

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

³ 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. Ogilive^{3,4}, Emily J. Woodworth^{1,2},

⁴ **Abstract**

- 5 1) DNA Barcoding has been successful for the rapid analysis of ecological assemblages. Although
- 6 barcoding in the plant kingdom barcoding has been more difficult than others, and hence may begin
- 7 to lag behind other kingdoms.
- 8 2) Here we test the utilization of Angiosperms 353 probes to barcode plant species found in pollen
- 9 loads collected from Queen Bumble Bees.
- 10 3) To verify the accuracy for this barcoding system we compared the data to museum species, obser-
- 11 vation studies, and species distribution modelling to identify likely candidate species.
- 12 4) By utilizing Species distribution modelling we allow users to create a regionally appropriate sequence
- 13 databases which may use increase the alignment algorithms minimizing need for large computational
- 14 power, and run time.
- 15 5) We show that the Angiosperms 353 probes, which are currently being used in the largest ever plant
- 16 systematic endeavor, offers significant promise to metagenomic approaches.
- 17 6) Understanding plants in ecological contexts and understandings of their synecology.

*Correspondence: reedbenkendorf2021@u.northwestern.edu

¹⁸ 1 | INTRODUCTION

¹⁹ The inability to reliably identify plants to terminal taxon can limit our understanding of ecosystem function
²⁰ and interactions (Bortolus (2008)). This is especially true for genera where identification down to species
²¹ can be useful as specific bioindicators; defining ecological and behavioral properties (e.g. different species
²² of Sagebrush- *Artemisia* L., Willows - *Salix* L., and Sedges - *Carex* L.) (Gage & Cooper (2013)). In these
²³ instances the lack of species level data can hinder our understanding of the breadth of habitat which some
²⁴ species occupy, and their interactions with other species. This can be further complicated by the fact that the
²⁵ identification of organisms to terminal taxon is also often mired by lack of diagnostic characters(e.g. flowers,
²⁶ fruits, roots or combinations thereof), an increasing lack of taxonomic experts (Hebert *et al.* (2003)) and
²⁷ the presence of cryptic species (Janzen *et al.* (2017), Oliver *et al.* (2009)). Taxonomic verification can
²⁸ also be limited by the fact that revisiting field sites to identify material using morphological or chemical
²⁹ approaches, can be resource intensive and often does not work. The current methods to ameliorate this
³⁰ situation include: ignoring these ecologically relevant levels of detail, revisiting plots as diagnostic material
³¹ becomes temporally available, seeking the assistance from taxonomic specialists, or the use of barcoding
³² using molecular techniques (CITE)

³³ Recently molecular barcoding (the identification of a sample from a single organism *e.g.* a piece of leaf),
³⁴ or metabarcoding (the identification of a sample containing a mix of organisms *e.g.* soil), have shown
³⁵ considerable promise in many taxa (Ruppert *et al.* (2019)). For plants the success is a little more mix,
³⁶ with the identification of certain clades using barcoding being quite successful (Kress (2017)), however for
³⁷ many other clades the results have been more elusive (Liu *et al.* (2014), Group *et al.* (2011), Coissac *et al.*
³⁸ (2012)), while metabarcoding has incurred additional challenges for the currently available barcodes (Li *et*
³⁹ *al.* (2015), Kress & Erickson (2007), Group *et al.* (2009), Coissac *et al.* (2012)). Particular challenges for
⁴⁰ the utilization of the high copy number barcodes (*e.g.* ITS2, *rbcL*, *matK*, *trnH-psbA*) include their rates of
⁴¹ divergence, gene tree conflict, and hybridization (Coissac *et al.* (2016), Fazekas *et al.* (2009)).

⁴² Currently the largest plant systematic endeavor ever undertaken,by the Royal Botanic Gardens Kew, the
⁴³ Plant and Fungal Tree of Life (PAFTOL) is approaching completion (Baker *et al.* (2021a)). This data
⁴⁴ set will contain hybridization capture (Hyb-Seq) data from at least one species in each genus of the plant
⁴⁵ kingdom,14,000 represented species, using the popular Angiosperms353 (A353) probes, which includes 353
⁴⁶ single-copy orthologous loci, (Baker *et al.* (2021a), Johnson *et al.* (2019)). These publicly available data
⁴⁷ serve to provide a taxonomically comprehensive backbone for plant metabarcoding. The A353 probes are
⁴⁸ currently being used in many other plant phylogenetic studies increasing the sampling depth of many clades

49 (Baker *et al.* (2021b)). Data from the 10kP project, which seeks to develop reference genomes from a
50 phylogenetically diverse suite of plants will contribute many more species upon it's intended completion,
51 slated for 2030. Similar projects such as the 'Darwin Tree of Life' which will sequence all described taxa in
52 Britain and Ireland, seek to sequence high amounts of genomes in geographic regions will contribute data
53 sets applicable to enormous spatial domains (Cheng *et al.* (2018), Life Project Consortium *et al.* (2022),
54 Lewin *et al.* (2022)). These data will promote the ability to apply metabarcoding to resolve a diverse array
55 of questions relevant to theoretical and applied ecology (Kress (2017), Hollingsworth *et al.* (2016)). However,
56 the application of metabarcoding still face challenges relating to the enormity of the genomic data sets and
57 the computational power required to process sequence data.

58 Herein we have resolved major components of the problems of identifying plant material without diagnostic
59 morphological character states using the A353 Hyb-Seq probes (Johnson *et al.* (2019)), and custom species
60 sequence databases derived via species distribution modelling, and temporal filtering.

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

74 As species interactions vary both in space and time (@~THOMSON1994 GEO?, CaraDonna *et al.* (2021)).
75 Contrasts in the flowering periods of many plant species, can provide an additional filter for identifying
76 material in certain types of metagenomic samples (Janzen (1967), Newstrom *et al.* (1994)). In high elevation
77 temperate regions, pollination interactions vary temporally and are characterized by high turnover in active
78 periods of species (CaraDonna *et al.* (2017)), however the overall shorter extent of the active growing
79 season in these systems results in the presence of few to any natural breaks, which reduces the utility of
80 these to operate as filters in the post-processing of sequence matches. Nonetheless, we work develop a

81 general approach which seems applicable to many areas which utilize the temporal dimension for classifying
82 sequences in metagenomic samples (but see Davis *et al.* (2022)).

83 We tested this metagenomic and informatics approach to determine to examine the foraging behaviour of
84 queen bumblebee, and compare this approach with direct observations and the pollen record, which has shown
85 some incongruity in several floral visitation networks involving smaller bodied fauna (Barker & Arceo-
86 Gomez (2021), Zhao *et al.* (2019), Alarcón (2010)). The assessment of the plant species compositions from
87 pollen is a desirable approach, with several potential applications, despite numerous potential complications
88 (Poronon *et al.* (2017), Bell *et al.* (2017), Sickel *et al.* (2015), Bell *et al.* (2019), Suchan *et al.* (2019),
89 Johnson *et al.* (2021)). The two foraging phases of the queen bumblebee life cycle are essential to 1) increase
90 their weight before diapause, and 2) increase their ovary weights while establishing their recently found
91 nests. Both of these time periods represent potential demographic bottlenecks in bumble bee populations
92 (Sarro *et al.* (2022), ...). Bumblebees are one of the only groups of insects with unequivocal quantitative
93 evidence for numerous population declines, while simultaneously serving as the most effective pollinators
94 in many temperate montane ecosystems (Cameron & Sadd (2020), Goulson *et al.* (2008), Williams (1982),
95 Colla *et al.* (2012), Bergman *et al.* (1996), Bingham & Orthner (1998), Grixti *et al.* (2009)). These
96 areas often represent the most diverse areas in the temperate and often offer the sole potential refugia for
97 multiple dimensions of biodiversity under climate change, whilst simultaneously experiencing the greatest
98 proportional changes in mean annual temperature (Brito-Morales *et al.* (2018), Pepin *et al.* (2022)). An
99 immediate understanding of how to manage previously overlooked keystone insect species, such as bumble
100 bees, is essential if the refugial potential of the temperate mountains are to be incorporated into their current
101 diversity (Loarie *et al.* (2009), Dobrowski & Parks (2016)).

102 2 | METHODS

103 2.1 Study System & Field Work

104 Observations and bee sample collection was conducted at The Rocky Mountain Biological Laboratory
105 (RMBL; 38°57.5" N, 106°59.3" W (WGS 84), 2900 m.a.s.l.), Colorado, USA (APPENDIX 1 for site informa-
106 tion), characterized by high-montane/subalpine Parkland vegetation communities. Pollinator observations
107 of *Bombus* Latreille spp. (Apidae Latreille) were conducted from May 29th - July 23rd of 2015 in six study
108 sites. Observations of *Bombus* foraging took place for one hour at each field site in three 50m transects, each
109 transect was in a major vegetation type (dry, and wet meadows, and Aspen forest), where all flowers abun-

110 dances were estimated and placed into bins of abundance. Corbiculae loads were, non-lethally, collected
111 once from all Queen individuals encountered.

112 2.2 | Floral Visitation

113 ... does this need more info and warrant a section independent of the above?

114 2.3 | Pollen Morphological identification

115 2.3.1 | Pollen Reference Library

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

123 We used Divisive Hierarchical Clustering techniques to determine which plant taxa were distinguishable via
124 light microscopy, and to develop a dichotomous key to pollen morphotypes. Ten readily discernible categorical
125 traits were collected from each specimen in the image collection. These traits were transformed using Gower
126 distances, and clustered using Divisive Hierarchical clustering techniques (Maechler *et al.* (2022)). Using
127 the cluster dendrogram, elbow plot, and heatmaps (Hennig (2020)), of these results morphological groups
128 of pollen which could not be resolved via microscopy were delineated, and a dichotomous key was prepared
129 (APPENDIX NO.). This key was then used to identify the pollen grains sampled from corbiculae loads to
130 morphotypes in a consistent manner.

131 2.3.2 | Pollen Corbiculae Loads

132 To prepare the pollen slides from corbiculae, all corbiculae loads were broken apart and rolled using dissection
133 needlepoints to increase heterogeneity of samples. *Cerca* 0.5mm² of pollen was placed onto a ~4mm² fuchsin
134 jelly cube (Beattie (1971)) atop a graticulated microscope slide, with 20 transects and 20 rows (400 quadrants)
135 (EMS, Hartfield, PA). The jelly was melted, with stirring, until pollen grains were homogeneously spread
136 across the microscope slide. Slides were sealed with Canada Balsam (Rublev Colours, Willits, CA) followed

137 by sealing with nail polish to prevent oxidation; all samples are noted in *APPENDIX 3*. To identify the pollen
138 present in corbiculae loads, light microscopy at 400x (Zeiss Axioscope A1) was used. In initial sampling in
139 three transects, each pollen grain was identified to morphotype and counted; an additional two transects
140 were scanned for morphotypes unique to that slide, if either transect contained an unique morphotype than
141 all grains in that transect were also identified and counted. Subsequent to the first round of sampling, non-
142 parametric species richness rarefaction curves (Oksanen *et al.* (2022)), and non-parametric species diversity
143 rarefaction curves were used to assess the completeness of sampling (Chao *et al.* (2014), Hsieh *et al.* (2020)).
144 Slides not approaching the asymptote of the rarefaction curve were then re-sampled, and analysed iteratively
145 for up to a total of seven transects *APPENDIX 2*.

146 **2.4 | Molecular Barcoding**

147 **2.4.1 | Species reference list**

148 **2.4.1.1 Spatial Analyses** We generate a short list of potential candidate species we downloaded from the
149 Botanical Information and Ecology Network ‘BIEN’ (Maitner (2022)) all records adjacent to the field sites
150 to develop an ecologically relevant list of vascular plant species, with expected biotic pollination, which may
151 be present at the study area. To reduce the list of species to include in the genomic sequence databases, we
152 then generated Species Distribution Models (SDMs) for these taxa to predict their distribution throughout
153 the study area.

154 In order to minimize the number of species for which SDM’s were to be generated, BIEN was queried at a
155 distance of up to 100km from our study area and all plant species records were downloaded. To account
156 for the stochasticity of botanical collecting and offset the number of records associated with the research
157 station, this data set was bootstrap re-sampled 250 times, with 90% of samples selected, to create a testing
158 data set. The median of the logistic regression assessing the probability of occurrence of a species record as
159 a function of distance from the study area was used as a threshold distance, under which, to include species
160 as candidates for distribution modelling.

161 **2.4.1.2 Distribution Modelling** We used all occurrence records from BIEN ($n = 23,919$) within a 50km
162 border of the Omernik level 3 ecoregion, which includes the study area (*No. 21 “Southern Rockies”*) to
163 construct the species distribution model (Omernik (1987)). These records were copied into two, initially
164 identical, sets, one for generating machine learning models (ML; Random Forest, and Boosted Regression

165 Tree's), and the other for Generalised Linear (GLM) and Generalized Additive Models (GAM) (Barbet-
166 Massin *et al.* (2012)). Ensembled predictions have been shown to outperform their constituent models, on
167 average, and to reduce the ecological signal to the analytical noise of individual runs (Araujo & New (2007)).
168 No single method of producing SDMs has been shown to universally outperform others when faced with
169 a large and diverse number of applications, in our case a great number of species with differing biologies
170 and ecologies (Elith* *et al.* (2006), Qiao *et al.* (2015)). In the spirit of these findings, multiple families of
171 models, which can be generated together as they have similar requirements regarding the number and ratios
172 of Presence to Absence records were ensembled together (Barbet-Massin *et al.* (2012)).

173 We then generated 4,029 absence points, locations where the focal taxon is anticipated missing, through a
174 random stratification of 19% of the land cover in the area and included them in (Land Management (2019)).
175 To achieve a larger absence data set, we generated 1,000 pseudo-absence records for each taxon by randomly
176 selecting coordinates located at least 10km away from an occurrence record. For ML models, these pseudo-
177 absences were reduced so that the ratio of presence to absence records were balanced (Barbet-Massin *et al.*
178 (2012)). To achieve this, we removed absence records inside of 10% of the mean sample value of any predictor
179 variable the presence records; the required number of absence records were then randomly sampled.

180 To predict the potential distribution of each species we used 26 environmental variables at 30m resolution,
181 six related to climate, five soil, four topographic, four related to cloud cover, with the remaining reflecting
182 assorted abiotic parameters (Wilson & Jetz (2016), Wang *et al.* (2016), Hengl *et al.* (2017), Robinson *et al.*
183 (2014)) (**APPENDIX 6**). **These publicly available data sets, were selected as they** For linear
184 regression models these predictors underwent both *vifstep* (theta = 10, max observations = 12,500) and
185 *vifcor* (theta = 0.7, max observations = 12,500) to detect highly correlated variables, and collinear features
186 were removed leaving 16 variables (Naimi *et al.* (2014)).

187 Modelling: Random Forest and Boosted Regression Trees, were sub sampled with 30% test and two replicates
188 each before weighted ensemble based on True Skill Statistics (tss) (Naimi & Araujo (2016)). Generalised
189 linear models (GLM) and Generalised additive models (GAM) with 30% sub sampling and three replicates
190 each were also ensembled using the tss (Naimi & Araujo (2016), @). TSS was chosen as the ensemble
191 criterion as it has been shown to work across a wide range of species occurrences prevalence (Allouche *et*
192 *al.* (2006)). The results of these models were extracted on a cell-by-cell basis to a polygon feature derived
193 from a minimum-spanning tree which encompasses the study sites, and species from either ensemble with
194 greater than 50% mean habitat suitability across all cells were considered present for further purposes (Prim
195 (1957)).

196 A total of 535 species were modelled using Generalized Linear Models and Generalized Additive Models and

197 534 species were modelled using Random Forest and Boosted Regression Trees. To evaluate the accuracy of
198 the species distribution models, additional presence records from GBIF ($n = 61,789$), and AIM ($n = 12,730$)
199 were used as test and training sets ($n = 74,519$) for logistic regression (Occdownload Gbif.Org (2021), Land
200 Management (2019)). Additional novel absence records were generated from the AIM data set to create a
201 data set where each species has balanced presence and absences. Eleven or more paired presence and absence
202 records were required for this testing, resulting in 334 species being included in the logistic regression (Mdn
203 = 110.0, $\bar{x} = 223.1$, max = 1568 record pairs used) with a 70% test split (Kuhn (2022)).

204 **2.4.2 | Temporal Analyses**

205 To estimate the duration of dates in which plant species were flowering weibull estimates of several pheno-
206 logical parameters all spatially modelled taxa were developed (Belitz *et al.* (2020), Pearse *et al.* (2017)).
207 Only BIEN records which occurred in the Omernik Level 4 Ecoregions within 15km of the study area ($n =$
208 5 Level 4 Ecoregions, or conditionally 6 ecoregions if enough records were not be found in the nearest 5),
209 and which were from herbarium records were included. To remove temporally irrelevant herbarium records,
210 i.e. material collected during times which flowering is impossible at the study area due to snow cover, we
211 used the SnowUS data set (Iler *et al.* (2021), Tran *et al.* (2019)) from 2000-2017 were analyzed for the first
212 three days of contiguous snow absence, and the first three days of contiguous snow cover in Fall. Herbarium
213 records after the 3rd quantile for melt, and the 1st quantile for snow cover of these metrics were removed.
214 Species with > 10 records had their weibull distributions generated for the date when 10% of individuals had
215 begun flowering, when 50% were flowering, and when 90% of individuals had flowered, we used the initiation
216 and cessation dates, respectively, as effective start and ends of flowering.

