Received: 23 April 2024 | Accepted: 5 September 2024

DOI: 10.1111/jpy.13512

RESEARCH ARTICLE

Population genetics of the freshwater red alga *Batrachospermum gelatinosum* (Rhodophyta) II: Phylogeographic analyses reveal spatial genetic structure among and within five major drainage basins in eastern North America

Roseanna M. Crowell[1] | **Sarah J. Shainker-Connelly**[2] | **Stacy A. Krueger-Hadfield**[2,3,4] | **Morgan L. Vis**[1]

[1]Department of Environmental and Plant Biology, Ohio University, Athens, Ohio, USA

[2]Department of Biology, University of Alabama at Birmingham, Birmingham, Alabama, USA

[3]Virginia Institute of Marine Science Eastern Shore Laboratory, Wachapreague, Virginia, USA

[4]William & Mary's Batten School of Coastal and Marine Science, Gloucester Point, Virginia, USA

Correspondence
Morgan L. Vis, Department of Environmental and Plant Biology, Ohio University, Athens, OH 45701, USA.
Email: vis-chia@ohio.edu

Funding information
Phycological Society of America Grant-in-Aid of Research; Ohio University Student Enhancement Award; Ohio Center for Ecology and Evolutionary Studies Fellowship; Ohio Center for Ecology and Evolutionary Studies Fellowship; Ohio University Graduate Student Senate Original Work Grant; Ohio University research incentive funds; UAB Department of Biology's Harold Martin Outstanding Student Development Award; UAB College of Arts and Sciences start-up funds; National Science Foundation CAREER Award (DEB-2141971); NSF GRFP (2020295779); UAB Blazer Fellowship

Editor: C. Pfister

Abstract

The freshwater red alga *Batrachospermum gelatinosum* has a well-documented distribution spanning historically glaciated and unglaciated eastern North America. This alga has no known desiccation-resistant propagule; thus, long-distance dispersal events are likely rare. We predicted strong genetic structure among drainage basins and admixture among sites within basins. We predicted greater genetic diversity at lower latitude sites because they likely serve as refugia and the origin of northward, post-Pleistocene range expansion. We used 10 microsatellite loci to investigate genetic diversity from 311 gametophytes from 18 sites in five major drainage basins: South Atlantic Gulf, Mid-Atlantic, Ohio River, Great Lakes, and Northeast. Our data showed strong genetic partitioning among drainage basins and among sites within basins, yet no isolation by distance was detected. Genetic diversity varied widely among sites and was not strictly related to latitude as predicted. The results from *B. gelatinosum* provide strong support that each stream site contributes to the unique genetic variation within the species, potentially due to limited dispersal and the prevailing reproductive mode of intragametophytic selfing. Simulations of migration suggested post-Pleistocene dispersal from the Mid-Atlantic. *Batrachospermum gelatinosum* potentially persisted in refugia that were just south of the ice margins rather than in the southernmost part of its range. Research of other taxa with similar ranges could determine whether these results are generally applicable for freshwater red algae. Nevertheless, these results from *B. gelatinosum* add to the growing literature focused on the patterns and genetic consequences of post-Pleistocene range expansion by eastern North American biota.

KEYWORDS
algae, approximate Bayesian computation, dispersal, glaciation, microsatellites, Pleistocene, refugium, river, stream

Abbreviations: DAPC, discriminant analysis of principal components; dNTP, deoxynucleotide triphosphate; eMLG, expected number of multilocus genotypes; GPS, Global Positioning System; HSD, honestly significant difference; PC, principal component; PCA, principal component analysis; PCR, polymerase chain reaction; *uh*, unbiased diversity.

