	88	19.60	28.38	Species
B	126	28.06	8.94	Species
C	158	35.19	47.95	Genus
D	11	2.45	1.33	Species
E	38	8.46	4.61	Family
F	2	0.45	0.07	Species
X	26	5.79	8.71	Species