217 **2.5.2 | Barcode references library**

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

220 **2.5.2.1 | Sampling Species for Barcoding** Using five years (2015-2020) of observational data on *Bombus*
221 Queen Bee foraging at these studies sites, we identified the plant taxa most frequently visited by Queens
222 across all years. We sequenced the 12 most commonly visited taxa twice using samples collected from one site
223 within the Gunnison Basin River Drainage and one individual collected from another more distal population.
224 In addition we included a congener - or a species from a closely related genus to serve as an outgroup for
225 all 12 taxa. In addition we sequenced another 15 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 typically using 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)).

2.5.2.2 | Plant Genomic DNA Extraction Plant genomic DNA was isolated from ~ 1 cm² 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).

2.5.2.3 | 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 (350µL , 100mM Tris-HCl, 50 mM EDTA, 50 mM NaCl, 10% SDS v/v., pH 7.5) was added followed by vortexing to allow dissolution of corbiculae. Pollen grains were then macerated with Kontes Pellet Pestles, and the tip of these washed with 130 µL of the SDS extraction buffer, samples were then incubated for 1 hour at 30°C. This was followed by the addition of 10% CTAB solution (450ul, of 20 mM Tris-Cl pH. 8.0, 1.4 M NaCl, 10 mM EDTA pH 7.5, 10% CTAB, 5% PVP, ~85% Deionized water) and RNase (10 uL of 10 mg/mL) and samples were incubated for 40 minutes at 37°C, on heat block (Multi-Blok, Thermo Fisher Scientific, Waltham Massachusetts) set to 40°C. After 20 minutes incubation, Proteinase K (15 µL of 20mg/ml) and DTT (12.5 µL of 1M in water) were added, and the samples were further incubated at 60°C for 1 hour. Samples were then incubated overnight at 40°C. 500 µL of Phenol-Chloroform-Isoamyl alcohol (25:24:1) were added, vortexed, and centrifuged at 10,000 rpm for 10 minutes and the aqueous phase was pipetted to a 1.5 ml centrifuge tube.

To precipitate the DNA, chilled Isopropyl alcohol & 3 mM Sodium acetate (5:1) equivalent to $\frac{2}{3}$ of the volume of sample were added, with 1 hour of chilling at -20°C, followed by 10 minutes of centrifuging at 13,000 rpm. The supernatant was pipetted to a new 1.5 ml centrifuge tube, and 70% EtOH (400 µL) were added before chilling at -20°C for 20 minutes followed by centrifugation at 13,000 rpm for 10 minutes. Both tubes were then washed with 75% EtOH (400 µL), inverted, centrifuged at 13,000 rpm for 4 minutes, and the solution discarded, then washed with 95% EtOH (400 µL) , inverted, centrifuged at 13,000 rpm for 4 minutes, and the solution discarded. Pellets were dried at room temperature overnight before resuspension in nuclease free H₂O. Extractions were assessed using a Nanodrop 2000 (Thermo Fisher Scientific) and Qubit

256 fluorometer (Thermo Fisher Scientific). DNA extracts were then cleaned using 2:1 v./v. Sera-Mag beads
257 (Cytiva, Little Chalfont, UK) to solute ratio following the manufacturer's protocol, eluted in 0.5x TE, and
258 the eluent allowed to reduce by half volume in ambient conditions. DNA was quantified using a Qubit
259 fluorometer.

260 **2.5.2.4 | Fragmentation, Library Preparation & Target Enrichment** Library preparation was
261 performed using the NEBNext Ultra II FS-DNA Library Prep Kit for Illumina (New England BioLabs,
262 Ipswich, Massachusetts, USA) using slightly modified manufacturers recommendation. Fragmentation was
263 performed 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
264 DNA. Adapter Ligation and PCR enrichment were performed with $\frac{1}{2}$ volumes, while cleanup of products
265 was performed using SPRI beads (Beckman Coulter, Indianapolis, Indiana, USA) and recommended volumes
266 of 80% v./v. ethanol washes. The exception was the herbarium specimens which were not fragmented and
267 only end repaired, with similar library preparation of all samples. Products were analysed on 4% agarose
268 gels, and a Qubit fluorometer. Libraries were pooled and enriched with the Angiosperms 353 probe kit V.4
269 (Arbor Biosciences myBaits Target Sequence Capture Kit) by following the manufacturer's protocol and
270 Brewer et al. 2019. Sequencing was performed using an Illumina mi-Seq with 150-bp end reads, (NUSeq
271 Core, Chicago, Illinois).

272 **2.6 | Computational Processes and Analyses.**

273 **2.6.1 | Reference Library Data Processing** Sequences were processed using Trimmomatic, which
274 removed sequence adapters, clipped the first 3 bp, discarding reads less than 36 bp, and removing reads
275 if their average PHRED score dropped beneath 20 over a window of 5 bp (Bolger & Giorgi (2014), Tange
276 (2021)). Contigs generated were mapped to a reference with HybPiper with using target files created by
277 M353 (Johnson *et al.* (2016), McLay *et al.* (2021)).

278 **2.6.2 | Sequence Identification** A custom Kraken2 database was created by downloading representative
279 species indicated as being present in the study area by the spatial analyses from the Sequence Read Archive
280 (SRA) NCBI (Wood *et al.* (2019)). These sequences were processed in the same manner as our novel
281 sequences. The Kraken2 database was built using default parameters. Kraken2 was run on sequences using
282 default parameters (*APPENDIX 5*). Following Kraken2, Bracken was used to classify sequences to terminal
283 taxa (Lu *et al.* (2017)). Results from both Kraken2 and Bracken, results were reclassified manually to
284 identify terminal taxa. For example, when only a single species of a genus was known in the study area, but

285 our database used a representative of another taxon in the genus, this species was coded as the result.

286 **2.6.3 | Identification of Sequence Matching Loci** A local NCBI database was built using the same
287 processed novel and downloaded sequences as the previous database (Camacho *et al.* (2009)).

288 **3 | RESULTS**

289 **3.1 | Floral Observations**

290 The six sites were surveyed once weekly from May 27-July 27 for a total of 52 hours from . A total of 723
291 queen-pollen foraging interactions were observed (range per bee species by week range = 1-18, $\bar{x} = 3.46$,
292 Mdn = 2), with a range of total observed interactions per bee species across this time period (min = 1, \bar{x}
293 = 59.08, Mdn = 19, max = 184). Plants varied widely in the number of interactions which they partook in
294 with each species of bee (range per plant species by week min = 1 - 20, $\bar{x} = 3.51$, Mdn = 2), with a range
295 of total observed interactions per plant species over this time period (min = 1, $\bar{x} = 20.26$, Mdn = 4, max =
296 141). The number of plant species which bees were observed interacting with varied more narrowly (range
297 = 1 - 18, $\bar{x} = 8$, Mdn = 6).

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

299 [Figure 1 about here.]

300 **3.1 | Spatial Analyses**

301 [Table 1 about here.]

302 [Table 2 about here.]

303 The median (25.009 km) of the logistic regression assessing the probability of occurrence of a species record as
304 a function of distance from the study area was used as a threshold distance to include species for distribution
305 modelling. A 2-sample test for equality of proportions with continuity correction ($X^2 = 13.254$, df
306 = 1, p-value = 0.000136, 95% CI 0.04-1.00) was used to test whether more of the records located in the
307 broad ecological sites present at the field station, between the distance of the median (25.009 km) to the
308 third quantile (ca 43.830 km) of the regression distance, were true presences at the field station. Including

309 these records would have resulted in modelling an additional 222 species distributions of which 30 are true
310 presences, these taxa were not modelled.

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

313 In the area of the minimum-spanning tree encompassing the field sites, of the 554 vascular plants with biotic
314 pollination syndromes, the 493 ML ensembles accurately predicted the presence of 362 (65.3%), incorrectly
315 predicted the presence of 64 (11.6%), incorrectly predicted 34 true presences (6.1%) as being absent, and
316 correctly predicted the true absence of 33 (6.0%). The balanced accuracy of the ensembled models is 0.627
317 (Sensitivity = 0.340, Specificity 0.914). Of the 554 vascular plants with biotic pollination syndromes, the
318 475 LM ensembles accurately predicted the presence of 286 (51.6%), incorrectly predicted the presence of
319 41 (14.3%), incorrectly predicted 93 true presences (16.8%) as being absent, and correctly predicted the
320 true absence of 55 (9.9%). The balanced accuracy of the ensembled models is 0.664 (Sensitivity = 0.573,
321 Specificity 0.754). Of the 554 vascular plants with biotic pollination syndromes in the flora 13 (2.3%) were
322 in the Orchid family and 41 (7.4%) are non-natives, both of which are restricted from the database, and can
323 only reduce the number of true predicted presences by roughly 10%.

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

329 3.2 | Microscopic Pollen identification

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

336 **SUMMARY REQUIRED - Number of species from number of families, and how many species**
337 **are identified to species versus number only to genera**

339 **3.3 | Metabarcoding Pollen Identification**

340 **3.3.1 | Spatial Analyses to identify candidate taxa**

341 The median (25.009 km) of the logistic regression assessing the probability of occurrence of a species record as
342 a function of distance from the study area was used as a threshold distance to include species for distribution
343 modelling. A 2-sample test for equality of proportions with continuity correction (χ^2 -squared = 13.254, df
344 = 1, p-value = 0.000136, 95% CI 0.04-1.00) was used to test whether more of the records located in the
345 broad ecological sites present at the field station, between the distance of the median (25.009 km) to the
346 third quantile (ca 43.830 km) of the regression distance, were true presences at the field station. Including
347 these records would have resulted in modelling an additional 222 species distributions of which 30 are true
348 presences, these taxa were not modelled.

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

351 In the area of the minimum-spanning tree encompassing the field sites, of the 554 vascular plants with biotic
352 pollination syndromes, the 493 ML ensembles accurately predicted the presence of 362 (65.3%), incorrectly
353 predicted the presence of 64 (11.6%), incorrectly predicted 34 true presences (6.1%) as being absent, and
354 correctly predicted the true absence of 33 (6.0%). The balanced accuracy of the ensembled models is 0.627
355 (Sensitivity = 0.340, Specificity 0.914). Of the 554 vascular plants with biotic pollination syndromes, the
356 475 LM ensembles accurately predicted the presence of 286 (51.6%), incorrectly predicted the presence of
357 41 (14.3%), incorrectly predicted 93 true presences (16.8%) as being absent, and correctly predicted the
358 true absence of 55 (9.9%). The balanced accuracy of the ensembled models is 0.664 (Sensitivity = 0.573,
359 Specificity 0.754). Of the 554 vascular plants with biotic pollination syndromes in the flora 13 (2.3%) were
360 in the Orchid family and 41 (7.4%) are non-natives, both of which are restricted from the database, and can
361 only reduce the number of true predicted presences by roughly 10%.

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

367 **3.3.2 | Temporal Analysis**

368 The first date of modeled snow melt in the Gothic area ($n = 17$, $\bar{x} = 137.9$, Mdn = 135, 3rd quartile =
369 151), and the first date of a consistent winter snow base ($n = 17$, $\bar{x} = 299.9$, Mdn = 300, 1st quartile =
370 291) from 2000-2017, were used as delimiters for the inclusions of herbarium records in modelling. Of the
371 439 species predicted likely present in the area via logistic regression, 332 species (64.4%) with more than 10
372 records in the focal level 4 ecoregions ($\bar{x} = 35.016$, Mdn = 35, max = 96) had weibull estimates calculated,
373 an additional 56 species (11.2%) with enough contributing records from the “Sedimentary Mid-Elevation
374 Forests”, a large ecoregion generally just beneath the elevation bands occupied by the five ecoregions around
375 the study area had weibull estimates also calculated ($\bar{x} = 13.868$, Mdn = 13, max = 24).

376 Only 58 of these 388 species ($n = 34.568$, Mdn = 31) were able to be compared to plot based observational
377 data from the long term (1974–2012) data set (CaraDonna *et al.* (2014)). Of these species relatively high
378 accord was observed between the long-term ground truthed data set, and the modelled species. There was
379 very strong evidence that the weibull estimates were positively associated with the observed onset ($r^2 =$
380 0.72, $p < 0.0001$, $\tau = 0.61$) and peak ($r^2 = 0.70$, $p < 0.0001$, $\tau = 0.65$) of flowering, and that the number
381 of herbarium samples had a moderate effect on the estimates ($p = 0.004$ and $p = 0.034$ respectively). There
382 was very strong evidence that the weibull estimates had a positive association with the observed cessation
383 of flowering ($r^2 = 0.4339$, $p < 0.0001$, $\tau = 0.489$), however there was no evidence that sample size had an
384 effect ($p = 0.349$). There was moderate evidence that the weibull estimates, with an effect of sample size,
385 had a weak positive association with the observed duration of flowering ($p = 0.0401$, $r^2 = 0.07$, $\tau = 0.17$).

386 [Figure 3 about here.]

387 **3.3.1 | Molecular analysis of corbiculae loads**

388 The 54 corbiculae loads had DNA extracted and underwent various steps towards hyb-seq, in the end a total
389 of 44 corbiculae samples were sequenced, 7,752,353 reads were recovered from sequencing. The number of
390 reads per sequence varied widely (range = 76 - 508,795, $\bar{x} = 176,189.8$, Mdn = 138,395). Of the possible 353
391 loci, the number which were recovered from each sample, and informative to BLAST were range = 24 - 353,
392 $\bar{x} = 305.5$, Mdn = 331. The number of reads per loci from across all samples had a range of 178 - 506,653,
393 $\bar{x} = 20,688$, Mdn = 12,616. **APPENDIX X Reads Per Loci.**

394 ... with samples 22, and 41 both having < 100 reads indicates, virtual failure of these records (REMOVE
395 from analyses).

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

402 [Table 3 about here.]

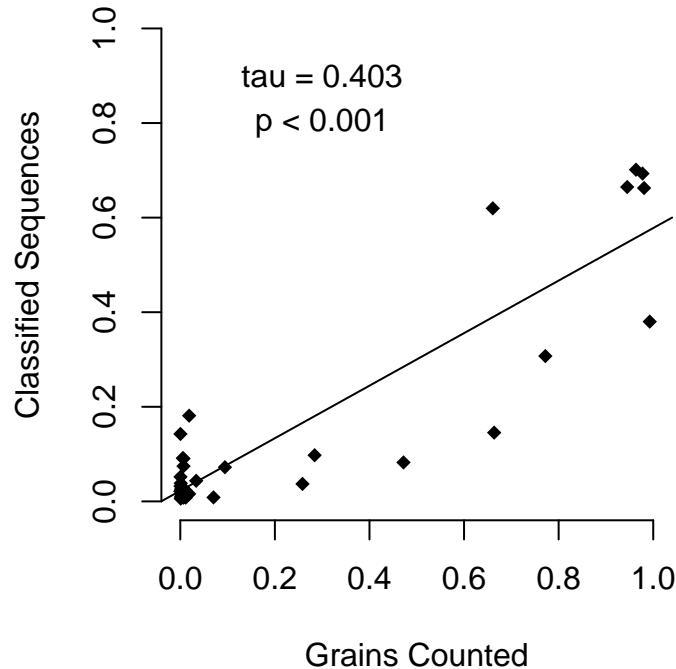
403 The initial classification of sequences which were made by BLAST were reviewed programmatically, using
404 predicted presence of the species (from spatial modelling), modelled flowering time (from temporal mod-
405 elling), and taxonomy (from existing sources). A sequential process was utilized which reassigned sequences
406 based on binary combinations of the factors above (Appendix XX). Given the relative sparsity of the number,
407 and relatedness, of species represented in the sequence database this was performed to: 1) Identify locally
408 present species represented by surrogates in the DB 2) Reduce false classifications of focal species 3) Identify
409 high confidence sequence matches. Of the top ten taxa which were identified by BLAST for the 438 distinct
410 records, 57.08% of the reads were classified to a species 46.39% representing of all classified reads, 39.04% of
411 the reads were classified to genus representing 46.39% of all classified reads, and 3.88% of the records were
412 classified to family 2.23% of all total sequences.

413 Of the 18 classifications which were assigned to genera without any species predicted by spatial analyses, were
414 investigated by hand after post-processing steps. These were all assigned via post-processing conditions (B:
415 10, E: 7, F: 1, APPENDIX XX). These were manually assigned to a variety of ranks, occasionally to genus -
416 10, and species - 6, by consultation of the alpha-taxonomic literature (Sadeghian *et al.* (2015), Sennikov &
417 Kurtto (2017), Rabeler & Wagner (2016), Pusalkar & Singh (2015), Moore & Bohs (2003), Weber (1998)).

418 To determine at which level species in pollen loads could be detected the results of light microscopy were
419 compared to the molecular results. The pollen samples contained three morphotypes which could readily
420 be identified via microscopy. Two of these mapped to the clades (Boraginaceae & Heliantheae Alliance),
421 and one to a Asteraceae less Heliantheae. Boraginaceae grains were detected in 92.9% of samples where the
422 proportion of target grains were between 0.01-1 ($n = 14$ Mdn = 0.572). Asteraceae type 1, non-helianthoids,
423 were detected in 20% of samples where the proportion of target grains were between 0.001-0.01 ($n = 5$ Mdn
424 = 0.002) Asteraceae type 2, Helianthoids, were detected in 62.5% of samples where the proportion of target
425 grains were between 0.001-0.01 ($n = 8$ Mdn = 0.003); however, Asteraceae were detected in 76.9% of samples
426 where the proportion of target grains were between 0.001-0.01 ($n = 13$ Mdn = 0.002). Both morphotypes

427 of Asteraceae pollen were detected in 100% of samples where the proportion of target grains were between
428 0.01-1 ($n = 3$ Mdn = 0.011), and Ericaceae were detected in 50% of samples where the proportion of target
429 grains were between 0.001-0.1 ($n = 2$ Mdn = 0.01).