INTRODUCTION

Population genetic methods, from DNA sequencing to microsatellite genotyping, have provided us with an opportunity to understand the processes in which species evolve across their ranges (de Meeûs et al., 2007; Tibayrenc & Ayala, 2002). They have also highlighted how dispersal capabilities can influence phylogeographic patterns (Lehrian et al., 2010; Suetsugu et al., 2023). Yet, recent theoretical work has shown that predictions made in diploid animals or diploid-dominant angiosperms cannot be applied to macroalgae with complex life cycles in which there are morphologically distinct phases, necessitating the development of appropriate predictions (e.g., Krueger-Hadfield & Hoban, 2016; Stoeckel et al., 2021). Nevertheless, for marine macroalgae, population genetic approaches have been used to trace the route of invasive species (e.g., Krueger-Hadfield et al., 2017; Voisin et al., 2005), describe their reproductive system (e.g., Engel et al., 2004; Guillemin et al., 2008; see also review in Krueger-Hadfield et al., 2021), characterize the relative proportions of sporophytic and gametophytic phases (e.g., Krueger-Hadfield et al., 2013; van der Strate et al., 2002), re-evaluate species relationships and radiations (e.g., Cánovas et al., 2011), and resolve phylogeographic patterns of recolonization (e.g., Díaz-Tapia et al., 2018; Provan et al., 2005). Yet, for freshwater red algae, these methods have been under-utilized, leaving us with a poor understanding of genetic variation across their ranges. Freshwater red algae may nevertheless provide unique insights and further our understanding of the influence of life cycle on genetic patterns and species distributions (Krueger-Hadfield et al., 2024), highlighting the importance of applying appropriate tools broadly across eukaryotes.

In marine red algae, the typically macroscopic sporophytes (tetrasporophyte) and gametophytes are physically separated and distinct "individuals." By contrast, for freshwater red algae in the Batrachospermales, a microscopic filamentous diploid phase, called the "chantransia," produces a physically attached macroscopic haploid gametophyte via vegetative meiosis (Sheath, 1984). This strategy is hypothesized to be an adaptation that allows proliferation in the unidirectional flow of streams, keeping a population anchored in a suitable habitat (Hambrook & Sheath, 1991; Sheath, 1984). The carposporophyte produces carpospores that are released and germinate to produce chantransia. Haploid–diploid life cycles have consequences for the way in which we model genetic variation using population genetic tools (Stoeckel et al., 2021), some of which are unique to freshwater reds (Krueger-Hadfield et al., 2024), and the partitioning of genetic variation within stream reaches (Shainker-Connelly et al., 2024).

Freshwater red algae are generally assumed to be poor dispersers, and long-distance dispersal events are presumed to be rare due to the lack of desiccation-resistant propagules (Hambrook & Sheath, 1991; Sheath & Hymes, 1980). Birds may be a long-distance dispersal agent carrying spores or fragments from one stream to another (Kristiansen, 1996), but there is no evidence that freshwater red algal spores or fragments of gametophytes with attached carposporophytes can survive long periods out of water. Hambrook and Sheath (1991) demonstrated experimentally that carposporophytes, once detached from gametophytes, could travel 5–35 m downstream. Then, presumably, the carpospores would be released, germinate, and colonize new habitat. However, after dispersal, many factors may constrain the establishment of a downstream population, including stream depth and turbidity, which would limit light for this photosynthetic organism attached to the benthos. Limited dispersal capabilities and the inability to survive in deeper parts of streams may lead to genetically distinct populations within and among stream reaches.

Batrachospermum gelatinosum is well studied and the most widespread freshwater red alga in North America (Sheath & Cole, 1992). In eastern North America, the geographic distribution of *B. gelatinosum* ranges from Louisiana, USA, in the south to Newfoundland, Canada, in the north—spanning the coastal plain, deciduous forest, hemlock-hardwood forest, and boreal forest biomes (Vis et al., 1996). The streams inhabited by *B. gelatinosum* have equally wide ranges of physical and chemical parameters in which they exist, including pH 4.8–8.5, specific conductance 10–490 μS. cm^{-1}, water temperature 0–27°C, and current velocity 0–126 $cm \cdot s^{-1}$ (Vis et al., 1996). This alga is morphologically variable with some trait variation linked to latitude, although sampled locations have contained few genetic differences among them as assessed by rDNA internal spacer regions I and II (Vis et al., 1996; Vis & Sheath, 1997). A subsequent study using sequence data from the mitochondrial *cox*1 gene (664 bp) and 193 gametophytes spanning 16 geographically widespread sites discovered only five haplotypes, with most gametophytes (90%) and sites (81%) belonging to one of those haplotypes (House et al., 2010). Only the three southernmost locations in Alabama, all geographically close (<155 km), had unique haplotypes and enough within-site genetic variation to support a southern glacial refugium. Part of the geographic area for *B. gelatinosum* comprises five major drainage basins: the Great Lakes, the Ohio River, the Northeast, the South Atlantic Gulf, and the Mid-Atlantic. The Great Lakes, the Northeast, and much of the Ohio River basins were historically glaciated and were recolonized when glaciers retreated after the Pleistocene. Dispersal of *B. gelatinosum* from

southern drainage basins likely occurred after glacial retreat, but it is unknown how many long-distance dispersal events occurred.

