bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Title: Old Meets New: Combining Herbarium Databases with Genetic Methods to Evaluate the Invasion Status of Baby's Breath (*Gypsophila paniculata*) in North America

Running Title: Lamar & Partridge: Tracing an Invasion

Sarah K. Lamar[1,2] and Charlyn G. Partridge[1]

[1]Robert B. Annis Water Resources Institute, Grand Valley State University

740 West Shoreline Drive, Muskegon, MI 49441 USA

[2]Corresponding author: Sarah K. Lamar

Robert B. Annis Water Resources Institute, Grand Valley State University

740 West Shoreline Drive, Muskegon, MI 49441 USA

Telephone: (616) 331-3749

Fax: (616) 331-3864

Email: lamarsarahk@gmail.com

Acknowledgements

The authors would like Hailee Leimbach-Maus and Emma Rice for their help in the field. Additionally, the authors would like to thank the Environmental Protection Agency – Great Lakes Restoration Initiative (C.G.P., Grant #00E01934), the Michigan Botanical Foundation, and Grand Valley State University for financial support.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Aim: This paper aims to inform our knowledge of common baby's breath's (*Gypsophila paniculata*) current population structure and invasion status using a combination of contemporary genetic methods and historical herbarium data.

Taxon: *Gypsophila paniculata* (Angiosperms: Eudicot, Caryophyllaceae)

Location: Samples were collected from seven locations spanning a portion of the plant's North American range: Washington, North Dakota, Minnesota, and Michigan, United States.

Methods: To analyze contemporary population structure, individuals of *G. paniculata* from 7 distinct sampling locations were collected and genotyped at 14 microsatellite loci. Population structure was inferred using both Bayesian and multivariate methods. To investigate *G. paniculata*'s invasion status, public herbarium databases were searched for mention of the species. Records were combined, resulting in a database of 307 herbarium collections dating from the late 1800's to current day. Using this database, invasion curves were created at different geospatial scales.

Results: Results of genetic analyses suggest the presence of at least two genetic clusters spanning our seven sampling locations. Sampling locations in Washington, North Dakota, Minnesota, and northwestern Michigan form one genetic cluster, distinct from our two more southern sampling locations in Michigan, which form a second cluster with increased relative genetic diversity. Invasion curves created for these two clusters show different time periods of invasion. An invasion curve created for North America suggests *G. paniculata*'s range may still be expanding.

Main conclusions: *Gypsophila paniculata* has likely undergone at least two distinct invasions in North America, and its range may still be expanding. Restricted genetic diversity seen across a wide geographic area could be a signature of limited seed distributors present during the early period of this garden ornamental's invasion.

Keywords: Baby's breath, genetic structure, *Gypsophila paniculata*, herbarium data, invasion history, invasive species

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Biological invasions are a growing concern in the era of global trade and transport. In the United States alone, there have been over 50,000 introductions of plant, animal, and microbe species into environments beyond their native range (Pimentel, Zuniga, & Morrison, 2005). These introductions can have dramatic impacts on native flora and fauna; roughly 42% of species listed on the Endangered Species Act are threatened by competition with invasives (Wilcove, Rothstein, Dubow, Phillips, & Losos, 1998). Of particular concern among invasive species are invasive weeds, a group that currently spreads across the United States at a rate of 700,000 ha/year (Pimentel et al., 2005). This rapid consumption of land by non-native species makes managing invasive weeds a priority for the preservation of native ecosystems and the native biota that inhabit them.

Many plant and animal species that are transported into new environments will not become problematic invaders, defined as species not native to an area whose range or abundance is increasing regardless of habitat (P. Pyšek, 1995; Williamson & Fitter, 1996). Non-native species that go on to become invasive in their new environments must survive transport, reproduce as a relatively small founding population, respond to potentially novel environmental stressors, and overcome the “lag phase” of an invasion (Larkin, 2012; Williamson & Fitter, 1996). This lag phase is characterized by a period of slow growth after initial introduction that, if overcome, can lead to a period of rapid population expansion before eventually plateauing as the new range is saturated (Mack et al., 2000). Despite the many barriers that species face on the road to becoming invasive, the impacts of these events are a growing cause for concern.

As the number of global invasion events increases, so does the importance of developing and implementing cost effective methods for studying invasion events. Invasion curves are one

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

such tool used to assess an invasive species' status and rate of spread (see Antunes & Schamp, 2017; Shih & Finkelstein, 2008). Invasion curves can offer researchers important insight to a species' lag time after introduction into new environments, providing valuable information associated with response time, geographic barriers to spread, and the efficacy of existing management strategies (Antunes & Schamp, 2017; Crooks, 2007). Because they are crafted using historical data, such as herbarium records, invasion curves are both cost effective and capable of offering important glimpses into the often-unnoticed lag phase of an invasion (Antunes & Schamp, 2017). Invasion curves have been used to recognize potential refuges for weed species (e.g. Lavoie, Jodoin, & De Merlis, 2007), identify major drivers of invasive species spread (e.g. Fuentes, Ugarte, Kühn, & Klotz, 2008; Petr Pyšek, Jarošík, Müllerová, Pergl, & Wild, 2008), and even help assess the efficacy of potential biocontrol agents (e.g. Boag & Eckert, 2013).

While invasion curves are useful for addressing many questions managers and researchers may have, they are limited by the constraints associated with herbarium records and survey data. To supplement these constraints, genetic analyses may be used to provide information concerning contemporary gene flow, adaptive potential, relatedness among invasive populations, and possible resistance to control efforts (e.g. Abdelkrim, Pascal, Calmet, & Samadi, 2005; Zalewski et al., 2010). Genetic analyses of invasive species has been used to identify potential barriers to migration (Haynes, Gilligan, Grewe, & Nicholas, 2009) and estimate the number of likely invasion events a species may have undergone (Meimberg et al., 2010) While this information can help improve our understanding of invasive science as a whole, it also has immediate benefits to managers; because distinct genetic populations have different

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

potential evolutionary trajectories, understanding the genetic structure of populations is critical for effective management (Moritz, 1994; Palsbøll, Bérubé, & Allendorf, 2007).

Gypsophila paniculata (common baby's breath) is a perennial forb native to the Eurasian steppe region (Darwent, 1975; Darwent & Coupland, 1966). *Gypsophila paniculata* is characterized by a taproot that can reach several meters deep, which is thought to help the plant to out-compete natives for limited resources in harsh environments (Darwent & Coupland, 1966). Though it does not produce floral primordia until at least its second year, *G. paniculata* can yield almost 14,000 seeds per growing season (Darwent & Coupland, 1966; Stevens, 1957). These seeds are small (86mg/100 seeds) and primarily distributed by wind forces; when plants reach senescence, they break off above the caudex and form tumbleweeds that spread seeds as they roll (Darwent & Coupland, 1966; Stevens, 1957).

Populations of *G. paniculata* were established in North America by the late 1880's, likely having been introduced due to its popularity in the garden and floral industries (Darwent & Coupland, 1966). According to the Early Detection and Distribution Mapping System, *G. paniculata* can now be found growing as an invasive species in 30 U.S. states (EDDMapS, 2019). It has been listed as a Class C (widespread noxious weed) in Washington and California and is considered a priority invasive by Michigan Department of Natural Resources (Emery & Doran, 2013; Michigan Department of Natural Resources, 2015; Swearingen & Bargeron, 2016). *Gypsophila paniculata* can form dense stands in the areas that it invades; in some parts of Sleeping Bear Dunes National Lakeshore, an invaded area in Michigan, *G. paniculata* forms as much as 75% of the vegetation present (Karamanski, 2000; Rice, 2018). These dense monocultures can have impacts on native plant, nematode, and arthropod communities,

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

potentially having ripple effects across the trophic system (Emery & Doran, 2013; Reid & Emery, 2018).

To help understand the invasion status of this problematic plant species, this study aims (1) to define the population structure of contemporary *G. paniculata* growing throughout a portion of its introduced range, and (2) to create invasion curves of *G. paniculata* to assess its current invasion status at different geospatial scales.

Methods

Study Sites and Contemporary Sample Collection

To investigate contemporary population structure of *G. paniculata*, tissue samples from five locations across the United States were collected in the summer of 2018: Petoskey, MI; Knife River Indian Villages National Historic Site, ND; Ottertail, MN; Chelan, WA; and Osborne Bay, WA (Figure 1, Table 1). Samples from two additional locations in Sleeping Bear Dunes National Lakeshore, MI and Arcadia Dunes, MI were collected in the summer of 2016 (Table 1) (Leimbach-Maus, Parks, & Partridge, 2018a). Leaf tissue was collected from 15-30 individuals per location (5-10 leaves per plant). Tissue samples were placed inside coin envelopes and stored in silica until DNA extraction. Individuals were collected for sampling by identifying a plant of any size separated from other sampled individuals by at least 2 meters, in efforts to minimize the likelihood of sampling closely related plants.

Microsatellite Analysis

For each sample (n=145), 0.25 g of dried leaf tissue was weighed out. DNA was extracted from tissue samples using a Qiagen DNeasy plant mini kit (QIAGEN, Hilde, Germany); manufacturer instructions were largely followed, apart from an extra wash with AW2 buffer. Extracted DNA was run through a Zymo OneStep PCR Inhibitor Removal Column (Zymo, Irvine, CA) twice.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Samples were amplified at 14 nuclear microsatellite loci identified as polymorphic and specific to *G. paniculata* (Leimbach-Maus, Parks, & Partridge, 2018b) (Table S1). PCR was conducted using a 5' fluorescently-labelled primer (6-FAM, PET, NED, or VIC) (Applied Biosystems, Foster City, CA) and an unlabeled reverse primer. Reaction mixtures consisted of 1x KCl buffer, 2.0-2.5 mM $MgCl_2$, 300 µM dNTP, 0.08 mg/mL BSA, 0.4 µM forward primer, 0.4 µM reverse primer, 0.25 units Taq polymerase, and 50 ng DNA template. The thermal cycling profile consisted of 5 minutes of denaturation at 94˚C, followed by 35 cycles of 94˚C for 1 minute, 1 minute of annealing at 62˚ (with the exception of locus BB_2888, see Table S1), 1 minute of extension at 72˚C, and a final elongation step of 10 minutes at 72˚C. PCR products were visualized on a 2% agarose gel using GelRed™ (Biotium, Freemont, CA) before multiplexing with consideration to dye color and allele size (Table S1). Genescan 500 LIZ size standard (Thermo Fisher Scientific, Waltham, MA) was added to multiplexed product with Hi-Di™ Formamide (Thermo Fisher Scientific, Waltham, MA) to aid in denaturing. Fragment analysis was conducted on an ABI3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA). Individuals were genotyped using the automatic binning procedure on GENEMAPPER v5 (Applied Biosystems, Foster City, CA) before being visually verified to reduce error. A subsample of 20 individuals were genotyped twice to ensure consistent allele scoring.

Exploratory Data Analysis

The presence of null alleles was investigated using MICRO-CHECKER v2.2.3; using this method, none were found (Van Oosterhout, Hutchinson, Wills, & Shipley, 2004). Data were screened using the 'STRATAG' package in the R statistical program v3.4.3 (Archer, Adams, & Schneiders, 2016; R Development Core Team, 2017) for any individual that was missing greater than 20% of

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

loci and any locus that was missing greater than 10% of individuals; on this basis, no data were removed.

Measures of Genetic Diversity

Linkage disequilibrium and a test for Hardy-Weinberg equilibrium were calculated using GENEPOP v4.6 with 1,000 batches of 1,000 Markov chain Monte Carlo iterations (Raymond & Rousset, 1995; Rousset, 2008). There was no significant deviation from linkage equilibrium across populations and no data were removed on this basis. Expected versus observed heterozygosity, number of private alleles, and Weir and Cockerham's population pairwise F_{ST} values were conducted using GENALEX v6.502 in Microsoft Excel (Peakall & Smouse, 2006, 2012; Weir & Cockerham, 1983). Inbreeding coefficient (F_{IS}) values were calculated in GENEPOP.