Correlation of Proportion Counted Grains and Sequence Reads



430
431 To detect whether the sequencing reads were semi-quantitative the subset of all pollen morphotypes distin-
432 guishable by microscopy were compared to the sequence reads. In all instances sequence reads were pooled
433 to the highest taxonomic rank associated with the morphotype, e.g. if both species of *Mertensia* Huth, or
434 one species and read only classified to genus were present in a sample, the reads were summed. The total
435 percentage of the ten most abundant grains per sample were then were then relativized to constitute the
436 entire sample.

437 The relationship between the number of pollen grains in a sample and the number of sequence reads is roughly
438 *curvilinear*, where grains which are present in trace amounts are overestimated by sequence counts, while
439 grains present in high amounts are underestimated. This is likely due to the proportion of high false positives
440 which occur in the classification process with NGS (BELL NOVEMBER 2021). There was evidence of a
441 strong correlation between the proportion of grains per morphotype and the number of sequences per group
442 (0.403, $p < 0.0001$, $n = 37$).

443 To ascertain the extent to which records of multiple species in a family, which were suspected to be sampling
444 artefacts occurred in molecular samples an index of similarity, ala jaccard, the affinity index was used

445 to assess co-occurrence (Mainali *et al.* (2022), Mainali & Slud (2022)). Numerous taxa from the family
446 Ranunculaceae Jussieu (*Caltha* L. sp., *Thalictrum* L. spp., *Trollius* L. sp., *Aquilegia* L. spp.), had α scores
447 which indicated that they are only present when a more common confamilial taxa *Delphinium barbeyi* (Huth)
448 Huth *nuttallianum* Pritz. were recorded. A similar relationship was observed in the Hydrophyllaceae R.Br.
449 with samples placed in *Nemophila* Nutt., which only occurred when the more abundant *Hydrophyllum* L.
450 species were present.

451 3.6 | Integrated Observational, Molecular, and Palynological Network

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

455 For example a many sequences which mapped to the Asteraceae family, but which was flagged by temporal
456 filters and is present in both *B. nevadensis* Cresson and *B. rufocinctus* Cresson pollen is most likely *Frasera*
457 Walter, which failed extractions for the reference library failed (APPENDIX XX). A similar likely mismatch
458 could be between what was fide molecular evidence as *Agastache pallidiflora* (A. Heller) Rydb. but where
459 feeding was infrequently observed on *Pedicularis* L., likely due to this entire order being represented by only
460 a single molecular reference species.

461 Situations where SDM's led to incorrect results at the species level are evident with classification to *Scabrethia*
462 *Scabra* (Hooker) W.A. Weber, this match almost certainly representing *Wyethia arizonica* A. Gray (Weber
463 (1998)), a taxon known to be visited by Queen bee's via our floral observations.

464 It is not unlikely that much of the difference in the results between the observational and molecular work
465 are attributable to the challenges in detecting rare events in these smaller sizes. For example, no more than
466 10 bee corbiculae loads per species were sequenced with the Mdn = 5.5 . . . , and the median of interactions
467 with the top 5 plant sizes constituted 0.9135611 of the top.

468 An expected inaccuracy of the classification scheme is in genus level placements, e.g. were *Epilobium* L.
469 (Onagraceae Juss.) spp. were classified. However, given the small size of their flowers in the study area, these
470 results more likely indicate that a species of *Chamaenerion* Seg. (a segregate genus) such as *C. angustifolium*
471 (L.) Scop. or *latifolium* (L.) Sweet is occasionally utilized, as it supported by limited palynology data.

472 Accordingly, combining the results of floral observations, and palynology, molecular sequencing - both pre
473 and post processing, we subjectively developed re-classifications of the contents of pollen grains . . .

474 4 | DISCUSSION

475 ~ **What we DEMONSTRATED** ~ We have demonstrated how the Angiosperms533 hyb-seq probes may
476 be used for plant barcoding in a metagenomic context (Johnson *et al.* (2019), Hollingsworth *et al.* (2016)).
477 This was exemplified in an ecologically relevant scenario, where the results have immediate implications
478 for natural history driven fundamental science and land management. The test pollen loads contained
479 a number of closely related taxa, some in notoriously morphologically difficult clades with rapid rates of
480 diversification (e.g. *Mertensia*, *Lupinus* L.), at naturally occurring proportions (Nevado *et al.* (2016), Nazaire
481 & Hufford (2014)). We incorporated spatial and temporal approaches for creating custom sequence databases
482 an approach which is readily applicable to any lab group with the capacity to perform next-generation
483 sequencing across the entirety of multiple continents, and which we expect to be highly beneficial in many
484 study areas. By combining insights from these novel approaches with an extensive observational field based
485 study we show how these methods may be applied to test a variety of hypotheses related to ecological
486 interactions.

487 ~ **What CHALLENGES we FACED** ~ The SDM's which we generated, with relatively few occurrence
488 records and few modelling iterations, performed beyond expectations, likely due to the utility of the predictor
489 variables and strong alignment of vegetation by orographic precipitation in the study area. However, we had
490 difficulties in evaluating our predictions in an operational context. We utilized the database query approach,
491 to only model species with a high probability of not being dispersal limited to the focal area, and focused
492 on a relevant subset of many of these species ranges to reduce the contributions of range wide adaptions
493 on habitat (Sork (2018), Joshi *et al.* (2001)). While the models worked well compared to both test, and
494 validation with external point data, moving from points to polygon features was more difficult. We were able
495 to compare our results to 1) a Flora, 2) lists of plants used by Bumble Bees at plots; the former inappropriate
496 in that it contained a great number of species which we sought to use modelling to reduce *e.g.* all strictly
497 alpine species, and the latter inappropriate in that it contained only species relevant to *Bombus* but had no
498 official 'absence' data. Further given the size of the minimum spanning tree (AREA???) which we extracted
499 points to, a formal floristic inventory would still be a time intensive process. Accordingly, we expect the
500 real results of our data lay somewhere in between these two evaluations; with an excess of species predicted
501 present (Dubuis *et al.* (2011), Calabrese *et al.* (2014), Pinto-Ledezma & Cavender-Bares (2021)), but few
502 enough that they lend themselves to metabarcoding. We observe that our models seemed very capable of
503 effectively identifying alpine species and removing them in binomial contexts. Difficulties in temporal models
504 related to variability in drivers of flowering phenology.

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

509 Results from palynological analyses show that several species of bee show near perfect fidelity to the genus
510 *Hydrophyllum* on a per visit basis... General results show high congruence between foraging and molecular
511 results, indicating that concerns regarding mismatch between observational networks need not persist with
512 *Bombus* studies...

513 ~ **WHAT this tells us about Bee foraging (APPLIED)** ~ Some foraging preferences of *Bombus*,
514 both at this field site and across a great many localities globally emerge from this work, which reiterates the
515 needs for land managers to maintain relatively high amounts of members of the Fabaceae, Boraginaceae, and
516 Ranunculaceae, in Western North American montane landscapes (Goulson *et al.* (2005), Goulson (2010),
517 Liang *et al.* (2021), Bontsutsnaja *et al.* (2021)). Numerous historic, and some ongoing, land management
518 practices reduce the ability of many landscapes to support stable populations of *Bombus*. Historic livestock
519 grazing was often associated with the targeted removal of many species of plants which are known to have
520 compounds toxic to cattle. In particular, the removal of locoweeds (Fabaceae: *Astragalus* L. & *Oxytropis* DC.)
521 and larkspurs (Ranunculaceae: *Delphinium*) were common across public lands administered by the United
522 States Forest Service (Ralphs & Ueckert (1988), Aldous (1919), Ralphs *et al.* (2003)). Further actions,
523 generally initiated by early settlers, involved the channelization and incising of streams, culling of beavers,
524 and leaving cattle concentrated on higher order stream banks for significant periods of time, all processes
525 which lower the water tables and reduced the extent of stream-associated [riverine] wetlands and the mesic
526 meadows fringes which provide habitat for many species of tall *Mertensia* (Boraginaceae, e.g. *M. ciliata* Torr.
527 G. Don.) widely distributed across Western North America, and to an extent *Delphinium barbeyi* and many
528 species of native *Trifolium* L. (Dahl (1990), Naiman *et al.* (1988), Belsky *et al.* (1999), Cooke & Reeves
529 (1976)). Fire suppression further resulted in the succession of many Aspen (*Populus tremuloides* Michx.)
530 groves to Conifer stands, decreasing the mosaic of age structured habitats in many landscapes, adversely
531 effects habitat for tall *Mertensia* species and several species of *Delphinium* (Brewen *et al.* (2021), Keane
532 (2002)). Finally the effects of Nitrogen deposition, especially given the West's rapidly growing population
533 still pose adverse effects on the abundance of a variety of species of Fabaceae at Urban-Rural interfaces
534 (see Stevens *et al.* (2018), Fenn *et al.* (2003)). Current solutions to these issues, involve targeted burns,
535 reintroduction of beavers and beaver habitat analogs, and the possibility of re-seeding a variety of 'locoweeds'

536 and ‘larkspurs’ in areas now seldom used, or only used for early, grazing. The highly enthusiastic response of
537 land managers, and homeowners, to plant *Asclepias* L., using genetically appropriate materials, to improve
538 Monarch Butterfly (*Danaus plexippus* L.) habitat provides an effective framework for the latter (Oberhauser
539 *et al.* (2015), Basey *et al.* (2015)).

540 ~ **WHERE we see spatial/temporal going** We have concerns regarding the number of persons training
541 to become and practice botany, and grave concerns regarding the funding mechanisms for floristic and field
542 based botanical research and for centralized authorities to produce consensus opinions on alpha taxonomy
543 (Prather *et al.* (2004b), Kramer & Havens (2015), Prather *et al.* (2004a), Crisci *et al.* (2020), Manzano
544 (2021), Stroud *et al.* (2022)). To reduce the effects of a low population density of botanists on the mainte-
545 nance of and production of flora’s and to foster meta-genomics across landscapes without field stations we
546 utilized Species Distribution Modelling to generate predictive species lists. In this proof of concept example
547 we performed several iterations of modelling runs, and several approaches (i.e. the ‘linear models’, and the
548 ‘machine learning’), which took notable amounts of compute power. We suspect the possible deleterious
549 nature of this endeavor may be reduced by: 1) more field surveying by crews will reduce the need to generate
550 as many species 2) fewer runs of models, 3) only running machine learning models which do not require an
551 explicitly process to reduce spatial autocorrelation. However, given the time required to perform all aspects
552 of a study, even our amount of computation was negligible. Further, we are very optimistic about the pos-
553 sibility for persons to perform these tasks, as mentioned we utilized roughly only one quarter of the records
554 which were digitally available for presence, and we suspect others will have enough records to perform this
555 process nearly anywhere else in the temperate. In certain scenarios modelling of predicted species via more
556 formally tailored S(tacked)-SDM or J(oint)-SDM approaches may be beneficial (Wilkinson *et al.* (2021),
557 Pinto-Ledezma & Cavender-Bares (2021), Schmitt *et al.* (2017)).

558 Tandem to the lack of continued expertise required to generate and maintain species lists, is the expertise
559 required to continue tracking when major phenological events occur in many plant species at relatively fine
560 scales or under novel climates. Knowledge of these events is currently limited to general time periods of only
561 a handful of phenological events and groups of organisms (e.g. flowering initiation, or trees) (Prather *et al.*
562 (2004a), Li *et al.* (2016)). While many programs and initiatives exist to collect phenological information on
563 subsets of easily identifiable charismatic species to detect major trends in phenology, these capture only a
564 subset of the extent diversity (Betancourt *et al.* (2005), Havens *et al.* (2007)). In many instances it appears
565 that while landscapes respond similarly to environmental variables which predict phenological responses,
566 that individual species vary widely in their responses to similar environmental cues, or respond to different

567 cues (Augspurger & Zaya (2020), Xie *et al.* (2015), Xie *et al.* (2018), CaraDonna *et al.* (2014)). As can
568 be seen here, predictions of when a single, major phenological event occurs is already data
569 limited, with sample size having an effect on the subset of species which we could even generate
570 weibull estimates for. ; check assumptions of model again and just do a quasi something A
571 more promising approach for the tropics may lay in circular statistics (Park *et al.* (2022)).

572 ~ WHERE we see MOLECULAR going The nearly complete Plant and Fungal Tree of Life (PAFTOL)
573 will provide a comprehensive phylogenetic backbone of the entire plant kingdom, and the inclusion of
574 A353 probes with lineage specific probe sets is common in producing massive genetic datasets (Baker *et*
575 *al.* (2021b)). We predict that the A353 probes which it is utilizing to work nearly immediately for DNA
576 barcoding of whole plant material, and that more elaborate validation studies in controlled metabarcoding
577 settings, utilizing existing experimental designs, will have favorable results (Bell *et al.* (2017), Bell *et al.*
578 (2019), Bell *et al.* (2021), Lamb *et al.* (2019)). In particular the harvesting of loci with more variation
579 in certain lineages, and or with more variable flanking regions, will prove promising for identifying closely
580 related plant material (CITE). We suspect that conserved reaches of genes resulted in the high amounts
581 of reads in somewhat obscure species. Given that the A353 loci are nuclear, single copy, and a variety are
582 present the possibility of identifying target loci for quantitative purposes is high, without continual PCR
583 enrichment is possible; this would align with relatively high efficacy of WGS (Lang *et al.* (2019), Peel *et al.*
584 (2019), Bell *et al.* (2021)). Recent evidence indicates that the potential for identifying nearly cryptic taxa
585 and even infra-specific inference, of either whole plant material, and perhaps in metagenomic context are
586 possible (Ottenlips *et al.* (2021), Wenzell *et al.* (2021), Loke *et al.* in prep, Slimp *et al.* (2021), Beck *et al.*
587 (2021)). We further believe that in synthetic phylogenetic trees - with incorporation of NGS backbones - will
588 allow in automatic reassignment of reads as a function of phylogenetic distance with measures of uncertainty
589 (Hinchliff *et al.* (2015), Smith & Brown (2018), Baker *et al.* (2021a)).

590 5 | CONCLUSION

591 We believe that the combination of spatial and temporal models, united and guided by localized natural
592 history knowledge, provides the essential components of a bayesian framework for approaching the coarse
593 elucidation of ecological interactions using DNA Barcoding. Herein we crudely utilized this thinking via
594 binary outcomes, should a species predicted be predicted present or not? Is it unequivocally flowering
595 or not? Myriad data show biological systems and ecological interactions have more variance than can be

596 reasonably discretely parsed. We expect that within a bayesian framework studies of pollinator behavior
597 may be enacted via this approach at a landscape level, e.g. the scale of an entire drainage basin such as the
598 Gunnison which is quickly becoming one of the worlds few model ecosystems. We hope that the promise of
599 A353 probes as tools for metabarcoding play a role in these endeavors.

600 **AUTHOR CONTRIBUTIONS:** R.C.B conducted botanical collections, conducted all molecular lab
601 work, lead all analyses, and writing. J.E.O conceived, designed, and conducted all ecological fieldwork,
602 assisted with analyses, and writing. E.J.W. prepared, imaged, and collected trait data on pollen reference
603 slides, and assisted with analysis of trait data and writing a dichotomous key. S.T. assisted with spatial
604 analyses and writing. P.J.C assisted with ecological analyses and writing. J.B.F. conceived, and designed all
605 lab work, analyses, and integration of approaches, assisted with writing, and secured funding for molecular
606 work.

607 **ACKNOWLEDGMENTS:** Nyree Zerega for assistance obtaining herbaria loans and accessioning our
608 collections at CHIC. Pat Herendeen for assistance with virtually all aspects of preparing pollen vouchers and
609 the identification process. Hilary Noble, Zoe Diaz-Martinez, Angela McDonnell, & Elena Loke for assistance
610 with genomic library preparation. Ian Breckheimer for sharing the SDM predictor variables. We thank
611 the curators at the following herbaria for supplying tissue: Ben Legler at Stillinger (ID), Charles (Rick)
612 Williams at Ray J. Davis (IDS), (B)Ernie Nelson at Rocky Mountain (RM); and the collectors: D. Knoke,
613 L. Brummer, J. Boyd, C. Davidson, I. Gilman, M. Kirkpatrick, S. McCauley, J. Smith, K. Taylor, & C.
614 Williams. David Giblin & Mare Nazaire for sharing relevant sections of an advanced draft of FNA V. 15.
615 The Bureau of Land Management is thanked as many plant specimens were collected by R.C.B as a partner
616 or contractor to the agency; Sarah Burnett and Lauren Price are thanked for sharing AIM data. Sanda and
617 New England Biotech are gratefully acknowledged for technical support and generously sharing samples.
618 T.C.H. Cole for sharing the Angiosperm Phylogeny 4 colour palette. The Program in Plant Biology and
619 Conservation is thanked for funding. The holdings of the following herbaria were essential for this project:
620 AK, ALTA, ASU, BABY, BC, BM, BMO, BOON, BRIT, CANB, CAS, CHSC, CM, CMN, CNS, COLO,
621 CONN, CS, CSU, DAV, DBG, DES, ENCB, F, FR, G, GH, GZU, IAC, K, KR, KSP, KSTC, KU, LD,
622 LOB, LSU, MA, MACF, MEL, MICH, MIL, MIN, MNHN, MO, MO, MT, MW, NCSC, NSW, NY, NYBG,
623 O, OBI, PI, RBG, RSA, SD, SDSU, SFV, TENN, TRT, UA, UAC, UAM, UAZ, UBC, UBC, UCR, UCS,
624 UCSB, UMO, UNM, UPS, US, USCH, USF, USU, UTEP, UWBM, V, VT, W, WSCO, WU, XAL, YPM,
625 Z.

