Identification and characterization of *Gypsophila paniculata* color morphs in Sleeping Bear Dunes National Lakeshore, MI, USA.

Marisa L. Yang[1] , Emma Rice[2], Hailee Leimbach-Maus[2], and Charlyn G. Partridge[2]

[1]Environmental Science Policy, and Management, University of California, Berkeley, 260 Mulford Hall, Berkeley, CA 94720, email: marisayang79@gmail.com

[2]Annis Water Resources Institute, 740 W. Shoreline Dr., Muskegon, MI 49441, email: riceemm@mail.gvsu.edu; hailee.pavisich@gmail.com

Corresponding author: Charlyn Partridge, Annis Water Resources Institute, Grand Valley State University, 740 W. Shoreline Dr., Muskegon, MI; 49441, email: partridc@gvsu.edu

Abstract

Background.

Gypsophila paniculata (baby's breath) is an invasive species found throughout much of the northwest United States and western Canada. Recently, plants exhibiting a different color morphology were identified within the coastal dunes along eastern Lake Michigan. The common baby's breath (*G. paniculata*) typically produces stems that are purple in color (purple morph), while the atypical morph has stems that are green-yellow (green-yellow morph). The purpose of this study was to characterize these newly identified morphs and determine if they are genetically distinct species from the common baby's breath in order to assess whether alternative management strategies should be employed to control these populations.

Methods. We sequenced two chloroplast regions, rbcL and matK, and one nuclear region, ITS2, from the purple morphs and green-yellow morphs collected from Sleeping Bear Dunes National Lakeshore, MI, USA (SBDNL). Sequences were aligned to the reference sequences from other *Gypsophila* species obtained from the Barcode of Life (BOLD) and GenBank databases. We also collected seeds from wild purple morph and wild green-yellow morph plants in SBDNL. We grew the seeds in a common garden setting and characterized the proportion of green-yellow individuals produced from the two color morphs after five-months of growth.

Results. Phylogenetic analyses based upon rbcL, matK, and ITS2 regions suggest that the two color morphs are not distinct species and they both belong to *G. paniculata*. Seeds collected from wild green-yellow morphs produced a significantly higher proportion of green-yellow individuals compared to the number produced by seeds collected from wild purple morphs. However, seeds collected from both color morphs produced more purple morphs than green-yellow morphs.

Discussion. Based upon these results, we propose that the two color morphs are variants of *G. paniculata*. Given the significant difference in the number of green-yellow morphs produced from the seeds of each morph type, we also suggest that this color difference has some genetic basis. We propose that current management continue to treat the two color morphs in a similar manner in terms of removal to prevent the further spread of this species.

Introduction

The Great Lakes sand dunes comprise the most extensive freshwater dune complex in the world, stretching over 1,000 km^2 in Michigan alone. Within northwest Michigan, the sand dunes ecosystem is vital both environmentally and economically. It is home to a number of threatened and endangered species, including piping plover (*Charadrius melodus*) and Pitcher's thistle (*Cirsium pitcheri).* Colonization of invasive species in this region has the potential to significantly alter the biological composition of these native communities (Leege and Murphy, 2001; Emery et al. 2013). One invasive species of significant concern is the perennial baby's breath (*Gypsophila paniculata*). In 2015, baby's breath was listed by the Michigan Department of Natural Resources (DNR) as a "priority" invasive species for detection and control in Michigan's northern lower peninsula (DNR, 2015). Since its colonization in the region it has spread along a 260 km stretch of the Michigan shoreline. Baby's breath produces a large taproot system that can extend down to 4 meters in depth, which likely helps it outcompete native vegetation for limited resources (Darwent & Coupland, 1966; Karamanski, 2000). In addition, while many of the vulnerable and endangered plant species in these areas are seed limited, (e.g., Pitcher's thistle produces approximately 50-300 seeds per plant total or 'per lifetime' (Bevill et al., 1999)), baby's breath can produce up to 14,000 seeds per plant annually (Stevens, 1957), effectively outcompeting native species in terms of overall yield. This has led to baby's breath composing approximately 50-80% of the ground cover in some areas (Karamanski, 2000; Emery et al., 2013).

One concern with current management efforts is that anecdotal evidence suggests there may be a new baby's breath variant within the Michigan dune system. In 2011 and 2012 The Nature Conservancy (TNC) removal crews reported baby's breath plants with different character

traits than what is commonly observed (TNC, 2014). The atypical morph has stems and leaves that are lighter in color and more yellow than the common *G. paniculata* purple morph (Figure 1a-c). The purple morph has a thick taproot (4-7 cm in diameter) just below the caudex that remains unbranched for approximately 60 – 100 cm (Darwent & Coupland, 1966). Severing just below the intersection of the caudex and the taproot is where manual removal efforts target to limit regrowth. However, TNC removal crews suggested that the atypical green-yellow morph's root system seemed to be more diffuse, making it harder to identify a primary taproot and thus, harder to sever without the potential for regrowth (TNC, 2014). Currently, these green-yellow morphs are treated with herbicide application (glyphosate) when observed; however, if this is a newly invaded baby's breath species and it continues to spread into areas where threatened or endangered species are present, removal methods will be a primary concern and alternative management strategies may need to be considered for these populations.

One of the first steps toward adapting current management strategies for this invasive is to identify whether the green-yellow morph is a genetically distinct species from the purple morph. While *G. paniculata* is the dominant invasive baby's breath species in northwest Michigan, a number of other species have been introduced to North America and, specifically, the Great Lakes region (Pringle, 1976, Voss & Reznicek, 2012). For example, *G. elegans*, *G. scorzonerifolia*, *G. muralis,* and *G. acutifolia* have been collected within Michigan (Pringle, 1976; Reznicek et al., 2011; Voss & Reznicek, 2012), and *G. perfoliata* is reported to have become naturalized in the United States (Pringle, 1976). *G. muralis* is an annual and has a very distinct morphology compared to the other *Gypsophila* species identified around the Great Lakes. It typically only reaches 5-20 cm in height, has linear leaves, and commonly produces white to pink flowers (Borkaudah, 1962). *G. elegans*, also an annual, is commonly sold in this