Genetic Structure

A principal coordinate analysis (PCoA) was conducted using a genetic distance matrix in GENALEX (Peakall & Smouse, 2006, 2012). Population clustering was analyzed in STRUCTURE v2.3.2(Pritchard, Stephens, & Donnelly, 2000) using an admixture model, both with and without *a priori* location information, and a burn-in length of 100,000 with 1,000,000 MCMC replicates after burn-in. Ten iterations were run for each *K* value (1-9). The number of genetic clusters was determined using the Evanno ΔK method (Evanno, Regnaut, & Goudet, 2005). Because ΔK is based on a rate of change, it does not evaluate *K*=1 and can be biased towards *K*=2 (Dupuis et al., 2017). Considering this, we also used discriminant analysis of principal components (DAPC) to support our STRUCTURE findings (Jombart, Devillard, & Balloux, 2010). DAPC separates variance into within-group and between-group categories and works to maximize cluster

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

discrimination; this analysis was conducted using the package 'adegenet' v2.1.1 in R (Jombart et al., 2010). Because retaining too many principal components (PC's) can lead to instability in cluster membership properties, a cross-validation was performed to inform the analysis of the optimal number of PC's. After cross-validation, 16 of 28 PC's and all eigenvalues were retained. An analysis of molecular variance (AMOVA) was run using 9,999 permutations in GENALEX to test how much variance could be explained by between-population and within-population variation.

Herbarium Invasion Curves

To create invasion curves for *G. paniculata* population clusters, public herbarium databases were searched for records of this species; species identification was visually confirmed when possible. Records that did not include location data (either GPS, county (U.S.) or regional municipality (Canada)) and year were discarded, resulting in 307 records from 65 North American institutions. All locality information was standardized to the county scale to reduce the risk of redundant specimen collection while maintaining adequate resolution (Antunes & Schamp, 2017). Earliest samples were found in the late 1890's-early 1900s in California, Michigan, Minnesota, and New York and this is consistent with the earliest times in which *G. paniculata* seeds were first being sold in the United States (1886), based on a search of the Henry G. Gilbert Nursery and Seed Trade Catalog Collection from the Biodiversity Heritage Library (https://www.biodiversitylibrary.org/).

To examine the invasion status of populations belonging to genetic clusters identified from our population genetics analysis, herbarium records were grouped according to desired geospatial scales (cumulative North America, genetic cluster 1, and genetic cluster 2). Only records for the first occurrence of *G. paniculata* in each county or regional municipality were

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

kept. Cumulative records for North America had 184 unique municipalities represented, while both genetic clusters had fewer unique localities (cluster 1 = 42, cluster 2 =16) and required log transformation for better visualization. Data were plotted as the cumulative number of localities invaded over time using the statistical program R.

Results

Measures of Genetic Diversity

Overall, the five western populations and northernmost Michigan population showed lower levels of genetic diversity compared with the two more southern populations in Michigan (Table 2). Pairwise comparisons yielded significant F_{ST} values between all populations; however, SBD-MI, AD-MI, and KR-ND showed comparatively high pairwise F_{ST} values compared to other populations (Table 3). F_{ST} values between CH-WA, OB-WA, and PS-MI were relatively low compared to other sample locations in this study, suggesting more limited genetic differentiation among these populations (Table 3).

Genetic Population Structure

Results of the Bayesian clustering analysis conducted in the program STRUCTURE suggest two population clusters (*K*=2), both from ΔK and Ln Pr (X|K) (Figure S1). Analysis was conducted both with and without prior sampling location; there was no observable difference between the two (without priors shown in Figure 2). Cluster 1 is comprised of sampling locations in North Dakota, Minnesota, Washington, and the northernmost site in Michigan; cluster 2 is comprised of the two more southern sites in Michigan (Figure 2). Overall, there is little admixture between the two groupings, with only few individuals in AD-MI showing any signs of genetic mixing.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Population structure was further analyzed with a PCoA based on a genotypic distance matrix. Population division along the primary principal coordinate accounted for 27.22% of variation present. Along this coordinate, the trends seen in STRUCTURE analysis were supported, with populations SBD-MI and AD-MI separating out from the remaining five populations (Figure 3). The secondary principal component suggests further separation may exist between SBD-MI and AD-MI (9.80% of variation present) if *K* is forced to 3. The grouping of CH-WA, OB-WA, OT-MN, KR-ND, and PS-MI into the same cluster is supported by this analysis.

DAPC's Bayesian Information Criterion suggested either 2 or 3 genetic clusters (Figure S2). Sampling locations in Arcadia Dunes, MI and Sleeping Bear Dunes, MI separated into distinct populations when *K* was pushed to 3, in order to investigate all cluster possibilities (Figure 4a). Individual membership to clusters is detailed in Figure 4c, which shows that cluster 1 is 82% comprised of individuals from SBD-MI, cluster 2 is 93% comprised of individuals from AD-MI, and cluster 3 has a relatively even contribution of individuals from CH-WA, OB-WA, KR-ND, OT-MN, and PS-MI. When individual distribution is viewed along the primary discriminant function, overlap between clusters 1 (SBD-MI) and 2 (AD-MI) is clearly visible (Figure 4b), while cluster 3 shows little to no overlap with clusters 1 or 2.

AMOVA results show that a significant amount of variation could be explained by differences among populations within regions (ϕ_{PR}= 0.229, $p < 0.001$) and by differences between our first region (CH-WA, OB-WA, KR-ND, OT-MN, PS-MI) and second region (SBD-MI and AD-MI) (ϕ_{RT} = 0.246, $p < 0.001$). However, most variation present was found within populations (ϕ_{PT} = 0.419, $p<0.001$).

Invasion Curves

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Invasion curves created using herbarium records, standardized to the scale of local municipality, were used to visualize the invasion stage (i.e. lag phase, expansion phase, or plateau phase) of *G. paniculata* at various geospatial scales (Figure 5). Records for North America slowly accumulate during the early periods of invasion (1890's) until roughly the 1940's, after which the number of records being collected in new localities begin to accumulate rapidly (Figure 5b). This likely represents the shift from the initial lag phase of invasion to the expansion phase. With no clear plateau being reached, the expansion phase of *G. paniculata* across the entirety of North America appears to continue. Considering records according to assignment with genetic cluster, initial collection for cluster 1 (WA, ND, MN, and PS-MI) is noted in the late 1890's, but few additional records were archived until the mid-1920's, when herbarium data for *G. paniculata* suggest an expansion of this population (Figure 5c). A plateau can be seen beginning in the mid 1990's when the curve of the line begins to taper. Records for genetic cluster 2 (Figure 5d) are comprised of collections from mid-southwest Michigan (defined as south of the Leelanau Peninsula, based on results from this study and a previous study conducted by Leimbach-Maus et al., 2018a). Rapid expansion began shortly after its first collection in the late 1940's, with the spread beginning to plateau around 1970. No discernable lag period is noted in the collection data for this cluster.

Discussion

Our data from populations of *G. paniculata* growing across a portion of its introduced range in North American reveals the presence of at least two distinct genetic clusters. The northernmost sampling location in Michigan (PS-MI) clustered with the four sampling locations located across North Dakota, Minnesota, and Washington, and separately from the two southernmost sampling locations in Michigan. When further structuring was explored, the two

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

MI locations (AD-MI and SBD-MI) separated out into their own genetic clusters, though overlap was clearly visible when viewing discriminant functions. The two more southern sampling locations in Michigan also had higher levels of genetic diversity than the other five sampling locations.

There are likely multiple factors contributing to the genetic patterns that we observe across these populations of invasive baby's breath. The increased levels of genetic diversity observed in the SBD-MI and AD-MI populations compared with the other sampled locations could be due to a combination of population size and connectivity. Populations located in SBD-MI tend to be much larger than other locations sampled in this study. Larger populations tend to be more robust to the effects genetic drift and can help resist the effects of inbreeding, helping to retain diversity within these populations (see Ellstrand & Elam, 2003). Another possible reason for the patterns found here is that sampling locations spread across the western U.S. are more isolated than the two southernmost Michigan locations, which may be contributing to lower levels of genetic diversity among these areas. Several sample locations (CH-WA, OT-MN) occur in relatively fragmented or space-limited environments, which may result in a lack of gene flow to other populations of *G. paniculata* growing nearby or prevent its spread altogether. The close geographic proximity between SBD-MI and AD-MI could also be maintaining some gene flow between these populations. However, many of our other sample locations with limited genetic diversity (OB-WA, PS-MI, KR-ND) were part of a contiguous landscape that was not obviously limiting to expansion.

One potential explanation for the distinct genetic clustering we observed with our data is that the populations of SBD-MI and AD-MI that were established in the 1940's could have been founded by individuals from the existing PS-MI population. SBD-MI and AD-MI could then

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

have significantly diverged from the initial source over the past 50 years. However, this scenario seems unlikely. Our data show that SBD-MI and AD-MI have higher levels of genetic variation compared to PS-MI and a number of private alleles were found in both SBD-MI and AD-MI that are not present in PS-MI. Additionally, chloroplast microsatellite data from a previous study (Leimbach-Maus et al., 2018a) show that the SBD-MI and AD-MI populations have distinct DNA haplotypes compared to the PS-MI population and other more northern Michigan populations not included in this study. The combination of these data suggest that SBD-MI and AD-MI are likely not the result of serial founding events from the source population of PS-MI.

A more likely explanation for the distinct patterns observed among our populations could be a signature of *G. paniculata*'s horticultural past. The earliest occurrences of *G. paniculata* populations across several different regions in the U.S. coincides with its initial introduction to N. America though seed sales. Based upon seed catalogs from the Biodiversity Heritage Library, *G. paniculata* was promoted as a garden ornamental as early as 1856 in the *Farmer's Promotion Book* (Reinhold, 1856). By 1868 at least two seed distributors (J.M. Thorburn & Co, NY and Hovey & Nichols, Chicago) were selling *G. paniculata* in their catalogs in New York and Chicago; the earliest herbarium records of *G. paniculata* collected in the United States were from CA (1907), MN (1896), MI (1913), and NY (1894) (Table S2). We hypothesize that when *G. paniculata* initially invaded N. America in the late 1890's there may have been little standing genetic diversity present in the garden cultivars being grown at the time. Additionally, the number of oversees distributors of seeds may have been further restricting possible diversity. These potential limitations to genetic diversity during the early periods of invasion are likely why some of our populations cluster together, despite the large geographic distances between them. According to herbarium records, populations of *G. paniculata* in SBD-MI and AD-MI

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were not established until the later 1940's, when *G. paniculata* had become a more popular garden ornamental. This increased popularity likely led to the number of seed distributors being greatly increased. We suggest then that the genetic patterns observed in this study among populations of *G. paniculata* are a signature of the horticultural past that helped facilitate its invasion into N. America.

One confounding factor to our genetic analyses is that tissue from the SBD-MI and AD-MI populations was collected two-years prior to the other locations (2016 compared to 2018), which could potentially impact our structure results. However, the same Leimbach-Maus et al. (2018a) study examining the genetic structure of population throughout west Michigan also found that baby's breath populations north of the Leelanau Peninsula (i.e., PS-MI) group in a cluster that is distinct from both SBD-MI and AD-MI. Samples from this study were all collected in the same year. Thus, the distinct clustering of PS-MI from the SBD-MI and AD-MI populations appears to be well supported.

Invasion curves created at multiple geospatial scales help assess the current invasion status of *G. paniculata* across its introduced range in North America. Herbarium records compiled for North America indicate that *G. paniculata* has likely not yet reached a plateau phase, and its range could still be expanding. When this larger invasion is viewed at a finer geospatial scale, additional trends become visible. Herbarium records collected from the geographic area of cluster 1 (Washington, North Dakota, Minnesota, and northwestern Michigan) show a lag period that ended in the 1920's as *G. paniculata* collection increased in new localities and its range began expanding. The invasion curve created for cluster 2 (Michigan south of the Leelanau Peninsula) shows that the expansion phase was already in process during the first collection period or shortly after, with little lag phase observed. Whether this is because

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

G. paniculata was present within the region prior to this period but not collected until the 1940's, or whether populations were not present in this area until the 1940's and began spreading rapidly shortly after introduction is unclear. Regardless, the expansion in this region was in progress in the mid 1940's, with a plateau in new localities invaded taking place around 1970. These distinct expansion phases could suggest at least two separate periods of invasion occurring across our sampled range, one expanding in the 1920's and another in the 1940's.