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

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

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

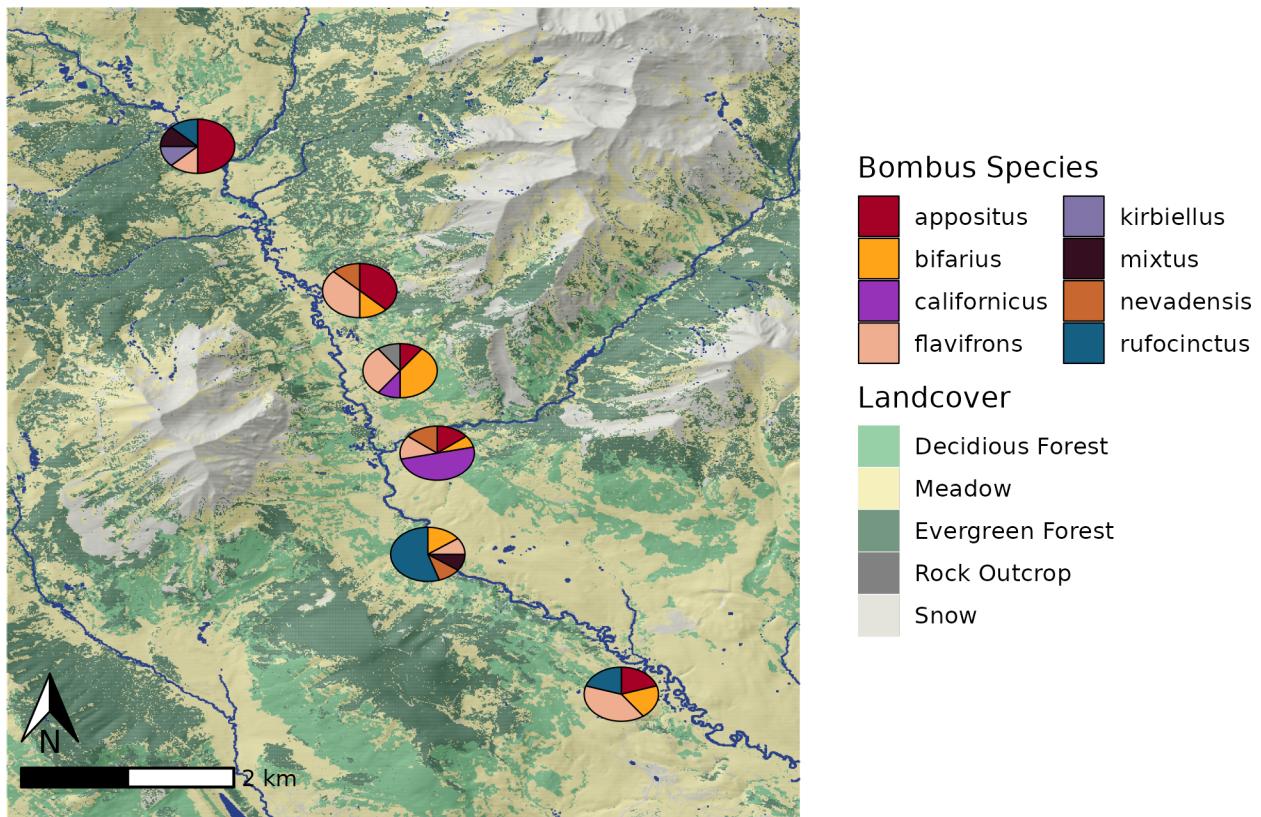
630 **ORCID**

631 Paul CaraDonna <https://orcid.org/0000-0003-3517-9090>
632 Jeremie Fant <https://orcid.org/0000-0001-9276-1111>
633 Jane Ogilvie <https://orcid.org/0000-0001-8546-0417>
634 Sophie Taddeo <https://orcid.org/0000-0002-7789-1417>

635 **References**

636 **Supporting**

Origins of Corbiculae Loads



Upper East River Valley, Colorado

639 Appendix 2 - Species Distribution Models Predictors

Layer	LM	Description	Source
1.	N	Mean annual cloudiness - MODIS	Wilson et al. 2016
2.	Y	Cloudiness seasonality 1 - MODIS	Wilson et al. 2016
3.	N	Cloudiness seasonality 2 - MODIS	Wilson et al. 2016
4.	Y	Cloudiness seasonality 3 - MODIS	Wilson et al. 2016
5.	N	Beginning of the frost-free period	Wang et al.
6.	N	Climatic moisture deficit	Wang et al.
7.	N	Degree-days above 5C from	Wang et al.
8.	N	Mean annual precipitation	Wang et al.
9.	Y	Mean annual precipitation as snow	Wang et al.
10.	Y	Temperature seasonality	Wang et al.
11.	Y	2015 Percent Grass/Herbaceous cover - MODIS	(MOD44B)
12.	Y	2015 Percent Tree cover from Landsat 7/8	(GLCF)
13.	Y	Soil probability of bedrock (R Horizon)	SoilGrids
14.	N	Soil organic carbon (Tonnes / ha)	SoilGrids
15.	N	Surface soil pH in H ₂ O	SoilGrids
16.	Y	Surface soil percent sand	SoilGrids
17.	Y	Soil USDA class	SoilGrids
18.	N	Topographic elevation	EarthEnv DEM
19.	Y	Topographic elevation, moving window.	EarthEnv DEM
20.	Y	Topographic percent slope	EarthEnv DEM
21.	Y	Topographic wetness index	EarthEnv DEM
22.	Y	Topographic aspect	EarthEnv DEM
23.	Y	Annual potential solar radiation computed	r.sun
24.	N	Estimated actual (w/-cloud) solar radiation	r.sun / Wilson et al. 2016
25.	Y	Log-transformed distance to surface water	Global Surface Water Explorer
26.	Y	Percent surface water	Global 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>Helianomeris 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>Mertenia ciliata</i> (James ex Torr.) G. Don	Boraginaceae	ID 1754185	S	Idaho, Valley	18.VI.2018	tba	979.3
<i>Mertenia 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>Campanula rotundifolia</i> L.	Campanulaceae	RMH 720600	P	Colorado, Gunnison	9.VII.1997	tba	38.9
<i>Lathyrus lanszwertii</i> Kellogg var. leucanthus (Ryd.) Dorn	Fabaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Lathyrus lanszwertii</i> Kellogg var. leucanthus (Ryd.) Dorn	Fabaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Lupinus argenteus</i> Pursh	Fabaceae	CHIC tba	P	Nevada, Pershing	29.V.2018	tba	971.2
<i>Lupinus argenteus</i> Pursh	Fabaceae	ISU 10387	P	Colorado, Gunnison	29.VI.2010	tba	0.2
<i>Lupinus bakeri</i> Greene	Fabaceae	ISU 10142	P	Colorado, Gunnison	15.VIII.2010	tba	2.6
<i>Vicia americana</i> Muhl. ex Willd.	Fabaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Vicia americana</i> Muhl. ex Willd. var. minor Hook.	Fabaceae	CHIC tba	S	Montana, Carbon	4.VII.2019	tba	10020.8
<i>Frasera speciosa</i> Douglas ex Griseb	Gentianaceae	RMH 721930	P	Colorado, Gunnison	20.VI.1997	tba	66.2
<i>Frasera speciosa</i> Douglas ex Griseb	Gentianaceae	RMH 719305	P	Colorado, Gunnison	7.VII.1997	tba	19.8
<i>Hydrophyllum capitatum</i> Douglas ex. Benth	Hydrophyllaceae	RMH tba	P	Colorado, Mesa	30.VI.2011	tba	64.6
<i>Hydrophyllum capitatum</i> Douglas ex. Benth	Hydrophyllaceae	RMH tba	P	Colorado, Delta	8.VI.2011	tba	65.3
<i>Hydrophyllum fendleri</i> (Gray) Heller	Hydrophyllaceae	ID 161100	P	Washington, Yakima	9.VI.2008	tba	1429.7
<i>Hydrophyllum fendleri</i> (Gray) Heller	Hydrophyllaceae	ID 164040	P	Idaho, Idaho	27.V.2009	tba	1014.4
<i>Agastache pallidiflora</i> (Heller) Rydberg	Lamiaceae	CHIC tba	S	Arizona, Coconino	17.VII.2020	tba	617.7
<i>Chamerion angustifolium</i> (L.) Holub	Lamiaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Delphinium barbeyi</i> (Huth) Huth	Ranunculaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Delphinium nuttallianum</i> Pritz.	Ranunculaceae	ID 166162	P	Idaho, Gem	15.VI.2011	tba	9825.5
<i>Delphinium nuttallianum</i> Pritz.	Ranunculaceae	ID 179376	P	Idaho, Gooding	29.IV.2017	tba	733.7
<i>Potentilla fruticosa</i> Pursh	Rosaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Potentilla fruticosa</i> Pursh	Rosaceae	CHIC tba	S	Colorado, Gunnison	18.VII.2020	tba	3.6
<i>Potentilla hippiana</i> Lehman.	Rosaceae	CHIC tba	S	New Mexico, Catron	15.VIII.2020	tba	573.8

(Continued on Next Page)

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

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

^a Accession includes both Herbarium and Accession number

^b Pres. refers to Preservation method. 'S' denotes silica gel dried, 'P' denotes pressed

^c All Localities are in the United States of America

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

(Continued on Next Page)

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.

⁶⁴⁷ 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)

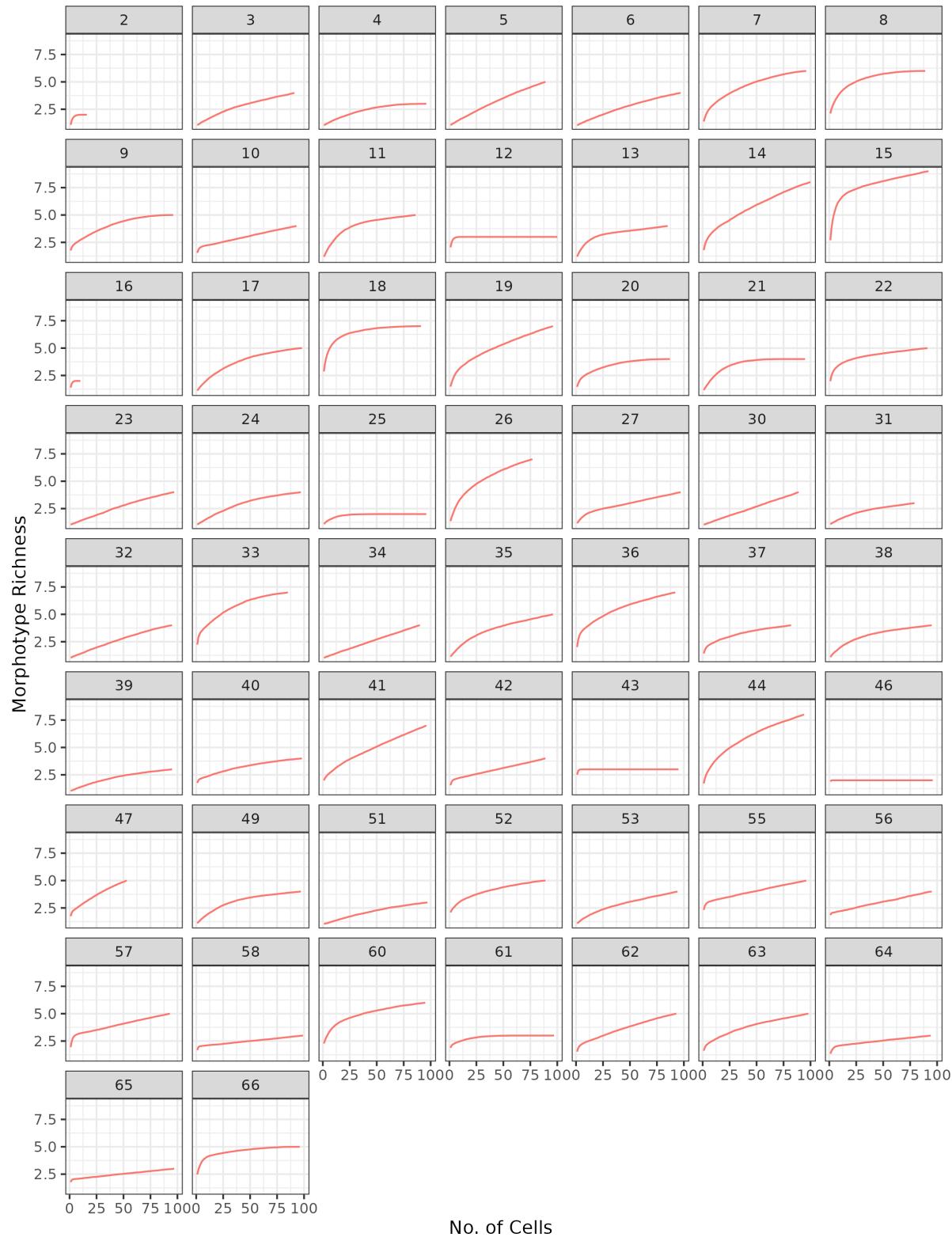
Morphotype A: Trifolium, Lupinus, Glycrrhiza, Mitella, Geum

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

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



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

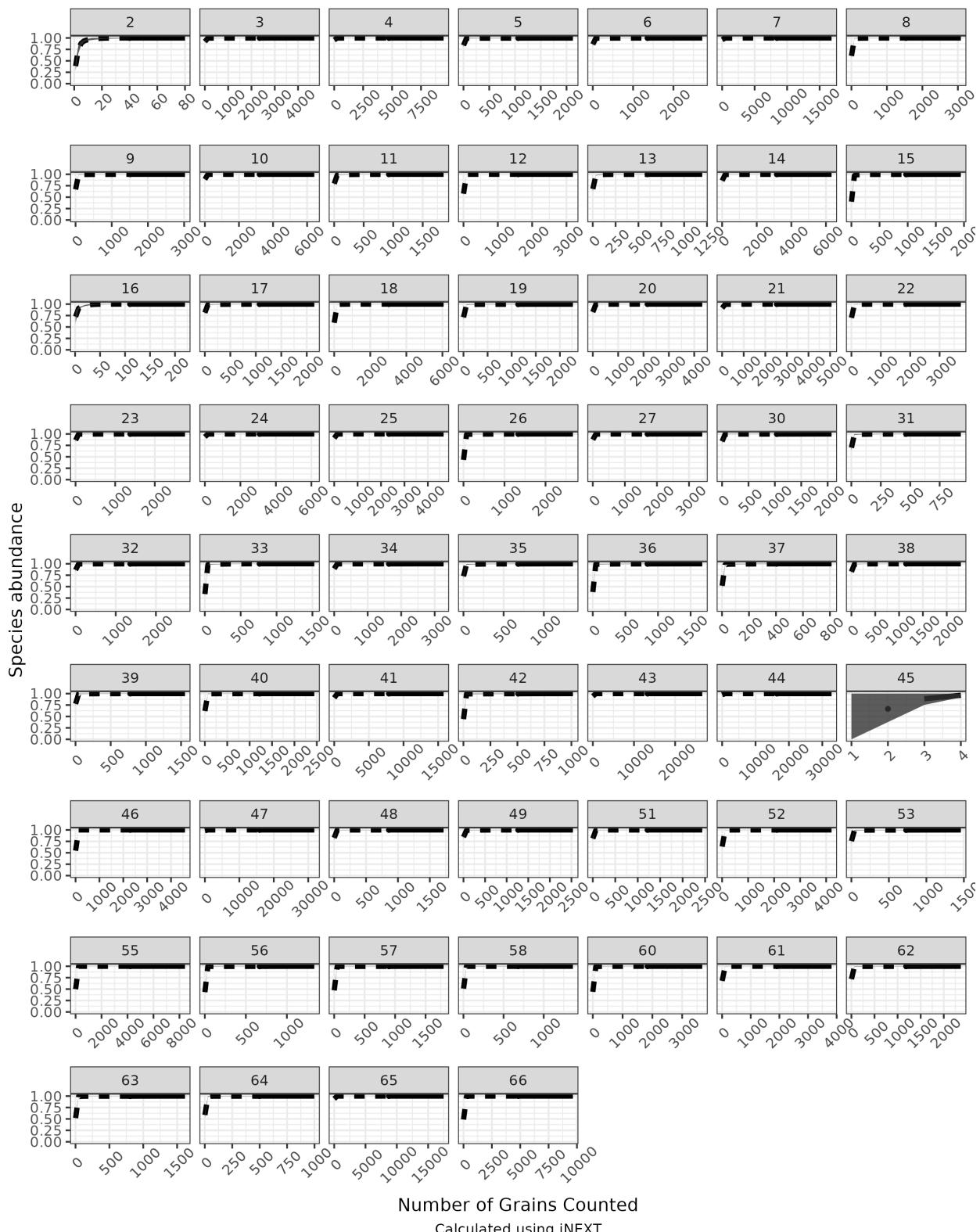


Table 1: All species present in the Reference Sequence Databases
(Kraken and BLAST)