One limitation in previous work on *Batrachospermum gelatinosum* population genetic structure is the lack of polymorphic markers, such as microsatellites. Recently, Crowell et al. (2024) described the development of 10 microsatellite loci that will enable the genotyping of *B. gelatinosum*. Moreover, we have observed that intragametophytic selfing (i.e., self-fertilization in monoicous gametophytes; Klekowski, 1969) dominates many sites across the latitudinal range of this alga, suggesting strong patterns of genetic differentiation across spatial scales (Shainker-Connelly et al., 2024). Here, we tested a suite of predictions about the population biology of *B. gelatinosum* in eastern North America: (1) Strong genetic partitioning among five major drainage basins will be observed due to limited dispersal, intragametophytic selfing, and disconnected drainage basins; (2) within drainage basins, there will be admixture among sites with gene flow between geographically proximate sites; (3) the higher latitude drainage basins were likely colonized postglaciation and will have low within-basin diversity based on founder events and have had little subsequent migration from unglaciated refugia at lower latitudes with higher genetic diversity; and (4) the lower latitude sites in the South Atlantic Gulf were the origin for dispersal to the other drainage basins after the Pleistocene glaciation, as suggested by previous single-gene sequencing (House et al., 2010). Studying phylogeographic patterns in freshwater reds will provide insights into passively dispersed organisms and add to the growing number of studies examining post-Pleistocene species migrations in eastern North America (Lyman & Edwards, 2022; Soltis et al., 2006).

MATERIALS AND METHODS

The current paper and Shainker-Connelly et al. (2024) used the same microsatellite loci and stream sites such that the sections on sample collection, DNA extraction, and microsatellite amplification are similar between the two papers, but data analyses differed due to the distinct foci of each study.

Sample collection

Batrachospermum gelatinosum gametophyte thalli (here after referred to as gametophytes) were collected at 18 locations in the eastern United States spanning five major drainage basins: South Atlantic Gulf (2 sites), Mid-Atlantic (1 site), Ohio River (5 sites), Northeast (4 sites), and Great Lakes (6 sites; Figure 1, Table 1). Since the gametophyte is macroscopic and readily sampled, it was the focus of the study instead of the microscopic chantransia, which we typically cannot see in the field. Sampling was undertaken during the boreal spring and summer in 2021 and 2022. For each site, we measured physiochemical parameters near the middle of the sampled stream reach and recorded GPS coordinates from photos taken at each stream. We measured pH and water temperature with an Oakton pH Testr 5 (Oakton), and specific conductivity was measured with an Oakton EC Testr low. At sites 1, 2, and 18, we used an Oakton PCTSTestr 50 Pocket Tester to measure pH, water temperature, and specific conductivity. We used a General Oceanics Mechanical Flow Meter (General Oceanics) to measure the current velocity, but at sites 1, 2, and 18, we used a Model FP111 flow probe (Global Water Instruments). We measured stream width, depth, and length of the area in which gametophytes were sampled. We estimated by eye the stream bed composition and water color as well as clarity. Percent canopy cover was estimated by eye, but at sites 1, 2, and 18, it was measured with a spherical densiometer (Forest Densiometer, Model A) following Lemmon (1956) and Lemmon (1957), but with one measurement instead of four. All physical and chemical characteristics recorded are in Appendix S1 in the Supporting Information: Table S1. At each site, we haphazardly collected 20–30 gametophytes, when possible, placed gametophytes into individual containers with stream water, and put the containers on ice. In the laboratory, gametophytes were cleaned of debris using forceps. The gametophytes were divided into small portions for mounting on herbarium paper as a voucher, and the larger portions were placed in silica gel for subsequent DNA extraction. Voucher specimens were deposited in the Bartley Herbarium Ohio University (Table 1).