This combination of genetic and herbarium data offers valuable insight into the invasion of a problematic weed across a large portion of its invaded range. Using genetic analyses, we were able to infer the likely number of distinct invasion events across a large geographic spread of invasive weed populations. Using the data gleaned from these analyses, we were then able to construct informed invasion curves that reveal trends that would otherwise have been obscured in the large pool of available data. This combination of genetic analyses as *a priori* information for the construction of herbarium-derived invasion curves proves a powerful method for extracting information on the invasion status of distinct invasion events, as well as maximizes the benefits of freely available data. In an era of increased invasions and dwindling conservation funding, the use of existing data in the most effective and informed way possible is paramount for the continued effective management of invasive species and increased understanding of invasion success.

In conclusion, this study offered insight to the population structure and invasion status of a *Gypsophila paniculata* in its introduced N. American range. Our data suggest that the distinct population clusters observed through genetic analyses are likely explained by the species' history as a horticultural species, a characteristic that facilitated its spread to the continent. When viewed in light of these genetic clusters, herbarium data further supported the presence of at least two

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

invasion events, evidenced by unique expansion phases across the species' range. Combining herbarium records with genetic analyses has provided a more complete analysis of the invasion history of this species, and this type of work would serve as a useful tool for characterizing the invasion status of other invasive populations.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure Legends

Figure 1. Sampling locations for assessing *G. paniculata* population structure used in this study; locations in Washington, North Dakota, and Minnesota are visualized in panel (a), locations in Michigan are visualized in panel (b). The Leelanau Peninsula is denoted by a black star.

Sampling location codes: Chelan, WA (CH-WA); Osborne Bay, WA (OB-WA); Knife River Historic Indian Villages, ND (KR-ND); Ottertail, MN (OT-MN); Petoskey State Park, MI (PS-MI); Sleeping Bear Dunes National Lakeshore, MI (SBD-MI); Arcadia Dunes, MI (AD-MI).

Figure 2. Results of Bayesian cluster analysis of *G. paniculata* genotyped at 14 microsatellite loci, performed using the program STRUCTURE (Pritchard et al. 2000). Each individual (n=145) is represented by a single column, with different colors indicating the likelihood of assignment to that cluster. Black lines delineate sampling location. Results suggest 2 population clusters (K=2). Locations are listed from west to east and north to south (MI). Sampling location codes: Chelan, WA (CH-WA); Osborne Bay, WA (OB-WA); Knife River Historic Indian Villages, ND (KR-ND); Ottertail, MN (OT-MN); Petoskey State Park, MI (PS-MI); Sleeping Bear Dunes National Lakeshore, MI (SBD-MI); Arcadia Dunes, MI (AD-MI).

Figure 3. Principal Coordinates Analysis (PCoA) of seven geographic populations of baby's breath (*G. paniculata*) genotyped at 14 microsatellite loci, based on a genotypic distance matrix, and performed in GenAlEx 6.502 (Peakall and Smouse, 2006,2012). Sampling location codes: Chelan, WA (CH-WA); Osborne Bay, WA (OB-WA); Knife River Historic Indian Villages, ND (KR-ND); Ottertail, MN (OT-MN); Petoskey State Park, MI (PS-MI); Sleeping Bear Dunes National Lakeshore, MI (SBD-MI); Arcadia Dunes, MI (AD-MI).

Figure 4. Discriminant analysis of principal components (DAPC) based on *G. paniculata* analyzed at 14 microsatellite loci and calculated in the '*adegenet'* package for R (Jombart et al., 2010). (a) Scatterplot showing both discriminant function axes and eigenvalues. Each point represents an individual (n=145). After cross validation, 16 of 28 PC's were retained. (b) Plot visualizing DAPC sample distribution on the primary discriminant function. (c) Individual assignment to clusters using all eigenvalues explained by the PCA.

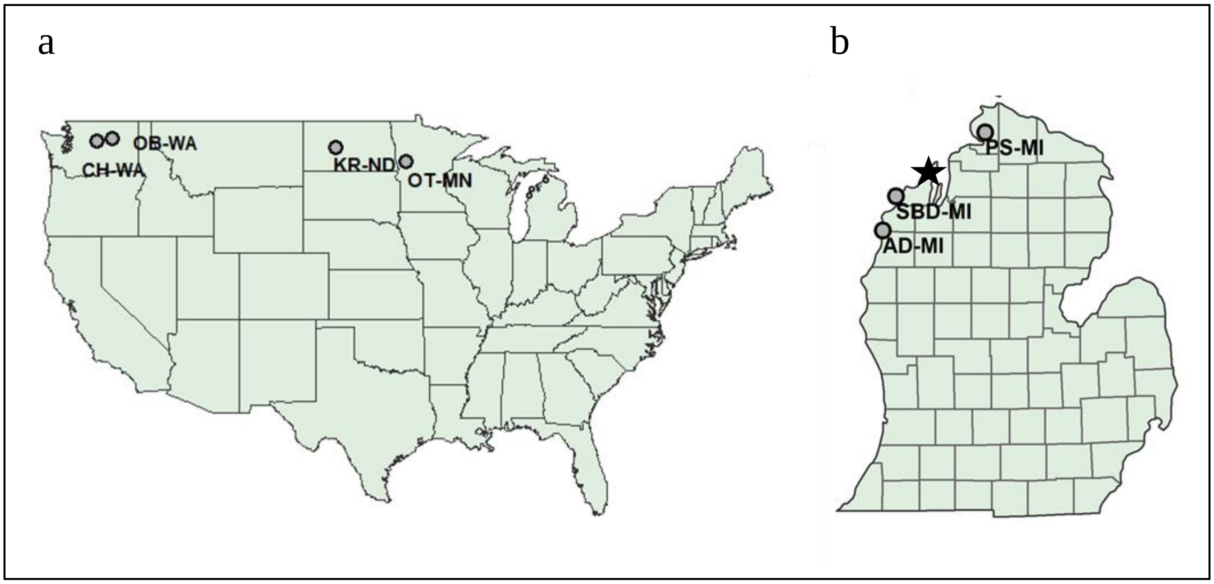
Sampling location codes: Chelan, WA (CH-WA); Osborne Bay, WA (OB-WA); Knife River Historic Indian Villages, ND (KR-ND); Ottertail, MN (OT-MN); Petoskey State Park, MI (PS-MI); Sleeping Bear Dunes National Lakeshore, MI (SBD-MI); Arcadia Dunes, MI (AD-MI).

Figure 5. Invasion curves created using herbarium data for *Gypsophila paniculata* collection in (b) North America, (c) genetic cluster 1, and (d) genetic cluster 2 (a gap in sample collection is evidence by the lack of points on the graph). An example invasion curve illustrating the three-stage invasion pathway typical of many invasions is visualized in panel (a).

Cluster assignment: (1) Washington, North Dakota, Minnesota, and northwest Michigan. (2) Michigan south of the Leelanau Peninsula.

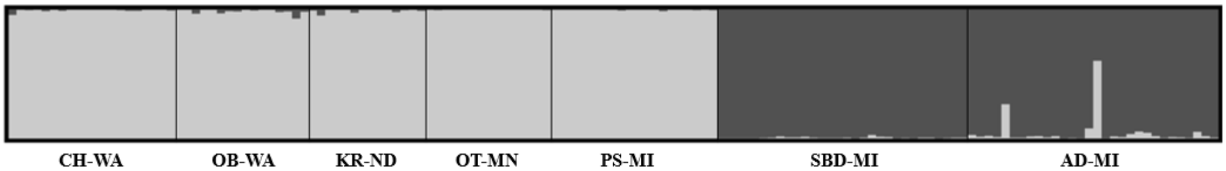
bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1



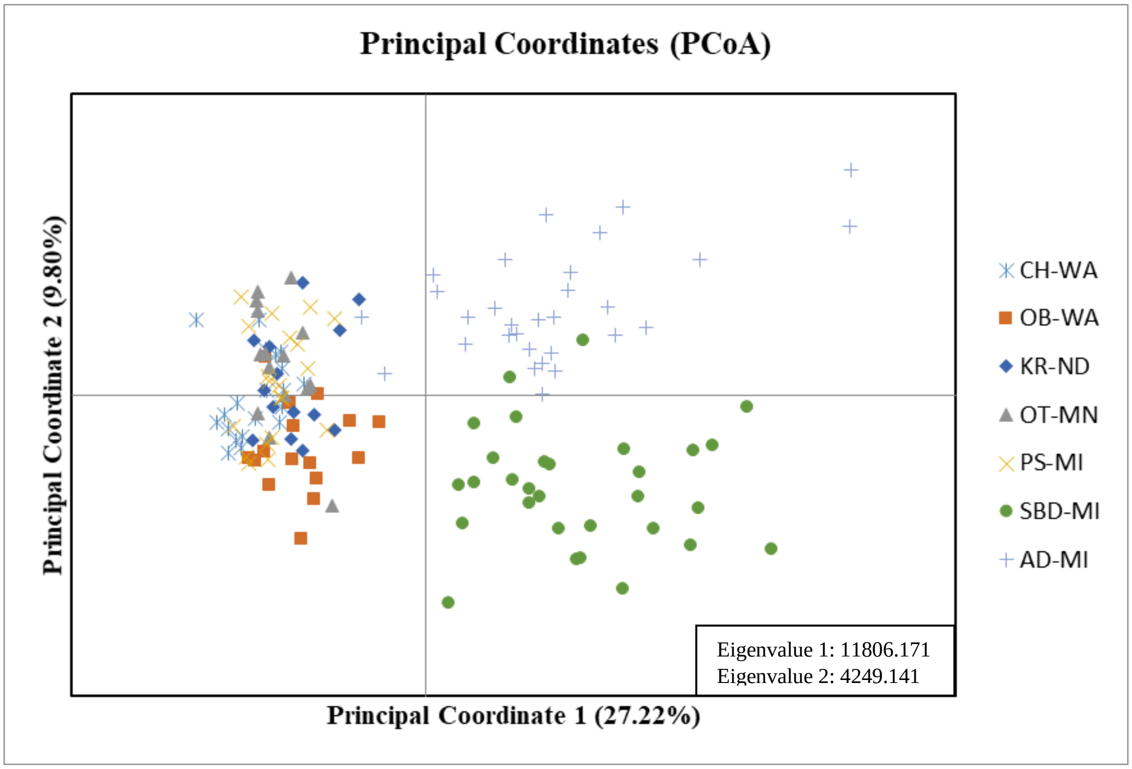
bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2