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 ecuadoriensis</i> <i>Erigeron grandiflorus</i> <i>Erigeron rosulatus</i> <i>Erigeron uniflorus</i> <i>Helianthella quinquenervis</i> <i>Heterotheca villosa</i> <i>Hieracium avilae</i> <i>Hieracium jubatum</i> <i>Hymenoxys hoopesii</i> <i>Leucanthemum graminifolium</i> <i>Microseris lindleyi</i> <i>Omalotheca supina</i> <i>Packera quercetorum</i> <i>Pseudognaphalium attenuatum</i> <i>Pseudognaphalium frigidum</i> <i>Pseudognaphalium lacteum</i> <i>Pseudognaphalium oxyphyllum</i> <i>Rudbeckia hirta</i> <i>Scabrethia scabra</i> <i>Senecio adenophyllus</i> <i>Senecio algens</i> <i>Senecio apolobambensis</i> <i>Senecio candollei</i> <i>Senecio chionogeton</i> <i>Senecio formosus</i> <i>Senecio funcii</i> <i>Senecio gilliesii</i> <i>Senecio humillimus</i> <i>Senecio nutans</i> <i>Senecio puchei</i> <i>Senecio rufescens</i> <i>Senecio spinosus</i> <i>Senecio tephrosioides</i>

(Continued on Next Page)

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

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

(Continued on Next Page)

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

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

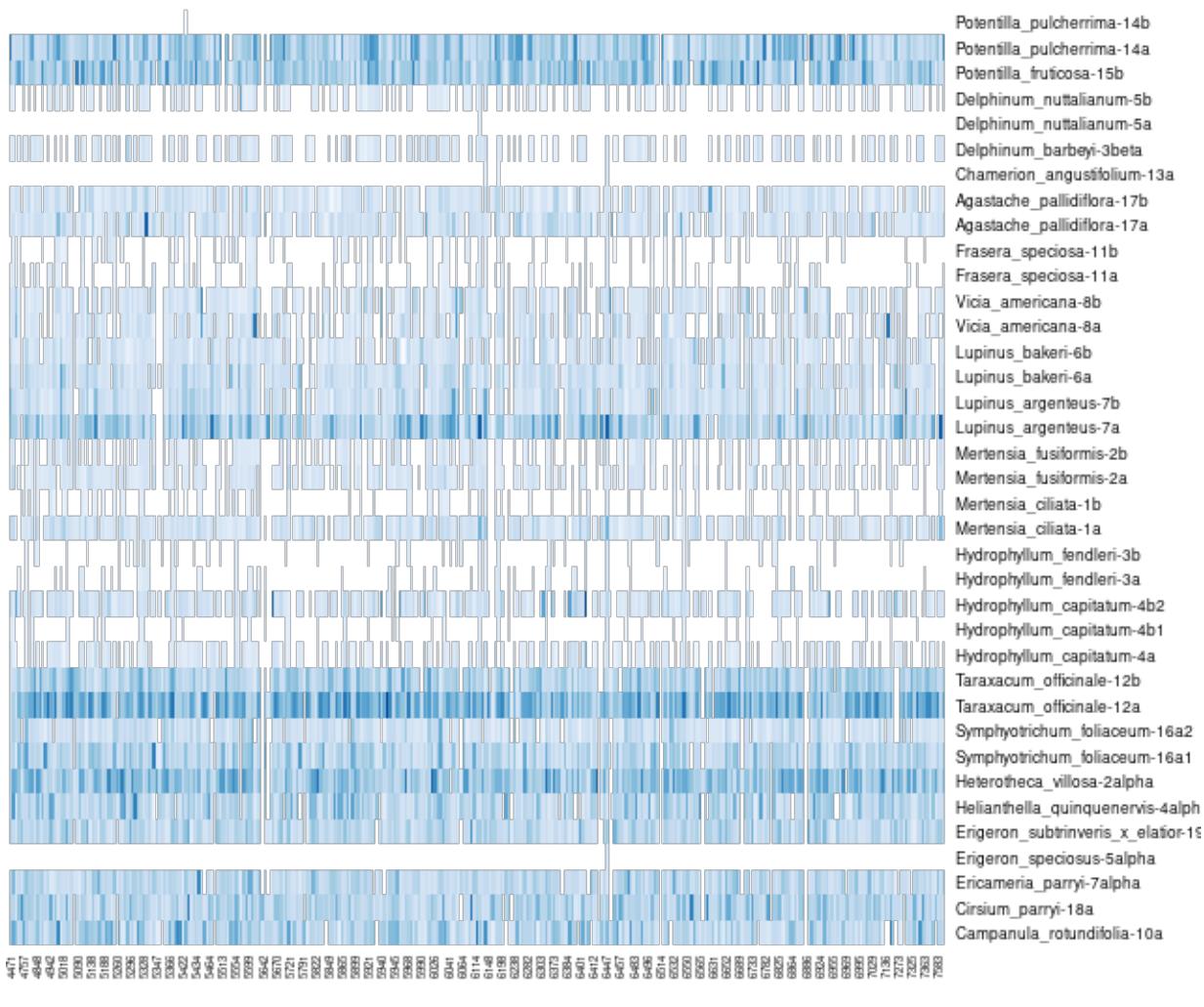
(Continued on Next Page)

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

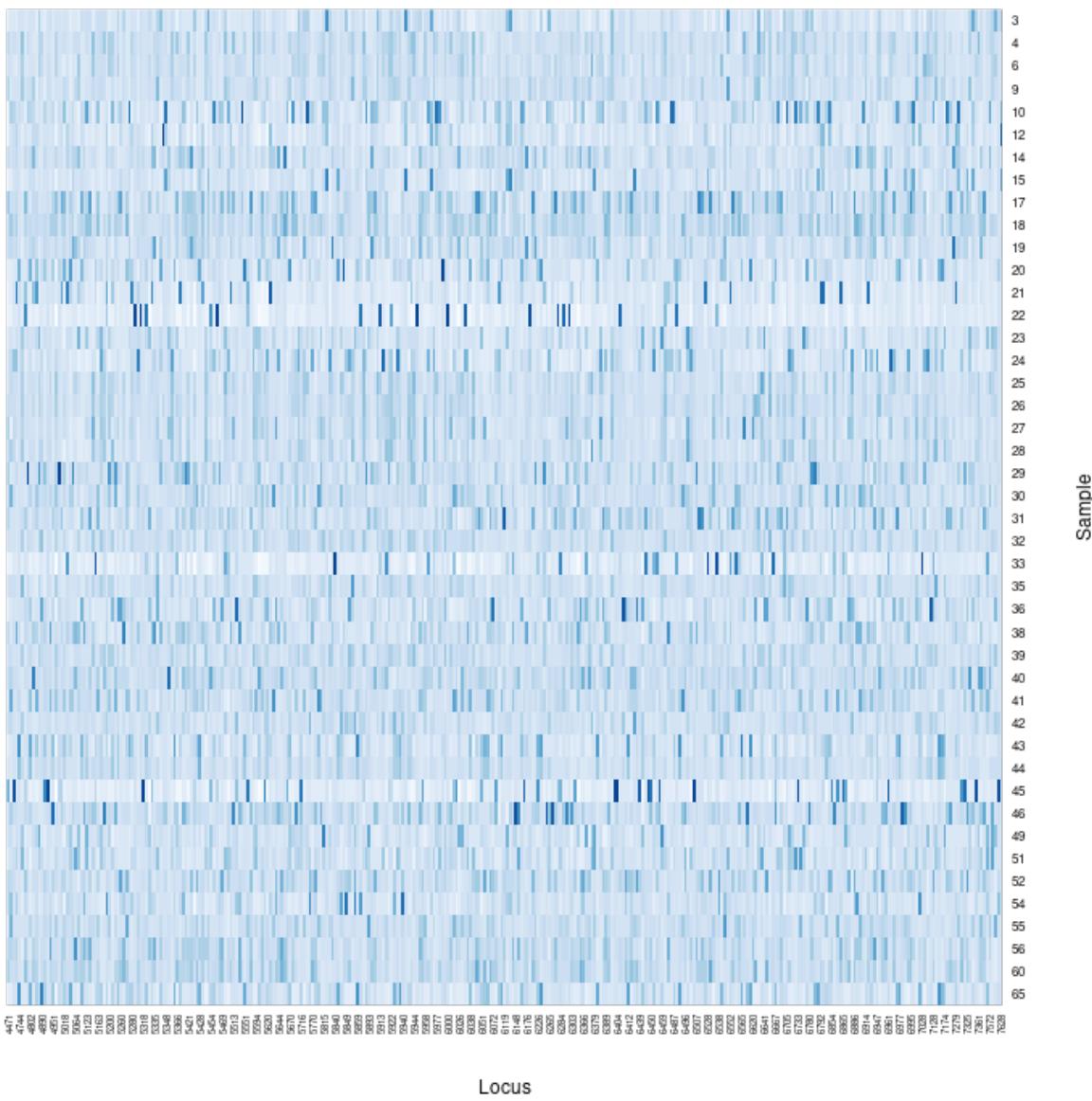
Order	Family	Taxon
		Saxifraga maderensis
		Saxifraga oppositifolia
		Saxifraga portosanctana
		Saxifraga x geum

x geum* \end{longtable}

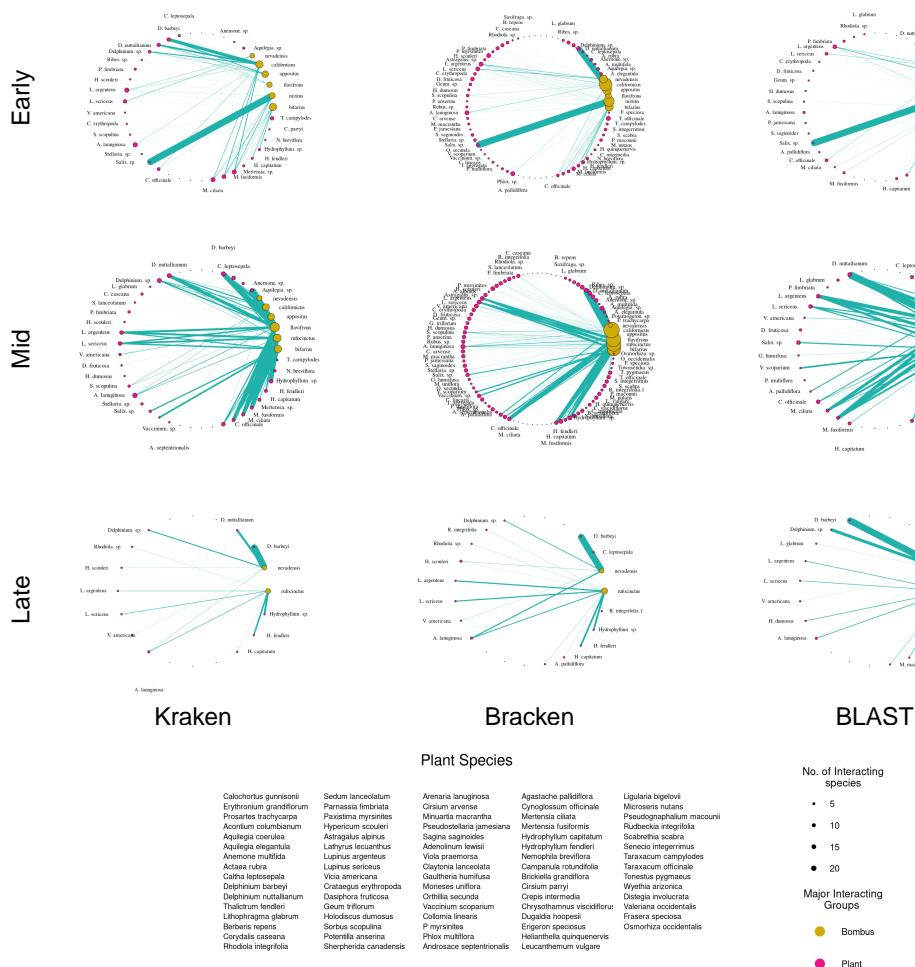
Loci & Nucleotides Returned per Reference Sample



Percent matched reads



Comparision of Foraging Patterns from Three Sequence Alignment Algorithms



667 Appendix XX - Models used for Species Distribution Model Ensembles

668 The two machine learning models utilize Ensemble learning.

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

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

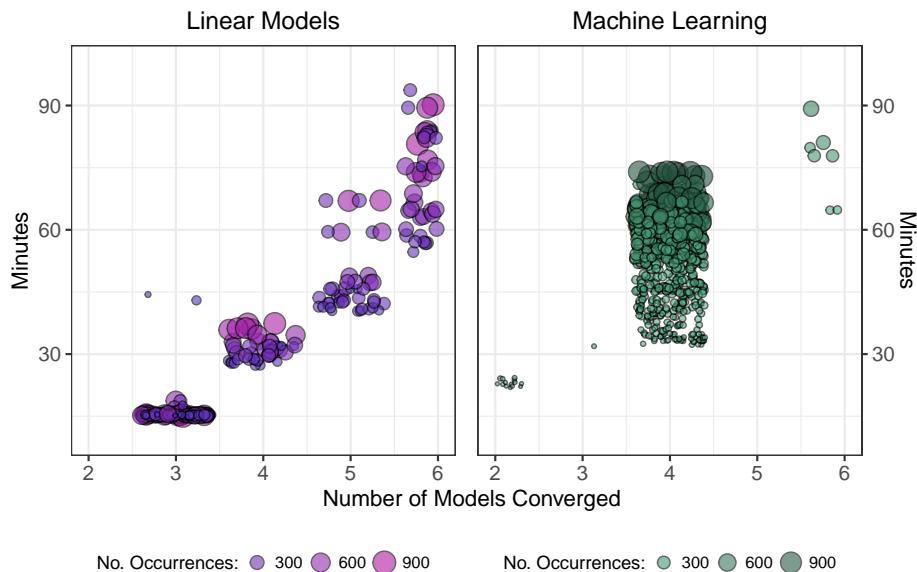
681 **Boosted Regression Tree (BRT)** (or Gradient Boosted tree) An initial tree is grown, and all other trees
682 are derived sequentially from it, as each new tree is grown the errors in responses from the last tree are
683 weighed more heavily so that the model focuses on selecting dependent variables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

685 **Bias** predictions from an algorithm are systematically in error due to being prejudiced for or against certain
686 results, due to assumptions during learning.

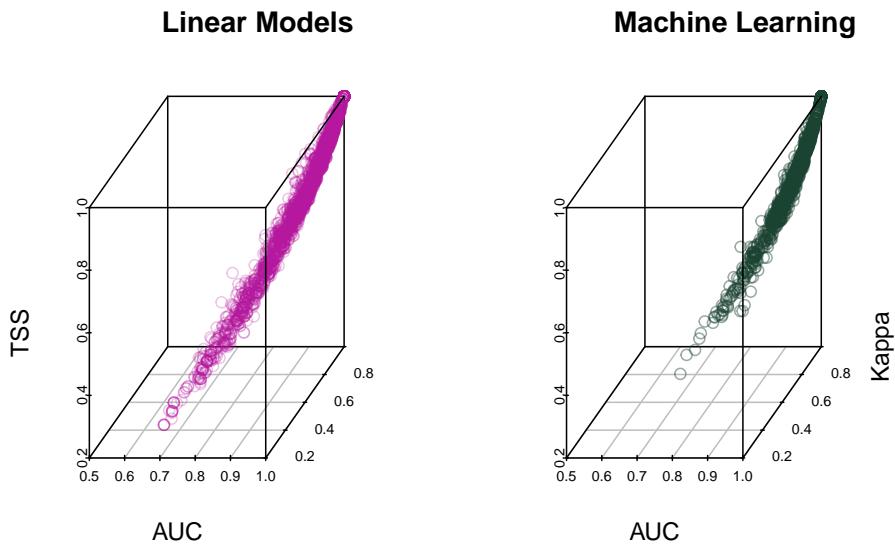
687 **Variance** errors in models due to an over-reliance and sensitivity of training to outliers in training data.

688 In general, Random Forest models have high bias and low variance, where boosted regressions trees have lower
689 bias and higher variance. Theoretically, the weaknesses and strengths of bootstrap aggregation (bagging) as
690 implemented by Random Forests are supplemented by the boosting.

Time Spent Fitting and Projecting Models onto Gridded Surfaces



Collectively it took 215 hours for all of the GLM and GAM to run, and for the converged models to be ensembled, and predicted onto a raster surface; it took 419 hours for the same process to be carried out for the Random Forest and Boosted Regression Tree models.



Results for each converged individual model which were then ensembled, using weights from the True Skill Statistic (TSS).

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.