DNA extraction and microsatellite amplification

Total genomic DNA was extracted from gametophytes using the Macherey-Nagel Nucleospin Plant II kit (Macherey-Nagel). The manufacturer's protocol was followed, except for the lysis step, in which the lysate was incubated at room temperature for 1 hour, and samples were eluted with 100 μL of molecular grade water (see also Krueger-Hadfield et al., 2013).

We used 10 of the microsatellite loci described in Crowell et al. (2024; Appendix S1: Table S2). Most loci were amplified in multiplex polymerase chain reactions (PCRs) with a final volume of 15 μL, which consisted of 2 μL of DNA, forward and reverse primers (Table S2), 1X Promega GoTaq® Flexi Buffer (Promega), 2 mM of $MgCl_2$, 250 μM of each dNTP

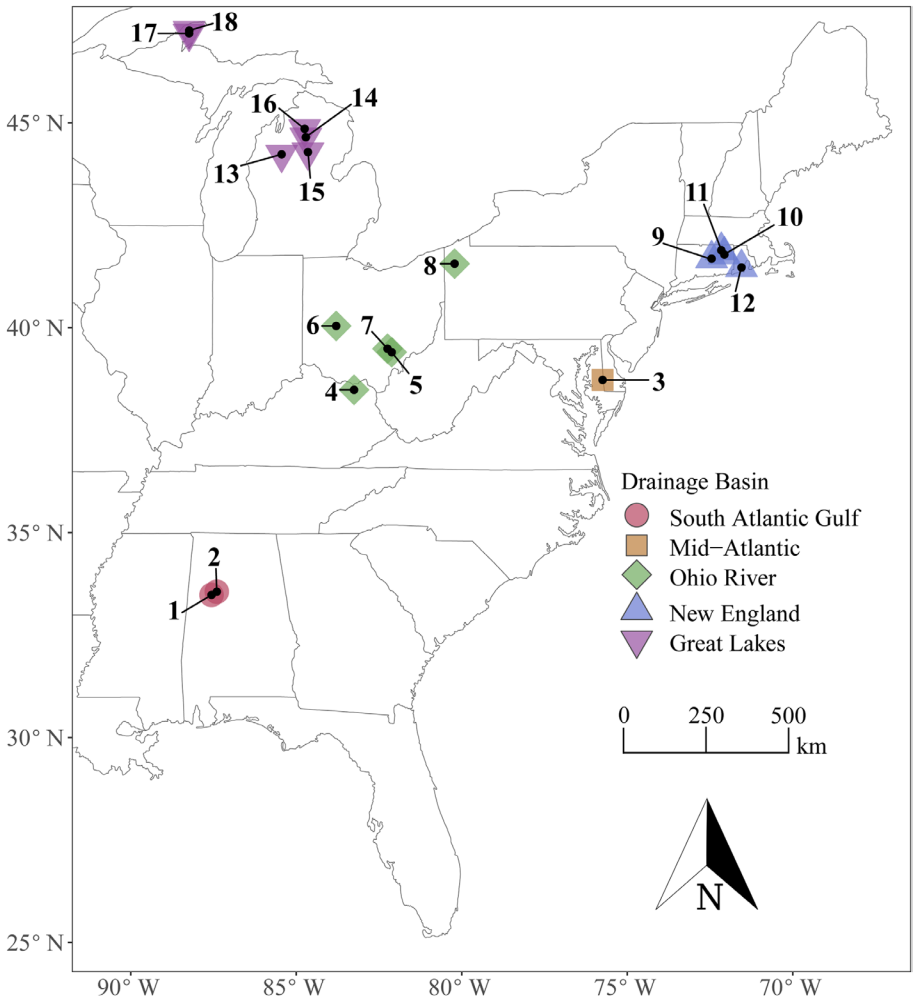

FIGURE 1 Map of *Batrachospermum gelatinosum* collection sites. Eighteen sites were sampled from five drainage basins in eastern North America. Symbols denote drainage basins. Site details provided in Table 1 and Table S1. [Color figure can be viewed at wileyonlinelibrary.com]