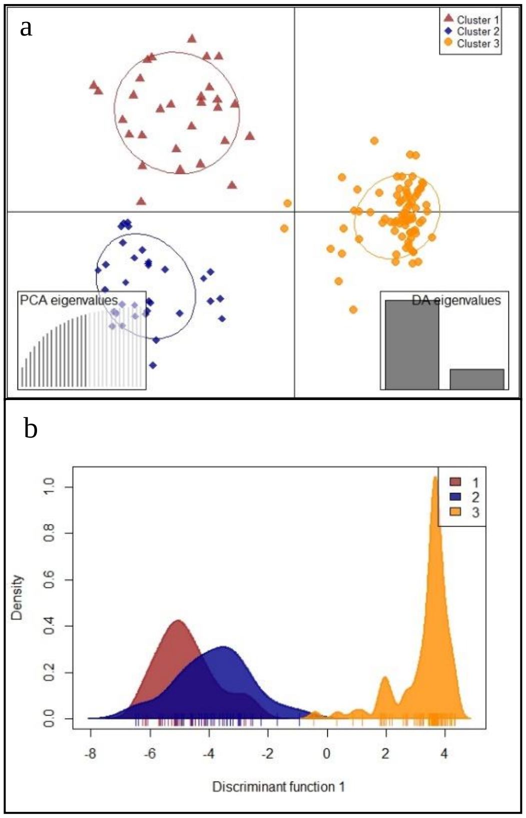
bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3



bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4



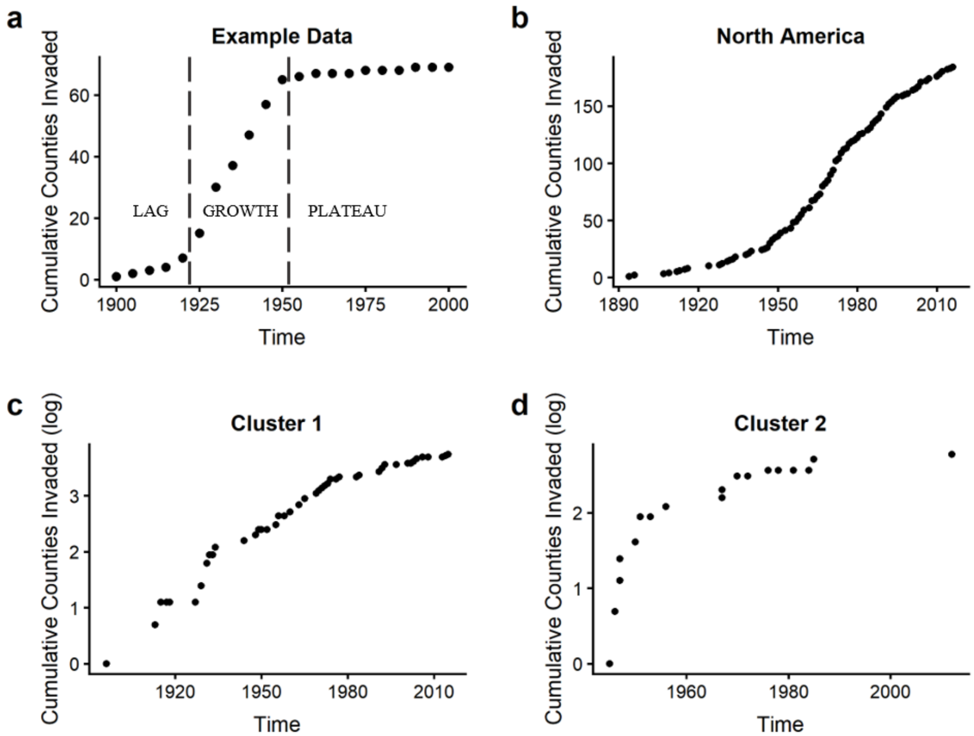
-

c

Population	Cluster 1	Cluster 2	Cluster 3
CH-WA	0	0	20
OB-WA	5	0	11
KR-ND	0	0	14
OT-MN	0	0	15
PS-MI	0	0	20
SBD-MI	28	2	0
AD-MI	1	27	2

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5



bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tables

Table 1

Table 1. Locations, dates, sample size, and geographic coordinates for analyzed samples of baby's breath (*G. paniculata*).

Sampling Location	Sampling Code	GPS Coordinates	Sampling Date	*n*
Chelan, WA	CH-WA	47.7421˚N 120.2177˚W	June 7-8, 2018	20
Osborne Bay, WA	OB-WA	47.9129˚N 119.0433˚W	June 7, 2018	16
Knife River Indian Villages National Historic Site, ND	KR-ND	47.3302˚N 101.3859˚W	June 6, 2018	14
Ottertail, MN	OT-MN	46.4627˚N 95.5733˚W	June 11, 2018	15
Petoskey State Park, MI	PS-MI	45.4037˚N 84.9121˚W	June 1, 2018	20
Dune Plateau, Sleeping Bear Dunes National Lakeshore, MI	SBD-MI	44.8731˚N 86.0585˚W	July, 2016	30
Arcadia Dunes, Sleeping Bear Dunes National Lakeshore, MI	AD-MI	44.5366˚N 86.2253˚W	July 8 and 15, 2016	30

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2

Table 2. Genetic diversity measures for seven *G. paniculata* sampling locations sequenced at 14 microsatellite (nSSR) loci.

	Sampling Locations						
	CH-WA	OB-WA	KR-ND	OT-MN	PS-MI	SBD-MI	AD-MI
Loci							
BB_21680							
N	20	15	14	15	19	30	30
N_A	4	3	1	2	2	3	3
H_O	0.400	0.267	0.000	0.467	0.474	0.500	0.700
H_E	0.599	0.646	0.000	0.370	0.491	0.549	0.555
F_{IS}	0.3377	0.5957	-	-0.2727	0.0357	0.0909	-0.2661
BB_6627							
N	20	16	14	15	20	30	30
N_A	1	1	1	1	1	2	2
H_O	0.000	0.000	0.000	0.000	0.000	0.500	0.467
H_E	0.000	0.000	0.000	0.000	0.000	0.503	0.472
F_{IS}	-	-	-	-	-	0.0068	0.0122
BB_3968							
N	20	16	14	15	20	30	30
N_A	1	1	1	1	2	4	2
H_O	0.000	0.000	0.000	0.000	0.150	0.367	0.133
H_E	0.000	0.000	0.000	0.000	0.219	0.421	0.183
F_{IS}	-	-	-	-	-0.0556	0.1320	0.2750
BB_5151							
N	20	16	14	15	20	30	28
N_A	2	2	2	1	2	2	2
H_O	0.150	0.063	0.357	0.000	0.100	0.467	0.179
H_E	0.142	0.063	0.389	0.000	0.097	0.499	0.508
F_{IS}	-0.0556	0.0000	0.0845	-	-0.0270	0.0667	0.6530
BB_4443							
N	20	16	14	15	20	30	30
N_A	1	3	4	1	4	9	5

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

H_O	0.000	0.563	0.429	0.000	0.450	0.767	0.567
H_E	0.000	0.558	0.516	0.000	0.562	0.771	0.675
F_{IS}	-	-0.0075	0.1746	-	0.2028	0.0052	0.1623
BB_31555							
N	20	16	13	15	20	30	30
N_A	1	2	2	2	1	4	3
H_O	0.000	0.500	0.462	0.400	0.000	0.600	0.467
H_E	0.000	0.484	0.443	0.460	0.000	0.624	0.554
F_{IS}	-	-0.0345	-0.0435	0.1340	-	0.0396	0.1603
BB_14751							
N	20	16	14	15	20	30	30
N_A	5	4	3	2	3	8	6
H_O	0.750	0.563	0.500	0.333	0.500	0.633	0.467
H_E	0.726	0.619	0.521	0.370	0.472	0.782	0.631
F_{IS}	-0.0345	0.0940	0.0421	0.1026	-0.0615	0.1933	0.2632
BB_3335							
N	20	16	14	13	19	30	30
N_A	2	3	1	2	3	7	6
H_O	0.000	0.000	0.000	0.538	0.368	0.667	0.600
H_E	0.097	0.492	0.000	0.508	0.534	0.831	0.721
F_{IS}	1.0000	1.0000	-	-0.0633	0.3505	0.2000	0.1707
BB_4258							
N	20	16	14	15	20	30	30
N_A	2	2	1	1	1	2	2
H_O	0.350	0.063	0.000	0.000	0.000	0.033	0.300
H_E	0.450	0.063	0.000	0.000	0.000	0.033	0.345
F_{IS}	0.2267	0.0000	-	-	-	0.0000	0.1329
BB_3913							
N	20	15	13	15	20	30	30
N_A	2	3	2	2	3	4	2
H_O	0.050	0.200	0.077	0.333	0.150	0.667	0.467
H_E	0.050	0.191	0.077	0.287	0.145	0.588	0.452
F_{IS}	0.0000	-0.0500	0.0000	-0.1667	-0.0364	-0.1373	-0.0331
BB_2888							
N	20	16	14	15	20	30	30
N_A	2	2	2	2	2	5	5

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

H_O	0.550	0.375	0.286	0.267	0.450	0.833	0.667
H_E	0.481	0.484	0.254	0.405	0.512	0.807	0.599
F_{IS}	-0.1484	0.23080	-0.1304	0.3488	0.1231	-0.0335	-0.1154
BB_5567							
N	19	16	14	15	20	30	30
N_A	3	3	2	3	3	4	5
H_O	0.842	0.563	0.500	0.600	0.550	0.667	0.767
H_E	0.681	0.599	0.495	0.549	0.612	0.614	0.728
F_{IS}	-0.2441	0.0625	-0.0111	-0.0957	0.1030	-0.0872	-0.0545
BB_7213							
N	20	16	14	15	19	30	30
N_A	1	1	2	2	2	3	3
H_O	0.000	0.000	0.286	0.467	0.105	0.500	0.667
H_E	0.000	0.000	0.254	0.370	0.102	0.575	0.644
F_{IS}	-	-	-0.1304	-0.2727	-0.0286	0.1317	-0.0366
BB_8681							
N	19	16	14	15	19	30	30
N_A	3	3	2	2	3	4	3
H_O	0.368	0.500	0.357	0.467	0.316	0.400	0.600
H_E	0.383	0.476	0.495	0.480	0.562	0.464	0.445
F_{IS}	0.0382	-0.0526	0.2857	0.0297	0.4447	0.1397	-0.3558

Notes: *N* number of individuals, N_A number of alleles per locus, H_O observed heterozygosity, H_E expected heterozygosity, F_{IS} inbreeding coefficient (Weir and Cockerham 1984). Sampling location codes: Chelan, WA (CH); Osborne Bay, WA (OB); Knife River Historic Indian Villages, ND (KR); Otter Tail, MN (OT); Petoskey State Park, MI (PS); Sleeping Bear Dunes National Lakeshore, MI (SBD); Arcadia Dunes, MI (AD).

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3

Table 3. Population pairwise F_{ST} (Weir and Cockerham, 1984) for *G. paniculata* populations using microsatellite data calculated in GenAlEx 6.502 (Peakall and Smouse, 2006, 2012) running 9,999 permutations. Darker colors indicate increasing (higher) values; all values are significant with p-values <0.05. Sampling location codes: Chelan, WA (CH-WA); Osborne Bay, WA (OB-WA); Knife River Historic Indian Villages, ND (KR-ND); Ottertail, MN (OT-MN); Petoskey State Park, MI (PS-MI); Sleeping Bear Dunes National Lakeshore, MI (SBD-MI); Arcadia Dunes, MI (AD-MI).

	CH-WA	OB-WA	KR-ND	OT-MN	PS-MI	SBD-MI	AD-MI
CH-WA	–						
OB-WA	0.077	–					
KR-ND	0.188	0.141	–				
OT-MN	0.124	0.104	0.194	–			
PS-MI	0.111	0.075	0.150	0.094	–		
SBD-MI	0.202	0.131	0.201	0.196	0.173	–	
AD-MI	0.192	0.153	0.188	0.168	0.160	0.070	–

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Literature Cited

Abdelkrim, J., Pascal, M., Calmet, C., & Samadi, S. (2005). Importance of assessing population genetic structure before eradication on invasive species: Examples from insular norway rat populations. *Conservation Biology*, 19(5), 1509–1518.

Antunes, P. M., & Schamp, B. (2017). Constructing standard invasion curves from herbarium data—toward increased predictability of plant invasions. *Invasive Plant Science and Management*, 10(04), 293–303.

Archer, F. I., Adams, P. E., & Schneiders, B. B. (2016). stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, 17(1), 5–11.

Boag, A. E., & Eckert, C. G. (2013). The effect of host abundance on the distribution and impact of biocontrol agents on purple loosestrife (*Lythrum salicaria*, Lythraceae). *Écoscience*, 20(1), 90–99.

Crooks, J. A. (2007). Lag times and exotic species: The ecology and management of biological invasions in slow-motion. *Écoscience*, 12(3), 316–329.

Darwent, A. L. (1975). The biology of Canadian weeds: 14. *Gypsophila paniculata* L. *The Biology of Canadian Weeds*, 55(14), 1049–1058.