93 region in commercial wildflower packets. It typically has a smaller taproot compared to *G.*
94 *paniculata,* and its coloration can be similar to that observed for the green-yellow morph. *G.*
95 *scorzonerifolia* and *G. acutifolia* are perennials and specimens of these species have been
96 collected in counties within the Great Lakes dune system that also contain *G. paniculata*
97 infestations (Voss, 1957; Pringle, 1976). Both *G. scorzonerifolia* and *G. acutifolia* have a deep
98 taproot and are similar in height to *G. paniculata,* and thus, can superficially resemble *G.*
99 *paniculata* (Voss, 1957; Pringle, 1976). However, both can be distinguished from *G. paniculata*
100 in that their leaves tend to be longer and wider, and the pedicels and calyces are glandular as
101 opposed to glabrous in *G. paniculata* (Voss, 1957; Pringle, 1976). Given the potential for these
102 other species to invade the fragile habitat of the Michigan dune system, the goal of this work was
103 to characterize the genetic relationship between the newly recognized green-yellow morph and
104 the common purple morph to determine if they are the same species.
105
106 **Methods and materials**
107 *DNA Extraction, Amplification, and Sequencing*
108 We collected leaf tissue from 1 green-yellow morph and 16 purple morphs in 2016 and
109 an additional 15 green-yellow morphs in 2017 from Sleeping Bear Dunes National Lakeshore
110 (SBDNL), Empire, MI, USA (specifically: 44.884941 N, 86.062111 W and 44.875302 N,
111 86.056821 W). Plant tissue collections were approved by the National Parks Service (permit ID
112 SLBE-2015-SCI-0013). Leaf tissue was dried in silica gel until DNA extractions could take
113 place. DNA was extracted using a Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Germany).
114 After extraction, the DNA samples were placed through Zymo OneStep PCR inhibitor removal
115 columns (Zymo, Irvine, CA) to remove any secondary metabolites that might inhibit PCR

116 amplification. The DNA for each sample was then quantified using a NanoDrop 2000
117 (ThermoFisher, Waltham, MA).
118 The DNA of green-yellow morphs and purple morphs was amplified at three genetic
119 regions: large subunit of the ribulose-bisphosphate carboxylase gene (rbcL), maturase K (matK),
120 and internal transcribed spacer 2 (ITS2). A combination of ITS region and matK have been used
121 to differentiate between other *Gypsophila* species (specifically, *G. elegans* and *G. repens*) in
122 previous studies (Fior et al., 2006). The rbcL region was amplified using rbcL 1F and rbcL 724R
123 primers (Fay et al., 1997), matK was amplified using matK 390F and matK 1440R primers (Fior
124 et al., 2006), and the ITS2 region was amplified using ITS2 2SF and ITS2 S3R primers (Chen et
125 al., 2010). PCR reactions for all loci consisted of 1X Taq Buffer, 2.0 mM $MgCl_2$, 0.3 μM dNTP,
126 0.08 mg/mL BSA, 0.4 μM forward primer, 0.4 μM reverse primer, and 0.5 units of Taq
127 polymerase in a 20 μL reaction volume. The thermal cycle protocols consisted of the following:
128 for rbcL, an initial denaturing step of 95ºC for 2 minutes, followed by 35 cycles of 94ºC for 1
129 minute, 55ºC for 30 seconds, and 72ºC for 1 minute. A final elongation step was performed at
130 72ºC for 7 minutes. For matK, the thermal profile consisted of 26 cycles of 94ºC for 1 minute,
131 48ºC for 30 seconds, and 72ºC for 1 minute, followed by a final elongation step at 72ºC for 7
132 minutes. For ITS2, an initial denaturing step of 95ºC for 2 minutes was applied, followed by 35
133 cycles of 95ºC for 30 seconds, 50ºC for 30 seconds, 72ºC for 1.5 minutes, and a final elongation
134 step of 72ºC for 8 minutes. Successful amplification was checked by running the PCR product on
135 a 2% agarose gel stained with ethidium bromide. PCR reactions were then cleaned using
136 ExoSAP-IT PCR Product Cleanup Reagent (ThermoFisher, Waltham, MA). Sequencing
137 reactions were performed with the forward and reverse primers for each of the three regions.
138 Sequencing reactions were cleaned using a Sephadex column (GE Healthcare Life Science,

139 Marlborough, MA) and sequenced on an ABI Genetic BioAnalyzer 3130xl (Applied Biosystems,
140 Foster City, CA). Out of the 16 green-yellow morphs a total of 13 were successfully sequenced
141 for rbcL, 13 were successfully sequenced for matK, and 14 were successfully sequenced for
142 ITS2. For the purple morphs a total of 15, 12, and 15 individuals were successfully sequenced
143 for rbcL, matK and ITS2, respectively.

144 Reference sequences for rbcL, matK, and ITS2 of other *Gypsophila spp*. were
145 downloaded from either the Barcode of Life Database (BOLD) (http://www.barcodeoflife.org) or
146 GenBank (https://www.ncbi.nlm.nih.gov/genbank/). We primarily focused on *Gypsophila*
147 species with reported occurrences within the United States, but also incorporated other species if
148 their information was available on BOLD. Sequences of the three regions were not always
149 available for the same species, thus for rbcL these included *G. paniculata*, *G. elegans*, *G.*
150 *fastigiata*, *G. scorzonerifolia*, *G. perfoliata*, and *G. murali*s. For matK the species included *G.*
151 *paniculata*, *G. elegans*, *G. fastigiata*, *G. scorzonerifolia*, *G. perfoliata, G. muralis*, *G. altissima*,
152 and *G. repens*. For the IST2 region, the species included *G. paniculata*, *G. elegans, G.*
153 *scorzonerifolia, G. perfoliata, G. repens,* and *G. acutifolia.* Because sequences for all three
154 regions were not available for all species, our merged phylogeny only contained *G. paniculata*,
155 *G. elegans, G. scorzonerifolia,* and *G. perfoliata* reference sequences. The accession numbers
156 and sequences for all reference species are provided in Supplemental Table 1. All FASTA files
157 corresponding to these data will be deposited in the Dryad database and sequences have been
158 submitted to GenBank (ITS2: MG385003-385031, matK: MG603322-603346, rbcL:
159 MG547346-547373).
160
161

162 *Alignment and Phylogenetic Analysis*

163 All successful sequences from our field samples, as well as sequences for other
164 *Gyposphila* species reference sequences obtained from BOLD or GenBank, were imported into
165 the program MEGA7 (version 7.0.14) (Kumar et al., 2016) and sequences for each of the three
166 regions were aligned both individually and with all sequences combined using Muscle (Edgar,
167 2004). The total number of base pairs aligned and analyzed for each region included: 427 base
168 pair (bp) for rbcL, 702 bp for matK, 201 bp for ITS2, and 1330 bp for the three regions
169 combined. All alignment parameters were kept at their default settings. Once aligned, we used
170 MEGA7 to identify the most appropriate substitution model (rbcL: Jukes-Cantor, matK: Tamura
171 3-parameter, ITS2: Jukes-Cantor, all genes combined: Tamura 3-parameter with gamma
172 distribution). We then created phylogenetic trees using a maximum-likelihood (ML) approach
173 with 500-replicated bootstrap analyses, as well as using neighbor joining, and parsimony models.
174 We also constructed a TCS haplotype network (Clement et al., 2002) based upon the combined
175 sequences using the statistical parsimony approach (Templeton et al., 1992) in the program
176 PopART (v 1.7, http://popart.tago.ac.nz).