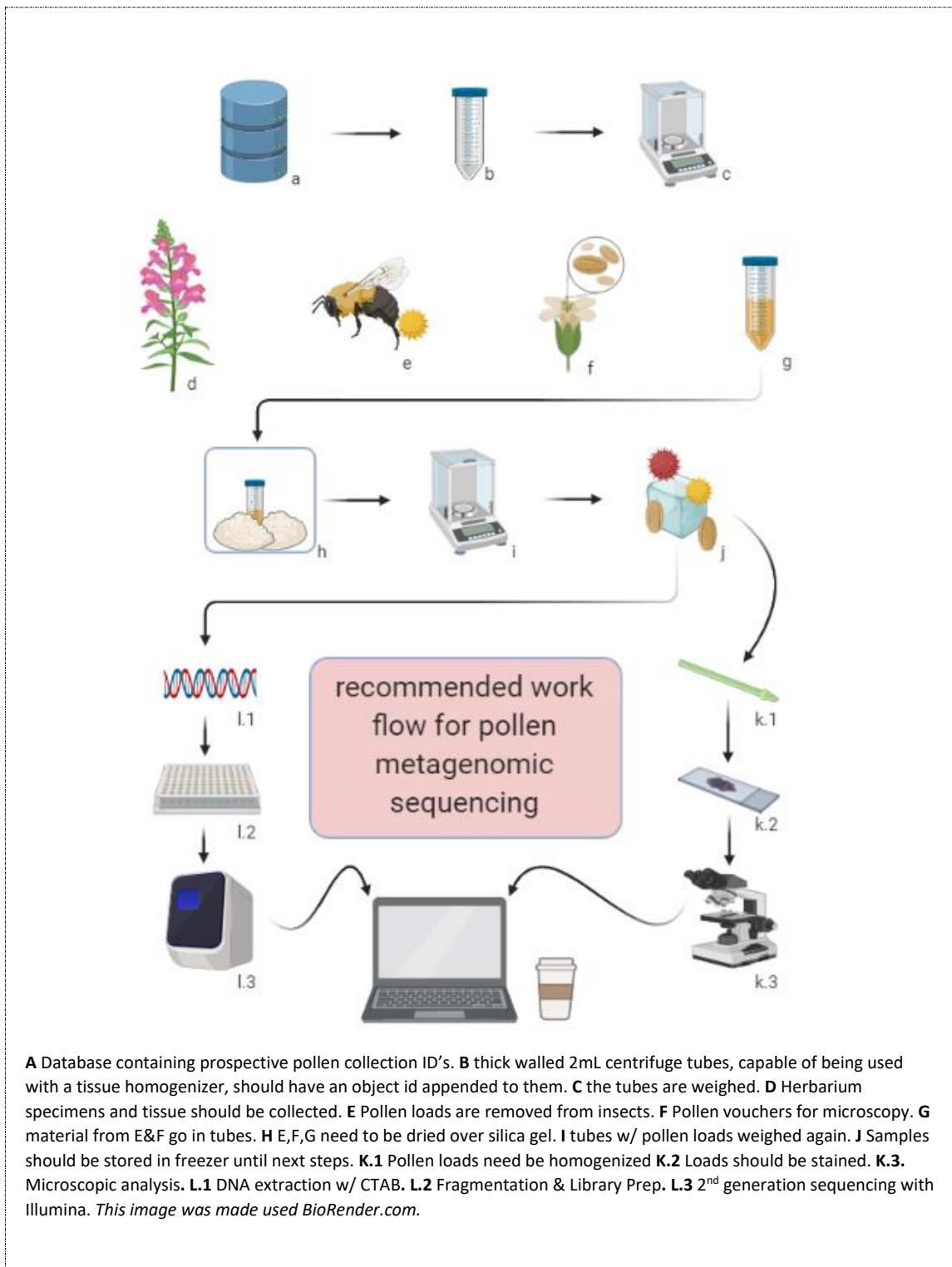
$$\begin{aligned} \text{Spatial} == 1 \& \text{ Temporal} == 1 \sim \mathbf{A} \\ \text{Spatial} == 1 \& \text{ Temporal} == 0 \& \text{ Congener} = 1 \sim \mathbf{B} \end{aligned}$$

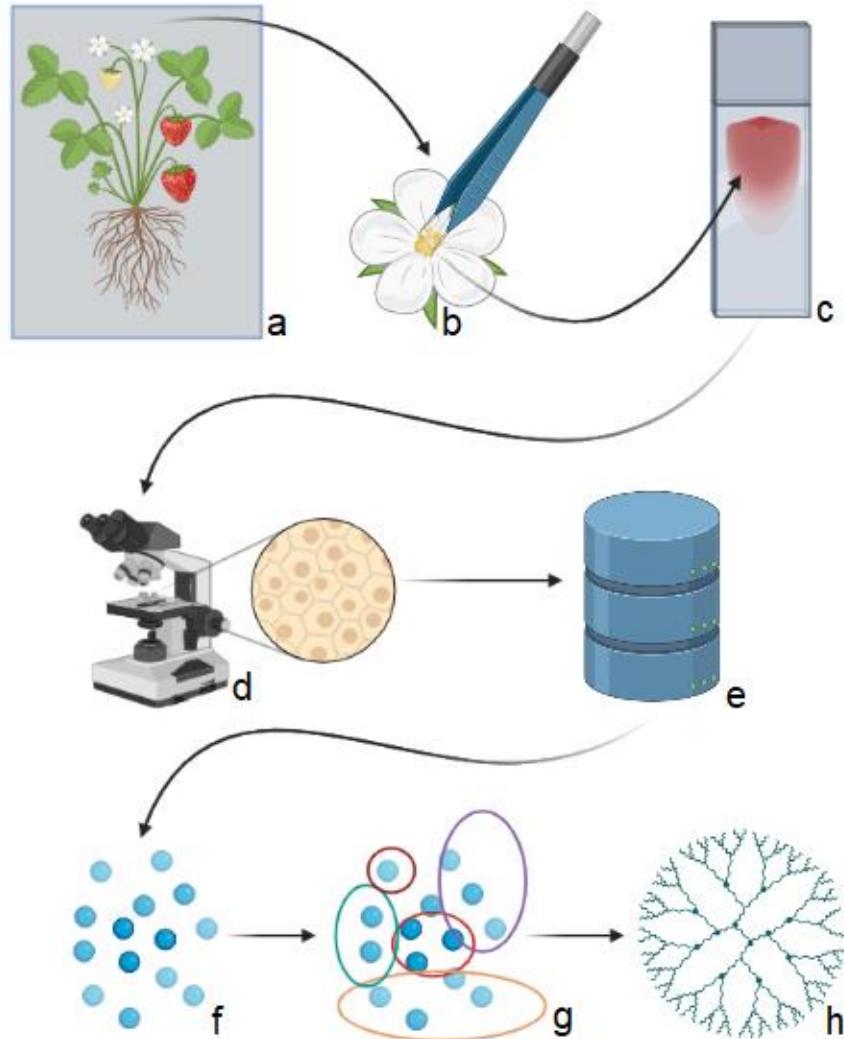
The temporal dimension is now buffered and a form of \mathbf{A} is employed
 $\text{Spatial} == 1 \& \text{ Temporal} +/- \text{ Buffer} == 1 \sim \mathbf{X}$

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

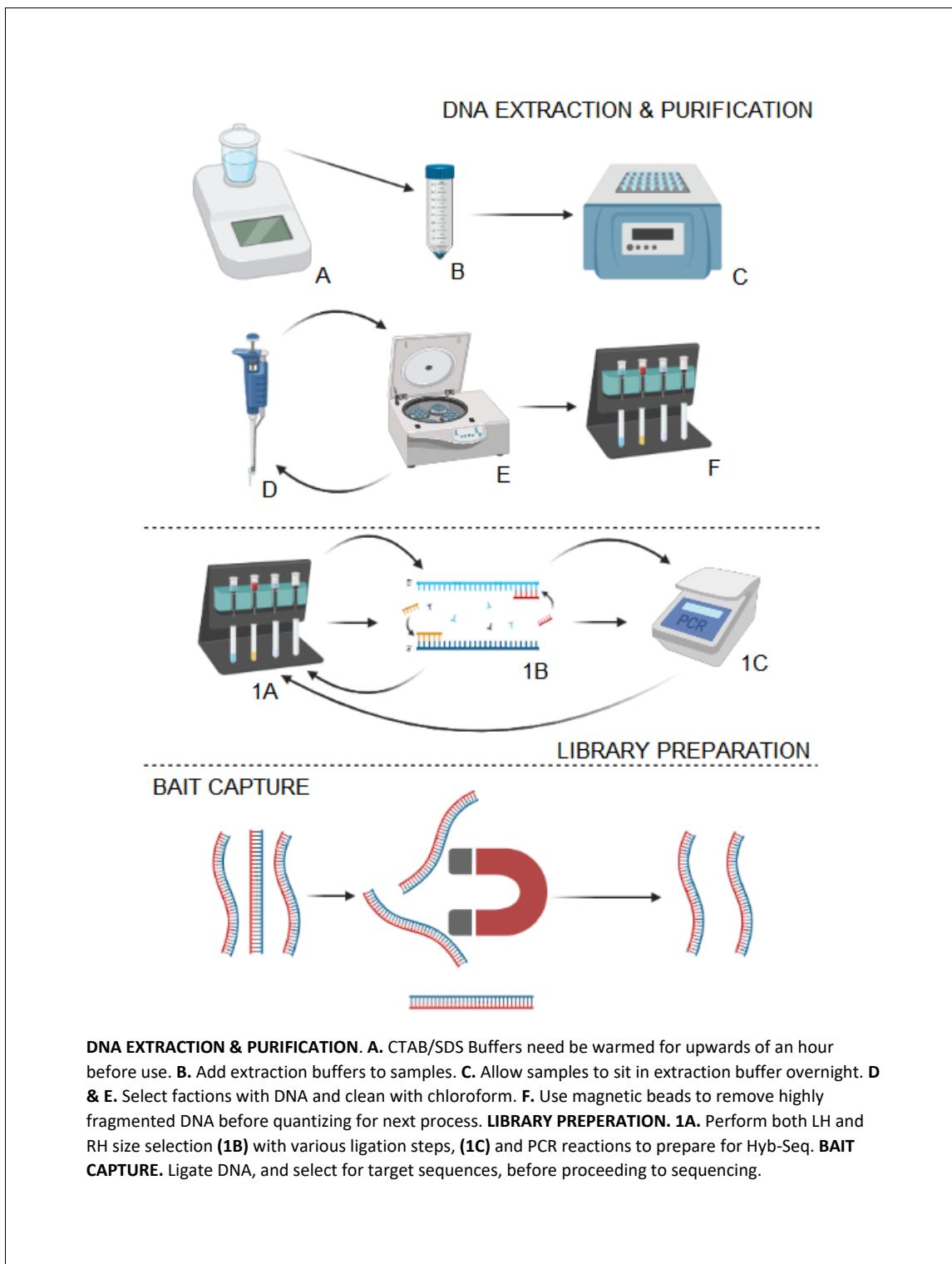
$$\begin{aligned} \text{Spatial} == 0 \& \text{ Temporal} == 0 \& \text{ Congener} == 1 \sim \mathbf{G} \\ \text{Spatial} == 0 \& \text{ Temporal} == 0 \& \text{ Congeners} == 1 \sim \mathbf{H} \\ \text{Spatial} == 0 \& \text{ Temporal} == 0 \& \text{ Confamilial} == 1 \sim \mathbf{I} \\ \text{Spatial}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

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 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

701 THIS SHOULD BE TURNED INTO A SMALLER PNG, AND HAVE THE NUMBER OF SEQUENCED
702 METAGENOMIC SAMPLES PLACED INTO THAT COLUMN AND INCLUDED IN TEXT ~~~ NEED
703 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

705 **References**

- 706 Ackerfield, J. (2015). *Flora of colorado*. BRIT Press Fort Worth.
- 707 Alarcón, R. (2010). Congruence between visitation and pollen-transport networks in a California plant–
708 pollinator community. *Oikos*, **119**, 35–44. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2009.17694.x>
- 710 Aldous, A.E. (1919). *Eradicating tall larkspur on cattle ranges in the national forest*. US Department of
711 Agriculture.
- 712 Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models:
713 Prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, **43**, 1223–1232.
- 714 Allred, K.W. & Ivey, R. (2012). Flora neomexicana III: An illustrated identification manual. *Lulu. com*.
- 715 Araujo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *Trends in ecology & evolution*,
716 **22**, 42–47.
- 717 Augspurger, C.K. & Zaya, D.N. (2020). Concordance of long-term shifts with climate warming varies among
718 phenological events and herbaceous species. *Ecological Monographs*, **90**, e01421.
- 719 Baker, W.J., Bailey, P., Barber, V., Barker, A., Bellot, S., Bishop, D., Botigué, L.R., Brewer, G., Carruthers,
720 T., Clarkson, J.J., Cook, J., Cowan, R.S., Dodsworth, S., Epitawalage, N., Françoso, E., Gallego, B.,
721 Johnson, M.G., Kim, J.T., Leempoel, K., Maurin, O., Mcginnie, C., Pokorny, L., Roy, S., Stone, M.,
722 Toledo, E., Wickett, N.J., Zuntini, A.R., Eiserhardt, W.L., Kersey, P.J., Leitch, I.J. & Forest, F. (2021a).
723 A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Systematic Biology*,
724 **71**, 301–319. Retrieved from <https://doi.org/10.1093/sysbio/syab035>
- 725 Baker, W., Dodsworth, S., Forest, F., Graham, S., Johnson, M., McDonnell, A., Pokorny, L., Tate, J., Wicke,
726 S. & Wickett, N. (2021b). Exploring Angiosperms353: An open, community toolkit for collaborative
727 phylogenomic research on flowering plants. *American Journal of Botany*, **108**.
- 728 Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012). Selecting pseudo-absences for species
729 distribution models: How, where and how many? *Methods in ecology and evolution*, **3**, 327–338.
- 730 Barker, D.A. & Arceo-Gomez, G. (2021). Pollen transport networks reveal highly diverse and temporally
731 stable plant–pollinator interactions in an Appalachian floral community. *AoB PLANTS*, **13**. Retrieved
732 from <https://doi.org/10.1093/aobpla/plab062>
- 733 Basey, A.C., Fant, J.B. & Kramer, A.T. (2015). Producing native plant materials for restoration: 10 rules
734 to collect and maintain genetic diversity. *Native Plants Journal*, **16**, 37–53.
- 735 Beattie, A. (1971). A technique for the study of insect-borne pollen. *The Pan-Pacific Entomologist*, **47**, 82.
- 736 Beck, J.B., Markley, M.L., Zielke, M.G., Thomas, J.R., Hale, H.J., Williams, L.D. & Johnson, M.G. (2021).

- 737 Are palmer's elm-leaf goldenrod and the smooth elm-leaf goldenrod real? The Angiosperms353 kit
738 provides within-species signal in solidago ulmifolia sl. *Systematic Botany*, **46**, 1107–1113.
- 739 Belitz, M.W., Larsen, E.A., Ries, L. & Guralnick, R.P. (2020). The accuracy of phenology estimators for use
740 with sparsely sampled presence-only observations. *Methods in Ecology and Evolution*, **11**, 1273–1285.
- 741 Bell, K.L., Burgess, K.S., Botsch, J.C., Dobbs, E.K., Read, T.D. & Brosi, B.J. (2019). Quantitative and
742 qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Molecular
743 Ecology*, **28**, 431–455.
- 744 Bell, K.L., Fowler, J., Burgess, K.S., Dobbs, E.K., Gruenewald, D., Lawley, B., Morozumi, C. & Brosi, B.J.
745 (2017). Applying pollen DNA metabarcoding to the study of plant–pollinator interactions. *Applications
746 in plant sciences*, **5**, 1600124.
- 747 Bell, K.L., Petit III, R.A., Cutler, A., Dobbs, E.K., Macpherson, J.M., Read, T.D., Burgess, K.S. & Brosi,
748 B.J. (2021). Comparing whole-genome shotgun sequencing and DNA metabarcoding approaches for
749 species identification and quantification of pollen species mixtures. *Ecology and evolution*, **11**, 16082–
750 16098.
- 751 Belsky, A.J., Matzke, A. & Uselman, S. (1999). Survey of livestock influences on stream and riparian
752 ecosystems in the western united states. *Journal of Soil and water Conservation*, **54**, 419–431.
- 753 Bergman, P., Molau, U. & Holmgren, B. (1996). Micrometeorological impacts on insect activity and plant
754 reproductive success in an alpine environment, swedish lapland. *Arctic and alpine research*, **28**, 196–202.
- 755 Betancourt, J.L., Schwartz, M.D., Breshears, D.D., Cayan, D.R., Dettinger, M.D., Inouye, D.W., Post, E.
756 & Reed, B.C. (2005). Implementing a US national phenology network.
- 757 Bingham, R.A. & Orthner, A.R. (1998). Efficient pollination of alpine plants. *Nature*, **391**, 238–239.
- 758 Bolger, A. & Giorgi, F. (2014). Trimmomatic: A flexible read trimming tool for illumina NGS data. *Bioin-
759 formatics*, **30**, 2114–2120.
- 760 Bontsutsnaja, A., Karise, R., Mand, M. & Smagghe, G. (2021). Bumble bee foraged pollen analyses in spring
761 time in southern estonia shows abundant food sources. *Insects*, **12**, 922.
- 762 Bortolus, A. (2008). Error cascades in the biological sciences: The unwanted consequences of using bad
763 taxonomy in ecology. *AMBIO: A journal of the human environment*, **37**, 114–118.
- 764 Brewen, C.J., Berrill, J.-P., Ritchie, M.W., Boston, K., Dagley, C.M., Jones, B., Coppoletta, M. & Burnett,
765 C.L. (2021). 76-year decline and recovery of aspen mediated by contrasting fire regimes: Long-unburned,
766 infrequent and frequent mixed-severity wildfire. *Plos one*, **16**, e0232995.
- 767 Brito-Morales, I., Molinos, J.G., Schoeman, D.S., Burrows, M.T., Poloczanska, E.S., Brown, C.J., Ferrier,
768 S., Harwood, T.D., Klein, C.J., McDonald-Madden, E. & others. (2018). Climate velocity can inform
769 conservation in a warming world. *Trends in ecology & evolution*, **33**, 441–457.