Darwent, A. L., & Coupland, R. T. (1966). Life History of *Gypsophila paniculata*. *Weeds*, 14(4), 313.

Dupuis, J. R., Janes, J. K., Malenfant, R. M., Gorrell, J. C., Andrew, R. L., Miller, J. M., & Cullingham, C. I. (2017). The K = 2 conundrum. *Molecular Ecology*, 26(14), 3594–3602.

EDDMapS. (2019). Early Detection & Distribution Mapping System. The University of Georgia - Center for Invasive Species and Ecosystem Health. Available online athttp://www.eddmaps.org/. Retrieved February 25, 2019, from https://www.eddmaps.org/distribution/usstate.cfm?sub=5682

Ellstrand, N. C., & Elam, D. R. (2003). Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. *Annual Review of Ecology and Systematics*, 24(1), 217–242.

Emery, S. M., & Doran, P. J. (2013). Presence and management of the invasive plant *Gypsophila paniculata* (baby's breath) on sand dunes alters arthropod abundance and community structure. *Biological Conservation*, 161, 174–181.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620.

Fuentes, N., Ugarte, E., Kühn, I., & Klotz, S. (2008). Alien plants in Chile: Inferring invasion periods from herbarium records. *Biological Invasions*, 10(5), 649–657.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Haynes, G. D., Gilligan, D. M., Grewe, P., & Nicholas, F. W. (2009). Population genetics and management units of invasive common carp *Cyprinus carpio* in the Murray-Darling Basin, Australia. *Journal of Fish Biology*, 75(2), 295–320.

Jombart, T., & Collins, C. (2015). *A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 2.0.0*.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11(94).

Karamanski, T. J. (2000). *A nationalized lakeshore: the creation and administration of Sleeping Bear Dunes National Lakeshore.*

Larkin, D. J. (2012). Lengths and correlates of lag phases in upper-Midwest plant invasions. *Biological Invasions*, 14(4), 827–838.

Lavoie, C., Jodoin, Y., & De Merlis, A. G. (2007). How did common ragweed (*Ambrosia artemisiifolia* L.) spread in Quebec? A historical analysis using herbarium records. *Journal of Biogeography*, 34(10), 1751–1761.

Leimbach-Maus, H. B., Parks, S. R., & Partridge, C. G. (2018a). Genetic structure of invasive baby's breath (*Gypsophila paniculata*) populations in a freshwater Michigan dune system. *Pre-Print*, 8042(440), 83–87.

Leimbach-Maus, H. B., Parks, S. R., & Partridge, C. G. (2018b). Microsatellite primer development for the invasive perennial herb *Gypsophila paniculata* (Caryophyllaceae). *Applications in Plant Sciences*, 6(12), e1203.

Mack, R. N., Simberloff, D., Lonsdale, W. M., Evans, H., Clout, M., & Bazzaz, F. A. (2000). Biotic invasions: causes, epidemiology, global consequences, and control. In *Ecological Applications* (Vol. 10).

Meimberg, H., Milan, N. F., Karatassiou, M., Espeland, E. K., McKay, J. K., & Rice, K. J. (2010). Patterns of introduction and adaptation during the invasion of *Aegilops triuncialis* (Poaceae) into Californian serpentine soils. *Molecular Ecology*, 19(23), 5308–5319.

Michigan Department of Natural Resources. (2015). *2015 Michigan Invasive Species Grant Program Handbook.*

Moritz, C. (1994). Defining ' Evolutionarily Significant Units .' *Trends in Ecology and Evolution*, 9(10), 373–375.

Palsbøll, P. J., Bérubé, M., & Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends in Ecology and Evolution*, 22(1), 11–16.

Peakall, R., & Smouse, P. E. (2006). genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*(1), 288–295.

Peakall, R., & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28(19), 2537–2539.

Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*,

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

52(3 SPEC. ISS.), 273–288.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus data. *Genetics*, 155, 945–959.

Pyšek, P. (1995). On the terminology used in plant invasion studies. In *Plant lnvasions - General Aspects and Special Problems* (pp. 71–81). Amsterdam, The Netherlands: SPB Academic Publishing.

Pyšek, Petr, Jarošík, V., Müllerová, J., Pergl, J., & Wild, J. (2008). Comparing the rate of invasion by *Heracleum mantegazzianum* at continental, regional, and local scales. *Diversity and Distributions*, 14(2), 355–363.

R Development Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raymond, M., & Rousset, F. (1995). An exact test for population differentiation. *Evolution*, 49(6), 1280–1283.

Reid, M. L., & Emery, S. M. (2018). Scale-dependent effects of *Gypsophila paniculata* invasion and management on plant and soil nematode community diversity and heterogeneity. *Biological Conservation*, *224*, 153–161.

Reinhold, C. G. (1856). *The farmer's promotion book*. Pittsburgh, PA: W.S. Haven.

Rice, E. (2018). Assessment of invasive *Gypsophila paniculata* control methods in the northwest Michigan dunes. *Masters Theses*, *888*.

Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8(1), 103–106.

Shih, J. G., & Finkelstein, S. A. (2008). Range dynamics and invasive tendencies in *Typha latifolia* and *Typha angustifolia* in eastern North. *Wetlands*, 28(1), 1–16.

Stevens, O. A. (1957). Weights of Seeds and Numbers Per Plant. *Weeds*, 5(1), 46–55.

Swearingen, J., & Bargeron, C. (2016). Invasive plant atlas of the United States. University of Georgia Center for Invasive Species and Ecosystem Health. Retrieved March 17, 2019, from invasiveplantatlas.org website: https://scholar.google.com/scholar?cluster=3299826492555989150&hl=en&as_sdt=80000005&sciodt=0,23#d=gs_cit&u=%2Fscholar%3Fq%3Dinfo%3AnhA5NbZWyy0J%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26scfhb%3D1%26hl%3Den

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4(3), 535–538.

Weir, B. S., & Cockerham, C. C. (1983). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358.

Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A., & Losos, E. (1998). Quantifying threats to imperiled species in the United States. *BioScience*, 48(8), 607–615.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Williamson, M., & Fitter, A. (1996). The varying success of invaders. *Ecology*, 77(6), 1661–1666.

Zalewski, A., Michalska-Parda, A., Bartoszewicz, M., Kozakiewicz, M., & Brzeziński, M. (2010). Multiple introductions determine the genetic structure of an invasive species population: American mink *Neovison vison* in Poland. *Biological Conservation*, 143(6), 1355–1363.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Biosketches

- Sarah K. Lamar's research interests rest at the intersection of molecular ecology and conservation biology, using genetic techniques to inform management and practical conservation. Her particular interests are related to 1) adaptation to environmental changes and 2) the genetic bases of behavior.
- Charlyn G. Partridge's work encompasses both field and laboratory studies and employs a variety of behavioral, molecular, and genetic techniques to address questions related to: invasive species management, sexual selection, and alternative mating strategies.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supporting/Supplemental Information

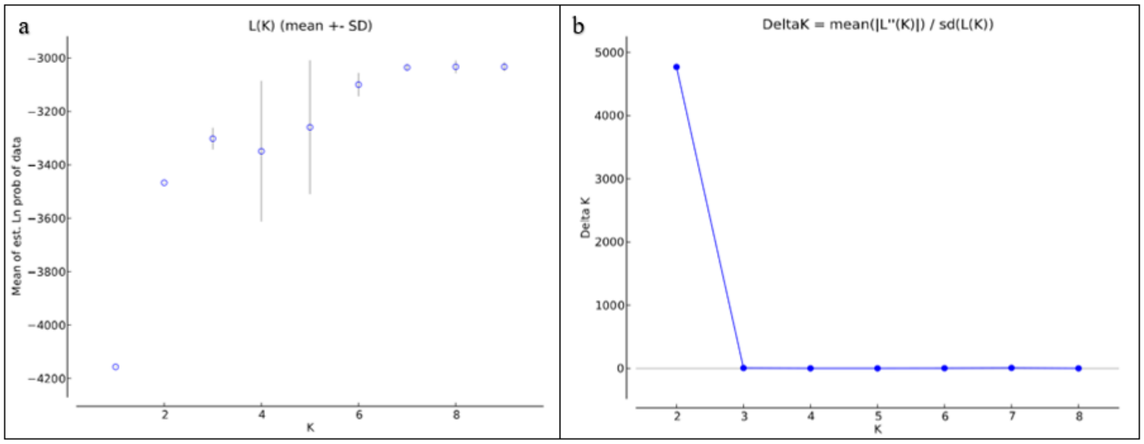


Figure S1. Bayesian cluster analysis of seven sampling locations of baby's breath (*G. paniculata*) genotyped at 14 microsatellite loci, gathered from the program STRUCTURE (Pritchard et al. 2000). (a) Mean L(K) (±SD) over 10 runs for each value of K (1-9). (b) Evanno's ΔK (Evanno et al., 2005) where the highest rate of change indicates the highest likelihood of cluster numbers. This analysis was conducted without prior sampling location information. Two genetic clusters were inferred.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

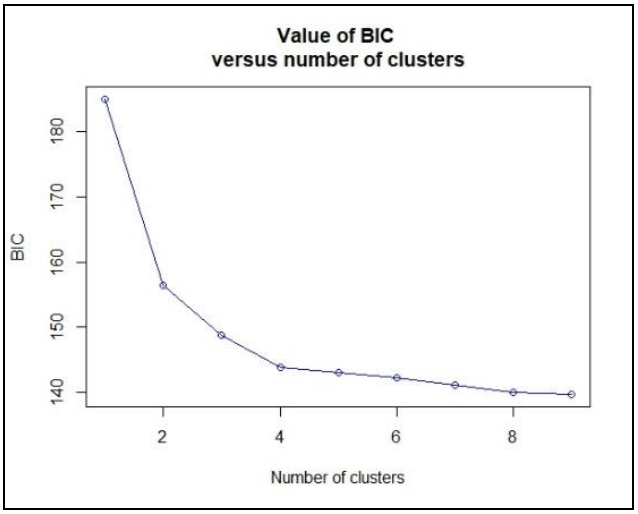


Figure S2. Bayesian Information Criterion for a DAPC of seven sampling locations of baby's breath (*G. paniculata*) genotyped at 14 microsatellite loci, created using the package '*adegenet'* in R (Jombart & Collins, 2015; Jombart et al., 2010). The inflection point suggests the supported amount of genetic clusters present; both a *K* of 2 and 3 were considered in analysis.

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S1. Details for 14 microsatellite (nSSR) loci specific to *G. paniculata* developed by Leimbach-Maus et al. (2018b) and used in this study.