177

178 *Color Morph Germination*

179 On May 8, 2018 we planted a total of 207 seeds collected from mature purple morphs
180 and 255 seeds collected from mature green-yellow morphs from SBDNL. For the purple morph
181 seeds, these were collected from a total of 14 plants that were sampled in 2016 (average 15 seeds
182 per plant) and 7 plants that were sampled in 2017 (average 1.7 seeds per plant). For the green-
183 yellow morphs, these seeds were collected from a total of 17 plants (15 seeds per plant) in 2017.
184 Plants were grown in the GVSU – Allendale greenhouse from May until August, 2018. The

185 greenhouse was on a 17 hour light/ 7 hour dark cycle. The average day temperature was 21ºC
186 and night temperature was 15 ºC. In August the plants were transported to the greenhouse at
187 AWRI-GVSU where they were allowed to grow until October 15, 2018. The greenhouse at
188 AWRI-GVSU has no external lighting source or temperature controls, and thus more closely
189 resembled seasonal day/night and temperature cycles. Plants were sampled after a decrease in
190 temperature occurred (from a high of 23 ºC on October 10, 2018 to a high of 11 ºC on October
191 15, 2018). Previous greenhouse observations have found that the differences between the purple
192 and green-yellow morphs can be best detected after a sudden drop in temperature (personal
193 observation, CGP). On October 15, we characterized the color of all individuals that successfully
194 germinated and survived over the 5-month period. We used a chi-square analysis in the R
195 statistical package (v3.5.1) to determine if germination success differed between seeds from the
196 two color morphs and whether the proportion of seeds that developed into green-yellow morphs
197 significantly differed between seeds collected from mature purple and mature green-yellow
198 plants.

199 **Results**

200 Our results indicate that the green-yellow morph identified in SBDNL is not a genetically
201 distinct species from the common purple found throughout SBDNL. The rbcL, matK, ITS2, and
202 combined dataset showed similar patterns with both the green-yellow morphs and the purple
203 morphs clustering together. The phylogenies constructed from rbcL and matK independently
204 show that the two color morphs cluster separately from *G. fastigata*, *G. elegan*, *G. muralis*, and
205 *G. repens*. For the rbcL locus, the relationship of the color morphs to *G. paniculata* and *G.*
206 *scorzonerifolia* was not resolved. Additionally, when we only examined the matK gene, the color
207 morphs clustered separately but within a clade that also included *G. altissima*, *G. scorzonerifolia*,

208 and *G. paniculata*. The ITS2 region was able to provide more resolution between *G. paniculata*,
209 *G. scorzonerifolia*, *G. acutifolia*, and the color morphs, with the color morphs clustering with *G.*
210 *paniculata* (Figure 2a-d; Supplemental Figure 1-8) and separately from the *G. scorzonerifolia*
211 and *G. acutifolia* clade. The same pattern was observed when all regions were analyzed together
212 (Figure 2d), with the exception that this phylogeny did not include *G. acutifolia*. In addition, the
213 TCS haplotype network shows that the purple and green-yellow morphs have shared haplotypes.
214 These two haplotypes are only one mutation away from one another and the *G. paniculata*
215 reference, while the next closest species, *G. scorzonerifolia* is 15 mutations away (Figure 3).
216 This further suggests that both color morphs are *G. paniculata*.
217 Of the regions analyzed for the green-yellow morphs, purple morphs, and reference
218 sequences, rbcL was the most conserved sequence with an overall mean genetic distance (*d)* =
219 0.004, followed by matK (d = 0.015) and ITS2 (d = 0.038). For the ITS2 region, there were six
220 purple morphs and one green-yellow morph that clustered together inside the *G. paniculata*
221 branch (Figure 2c & 2d). Further examination of the electropherograms for these individuals
222 show that they are likely heterozygous at position 138 of our aligned sequence and amplification
223 bias of the 'A' single nucleotide polymorphism (SNP) over the allele containing the 'G' SNP is
224 driving this pattern.
225
226 *Color Morph Germination*
227 Out of the 207 seeds that were collected from mature purple morphs and planted in the
228 greenhouse, 82 successfully germinated and survived over the 5 month period (39.6%). Out of
229 these 82 plants, only one green-yellow morph was produced (1.2%), while the remaining seeds
230 all produced purple morphs. Out of the 255 seeds collected from mature green-yellow morphs

231 and planted in the greenhouse, 105 successfully germinated and survived over the 5 month
232 period (41.2%). This was not significantly different than the proportion of plants that
233 successfully germinated from the purple morph seeds ($\chi^2 = 0.06$, df = 1, p = 0.81). Of the 105
234 successfully germinated seeds from the green-yellow morph plants, 12 developed into green-
235 yellow morph plants (11.4%), 91 developed into purple morphs plants (86.7%), and two plants
236 could not be determined (they appeared to be green-yellow morphs but displayed some dark
237 spots on the stem). The proportion of seeds that produced green-yellow individuals significantly
238 differed between seeds collected from mature green-yellow morphs and seeds collected from
239 mature purple morphs ($\chi^2 = 5.9$, df = 1, p = 0.015).
240
241 **Discussion**
242 Overall, our data suggest that the green-yellow morph is not a genetically distinct species
243 from the purple morph, and that both morphs are *G. paniculata*. For all molecular markers used,
244 the green-yellow and the purple color morphs grouped together. RbcL, matK, and ITS2 are
245 common 'barcode' genes used to delineate plant species (Newmaster et al., 2006; Group et al.,
246 2009; Chen et al., 2010; Yao et al., 2010; Stoeckle et al., 2011) and when used in combination
247 they provided adequate resolution to separate out the *Gypsophila* species included in this study.
248 In our data set, rbcL and matK worked well to separate our color morphs from *G. elegans*, *G.*
249 *muralis* and both of these species have been reported to occur in the Great Lakes region
250 (Reznicek et al., 2011; Voss & Reznicek 2012). While the morphology of *G. muralis* is very
251 distinct from the color morphs in SBDNL, it was initially thought that *G. elegans* shared some
252 similar traits to that originally described by the TNC removal crews and was a potential

253 candidate species for the green-yellow color morph. Based upon these results, this is clearly not
254 the case.

255 The phylogeny based on ITS2 region and the combined sequences provided the best
256 resolution for assigning the relationship of our *Gypsophila* species. Like the rbcL and matK
257 phylogenies, all the purple and green yellow morphs grouped together. For this region, the color
258 morphs also grouped within the same clade as the *G. paniculata* reference sequence. While *G.*
259 *scorzonerifolia* and *G. acutifolia* have also been recorded in the Great Lakes regions (Pringle,
260 1976), and have a similar general phenotype as *G. paniculata*, these species were clearly within a
261 distinct clade that was separate from the two color morphs. Similarly, while *G. perfoliata* has
262 been reported to be naturalized in North America (Pringle, 1976), it grouped outside of the *G.*
263 *paniculata* and color morph cluster.