(Promega), 1 mg · mL^{-1} of BSA, and 1.0 U of Promega GoTaq® Flexi DNA Polymerase (Promega). For locus Bgel_056, as well as reruns for any individual loci that did not amplify after the first attempt in multiplex, we employed simplex PCRs. The final volume was 15 μL, with the same amounts of each reagent as listed previously. For multiplex and simplex PCRs, the following program was used: 95°C for 2 min, followed by 35 cycles of 95°C for 30 s, 59°C for 30 s, and 72°C for 30 s, and a final extension of 72°C for 5 min. We diluted 1.5 μL PCR product in 9.7 μL HiDi formamide (Applied Biosystems) and 0.30 μL GS 500 LIZ (Applied Biosystems) and used an Applied Biosystems 3730xl at the Heflin Center for Genomic Sciences at the University of Alabama at Birmingham to perform fragment analysis. We scored alleles using the microsatellite plugin in Geneious Prime v.2022.2.2 (Biomatters, Ltd.). We then employed TANDEM (Matschiner & Salzburger, 2009) to bin alleles while reducing rounding errors and manually checked each bin (Appendix S1: Table S3).

Data analyses

After several PCR attempts, any gametophytes that had three or more loci with no amplification were excluded from analyses. (See Table 1 for final number of gametophytes per stream segment.) We estimated null allele frequency by calculating the percent of gametophytes that did not amplify at each locus after several PCR attempts, discounting technical errors (Krueger-Hadfield et al., 2011; Appendix S1: Table S4).

Since the five drainage basins represent a large geographic area, we tested for significant differences in the physical and chemical variables among drainage basins using one-way analysis of variance (ANOVA) for data with a normal distribution and Kruskal–Wallis test for data with nonnormal distribution using the functions aov and kruskal.test, respectively, in R version 4.3.1 (R Core Team 2023; Appendix S1: Table S5). We employed Tukey's honestly significant difference (HSD) using the function TukeyHSD in R version 4.3.1 (R Core Team 2023; Table S5) to further

TABLE 1 Site information for *Batrachospermum gelatinosum* collections.

Drainage basin	Site number	Site abbreviation	Number of gametophytes	Location	Latitude, longitude	Date	Herbarium voucher
South Atlantic Gulf	1	AL-CRC	28	Alabama, Cripple Creek	33.492530, −87.562633	02 May 2022	BHO A-2037
	2	AL-YEC	22	Alabama, Yellow Creek	33.572000, −87.403000	02 May 2022	BHO A-2038
Mid-Atlantic	3	MD-HOU	25	Maryland, Houston Branch	38.737220, −75.747900	19 June 2022	BHO A-1849
Ohio River	4	KY-CDK	19	Kentucky, Kinniconick Creek	38.497000, −83.256723	11 May 2021	BHO A-1716
	5	OH-BBC	6	Ohio, Big Bailey Creek	39.415676, −82.119042	10 May 2021	BHO A-1715
	6	OH-BOG	30	Ohio, Cedar Bog	40.055862, −83.796495	12 May 2021	BHO A-1721
	7	OH-MCC	29	Ohio, Monday Creek	39.500500, −82.246300	5 May 2022	BHO A-1814
	8	PA-COT	25	Pennsylvania, Conneaut Outlet	41.574806, −80.218771	30 April 2022	BHO A-1811
Northeast	9	CT-BLR	4	Connecticut, Black Ledge River	41.697781, −72.454272	12 April 2022	BHO A-1796
	10	CT-FLB	11	Connecticut, Fuller Brook	41.798208, −72.069058	12 April 2022	BHO A-1805
	11	CT-IRM	3	Connecticut, Chism Brook at Iron Mine Lane	41.901175, −72.161356	12 April 2022	BHO A-1799
	12	RI-CPR	11	Rhode Island, Chipuxet River	41.482500, −71.551111	11 April 2022	BHO A-1803
Great Lakes	13	MI-CLI	16	Michigan, Cadillac Lake Inflow	44.243620, −85.444440	14 May 2022	BHO A-1832
	14	MI-CUT	9	Michigan, Cut River	44.659720, −84.713090	12 May 2022	BHO A-1823
	15	MI-HLK	29	Michigan, Knappens Creek at Houghton Lake	44.298437, −84.649343	12 May 2022	BHO A-1826
	16	MI-KCL	15	Michigan, Kolkee Creek at Lake Lynn	44.866490, −84.748380	14 May 2022	BHO A-1833
	17	MI-TRB	17	Michigan, Traverse River at Big Traverse Road	47.195299, −88.239869	26 July 2022	BHO A-1923
	18	MI-TRM	12	Michigan, Traverse River at Mohawk Gay Road	47.262585, −88.237123	11 May 22	BHO A-1887

Note: Site numbers as in Figure 1.