Locus	Primer Sequence (5' - 3')	Repeat Motif	Allele Size Range (bp)	Annealing Temperature (˚C)	Fluorescent Label	Multiplex	GenBank Accession no.
BB_21680	F: ACTACACACAGACTCGATCCTC R: CTTTGATTGTTTGGTGTAAGTTGC	$(AAG)_5$	199-218	62	PET	PS1	MH704705
BB_6627	F: CAAACTCAACCAACCAGACACC R: CACCTCAGCAACAACAGAGTG	$(AAAC)_5$	151-155	62	FAM	PS1	MH704715
BB_3968	F: CATGGAGGACAATGAGAAGACG R: ACGGTGGTAATGAAGTTTGGTG	$(AGG)_6$	207-219	62	FAM	PS2	MH704706
BB_5151	F: TCCACCTTATAACTCACCACCC R: TGAGGAAGGATAACAGCTCTCG	$(ACC)_5$	205-210	62	PET	PS2	MH704712
BB_4443	F: TAGGGTGGGTGCTTGTACTAAC R: AAAGTGGTGCTGCAGAAGAATC	$(AAG)_{16}$	171-211	62	NED	PS2	MH704704
BB_31555	F: TGTATAACTGAGATAACCCAGACG R: TTGTTACCTTGTTCCGGCAAAG	$(AC)_7$	150-156	62	VIC	PS2	MH704716
BB_14751	F: CCTCAAACCCTAACAATGCTCC R: TCAGCCGATCCTCTAACACG	$(AAG)_{12}$	201-223	62	FAM	PS3	MH704713
BB_3335	F: TCCACCAAACTCTTAAACTGCC R: CACAGACACAAAGGATCCAACC	$(AGG)_5$	215-244	62	NED	PS3	MH704701
BB_4258	F: TCACAAGAGGCCCAATTTCTTC R: ACTTGAACCCGAACCTATACCC	$(AAT)_5$	178-195	62	VIC	PS3	MH704714
BB_3913	F: GGCTGTCGGGTAATAAACACAG R: TCCCAACTCAAGTCATAGCCTAG	$(ACAG)_5$	159-171	62	PET	PS3	MH704702
BB_2888	F: CTTCATTCATGTACAAGAGCGC R: AGAACTGGCTATGGATCGAAATG	$(AC)_{16}$	219-232	63	FAM	PS4	MH704709
BB_5567	F: GGCTAGGGAAAGTAGGAAGACC R: CGTGTCCTGTTTCTCCATGATC	$(AAT)_5$	198-222	62	VIC	PS4	MH704703
BB_7213	F: TTGCATTCCCACCATTTCATCC R: AGCCAACCTCGTATTAATTGCC	(AC_7)	161-167	62	PET	PS4	MH704708
BB_8681	F: ATCTCCAGTTTCCGTGATTTGC R: TACGTCACAAGAGCTTTCAACC	$(ACC)_8$	204-222	62	NED	PS4	MH704710

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S2. Details for *G. paniculata* herbarium records used in this study.

Institution	Catalog #	Collection Date	GPS Coordinates	Location Information Provided	State/Province
Arizona State Univ. Vascular Plant Herbarium	ASU0080637	7/31/2013	41.301038, -105.570631	Laramie Basin	WY(USA)
B. A. Bennett Herbarium, Yukon Government	BABY-0160	7/9/1991	49.5833, -119.65	Summerland	BC(CA)
B. A. Bennett Herbarium, Yukon Government	BABY-6662	7/26/2008	49.18418, -119.535292	Osoyoos	BC(CA)
Boise District Bureau of Land Management	1461	10/12/1995		9 km NE of Weiser	ID(USA)
Boise State Univ., Snake River Plains Herbarium	49505	8/1/1972		Cardston	AB(CA)
Boise State Univ., Snake River Plains Herbarium	35193	7/1/2007	43.724433, -115.604067	Loftus Hot Springs	ID(USA)
Boise State Univ., Snake River Plains Herbarium	54162	8/23/2013	47.702411, -116.802719	Coeur d'Alene	ID(USA)
Brigham Young Univ., S.L Welsh Herbarium	BRYV0140072	6/22/2012	40.23994, -109.01077	Dinosaur, Rio Blanco	CO(USA)
Brigham Young Univ., S.L Welsh Herbarium	BRYV0092109	7/31/2011	46.00617, -112.61569	Silver Bow	MT(USA)
Brigham Young Univ., S.L Welsh Herbarium	BRYV0030863	8/15/2011	40.38927, -109.79833	Uintah	UT(USA)
Canadian Museum of Nature	CAN 450828	8/22/1980	43.533333, -79.633333	Mississauga Lorne Park	ON(CA)
Carnegie Museum of Nat. History Herbarium	CM195622	6/30/1956	44.686204, -85.512464	7.5mi SE of Traverse City	MI(USA)
Carnegie Museum of Nat. History Herbarium	CM462845	7/26/1967		Little Manistee River Crossing on Route 37	MI(USA)
Carnegie Museum of Nat. History Herbarium	CM195621	7/8/1966		10mi W of Coronport	SK(CA)
Central Michigan Univ.	CMC00019957	7/27/2015		Beaver Island, Whiskey Point lighthouse, St James	MI(USA)
Clemson Univ. Herbarium	6157	7/6/1928		Anderson	SC(USA)
Colorado State Univ. Herbarium	9072	8/20/1974	37.438, -105.7597		CO(USA)
Colorado State Univ. Herbarium	48075	7/22/1982	40.6796, -107.4408	Moffat County	CO(USA)
Colorado State Univ. Herbarium	71428	8/14/1984	40.9955, -104.9148	Weld County	CO(USA)
Colorado State Univ. Herbarium	72900	7/15/1989	40.5684, -105.0267	Fort Collins	CO(USA)
Consortium of California Herbaria	UC1714554	8/1907		Cisco, Placer	CA(USA)
Consortium of California Herbaria	UC455027	9/24/1909	35.30012, -120.66232	San Luis Obispo	CA(USA)
Consortium of California Herbaria	CASBOTBC388473	7/1912		Yrkeka, Siskiyou	CA(USA)
Consortium of California Herbaria	UCD98413	7/25/1950		Dorris, Siskiyou	CA(USA)
Consortium of California Herbaria	CDA3427	7/29/1953		McDoel, Tule Lake, Siskiyou	CA(USA)
Consortium of California Herbaria	CDA3425	7/16/1963	40.32005, -120.53503	Janesville, Lassen	CA(USA)
Consortium of California Herbaria	CASBOTBC388470	6/29/1967		SW part of Weed, Siskiyou	CA(USA)
Consortium of California Herbaria	CDA3428	6/23/1971		Benton Station, Mono County	CA(USA)
Consortium of California Herbaria	CDA3429	10/6/1971		1mi N of Janesville, Lassen	CA(USA)
Consortium of California Herbaria	CDA3426	5/15/1972		Orosi, Tulare	CA(USA)
Consortium of California Herbaria	UCSB39545	9/2/1981	34.42200, -119.79500	Santa Barbara	CA(USA)
Consortium of California Herbaria	CDA34391	6/17/1987		Janesville, Lassen	CA(USA)
Consortium of California Herbaria	CDA35529	8/15/1991		Stanislaus	CA(USA)
Consortium of California Herbaria	RSA719893	7/29/2006	40.31370, -120.53863	Janesville, Lassen	CA(USA)
Consortium of California Herbaria	RSA820288	5/8/2014	33.36120, -117.32250	Camp Pendelton North	CA(USA)
Eastern Michigan Herbarium	EMC010873	7/28/1976		Lapeer	MI(USA)
Eastern Michigan Herbarium	EMC010872	8/1894		Geneva	NY(USA)
Gouvernement du Québec	QUE0139003	7/7/1960		Rimouski	QC(CA)
Harvard Univ. Herbarium	691948	6/30/1938		Danbury, CT	CT(USA)
Harvard Univ. Herbarium	691945	8/10/1916		Westmore, Maine	ME(USA)
Harvard Univ. Herbarium	691946	7/18/1967		Burlington, VT	VT(USA)
Harvard Univ. Herbarium	691947	8/5/1967		Colchester, VT	VT(USA)
Hope College	HCHM01972	7/24/1978		West end of Crystal Lake, Benzie	MI(USA)
Illinois Natural History Survey	7546	6/17/1939		Starved Rock Park, La Salle	IL(USA)
Illinois Natural History Survey	93696	7/13/1963		Kankakee	IL(USA)
Illinois Natural History Survey	158788	6/18/1977		Mason County	IL(USA)
Illinois State Museum Herbarium Collection	14598	7/9/1940		Winnebago County	IL(USA)
Illinois State Museum Herbarium Collection	53098	6/23/1957		Mason County	IL(USA)
Illinois State Museum Herbarium Collection	57134	6/21/1959		Cook County	IL(USA)
iNaturalist Observations		4/12/2016	36.032, -90.44027	Greene	AR(USA)
Intermountain Herbarium	UTC00212261	9/12/1975		West end of Craig	CO(USA)
Intermountain Herbarium	UTC00110332	7/24/1958		Logan. Cache	UT(USA)
Intermountain Herbarium	UTC00240481	10/14/2004	37.9085, -111.3768	Garfield	UT(USA)
Kathryn Kalmbach Herbarium	KHD00013339	8/4/1975	39.740063, -105.512601	Clear Creek County	CO(USA)
Kathryn Kalmbach Herbarium	KHD00013340	7/18/1981	39.547561, -105.093572	Littleton	CO(USA)

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kathryn Kalmbach Herbarium	KHD00027068	8/12/2010	40.989279, -105.009321	Larimer County	CO(USA)
Klamath National Forest Herbarium		7/26/1978		Klamath Nat'l Forest, Siskiyou	CA(USA)
Louisiana State Univ., Shirley C. Tucker Herbarium	LSU00080268	8/4/1972		Custer	SD(USA)
Minot State Univ.	889	8/13/1963	48.2618, -101.4468	Burlington	ND(USA)
Missouri Botanical Garden	1663185	7/20/1987		Grand Junction, Mesa	CO(USA)
Missouri Botanical Garden	744953	8/17/1991	39.75, -105.66666		CO(USA)
Montana State Univ.	51309	7/26/1956	48.20178, -114.314	Kalispell	MT(USA)
Montana State Univ.	57531	7/20/1959	45.712572, -111.04224	Bozeman	MT(USA)
Montana State Univ.	60068	8/1/1960	45.981500, -112.519000	Deer Lodge, Silver Bow	MT(USA)
Montana State Univ.	63044	7/14/1967	46.988237, -114.18249	Missoula	MT(USA)
Montana State Univ.	63273	7/19/1967	48.77472, -104.56194	Plentywood	MT(USA)
Montana State Univ.	64029	7/19/1968	45.65579, -111.87232	Madison County	MT(USA)
Montana State Univ.	78364	7/19/1969	45.754509, -111.05906	Bozeman	MT(USA)
Montana State Univ.	65746	7/10/1970	48.7925, -105.42028	Scobey	MT(USA)
Montana State Univ.	65961	7/19/1971	47.71667, -104.15583	Sidney	MT(USA)
Montana State Univ.	125437	7/22/1999	47.574800, -112.338200	Teton County	MT(USA)
Montana State Univ.	78365	7/16/2001		Eddy Flat, Sanders	MT(USA)
Montana State Univ.	78564	6/24/2003	46.19389, -104.36944	Baker	MT(USA)
Montana State Univ.	79545	7/14/2005	46.596034, -112.02693	Helena	MT(USA)
Montana State Univ.	82151	7/31/2008	47.774443, -112.33899	Teton County	MT(USA)
Morton Arboretum	0013059MOR	7/5/1974		Kane	IL(USA)
Morton Arboretum	0013060MOR	7/12/1992		St. Joseph	IN(USA)
Muhlenberg College	MCA0012438	8/9/1963		Lehigh, West Bethlehem	PA(USA)
Muhlenberg College	MCA0012437	6/18/1964		Lehigh, West Bethlehem	PA(USA)
Muhlenberg College	MCA0012436	6/28/1964		Lehigh, West Bethlehem	PA(USA)
Murray State Univ. Herbarium	12357R	8/12/1972	46.699720, -92.001390	South Range	WI(USA)
National Museum of CA, Flora of New Brunswick	50157	8/5/2010	46.50, -66.75	Lawrence, New Brunswick	NB(CA)
Nevada Dept. of Agriculture Herbarium	NDOA0085	9/9/1967		Washoe, 6mi S of Reno	NV(USA)
Nevada Dept. of Agriculture Herbarium	NDOA0082	6/23/1976		Washoe, Stewart Indian Colony, Carson City	NV(USA)
New York Botanical Garden	446359	8/27/1982		Mono County	CA(USA)
New York Botanical Garden	1104462	7/1/2007	43.724433, -115.604067	Loftus Hot Springs	ID(USA)
New York Botanical Garden	88097	8/13/1997	42.65, -103.98	Bowen	NE(USA)
New York Botanical Garden	446361	7/9/1978		Mottsville Cemetery, Douglas County	NV(USA)
New York Botanical Garden	446357	8/7/1986		White Pine County	NV(USA)
New York Botanical Garden	446362	7/21/1973		Pine Valley Campground, Washington County	UT(USA)
New York Botanical Garden	446360	7/17/1984		Washington County	UT(USA)
New York Botanical Garden	446358	8/13/1991	42.8732, -109.8512	Pinedale	WY(USA)
New York Botanical Garden	1192083	7/20/2001	43.3064, -110.6775	Jackson	WY(USA)
Northern KY U, John W. Theiret Herbarium	31973000024234	7/10/1976		Emmett	MI(USA)
Northern KY U, John W. Theiret Herbarium	31973000024236	8/5/1967		Chittenden	VT(USA)
OAC Herbarium	41438	8/10/1967		4mi E of Okotoks	AB(CA)
OAC Herbarium	25058	8/6/1962		1mi E of Fishe	SK(CA)
OAC Herbarium	40625	8/29/1963		Regina	SK(CA)
OAC Herbarium		9/17/1965		Eastend	SK(CA)
OAC Herbarium	58551	8/12/1986		Regina	SK(CA)
Oregon State Univ.	OSC241930	8/23/2013		Kootenai	ID(USA)
Oregon State Univ.	OSC233030	7/31/2011		Silver Bow	MT(USA)
Oregon State Univ.	OSC90946	11/5/1956	42.225, -121.7806	Klamath Falls	OR(USA)
Oregon State Univ.	OSC130826	7/28/1969	43.5864, -119.0531	Burns	OR(USA)
Oregon State Univ.	OSC130826	7/1969	43.5864, -119.0531	Burns	OR(USA)
Oregon State Univ.	OSC215439	9/2/2005	44.1461, -121.3322	Bend	OR(USA)
Oregon State Univ.	OSC241888	7/2012	43.5438, -119.084	Hines	OR(USA)
Pacific Lutheran Univ.	963	7/23/1972		E of Parkland, Pierce	WA(USA)
Pacific Northwest National Library	PNNL00903	7/20/1984		Hanford, Benton	WA(USA)
Pacific Northwest National Library	PNNL00902	6/14/1993		Benton	WA(USA)
PNW Herbarium, Western Washington Univ.	8621	1963		Sand Hills Region, 75mi S of Fargo	ND(USA)
PNW Herbarium, Western Washington Univ.	15487	7/18/1971		Winthrop, Okanogan	WA(USA)
Portland State Univ.	16759	7/19/1974	44.056012, -121.31584	Bend	OR(USA)
R. L. McGregor Herbarium	238214	7/21/1989		W. Moosejaw	SK(CA)
Robert F. Hoover Herbarium, Cal Poly State Univ.	5327	8/27/1963		Siskiyou	CA(USA)
Robert F. Hoover Herbarium, Cal Poly State Univ.	59388	7/22/1966	34.60500, -120.41700	Santa Barbara	CA(USA)