264 Our greenhouse germination study showed that seeds collected from mature green-yellow
265 morphs produced a significantly higher proportion of green-yellow individuals than seeds
266 collected from mature purple morphs. Of the seeds collected from mature green-yellow morphs
267 11% resulted in green-yellow morphs, while only 1% of seeds from mature purple morphs
268 resulted in green-yellow morphs. However, seeds from both color morphs primarily produced
269 purple morphs. The mechanism driving the color difference between the purple and green-
270 morphs is currently unknown. Within SBDNL, the purple morph is the most common form, with
271 green-yellow individuals found interspersed in a couple of locations throughout the dunes
272 (personal observation, CGP/HLM/ER). The largest observed group of green-yellow morphs
273 consists of a few hundred plants clumped within approximately an acre-sized area and
274 interspersed throughout large groups of purple morphs. Based upon the dispersal patterns of the
275 two morphs throughout the dunes, and our germination results, the color difference observed

276 does not appear to be solely environmentally driven, and likely has a genetic component.
277 Potential candidate genes that could be influencing these color differences include those involved
278 in the anthocyanin pathway, which influences red – purple coloration in a number of plants
279 (Asen et al., 1972; de Pascual-Teresa et al., 2002; Abdel-Aal et al., 2006). In addition,
280 anthocyanin can rapidly accumulate in the shoots of plants following cold exposure (Leng et al.,
281 2000), and we have observed an increase in the amount of purple coloration in the purple morphs
282 as temperatures decrease (personal observation, CGP). Further work will begin to elucidate the
283 specific mechanism influencing this color difference in invasive *G. paniculata* populations, as
284 well as to explore whether this color variation drives functional differences between the morphs.
285 Taken together, these data show that the purple and green-yellow morph within SBDNL
286 are the same species, and that species is *G. paniculata*. One concern with the green-yellow
287 morph initially noted by TNC removal crews was that the taproot tended to be more diffuse than
288 the purple morph, potentially making manual removal of these plants less effective. However, we
289 have not noted differences in the taproot structure between these two morphs when grown under
290 controlled conditions (Figure 4). Additionally, our lab's personal observations (CGP, HLM) in
291 the field have not found any indication that large differences in root structure occur between
292 mature plants of the two color morphs. Therefore, current management approaches for these
293 populations should be maintained to control the further spread of *G. paniculata* throughout the
294 Michigan coastal dune system.
295
296 **Conclusions**
297 Our data show that both the purple and green-yellow color morph of baby's breath in
298 Sleeping Bear Dunes National Lakeshore are *G. paniculata* and the observed color differences

299 likely have some genetic basis. Based on this current information, we recommend that these
300 color morphs continue to be managed in a similar manner and that distinct management
301 strategies do not need to be established at this time.
302

303 **Acknowledgements**

304 We would like to thank Shaun Howard from The Nature Conservancy for his help in identifying
305 the green-yellow morph and Benjamin Giffin for his help with sequencing analysis. We would
306 also like to thank Kurt Thompson and Doug Haywick for assisting with the figure construction,
307 and Alexis Hoskins and Doug Haywick for helping with planting. Funding for this project was
308 provided through an EPA-Great Lake Restoration Initiative Grant and an NSF- Research
309 Experience for Undergraduates Grant.
310

311 **References**

312 Abdel-Aal ESM, Young JC, Rabalski I. 2006. Anthocyanin composition in black, blue, pink,
313 purple, and red cereal grains. *Journal of Agriculture Food Chemistry* 54(13): 4696-4704.
314 Asen S, Stewart RN, Norris KH. 1972. Co-pigmentation of anthocyanins in plant tissues and its
315 effect on color. *Phytochemistry* 11(3): 1139-1144
316 Barkoudah YI. 1962. Revision of Gypsophila, Bolanthus, Ankyropetalum and Phryna. *Wentia* 9:
317 1-203.
318 Bevill RL, Louda SM, Stanforth LM. 1999. Protection from natural enemies in managing rare
319 plant species. *Conservation Biology* 13(6): 1323-1331.
320 Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li X, Jai X,
321 Lin Y, Leon C. 2010. Validation of the ITS2 region as a novel DNA barcode for

322 identifying medicinal plant species. *PloS One* **5**(1): e8613. doi:
323 10.1371/journal.pone.0008613L
324 Clement M, Snell Q, Walker P, Pondosa D, Crandall KA. 2002. TCS: Estimating gene
325 genealogies. Parallel and Distributed Processing Symposium, International Proceedings,
326 2, 184.
327 de Pascual-Teresa S, Santos-Buelga C, Rivas-Gonzalo JC. 2002. LC–MS analysis of
328 anthocyanins from purple corn cob. *Journal of the Science of Food and*
329 *Agriculture* 82(9): 1003-1006.
330 Darwent AL, Coupland RT. (1966). Life history of *Gypsophila paniculata*. *Weeds* 14(4): 313-
331 318.
332 Department of Natural Resources (DNR). 2015. Michigan Invasive Species Grant Program
333 Handbook. http://www.michigan.gov/documents/dnr/2015-MISGP-
334 hndbk_491809_7.pdf?source=govdelivery
335 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
336 throughput. *Nucleic Acids Research* 32(5): 1792-1797.
337 Emery SM, Doran PJ, Legge JT, Kleitch M, Howard S. 2013. Aboveground and belowground
338 impacts following removal of the invasive species baby's breath (*Gypsophila paniculata*)
339 on Lake Michigan sand dunes. *Restoration Ecology* 21(4): 506-514.
340 Fay MF, Swensen SM, Chase MW. 1997. Taxonomic affinities of *Medusagyne oppositifolia*
341 (Medusagynaceae). *Kew Bulletin* 52(1): 111-120.
342 Fior S, Karis PO, Casazza G, Minuto L, Sala F. 2006. Molecular phylogeny of the
343 Caryophyllaceae (Caryophyllales) inferred from chloroplast matK and nuclear rDNA ITS
344 sequences. *American Journal of Botany* 93(3): 399-411.