- 770 Brosi, B.J. & Briggs, H.M. (2013). Single pollinator species losses reduce floral fidelity and plant reproductive
771 function. *Proceedings of the National Academy of Sciences*, **110**, 13044–13048.
- 772 Calabrese, J.M., Certain, G., Kraan, C. & Dormann, C.F. (2014). Stacking species distribution models and
773 adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, **23**, 99–112.
- 774 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T.L. (2009).
775 BLAST+: Architecture and applications. *BMC bioinformatics*, **10**, 1–9.
- 776 Cameron, S.A. & Sadd, B.M. (2020). Global trends in bumble bee health. *Annual review of entomology*, **65**,
777 209–232.
- 778 CaraDonna, P.J., Burkle, L.A., Schwarz, B., Resasco, J., Knight, T.M., Benadi, G., Bluthgen, N., Dormann,
779 C.F., Fang, Q., Frund, J. & others. (2021). Seeing through the static: The temporal dimension of
780 plant–animal mutualistic interactions. *Ecology Letters*, **24**, 149–161.
- 781 CaraDonna, P.J., Iler, A.M. & Inouye, D.W. (2014). Shifts in flowering phenology reshape a subalpine plant
782 community. *Proceedings of the National Academy of Sciences*, **111**, 4916–4921.
- 783 CaraDonna, P.J., Petry, W.K., Brennan, R.M., Cunningham, J.L., Bronstein, J.L., Waser, N.M. & Sanders,
784 N.J. (2017). Interaction rewiring and the rapid turnover of plant–pollinator networks. *Ecology letters*,
785 **20**, 385–394.
- 786 Chao, A., Gotelli, N.J., Hsieh, T.C., Sande, E.L., Ma, K.H., Colwell, R.K. & Ellison, A.M. (2014). Rarefac-
787 tion and extrapolation with hill numbers: A framework for sampling and estimation in species diversity
788 studies. *Ecological Monographs*, **84**, 45–67.
- 789 Cheng, S., Melkonian, M., Smith, S.A., Brockington, S., Archibald, J.M., Delaux, P.-M., Li, F.-W., Melko-
790 nian, B., Mavrodiev, E.V., Sun, W., Fu, Y., Yang, H., Soltis, D.E., Graham, S.W., Soltis, P.S., Liu,
791 X., Xu, X. & Wong, G.K.-S. (2018). 10KP: A phylogenetic genome sequencing plan. *GigaScience*, **7**.
792 Retrieved from <https://doi.org/10.1093/gigascience/giy013>
- 793 Coissac, E., Hollingsworth, P.M., Lavergne, S. & Taberlet, P. (2016). From barcodes to genomes: Extending
794 the concept of DNA barcoding.
- 795 Coissac, E., Riaz, T. & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants
796 and animals. *Molecular ecology*, **21**, 1834–1847.
- 797 Colla, S.R., Gadallah, F., Richardson, L., Wagner, D. & Gall, L. (2012). Assessing declines of north american
798 bumble bees (*bombus* spp.) Using museum specimens. *Biodiversity and Conservation*, **21**, 3585–3595.
- 799 Cooke, R.U. & Reeves, R.W. (1976). *Arroyos and environmental change in the american south-west*. Claren-
800 don Press.
- 801 Crisci, J.V., Katinas, L., Apodaca, M.J. & Hoch, P.C. (2020). The end of botany. *Trends in Plant Science*,
802 **25**, 1173–1176.

- 803 Cronquist, A., Holmgren, A.H., Holmgren, N.H., Reveal, J.L., Holmgren, P.K., Barneby, R & others.
804 (1977+). *Intermountain flora. Vascular plants of the intermountain west, USA volume six. The mono-*
805 *cotyledons.* Columbia University.
- 806 Dahl, T.E. (1990). *Wetlands losses in the united states, 1780's to 1980's.* US Department of the Interior,
807 Fish; Wildlife Service.
- 808 Davis, C.C., Lyra, G.M., Park, D.S., Asprino, R., Maruyama, R., Torquato, D., Cook, B.I. & Ellison, A.M.
809 (2022). New directions in tropical phenology. *Trends in Ecology & Evolution.*
- 810 Dobrowski, S.Z. & Parks, S.A. (2016). Climate change velocity underestimates climate change exposure in
811 mountainous regions. *Nature Communications*, **7**, 1–8.
- 812 Doyle, J.J. & Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue.
813 *Phytochemical Bulletin*, **19**, 11–15.
- 814 Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J.-P. & Guisan, A. (2011). Predicting spatial
815 patterns of plant species richness: A comparison of direct macroecological and species stacking modelling
816 approaches. *Diversity and Distributions*, **17**, 1122–1131.
- 817 Elith*, J., H. Graham*, C., P. Anderson, R., Dudik, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann,
818 F., R. Leathwick, J., Lehmann, A. & others. (2006). Novel methods improve prediction of species'
819 distributions from occurrence data. *Ecography*, **29**, 129–151.
- 820 Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Percy, D.M., Graham, S.W., Barrett, S.C., Newmaster,
821 S.G., Hajibabaei, M. & Husband, B.C. (2009). Are plant species inherently harder to discriminate than
822 animal species using DNA barcoding markers? *Molecular Ecology Resources*, **9**, 130–139.
- 823 Fenn, M.E., Baron, J.S., Allen, E.B., Rueth, H.M., Nydick, K.R., Geiser, L., Bowman, W.D., Sickman, J.O.,
824 Meixner, T., Johnson, D.W. & others. (2003). Ecological effects of nitrogen deposition in the western
825 united states. *BioScience*, **53**, 404–420.
- 826 Flora of North America Editorial Committee, eds. (1993+). *Flora of north america north of mexico [online].*
827 Oxford University Press on Demand.
- 828 Frase, Barbara A. & Buck, P. (2007). Vascular Plants of the Gothic Area. Retrieved from https://www.digitalrmb.org/wp-content/uploads/2016/05/vascularplantlist_20071.pdf
- 830 Gage, E. & Cooper, D.J. (2013). Historical range of variation assessment for wetland and riparian ecosystems,
831 u.s. Forest service rocky mountain region
- 832 Goulson, D. (2010). *Bumblebees: Behaviour, ecology, and conservation.* Oxford University Press on Demand.
- 833 Goulson, D., Hanley, M.E., Darvill, B., Ellis, J. & Knight, M.E. (2005). Causes of rarity in bumblebees.
834 *Biological conservation*, **122**, 1–8.
- 835 Goulson, D., Lye, G. & Darvill, B. (2008). The decline and conservation of bumblebees. *Annual review of*

- 836 *entomology*, **53**, 191–208.
- 837 Grixti, J.C., Wong, L.T., Cameron, S.A. & Favret, C. (2009). Decline of bumble bees (*bombus*) in the north
838 american midwest. *Biological conservation*, **142**, 75–84.
- 839 Group, C.P.W., Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., Bank,
840 M. van der, Chase, M.W., Cowan, R.S., Erickson, D.L. & others. (2009). A DNA barcode for land
841 plants. *Proceedings of the National Academy of Sciences*, **106**, 12794–12797.
- 842 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.,
843 Chen, S.-L. & others. (2011). Comparative analysis of a large dataset indicates that internal transcribed
844 spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National
845 Academy of Sciences*, **108**, 19641–19646.
- 846 Havens, K., Vitt, P., Schwarz, J., Orr, B. & Crimmins, T. (2007). Chicago botanic garden's conservation
847 and outreach efforts on climate change. *BGjournal*, **4**, 13–16.
- 848 Hebert, P.D., Cywinska, A., Ball, S.L. & DeWaard, J.R. (2003). Biological identifications through DNA
849 barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, 313–321.
- 850 Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A.,
851 Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B. & others. (2017). SoilGrids250m:
852 Global gridded soil information based on machine learning. *PLoS one*, **12**, e0169748.
- 853 Hennig, C. (2020). *Fpc: Flexible procedures for clustering*. Retrieved from <https://CRAN.R-project.org/>
854 package=fpc
- 855 Hinchliff, C.E., Smith, S.A., Allman, J.F., Burleigh, J.G., Chaudhary, R., Coghill, L.M., Crandall, K.A.,
856 Deng, J., Drew, B.T., Gazis, R. & others. (2015). Synthesis of phylogeny and taxonomy into a compre-
857 hensive tree of life. *Proceedings of the National Academy of Sciences*, **112**, 12764–12769.
- 858 Hitchcock, C.L. & Cronquist, A. (2018). *Flora of the pacific northwest: An illustrated manual*. University
859 of Washington Press.
- 860 Hollingsworth, P.M., Li, D.-Z., Bank, M. van der & Twyford, A.D. (2016). Telling plant species apart with
861 DNA: From barcodes to genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*,
862 **371**, 20150338.
- 863 Hsieh, T.C., Ma, K.H. & Chao, A. (2020). *iNEXT: Interpolation and extrapolation for species diversity*.
864 Retrieved from http://chao.stat.nthu.edu.tw/wordpress/software_download/
- 865 Iler, A.M., Humphrey, P.T., Ogilvie, J.E. & CaraDonna, P.J. (2021). Conceptual and practical issues limit
866 the utility of statistical estimators of phenological events. *Ecosphere*, **12**, e03828.
- 867 Janzen, D.H. (1967). Synchronization of sexual reproduction of trees within the dry season in central america.
868 *Evolution*, **21**, 620–637.

- 869 Janzen, D.H., Burns, J.M., Cong, Q., Hallwachs, W., Dapkey, T., Manjunath, R., Hajibabaei, M., Hebert,
870 P.D. & Grishin, N.V. (2017). Nuclear genomes distinguish cryptic species suggested by their DNA
871 barcodes and ecology. *Proceedings of the National Academy of Sciences*, **114**, 8313–8318.
- 872 *Jepson flora project*. (2020).
- 873 Johnson, M.D., Fokar, M., Cox, R.D. & Barnes, M.A. (2021). Airborne environmental DNA metabarcoding
874 detects more diversity, with less sampling effort, than a traditional plant community survey. *BMC Ecology
875 and Evolution*, **21**, 1–15.
- 876 Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J. & Wickett,
877 N.J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput
878 sequencing reads using target enrichment. *Applications in plant sciences*, **4**, 1600016.
- 879 Johnson, M.G., Pokorny, L., Dodsworth, S., Botigue, L.R., Cowan, R.S., Devault, A., Eiserhardt, W.L.,
880 Epitawalage, N., Forest, F., Kim, J.T. & others. (2019). A universal probe set for targeted sequencing
881 of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic biology*,
882 **68**, 594–606.
- 883 Joshi, J., Schmid, B., Caldeira, M., Dimitrakopoulos, P., Good, J., Harris, R., Hector, A., Huss-Danell, K.,
884 Jumpponen, A., Minns, A. & others. (2001). Local adaptation enhances performance of common plant
885 species. *Ecology Letters*, **4**, 536–544.
- 886 Keane, R.E. (2002). Cascading effects of fire exclusion in rocky mountain ecosystems: A literature review.
- 887 Kramer, A.T. & Havens, K. (2015). Report in brief: Assessing botanical capacity to address grand challenges
888 in the united states. *Natural Areas Journal*, **35**, 83–89.
- 889 Kress, W.J. (2017). Plant DNA barcodes: Applications today and in the future. *Journal of systematics and
890 evolution*, **55**, 291–307.
- 891 Kress, W.J. & Erickson, D.L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL
892 gene complements the non-coding trnH-psbA spacer region. *PLoS one*, **2**, e508.
- 893 Kuhn, M. (2022). *Caret: Classification and regression training*. Retrieved from <https://CRAN.R-project.org/package=caret>
- 894 Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G. & Taylor, M.I. (2019). How quantitative is
895 metabarcoding: A meta-analytical approach. *Molecular ecology*, **28**, 420–430.
- 896 Land Management, B. of. (2019). U.s. Department of interior bureau of land management, BLM - assess-
897 ment, inventory, and monitoring (AIM) terrestrial indicators raw dataset. Retrieved from [https://gbp-
899 blm-egis.hub.arcgis.com/pages/aim](https://gbp-
898 blm-egis.hub.arcgis.com/pages/aim)
- 900 Lang, D., Tang, M., Hu, J. & Zhou, X. (2019). Genome-skimming provides accurate quantification for pollen
901 mixtures. *Molecular Ecology Resources*, **19**, 1433–1446.

- 902 Lesica, P., Lavin, M. & Stickney, P.F. (2012). *Manual of montana vascular plants*. Brit Press.
- 903 Lewin, H.A., Richards, S., Aiden, E.L., Allende, M.L., Archibald, J.M., Bálint, M., Barker, K.B., Baumgart-
904 ner, B., Belov, K., Bertorelle, G., Blaxter, M.L., Cai, J., Caperello, N.D., Carlson, K., Castilla-Rubio,
905 J.C., Chaw, S.-M., Chen, L., Childers, A.K., Coddington, J.A., Conde, D.A., Corominas, M., Crandall,
906 K.A., Crawford, A.J., DiPalma, F., Durbin, R., Ebenezer, T.E., Edwards, S.V., Fedrigo, O., Fliceck, P.,
907 Formenti, G., Gibbs, R.A., Gilbert, M.T.P., Goldstein, M.M., Graves, J.M., Greely, H.T., Grigoriev,
908 I.V., Hackett, K.J., Hall, N., Haussler, D., Helgen, K.M., Hogg, C.J., Isobe, S., Jakobsen, K.S., Janke,
909 A., Jarvis, E.D., Johnson, W.E., Jones, S.J.M., Karlsson, E.K., Kersey, P.J., Kim, J.-H., Kress, W.J.,
910 Kuraku, S., Lawniczak, M.K.N., Leebens-Mack, J.H., Li, X., Lindblad-Toh, K., Liu, X., Lopez, J.V.,
911 Marques-Bonet, T., Mazard, S., Mazet, J.A.K., Mazzoni, C.J., Myers, E.W., O'Neill, R.J., Paez, S.,
912 Park, H., Robinson, G.E., Roquet, C., Ryder, O.A., Sabir, J.S.M., Shaffer, H.B., Shank, T.M., Sherkow,
913 J.S., Soltis, P.S., Tang, B., Tedersoo, L., Uliano-Silva, M., Wang, K., Wei, X., Wetzer, R., Wilson,
914 J.L., Xu, X., Yang, H., Yoder, A.D. & Zhang, G. (2022). The earth BioGenome project 2020: Starting
915 the clock. *Proceedings of the National Academy of Sciences*, **119**, e2115635118. Retrieved from
916 <https://www.pnas.org/doi/abs/10.1073/pnas.2115635118>
- 917 Liang, H., Zhao, Y.-H., Rafferty, N.E., Ren, Z.-X., Zhong, L., Li, H.-D., Li, D.-Z. & Wang, H. (2021).
918 Evolutionary and ecological factors structure a plant–bumblebee network in a biodiversity hotspot, the
919 himalaya–hengduan mountains. *Functional Ecology*, **35**, 2523–2535.
- 920 Life Project Consortium, D.T. of, Blaxter, M., Mieszkowska, N., Palma, F.D., Holland, P., Durbin, R.,
921 Richards, T., Berriman, M., Kersey, P., Hollingsworth, P., Wilson, W., Twyford, A., Gaya, E., Lawniczak,
922 M., Lewis, O., Broad, G., Howe, K., Hart, M., Fliceck, P. & Barnes, I. (2022). Sequence locally, think glob-
923 ally: The darwin tree of life project. *Proceedings of the National Academy of Sciences*, **119**, e2115642118.
924 Retrieved from <https://www.pnas.org/doi/abs/10.1073/pnas.2115642118>
- 925 Li, X., Jiang, L., Meng, F., Wang, S., Niu, H., Iler, A.M., Duan, J., Zhang, Z., Luo, C., Cui, S. & others.
926 (2016). Responses of sequential and hierarchical phenological events to warming and cooling in alpine
927 meadows. *Nature Communications*, **7**, 1–8.
- 928 Liu, J., Shi, L., Han, J., Li, G., Lu, H., Hou, J., Zhou, X., Meng, F. & Downie, S.R. (2014). Identification
929 of species in the angiosperm family apiaceae using DNA barcodes. *Molecular ecology resources*, **14**,
930 1231–1238.
- 931 Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y. & Chen, S. (2015). Plant DNA barcoding: From gene
932 to genome. *Biological Reviews*, **90**, 157–166.
- 933 Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009). The velocity of
934 climate change. *Nature*, **462**, 1052–1055.