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Robert F. Hoover Herbarium, Cal Poly State Univ.	74231	7/25/1991		Albany	WY(USA)
Rocky Mountain Herbarium		8/19/2011	43.63266, -113.29578	Arco	ID(USA)
Rocky Mountain Herbarium	166182	7/18/1934	48.3818, -114.0832		MT(USA)
Rocky Mountain Herbarium	454655	8/1/1960	45.9815, -112.519	Butte	MT(USA)
Rocky Mountain Herbarium	100847	9/3/1924	44.0418, -103.1309	Green Valley	SD(USA)
Rocky Mountain Herbarium	118304	7/10/1929	44.0748, -103.2221	Rapid City	SD(USA)
Rocky Mountain Herbarium	268305	7/12/1962		Platte	WY(USA)
Rocky Mountain Herbarium	322704	7/9/1978	44.4646, -105.5809	Campbell County	WY(USA)
Rocky Mountain Herbarium	329531	7/25/1980	41.1248, -104.8767	Laramie County	WY(USA)
Rocky Mountain Herbarium	329531	7/25/1980	41.1248, -104.8767	Laramie County	WY(USA)
Rocky Mountain Herbarium	361100	7/29/1982	44.4935, -109.2042	Buffalo Bill Reservoir	WY(USA)
Rocky Mountain Herbarium	524271	7/18/1983	44.4128, -105.55889	Campbell County	WY(USA)
Rocky Mountain Herbarium	609437	7/7/1984	44.4618, -109.483	Park County	WY(USA)
Rocky Mountain Herbarium	389281	6/21/1987	43.1061, -108.6264	Wind River Reservation	WY(USA)
Rocky Mountain Herbarium	704685	7/8/1994	41.5886, -104.9877	Laramie County	WY(USA)
Rocky Mountain Herbarium	783428	6/7/1995	43.7922, -108.3447	Hot Springs County	WY(USA)
Rocky Mountain Herbarium	600758	8/1/1995	42.7611, -104.4461	Lusk	WY(USA)
Rocky Mountain Herbarium	600758	8/1/1995	42.7611, -104.4461	Lusk	WY(USA)
Rocky Mountain Herbarium	653597	8/13/1997	44.4645, -109.406	Park County	WY(USA)
Rocky Mountain Herbarium		8/15/1998	41.2105, -106.7877	Encampment	WY(USA)
Royal British Columbia Museum	V075731	7/29/1964	50.019722, -113.582778	Claresholm	AB(CA)
Royal British Columbia Museum	V020711	6/27/1947	49.616667, -115.633333	Windermere	BC(CA)
Royal British Columbia Museum	V034707	7/10/1958	50.466667, -115.983333	Windermere	BC(CA)
Royal British Columbia Museum	V051012	7/15/1964		Spences Bridge, Thompson-Okanagan	BC(CA)
Royal British Columbia Museum	V134152	7/18/1964	49.616667, -115.633333	East Kootenay	BC(CA)
Royal British Columbia Museum	V170481	8/20/1966	49.616667, -115.633333	Windermere	BC(CA)
Royal British Columbia Museum	V060882A	7/8/1972	50.233333, -119.216667	Coldstream	BC(CA)
Royal British Columbia Museum	V104347	7/23/1972		Kamloops, Princeton, Thompson-Okanagan	BC(CA)
Royal British Columbia Museum	V104242	7/20/1975	49.350000, -120.066667	Okanagan-Similkameen	BC(CA)
Royal British Columbia Museum	V109355	8/16/1975		Cathedral Provincial Park, Thompson-Okanagan	BC(CA)
Royal British Columbia Museum	V126087	6/13/1984	49.183333, -119.550000	Oliver	BC(CA)
Royal British Columbia Museum	V181785	9/11/1989	49.083333, -119.516667	Haynes Lease Ecological Reserve	BC(CA)
Royal British Columbia Museum	V180190	9/12/1989	50.750000, -121.000000	Thompson-Nicola	BC(CA)
Royal British Columbia Museum	V179656	7/13/1991	49.233333, -119.820000	Okanagan-Similkameen	BC(CA)
Royal British Columbia Museum	V201741	7/29/2007	48.458333, -123.497222	Victoria	BC(CA)
San Juan College Herbarium	49926	6/10/1989		Salmon Ruins, San Juan	NM(USA)
Snow College Herbarium	EPHR 000496	4/18/1977		Provo	UT(USA)
South Dakota State U Herbarium	7569	7/29/1993		Eddy	ND(USA)
U of Minnesota, Bell Museum	108748	6/23/1896	44.984523, -93.177092	Falcon Heights	MN(USA)
Univ. of Alaska, Anchorage	4108	7/29/2004		Anchorage Quad	AK(USA)
Univ. of Alberta Museums	127283	7/27/1967	53.55, -113.5	Edmonton	AB(CA)
Univ. of Alberta Museums	127096	7/1/2010	53.101817, -111.5652	Kinsella	AB(CA)
Univ. of British Columbia, Beaty Herbarium	V155368	6/30/1958		Macleoud, Champ Vague	AB(CA)
Univ. of British Columbia, Beaty Herbarium	V155368	6/30/1958		Champ Vague, Macleod	AB(CA)
Univ. of British Columbia, Beaty Herbarium	V95748	1933	50, -119	Shuswap Lake, Sorrento	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V7797	6/21/1938		Erickson	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V7796	6/27/1947		Fort Steele	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V72193	9/15/1950	49.616667, -115.616667	Fort Steele	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V72193	9/15/1950	49.616667, -115.61666	East Kootenay	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V111996	7/18/1964	49.616667, -115.61666	Fort Steele	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V140005	7/2/1972	50, -121	Ashcroft	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V190420	6/19/1986	50.750000, -121.983333	Lillooet	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V195448	8/13/1988	49.283333, -122.75	Coquitlam	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V218663	9/12/1989	50.75, -121	Thompson-Okanagan, Walhachin	BC(CA)
Univ. of British Columbia, Beaty Herbarium	V7795	7/8/1933	48.573336, -118.08704	Columbia River Valley, Northport	WA(USA)
Univ. of CA, Riverside Plant Herbarium	UCR-11266	8/2/1970			BC(CA)
Univ. of Colorado Museum of Natural History	226746	7/10/1924	38.8338819, -104.8213631	Colorado Springs	CO(USA)
Univ. of Colorado Museum of Natural History	226753	7/14/1949	40.2082377, -105.1638622	Longmont	CO(USA)
Univ. of Colorado Museum of Natural History	226613	9/12/1975	40.5139078, -107.5587807	Craig	CO(USA)
Univ. of Colorado Museum of Natural History	226738	7/18/1981		Jefferson	CO(USA)
Univ. of Colorado Museum of Natural History	226605	7/20/1987		Grand Junction, Mesa	CO(USA)