345 Group CPW, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der
346 Bank M, Chase MW, Cowan RS, Erickson DL, Fazekas AJ. 2009. A DNA barcode for
347 land plants. *Proceedings of the National Academy of Sciences 106*(31): 12794-12797.
348 Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In*:* Munrio HN, ed. *Mammalian*
349 *Protein Metabolism*. New York: Academic Press. pp. 21-132
350 Karamanski TJ. 2000. Nationalized Lakeshore: The Creation and Administration of Sleeping
351 Bear Dunes National Lakeshore. Online Book. National Park Service, Department of the
352 Interior. http://www.cr.nps.gov/history/online_books/slbe/
353 Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis
354 version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33(7): 1870-1874.
355 Lee SC, Wang CH, Yen CE, Chang C. 2017. DNA barcode and identification of the varieties and
356 provenances of Taiwan's domestic and imported made teas using ribosomal internal
357 transcribed spacer 2 sequences. *Journal of Food and Drug Analysis* 25(2): 260-274. doi:
358 10.1016/j.jfda.2016.06.008
359 Leege LM, Murphy PG. 2001. Ecological effects of the non-native *Pinus nigra* on sand dune
360 communities. *Canadian Journal of Botany* 79(4): 429-437.
361 Leng P, Itamura H, Yamamura H, & Deng XM. 2000. Anthocyanin accumulation in apple and
362 peach shoots during cold acclimation. *Scientia Horticulturae*, *83*(1): 43-50.
363 Newmaster SG, Fazekas AJ, Ragupathy S. 2006. DNA barcoding in land plants: evaluation of
364 rbcL in a multigene tiered approach. *Botany 84*(3): 335-341.
365 Pringle JS. 1976. *Gypsophila scorzonerifolia* (Caryophyllaceae), a naturalized species in the
366 Great Lakes region. *Michigan Botanist* 15: 215-219.

367 Reznicek AA, Voss EG, Walters BS. February. 2011. *Michigan Flora Online*. University of
368 Michigan. Web. March 6, 2019. https://michiganflora.net/genus.aspx?id=Gypsophila.
369 Stevens OA. 1957. Weights of seeds and numbers per plant. *Weeds* 5: 46-55.
370 Stoeckle MY, Gamble CC, Kirpekar R, Young G, Ahmed S, Little DP. 2011. Commercial teas
371 highlight plant DNA barcode identification successes and obstacles. *Scientific reports 1*:
372 42.
373 Tamura K. 1992. Estimation of the number of nucleotide substitutions when there are strong
374 transition-transversion and G + C-content biases. *Molecular Biology and Evolution* 9(4):
375 678-687.
376 Templeton AR, Crandall KA, Sing CF. 1992. A cladistics analysis of phenotypic association
377 with haplotypes inferred from restriction endonuclease mapping and DNA sequence data.
378 III. Cladogram estimation. *Genetics* 132: 619-633.
379 The Nature Conservancy (TNC). 2014. Lake Michigan Coastal Dunes Restoration Project. 2013
380 Field Season Report. The Nature Conservancy and Partners, Lansing, Michigan
381 Voss, EG. 1957. Observations on the Michigan flora—VI. Distribution records of some
382 angiosperms new, rare, or misinterpreted in the state. *Brittonia 9*(2): 83-101.
383 Voss EG, Reznicek, AA. 2012. *Field manual of Michigan flora*. Ann Arbor: University of
384 Michigan Press.
385
386 Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao H, & Chen S. 2010. Use
387 of ITS2 region as the universal DNA barcode for plants and animals. *PloS One 5*(10):
388 e13102.
389
390

391 **Figure Legends**

392 Figure 1: (A) Green-yellow morph, (B) common purple morph and (C) stem of the green-yellow
393 and purple baby's breath morph found in Sleeping Bear Dunes National Lakeshore.

394

395 Figure 2: Phylogenetic analysis of the purple and green-yellow baby's breath color morphs in
396 relationship to other *Gypsophila* species. All evolutionary histories were inferred using
397 maximum likelihood methods. (A) Phylogeny based on the rbcL region, (B) phylogeny based on
398 the matK region, (C) phylogeny based on the ITS2 region, (D) Phylogeny based on rbcL, matK,
399 and ITS2 combined. For rbcL and ITS2 we used a Jukes Cantor (JC) model of molecular
400 evolution (Jukes and Cantor, 1969). For matK we used a Tamura 3-parameter (T92) model of
401 molecular evolution with uniform distribution, and for the combined data set we used the T92
402 model of molecular evolution with a gamma distribution (Tamura 1992).

403

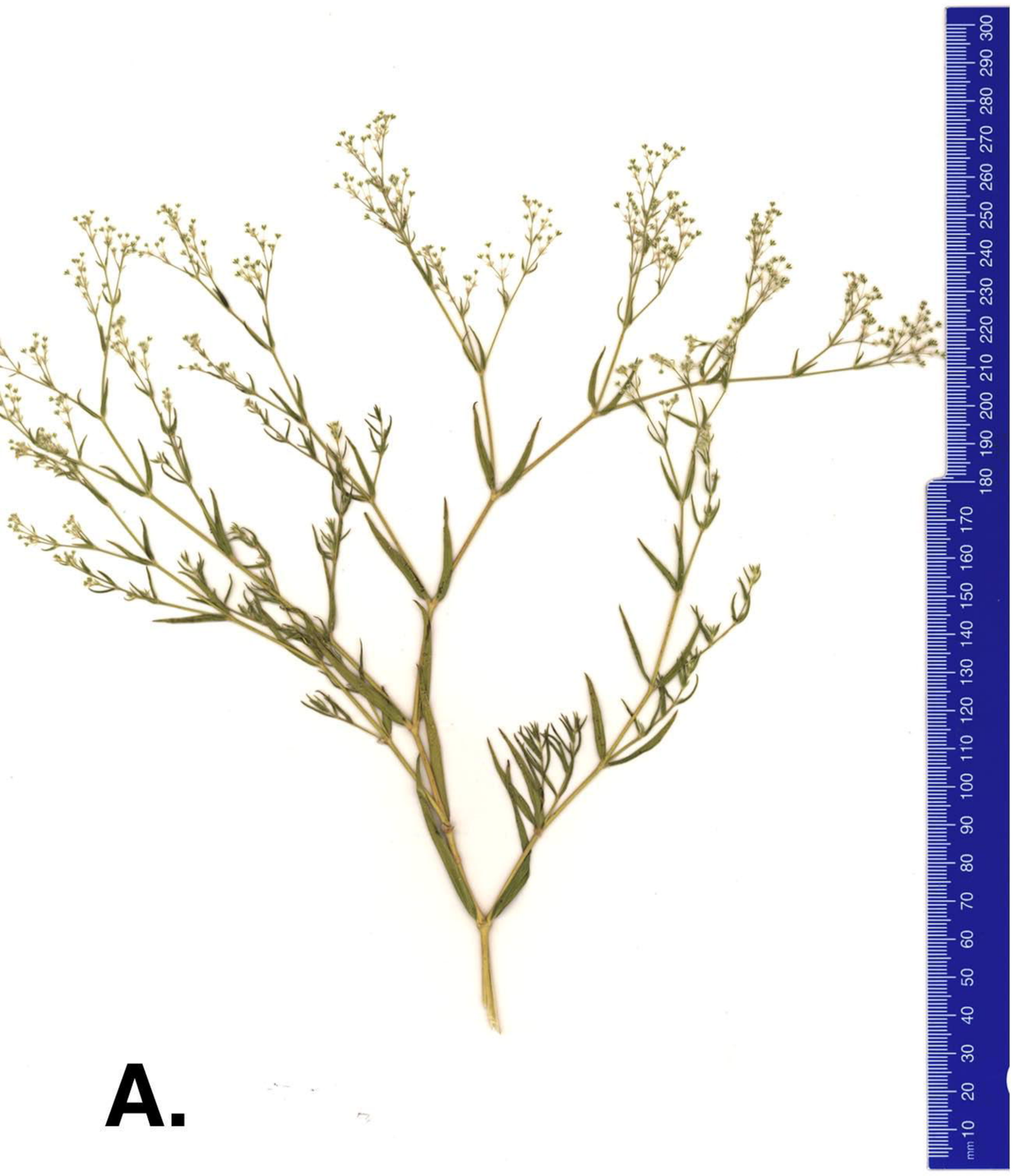
404 Figure 3: A TCS haplotype network based on rbcL, matK and ITS2 combined for the purple and
405 green-yellow baby's breath color morphs and the *G. paniculata*, *G. elegans*, *G. perfoliata*, and
406 *G. scorzonerifolia* reference sequences. The size of the ovals correspond to the haplotype
407 frequency. The hash marks represent the number of mutations between each haplotype.