- 935 Lu, J., Breitwieser, F.P., Thielen, P. & Salzberg, S.L. (2017). Bracken: Estimating species abundance in
936 metagenomics data. *PeerJ Computer Science*, **3**, e104.
- 937 Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2022). *Cluster: Cluster analysis basics*
938 and extensions. Retrieved from <https://CRAN.R-project.org/package=cluster>
- 939 Mainali, K. & Slud, E. (2022). *CooccurrenceAffinity: Affinity in cooccurrence data*.
- 940 Mainali, K.P., Slud, E., Singer, M.C. & Fagan, W.F. (2022). A better index for analysis of co-occurrence
941 and similarity. *Science Advances*, **8**, eabj9204.
- 942 Maitner, B. (2022). *BIEN: Tools for accessing the botanical information and ecology network database*.
943 Retrieved from <https://CRAN.R-project.org/package=BIEN>
- 944 Manzano, S. (2021). Flippant attitudes towards plant identification jeopardize early career botanists. *Trends*
945 in *Plant Science*, **26**, 987–988.
- 946 McLay, T.G., Birch, J.L., Gunn, B.F., Ning, W., Tate, J.A., Nauheimer, L., Joyce, E.M., Simpson, L.,
947 Schmidt-Lebuhn, A.N., William J & others. (2021). New targets acquired: Improving locus recovery
948 from the Angiosperms353 probe set. *Applications in plant sciences*, **9**.
- 949 Mohlenbrock, R.H. (2002). *Vascular flora of illinois*. SIU Press.
- 950 Moore, A.J. & Bohs, L. (2003). An ITS phylogeny of balsamorhiza and wyethia (asteraceae: heliantheae).
951 *American Journal of Botany*, **90**, 1653–1660.
- 952 Naiman, R.J., Johnston, C.A. & Kelley, J.C. (1988). Alteration of north american streams by beaver.
953 *BioScience*, **38**, 753–762.
- 954 Naimi, B. & Araujo, M.B. (2016). Sdm: A reproducible and extensible r platform for species distribution
955 modelling. *Ecography*, **39**, 368–375.
- 956 Naimi, B., Hamm, N. a.s., Groen, T.A., Skidmore, A.K. & Toxopeus, A.G. (2014). Where is positional
957 uncertainty a problem for species distribution modelling. *Ecography*, **37**, 191–203.
- 958 Nazaire, M. & Hufford, L. (2014). Phylogenetic systematics of the genus mertensia (boraginaceae). *Systematic*
959 *Botany*, **39**, 268–303.
- 960 Nevado, B., Atchison, G.W., Hughes, C.E. & Filatov, D.A. (2016). Widespread adaptive evolution during
961 repeated evolutionary radiations in new world lupins. *Nature communications*, **7**, 1–9.
- 962 Newstrom, L.E., Frankie, G.W. & Baker, H.G. (1994). A new classification for plant phenology based on
963 flowering patterns in lowland tropical rain forest trees at la selva, costa rica. *Biotropica*, **26**, 141–159.
- 964 Oberhauser, K.S., Nail, K.R. & Altizer, S. (2015). *Monarchs in a changing world: Biology and conservation*
965 of an iconic butterfly. Cornell University Press.
- 966 Occdownload Gbif.Org. (2021). Occurrence download. Retrieved from <https://www.gbif.org/occurrence/>
967 download/0206948-200613084148143

- 968 Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos,
969 P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D.,
970 Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly,
971 M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M.-H., Ribeiro Cunha,
972 E., Smith, T., Stier, A., Ter Braak, C.J.F. & Weedon, J. (2022). *Vegan: Community ecology package*.
973 Retrieved from <https://CRAN.R-project.org/package=vegan>
- 974 Oliver, P.M., Adams, M., Lee, M.S., Hutchinson, M.N. & Doughty, P. (2009). Cryptic diversity in vertebrates:
975 Molecular data double estimates of species diversity in a radiation of australian lizards (*diplodactylus*,
976 *gekkota*). *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2001–2007.
- 977 Omernik, J.M. (1987). Ecoregions of the conterminous united states. *Annals of the Association of American
978 geographers*, **77**, 118–125.
- 979 Ottenlips, M.V., Mansfield, D.H., Buerki, S., Feist, M.A.E., Downie, S.R., Dodsworth, S., Forest, F., Plun-
980 kett, G.M. & Smith, J.F. (2021). Resolving species boundaries in a recent radiation with the An-
981 giosperms353 probe set: The lomatium packardiae/l. *Anomalum* clade of the l. *Triternatum* (apiaceae)
982 complex. *American journal of botany*, **108**, 1217–1233.
- 983 Park, D.S., Lyra, G.M., Ellison, A.M., Maruyama, R.K.B., Torquato, D. dos R., Asprino, R.C., Cook, B.I. &
984 Davis, C.C. (2022). Herbarium records provide reliable phenology estimates in the understudied tropics.
985 *Journal of Ecology*.
- 986 Pearse, W.D., Davis, C.C., Inouye, D.W., Primack, R.B. & Davies, T.J. (2017). A statistical estimator
987 for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution*, **1**,
988 1876–1882.
- 989 Peel, N., Dicks, L.V., Clark, M.D., Heavens, D., Percival-Alwyn, L., Cooper, C., Davies, R.G., Leggett,
990 R.M. & Yu, D.W. (2019). Semi-quantitative characterisation of mixed pollen samples using MinION
991 sequencing and reverse metagenomics (RevMet). *Methods in Ecology and Evolution*, **10**, 1690–1701.
- 992 Pepin, N., Arnone, E., Gobiet, A., Haslinger, K., Kotlarski, S., Notarnicola, C., Palazzi, E., Seibert, P.,
993 Serafin, S., Schöner, W. & others. (2022). Climate changes and their elevational patterns in the mountains
994 of the world. *Reviews of geophysics*, **60**, e2020RG000730.
- 995 Pinto-Ledezma, J.N. & Cavender-Bares, J. (2021). Predicting species distributions and community compo-
996 sition using satellite remote sensing predictors. *Scientific Reports*, **11**, 1–12.
- 997 Pornon, A., Andalo, C., Burrus, M. & Escaravage, N. (2017). DNA metabarcoding data unveils invisible
998 pollination networks. *Scientific Reports*, **7**, 1–11.
- 999 Prather, L.A., Alvarez-Fuentes, O., Mayfield, M.H. & Ferguson, C.J. (2004a). Implications of the decline in
1000 plant collecting for systematic and floristic research. *Systematic Botany*, **29**, 216–220.

- 1001 Prather, L.A., Alvarez-Fuentes, O., Mayfield, M.H. & Ferguson, C.J. (2004b). The decline of plant collecting
1002 in the united states: A threat to the infrastructure of biodiversity studies. *Systematic Botany*, **29**, 15–28.
- 1003 Prim, R.C. (1957). Shortest connection networks and some generalisations. *Bell System Technical Journal*,
1004 **36**, 1389–1401.
- 1005 Pusalkar, P.K. & Singh, D.K. (2015). Taxonomic rearrangement of arenaria (caryophyllaceae) in indian
1006 western himalaya. *Journal of Japanese Botany*, **90**, 77–91.
- 1007 Pyke, G.H. (1982). Local geographic distributions of bumblebees near crested butte, colorado: Competition
1008 and community structure. *Ecology*, **63**, 555–573.
- 1009 Qiao, H., Soberon, J. & Peterson, A.T. (2015). No silver bullets in correlative ecological niche modelling:
1010 Insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and*
1011 *Evolution*, **6**, 1126–1136.
- 1012 Rabeler, R.K. & Wagner, W.L. (2016). New combinations in odontostemma (caryophyllaceae). *PhytoKeys*,
1013 77.
- 1014 Ralphs, M.H. & Ueckert, D.N. (1988). Herbicide control of locoweeds: A review. *Weed Technology*, **2**,
1015 460–465.
- 1016 Ralphs, M., Woolsey, L. & Bowns, J. (2003). Mechanism by which ammonium fertilizers kill tall larkspur.
1017 *Rangeland Ecology & Management/Journal of Range Management Archives*, **56**, 524–528.
- 1018 Robinson, N., Regetz, J. & Guralnick, R.P. (2014). EarthEnv-DEM90: A nearly-global, void-free, multi-
1019 scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. *ISPRS Journal of*
1020 *Photogrammetry and Remote Sensing*, **87**, 57–67.
- 1021 Ruppert, K.M., Kline, R.J. & Rahman, M.S. (2019). Past, present, and future perspectives of environmental
1022 DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global
1023 eDNA. *Global Ecology and Conservation*, **17**, e00547.
- 1024 Sadeghian, S., Zarre, S., Rabeler, R.K. & Heubl, G. (2015). Molecular phylogenetic analysis of arenaria
1025 (caryophyllaceae: Tribe arenarieae) and its allies inferred from nuclear DNA internal transcribed spacer
1026 and plastid DNA rps16 sequences. *Botanical Journal of the Linnean Society*, **178**, 648–669.
- 1027 Sarro, E., Tripodi, A. & Woodard, S.H. (2022). Bumble bee (*bombus vosnesenskii*) queen nest searching
1028 occurs independent of ovary developmental status. *Integrative Organismal Biology*, **4**, obac007.
- 1029 Schmitt, S., Pouteau, R., Justeau, D., De Boissieu, F. & Birnbaum, P. (2017). Ssdm: An r package to
1030 predict distribution of species richness and composition based on stacked species distribution models.
1031 *Methods in Ecology and Evolution*, **8**, 1795–1803.
- 1032 Sennikov, A.N. & Kurtto, A. (2017). A phylogenetic checklist of sorbus sl (rosaceae) in europe.
- 1033 Sickel, W., Ankenbrand, M.J., Grimmer, G., Holzschuh, A., Hartel, S., Lanzen, J., Steffan-Dewenter, I. &

- 1034 Keller, A. (2015). Increased efficiency in identifying mixed pollen samples by meta-barcoding with a
1035 dual-indexing approach. *BMC ecology*, **15**, 1–9.
- 1036 Slimp, M., Williams, L.D., Hale, H. & Johnson, M.G. (2021). On the potential of Angiosperms353 for
1037 population genomic studies. *Applications in Plant Sciences*, **9**.
- 1038 Smith, S.A. & Brown, J.W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American journal*
1039 *of botany*, **105**, 302–314.
- 1040 Sork, V.L. (2018). Genomic studies of local adaptation in natural plant populations. *Journal of Heredity*,
1041 **109**, 3–15.
- 1042 Stevens, C.J., David, T.I. & Storkey, J. (2018). Atmospheric nitrogen deposition in terrestrial ecosystems: Its
1043 impact on plant communities and consequences across trophic levels. *Functional Ecology*, **32**, 1757–1769.
- 1044 Stroud, S., Fennell, M., Mitchley, J., Lydon, S., Peacock, J. & Bacon, K.L. (2022). The botanical education
1045 extinction and the fall of plant awareness. *Ecology and Evolution*, **12**, e9019.
- 1046 Suchan, T., Talavera, G., Saez, L., Ronikier, M. & Vila, R. (2019). Pollen metabarcoding as a tool for
1047 tracking long-distance insect migrations. *Molecular Ecology Resources*, **19**, 149–162.
- 1048 Tange, O. (2021). GNU parallel 20220322 (savannah). Retrieved from <https://doi.org/10.5281/zenodo.6377950>
- 1049 Tran, H., Nguyen, P., Ombadi, M., Hsu, K., Sorooshian, S. & Qing, X. (2019). A cloud-free MODIS snow
1050 cover dataset for the contiguous united states from 2000 to 2017. *Scientific data*, **6**, 1–13.
- 1051 Wang, T., Hamann, A., Spittlehouse, D. & Carroll, C. (2016). Locally downscaled and spatially customizable
1052 climate data for historical and future periods for north america. *PloS one*, **11**, e0156720.
- 1053 Weber, W. (1998). New names and combinations in asteraceae: Heliantheae-ecliptinae. *Phytologia*, **85**,
1054 19–21.
- 1055 Wenzell, K.E., McDonnell, A.J., Wickett, N.J., Fant, J.B. & Skogen, K.A. (2021). Incomplete reproductive
1056 isolation and low genetic differentiation despite floral divergence across varying geographic scales in
1057 castilleja. *American Journal of Botany*, **108**, 1270–1288.
- 1058 Wilkinson, D.P., Golding, N., Guillera-Arroita, G., Tingley, R. & McCarthy, M.A. (2021). Defining and
1059 evaluating predictions of joint species distribution models. *Methods in Ecology and Evolution*, **12**, 394–
1060 404.
- 1061 Williams, P.H. (1982). The distribution and decline of british bumble bees (*bombus latr.*). *Journal of*
1062 *Apicultural Research*, **21**, 236–245. Retrieved from <https://doi.org/10.1080/00218839.1982.11100549>
- 1063 Wilson, A.M. & Jetz, W. (2016). Remotely sensed high-resolution global cloud dynamics for predicting
1064 ecosystem and biodiversity distributions. *PLoS biology*, **14**, e1002415.
- 1065 Wood, D.E., Lu, J. & Langmead, B. (2019). Improved metagenomic analysis with kraken 2. *Genome biology*,

- 1067 **20**, 1–13.
- 1068 Xie, Y., Wang, X. & Silander Jr, J.A. (2015). Deciduous forest responses to temperature, precipitation, and
1069 drought imply complex climate change impacts. *Proceedings of the National Academy of Sciences*, **112**,
1070 13585–13590.
- 1071 Xie, Y., Wang, X., Wilson, A.M. & Silander Jr, J.A. (2018). Predicting autumn phenology: How deciduous
1072 tree species respond to weather stressors. *Agricultural and Forest Meteorology*, **250**, 127–137.
- 1073 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,
1074 D.-Z. & Wang, H. (2019). The topological differences between visitation and pollen transport networks:
1075 A comparison in species rich communities of the himalaya–hengduan mountains. *Oikos*, **128**, 551–562.
1076 Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/oik.05262>

1077 **List of Figures**

1078 1	Number of the ten most commonly visited plants which are also in the top ten most common 1079 sequences	67
1080 2	Phylogenetic tree of all biotically pollinated plant genera in the study area. The innermost 1081 ring indicates every genus which Queen Bee's were observed to visit. The intermediate ring 1082 indicates that at least a single morphological pollen voucher slide was prepared for a member 1083 of the genus. The outermost ring indicates that sequence data were available for at least a 1084 member of that genus. Branch colors follow APG 4.	68
1085 3	Modelled dates of when major flowering events occurred	69

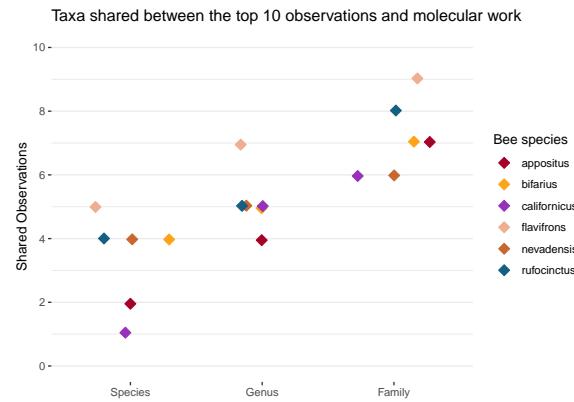
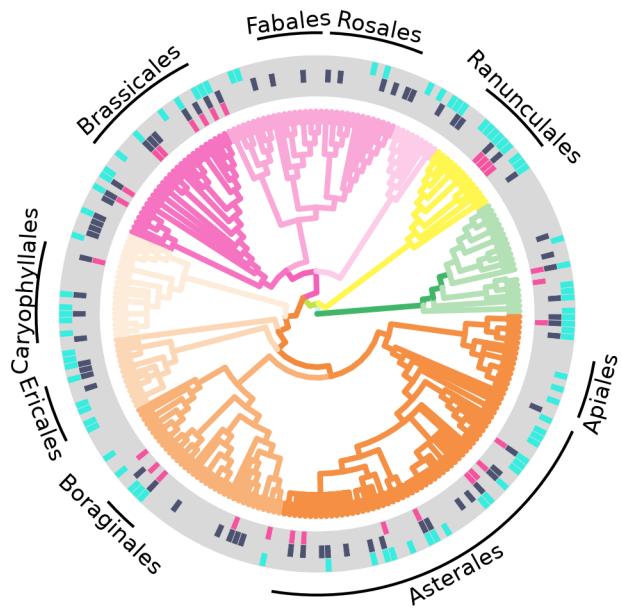


Figure 1: Number of the ten most commonly visited plants which are also in the top ten most common sequences

Biotically pollinated plant genera with morphological or molecular data



Status lacking observed sequenced slide

Figure 2: 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.

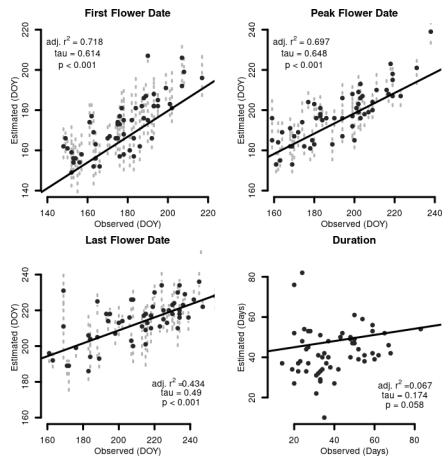


Figure 3: Modelled dates of when major flowering events occurred

₁₀₈₆ **List of Tables**

₁₀₈₇ 2	Logistic regression assessing accuracy of SDMs	71
₁₀₈₈ 3	Species Distribution Modeling evaluation contingency table	72
₁₀₈₉ 4	Post classification of Sequences via Taxonomy and Ecology	73

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	100	22.83	32.96	Species
B	108	24.66	9.35	Species
C	171	39.04	46.39	Genus
D	5	1.14	0.45	Species
E	17	3.88	2.23	Family
F	7	1.60	0.58	Species
X	30	6.85	8.03	Species