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Univ. of Colorado Museum of Natural History	226621	7/26/1989	40.0583166, -106.3755897	Kremmling	CO(USA)
Univ. of Colorado Museum of Natural History	964601	7/17/2009		Boulder	CO(USA)
Univ. of Colorado Museum of Natural History	1806686	6/22/2012	40.239944, -109.010778	Dinosaur	CO(USA)
Univ. of Idaho	19791	7/7/1940		4 miles S of Rathdrum, Kootenai	ID(USA)
Univ. of Idaho	90992	8/23/1986	42.051365, -111.39631	Bear Lake County	ID(USA)
Univ. of Idaho	164982	7/31/2011		Rocker, Silver Bow	MT(USA)
Univ. of Idaho	92071	8/7/1986		Egan Range, White Pine	NV(USA)
Univ. of Idaho	106119	8/13/1991		Penedale, Sublette	WY(USA)
Univ. of Lethbridge		6/27/1958		District de Medicine Hat	AB(CA)
Univ. of Lethbridge		8/4/1968		Fort Macleod	AB(CA)
Univ. of Lethbridge		7/21/1991	49, -111.95	Coutts, Warner	AB(CA)
Univ. of Manitoba Herbarium	49813	8/8/1948	49.595000, -99.683889	Wawanesa	MB(USA)
Univ. of Manitoba Herbarium	19178	08/06/1951	50.669167, -100.811111	Rossburn	MB(USA)
Univ. of Manitoba Herbarium	19180	7/24/1953	49.595000, -99.683889	Wawanesa	MB(USA)
Univ. of Manitoba Herbarium	19181	7/26/1953	49.667000, -99.960000	Brandon	MB(USA)
Univ. of Manitoba Herbarium	19179	8/3/1953	49.268000, -100.996000	Melita	MB(USA)
Univ. of Manitoba Herbarium	58123	8/6/1955	50.702600, -96.530400	Victoria Beach	MB(USA)
Univ. of Manitoba Herbarium	25162	7/18/1971	51.653, -100.4594	Near Garland	MB(USA)
Univ. of Manitoba Herbarium	27282	8/5/1972	50.702600, -96.530400	Reader Lake, The Pas	MB(USA)
Univ. of Manitoba Herbarium	42972	8/24/1979	50.816667, -100.368056	Whitewater Lake Camp Area	MB(USA)
Univ. of Manitoba Herbarium	45440	9/10/1982	53.938333, -101.341944	Reader Lake, The Pas	MB(USA)
Univ. of Manitoba Herbarium	44054	7/23/1985	49.67, -96.65	Richer	MB(USA)
Univ. of Manitoba Herbarium	25162	7/9/1986	50.020700, -100.440000	15km W of Rivers	MB(USA)
Univ. of Manitoba Herbarium	70727	6/29/1988	49.805556, -99.641667	CFB Shilo	MB(USA)
Univ. of Manitoba Herbarium	49771	7/18/1989	49.837042, -99.594542	Shilo M.R.	MB(USA)
Univ. of Manitoba Herbarium	57567	7/28/1993	49.763333, -99.676944	CFB Shilo	MB(USA)
Univ. of Manitoba Herbarium	71893	7/15/1994	50.525000, -96.583333	S of Grand Beach	MB(USA)
Univ. of Michigan Herbarium		7/16/1913		Emmet County	MI(USA)
Univ. of Michigan Herbarium		8/15/1915		Cheboygan County	MI(USA)
Univ. of Michigan Herbarium		7/13/1917		Cheboygan County	MI(USA)
Univ. of Michigan Herbarium	MICH1314642	07/23/1918		Cheboygan County	MI(USA)
Univ. of Michigan Herbarium	MICH1314641	08/12/1927		Cheboygan County	MI(USA)
Univ. of Michigan Herbarium		7/19/1945		Grand Rapids	MI(USA)
Univ. of Michigan Herbarium		7/7/1946		Washtenaw County	MI(USA)
Univ. of Michigan Herbarium		7/20/1947		Oakland County	MI(USA)
Univ. of Michigan Herbarium		8/5/1947		Leelanau County	MI(USA)
Univ. of Michigan Herbarium	MICH1314648	9/5/1949		Houghton County	MI(USA)
Univ. of Michigan Herbarium		7/13/1950		Jackson County, Leoni TWP	MI(USA)
Univ. of Michigan Herbarium	MICH1475294	07/25/1950		Emmet County	MI(USA)
Univ. of Michigan Herbarium		7/14/1951		Macomb County	MI(USA)
Univ. of Michigan Herbarium	MICH1314639	08/16/1951		Benzie County	MI(USA)
Univ. of Michigan Herbarium	MICH1314646	07/24/1952		Emmet County	MI(USA)
Univ. of Michigan Herbarium	MICH1314640	08/11/1953		Leelanau County	MI(USA)
Univ. of Michigan Herbarium	MICH1314635	06/30/1956		Grand Traverse County	MI(USA)
Univ. of Michigan Herbarium		7/13/1956		Wayne County	MI(USA)
Univ. of Michigan Herbarium		7/26/1967		Lake	MI(USA)
Univ. of Michigan Herbarium		8/4/1969		Schoolcraft County	MI(USA)
Univ. of Michigan Herbarium		6/12/1970		Newaygo County	MI(USA)
Univ. of Michigan Herbarium		6/12/1970		Wexford County	MI(USA)
Univ. of Michigan Herbarium		07/09/1971		Emmet County	MI(USA)
Univ. of Michigan Herbarium		7/10/1972		Oakland County	MI(USA)
Univ. of Michigan Herbarium	MICH1314634	07/28/1974		Schoolcraft County	MI(USA)
Univ. of Michigan Herbarium		7/24/1978		Benzie County	MI(USA)
Univ. of Michigan Herbarium		9/13/1981		Wexford County	MI(USA)
Univ. of Michigan Herbarium	MICH1314637	07/17/1983		Emmet County	MI(USA)
Univ. of Michigan Herbarium	MICH1314649	07/06/1984		Benzie County	MI(USA)
Univ. of Michigan Herbarium	MICH1314649	7/6/1984		Benzie County	MI(USA)
Univ. of Michigan Herbarium	MICH1314638	7/23/1984		Rogers TWP, Presque Isle County	MI(USA)
Univ. of Michigan Herbarium	MICH1314647	07/24/1984		Leelanau County	MI(USA)
Univ. of Michigan Herbarium		8/26/1984		Shiawassee County, Perry TWP	MI(USA)
Univ. of Michigan Herbarium	MICH1314636	6/24/1985		Lenawee CO, Raisin TWP	MI(USA)
Univ. of Michigan Herbarium	MICH1314645	07/29/1985		Benzie County	MI(USA)
Univ. of Michigan Herbarium	MICH1314644	07/14/1991		Crawford	MI(USA)
Univ. of Michigan Herbarium	MICH1314643	07/14/1991		Antrim County	MI(USA)
Univ. of Michigan Herbarium		7/30/1997		Schoolcraft County	MI(USA)

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Univ. of Michigan Herbarium		9/12/2004		Lakefield TWP, Luce	MI(USA)
Univ. of Michigan Herbarium		7/9/2008		Pellston, Emmet County	MI(USA)
Univ. of Michigan Herbarium		7/15/2012		Ludington State Park, Mason County	MI(USA)
Univ. of Minnesota, Bell Museum	353430	7/14/1934	46.862181, -94.766121	Nevis	MN(USA)
Univ. of Minnesota, Bell Museum	396865	7/30/1948	47.473563, -94.880277	Bemidji	MN(USA)
Univ. of Minnesota, Bell Museum	554063	6/2/1955	44.759815, -95.421672	Hawk Creek Township	MN(USA)
Univ. of Minnesota, Bell Museum	594503	8/4/1958	46.922181, -95.058632	Park Rapids	MN(USA)
Univ. of Minnesota, Bell Museum	568923	7/15/1960	45.428063, -93.203997	Athens Township	MN(USA)
Univ. of Minnesota, Bell Museum	584548	7/8/1963	46.922181, -95.058632	Park Rapids	MN(USA)
Univ. of Minnesota, Bell Museum	590432	9/1/1965	47.282797, -95.212519	Itasca Township	MN(USA)
Univ. of Minnesota, Bell Museum	687348	6/27/1977	47.231866, -93.522768	Grand Rapids	MN(USA)
Univ. of Minnesota, Bell Museum	690447	7/1/1977	45.695467, -94.172414	Rice	MN(USA)
Univ. of Minnesota, Bell Museum	473827	7/13/1992	46.5338898, -94.7980576	Bullard Township	MN(USA)
Univ. of Minnesota, Bell Museum	460210	7/28/1992	47.229024, -94.633282	Cass County	MN(USA)
Univ. of Minnesota, Bell Museum	838412	8/7/1993	46.32431, -92.83477	Willow River Reservoir	MN(USA)
Univ. of Minnesota, Bell Museum	479774	7/16/2001	47.058131, -95.180836	Two Inlets	MN(USA)
Univ. of Minnesota, Bell Museum	920949	8/1/2003	46.129324, -94.720884	Turtle Creek Township	MN(USA)
Univ. of Minnesota, Bell Museum	440927	7/27/2004	46.199367, -94.38887	Morrison County	MN(USA)
Univ. of Minnesota, Bell Museum	907085	6/29/2006	46.933898, -95.351552	Carsonville Township	MN(USA)
Univ. of Minnesota, Bell Museum	924746	7/23/2008	47.1975, -94.9922222	Lake George	MN(USA)
Univ. of Mississippi, Thomas M. Pullen Herbarium	MISS0022741	7/13/1968		Flathead (2.5mi N of Bigfork)	MT(USA)
Univ. of Mississippi, Thomas M. Pullen Herbarium	MISS0022740	7/24/1971		McHenry (9.5mi N of Butte)	ND(USA)
Univ. of Montana	46057	7/22/1948	47.887446, -114.117614	Flathead Lake	MT(USA)
Univ. of Montana	52322	7/15/1956	46.872702, -113.986498	Missoula	MT(USA)
Univ. of Montana	66308	7/21/1968	48.063287, -114.072613	Bigfork	MT(USA)
Univ. of Montana	66309	7/26/1968	46.592712, -112.036109	Helena	MT(USA)
Univ. of Montana	136889	8/12/1970	46.828900, -111.820900		MT(USA)
Univ. of Montana	136890	8/12/1970	46.828900, -111.820900		MT(USA)
Univ. of Montana	75494	7/15/1973	46.233333, -114.18333	Ravalli County	MT(USA)
Univ. of Nevada Herbarium	20511	7/4/1970		Washoe, Stead	NV(USA)
Univ. of Nevada Herbarium	13998	7/9/1978		Douglas	NV(USA)
Univ. of North Carolina Chapel Hill Herbarium	NCU00100830	7/26/1989		Grand	CO(USA)
Univ. of North Carolina Chapel Hill Herbarium	NCU00100825	6/2/1955		Rennville	MN(USA)
Univ. of North Carolina Chapel Hill Herbarium	NCU00100834	7/22/1948		Lake	MT(USA)
Univ. of North Carolina Chapel Hill Herbarium	NCU00100829	7/18/1970		Morton	ND(USA)
Univ. of North Carolina Chapel Hill Herbarium	NCU00100815	6/19/1947		Lawrence	SD(USA)
Univ. of North Carolina Chapel Hill Herbarium	NCU00100820	8/4/1972		Custer	SD(USA)
Univ. of Puget Sound	8729	7/17/1973		Maryhill, Klickitat	WA(USA)
Univ. of Washington Herbarium	229695	8/1/1965		Center of Missoula	MT(USA)
Univ. of Washington Herbarium	186632	11/2/1956	42.225000, -121.780600	Klamath Falls, Klamath	OR(USA)
Univ. of Washington Herbarium	18742JWT	6/25/1931	47.83556, -120.04917	Chelan	WA(USA)
Univ. of Washington Herbarium	18741	8/16/1931	48.626036, -119.46626	Okanogan	WA(USA)
Univ. of Washington Herbarium	107930	7/2/1932	48.91611, -117.78056	Northport	WA(USA)
Univ. of Washington Herbarium	107960	6/17/1944	46.73139, -117.17861	Pullman	WA(USA)
Univ. of Washington Herbarium	173712	7/2/1952	47.7675, -117.35389	Mead	WA(USA)
Univ. of Washington Herbarium	242422	9/1/1969	48.85056, -117.38972	Metaline	WA(USA)
Univ. of Washington Herbarium	368375	7/31/2006	48.680278, -120.882500	Whatcom County	WA(USA)
Univ. of Washington Herbarium	413958	5/16/2013		Chelan	WA(USA)
Univ. of Washington Herbarium		6/15/2014	47.913005, -119.045792	Grant County	WA(USA)
Univ. of Washington Herbarium	399061	6/23/2014	47.816880, -119.975560	Chelan	WA(USA)
Univ. of Washington Herbarium	365203	9/19/2002	48.105000, -119.780000	Okanogan	WA(USA)
Univ. of WI-Madison, WI State Herbarium	v0025191WIS	7/10/1959		Adams	WI(USA)
Univ. of WI-Madison, WI State Herbarium	v0025200WIS	7/28/1960		Marinette	WI(USA)
Univ. of WI-Madison, WI State Herbarium	v0025204WIS	6/24/1964		Waupaca (2mi SSW of Rural)	WI(USA)
Univ. of WI-Madison, WI State Herbarium	v0025197WIS	9/12/1972		Wisconsin Point, Douglas	WI(USA)
Univ. of WI-Madison, WI State Herbarium	v0025192WIS	7/25/1975		Ashland	WI(USA)
Univ. of WI-Madison, WI State Herbarium	v0025203WIS	7/24/1981		Oconto	WI(USA)
Washington State Univ., Marion Ownbey Herbarium	49713	1/2/1929	47.658890, -117.425000	Spokane	WA(USA)
Washington State Univ., Marion Ownbey Herbarium	76350	7/26/1931	48.098330, -119.733060	Okanogan	WA(USA)
Washington State Univ., Marion Ownbey Herbarium	241720	6/2/1956	42.230560, -121.798330	Klamath Falls, Klamath	WA(USA)
Washington State Univ., Marion Ownbey	334226	7/10/1973	48.541878, -120.378890	Okanogan	WA(USA)

bioRxiv preprint doi: https://doi.org/10.1101/686691; this version posted June 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Herbarium					
Washington State Univ., Marion Ownbey Herbarium	333550	9/15/1974	46.323890, -117.971390	Dayton	WA(USA)
Washington State Univ., Tri-Cities	WS-TC-00115	1965		Yakima River, Benton	WA(USA)
Western IL Univ, R.M. Myers Herbarium	MWI00015585	6/18/1977		Bath, Mason	IL(USA)