408

409 Figure 4: (A) Green-yellow morph, and (B) common purple morph after 5-months in the GVSU
410 greenhouse. Note the similarity in taproot structure between the two plants.

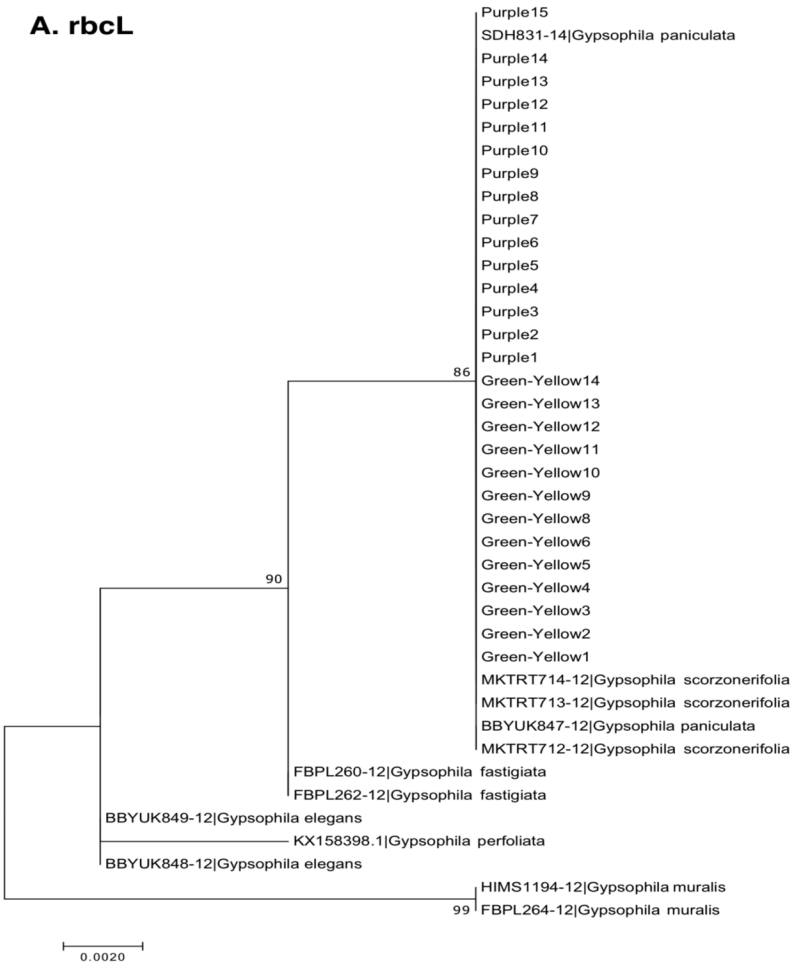
411

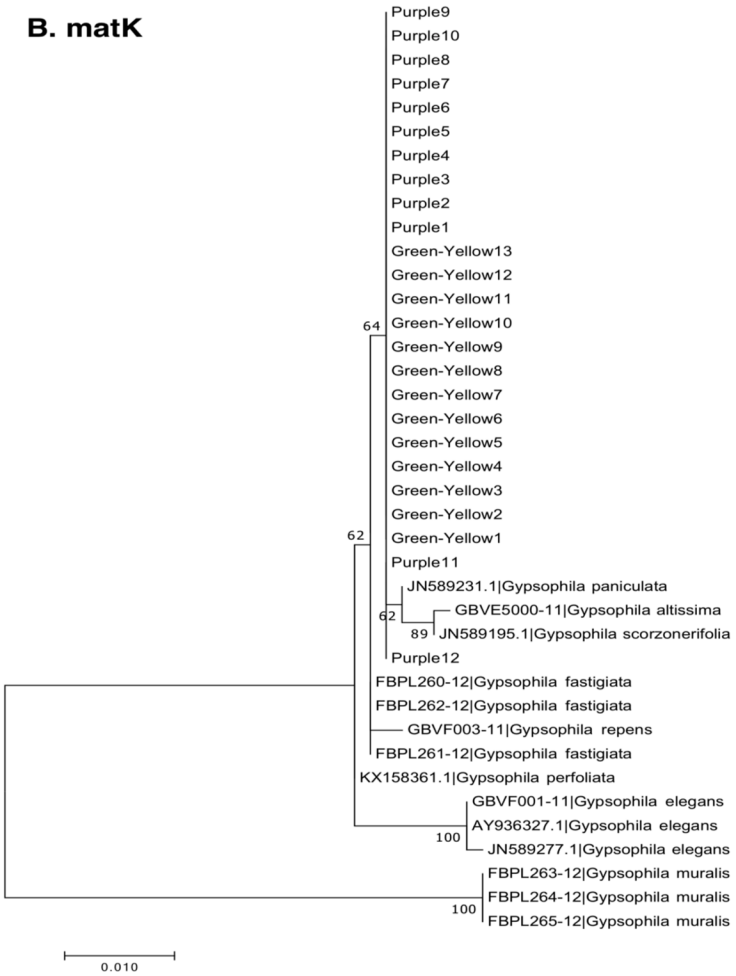
412
413

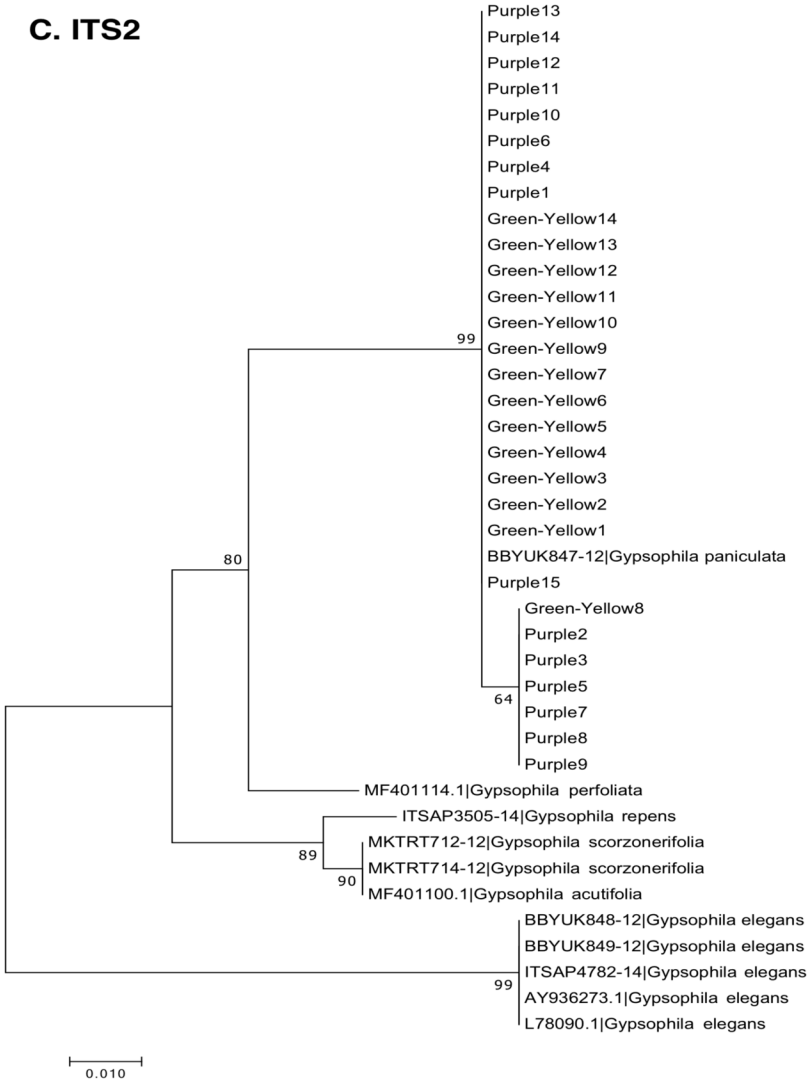

A.

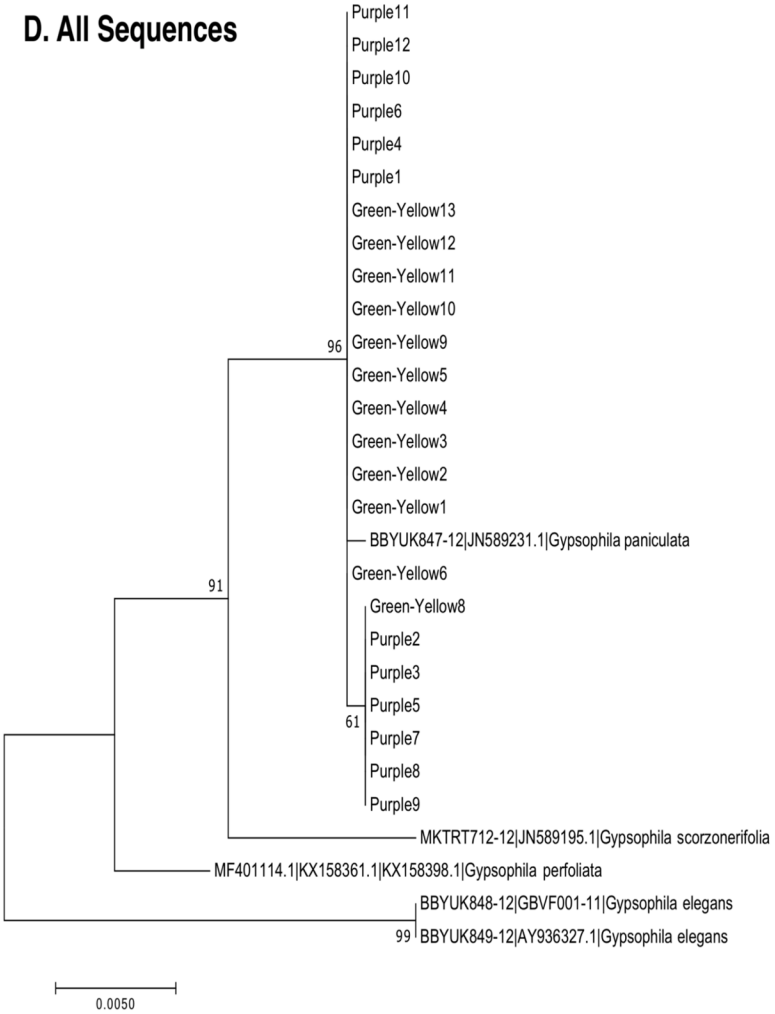
B.

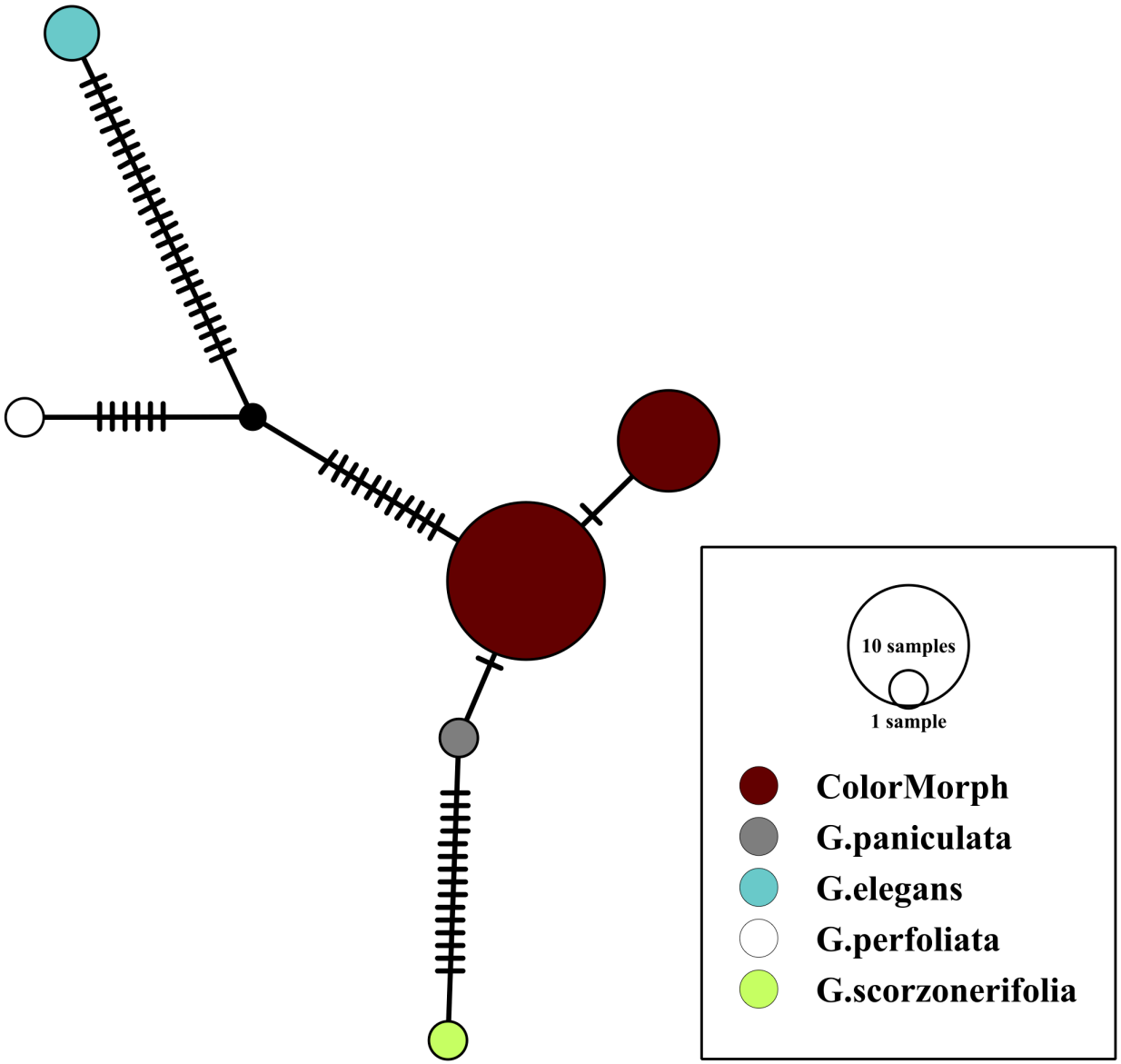
C.


A. rbcL
Purple15
SDH831-14|Gypsophila paniculata
Purple14
Purple13
Purple12
Purple11
Purple10
Purple9
Purple8
Purple7
Purple6
Purple5
Purple4
Purple3
Purple2
Purple1
86
Green-Yellow14
Green-Yellow13
Green-Yellow12
Green-Yellow11
Green-Yellow10
Green-Yellow9
Green-Yellow8
Green-Yellow6
Green-Yellow5
90
Green-Yellow4
Green-Yellow3
Green-Yellow2
Green-Yellow1
MKTRT714-12|Gypsophila scorzonerifolia
MKTRT713-12|Gypsophila scorzonerifolia
BBYUK847-12|Gypsophila paniculata
MKTRT712-12|Gypsophila scorzonerifolia
FBPL260-12|Gypsophila fastigiata
FBPL262-12|Gypsophila fastigiata
BBYUK849-12|Gypsophila elegans
KX158398.1|Gypsophila perfoliata
BBYUK848-12|Gypsophila elegans
HIMS1194-12|Gypsophila muralis
99
FBPL264-12|Gypsophila muralis
0.0020


B. matK
Purple9
Purple10
Purple8
Purple7
Purple6
Purple5
Purple4
Purple3
Purple2
Purple1
Green-Yellow13
Green-Yellow12
Green-Yellow11
64
Green-Yellow10
Green-Yellow9
Green-Yellow8
Green-Yellow7
Green-Yellow6
Green-Yellow5
Green-Yellow4
Green-Yellow3
Green-Yellow2
62
Green-Yellow1
Purple11
JN589231.1|Gypsophila paniculata
62
GBVE5000-11|Gypsophila altissima
89
JN589195.1|Gypsophila scorzonerifolia
Purple12
FBPL260-12|Gypsophila fastigiata
FBPL262-12|Gypsophila fastigiata
GBVF003-11|Gypsophila repens
FBPL261-12|Gypsophila fastigiata
KX158361.1|Gypsophila perfoliata
GBVF001-11|Gypsophila elegans
AY936327.1|Gypsophila elegans
100
JN589277.1|Gypsophila elegans
FBPL263-12|Gypsophila muralis
FBPL264-12|Gypsophila muralis
100
FBPL265-12|Gypsophila muralis
0.010


C. ITS2
Purple13
Purple14
Purple12
Purple11
Purple10
Purple6
Purple4
Purple1
Green-Yellow14
Green-Yellow13
Green-Yellow12
Green-Yellow11
Green-Yellow10
99
Green-Yellow9
Green-Yellow7
Green-Yellow6
Green-Yellow5
Green-Yellow4
Green-Yellow3
Green-Yellow2
Green-Yellow1
80
BBYUK847-12|Gypsophila paniculata
Purple15
Green-Yellow8
Purple2
Purple3
Purple5
64
Purple7
Purple8
Purple9
MF401114.1|Gypsophila perfoliata
ITSAP3505-14|Gypsophila repens
MKTRT712-12|Gypsophila scorzonerifolia
89
MKTRT714-12|Gypsophila scorzonerifolia
90
MF401100.1|Gypsophila acutifolia
BBYUK848-12|Gypsophila elegans
BBYUK849-12|Gypsophila elegans
ITSAP4782-14|Gypsophila elegans
99
AY936273.1|Gypsophila elegans
L78090.1|Gypsophila elegans
0.010


D. All Sequences
Purple11
Purple12
Purple10
Purple6
Purple4
Purple1
Green-Yellow13
Green-Yellow12
Green-Yellow11
Green-Yellow10
96
Green-Yellow9
Green-Yellow5
Green-Yellow4
Green-Yellow3
Green-Yellow2
Green-Yellow1
BBYUK847-12|JN589231.1|Gypsophila paniculata
Green-Yellow6
91
Green-Yellow8
Purple2
Purple3
Purple5
61
Purple7
Purple8
Purple9
MKTRT712-12|JN589195.1|Gypsophila scorzonerifolia
MF401114.1|KX158361.1|KX158398.1|Gypsophila perfoliata
BBYUK848-12|GBVF001-11|Gypsophila elegans
99
BBYUK849-12|AY936327.1|Gypsophila elegans
0.0050


10 samples
1 sample
ColorMorph
G.paniculata
G.elegans
G.perfoliata
G.scorzonerifolia

A
Gpyel 18-J

B
bbp220A