Abbreviation: BHO, Bartley Herbarium, Ohio University.

TABLE 2 Summary statistics calculated based on nine microsatellite loci.

Drainage basin	Site number	Site abbreviation	Number of gametophytes	*uh*	Mean allelic richness	eMLGs	Site private alleles
South Atlantic Gulf	1	AL-CRC	28	0.096	1.285	4.49	4
	2	AL-YEC	22	0.010	1.030	1.45	1
Mid-Atlantic	3	MD-HOU	25	0.075	1.157	2.65	
Ohio River	4	KY-CDK	19	0.341	1.949	5.74	3
	5	OH-BBC	6	0.000	1.000	1	0
	6	OH-BOG	30	0.014	1.041	1.56	1
	7	OH-MCC	29	0.015	1.046	1.69	2
	8	PA-COT	25	0.053	1.126	2.39	1
Northeast	9	CT-BLR	4	–	–	–	3
	10	CT-FLB	11	0.162	1.367	5.73	4
	11	CT-IRM	3	–	–	–	1
	12	RI-CPR	11	0.275	1.582	6.64	4
Great Lakes	13	MI-CLI	16	0.026	1.069	1.87	3
	14	MI-CUT	9	0.354	2.029	8	0
	15	MI-HLK	29	0.422	1.947	7.98	1
	16	MI-KCL	15	0.029	1.089	2.33	2
	17	MI-TRB	17	0.357	1.958	6.29	5
	18	MI-TRM	12	0.059	1.157	3.67	1

Note: Dash indicates the statistic was not calculated for sites with less than six gametophytes.
Abbreviations: eMLGs, expected number of multilocus genotypes; number of private alleles, grouped by drainage basin; *uh*, unbiased diversity.

explore which drainage basins were significantly different. We visualized the combined effects of physical and chemical variables among drainage basins using a principal components analysis (PCA) with the function prcomp in the stats package in R version 4.3.1 (Becker et al., 1988; Mardia et al., 1979; Venables & Ripley, 2002; Appendix S1: Figure S1).

We computed several summary statistics to better understand population-level genetic differences among and within drainage basins. We removed sites for which we had sampled fewer than six gametophytes (Tables 1 and 2). The average unbiased diversity (*uh*) was calculated using GenAlEx version 6.51b2 (Smouse et al., 2017). Mean allelic richness was calculated using rarefaction in the R package PopGenReport, version 3.0.7 employing the function allel.rich (Adamack & Gruber, 2014; El Mousadik & Petit, 1996; Gruber & Adamack, 2015). The number of expected multilocus genotypes (eMLGs) and the number of private alleles were estimated in the R package poppr version 2.9.4 (Kamvar et al., 2014, 2015).

We used discriminant analysis of principal components (DAPC; Jombart et al., 2010) implemented in the R package adegenet version 2.1.10 (Jombart, 2008; Jombart & Ahmed, 2011) to assess relationships of multilocus genotypes among (Figure 2) and within (Figure 3) drainage basins. This method avoids making strong assumptions about the underlying genetic model, generating a PCA based on a priori groups, which are then used as variables in a discriminant analysis (Jombart et al., 2009, 2010). We performed the DAPC on the training set (90% of our data) with an increasing number of principal components (PCs). This training set was used to predict the group membership of the excluded individuals (validation set), which was used to identify the optimal number of PCs to retain. Using the xvalDAPC function for cross-validation, we retained the following number of PCs for our subsequent analyses: 23 PCs among the five drainage basins (Appendix S1: Figure S2), 20 PCs within the Great Lakes basin (Appendix S1: Figure S3), five PCs in the Northeast Basin (Figure S3), and four PCs in the Ohio River basin (Figure S3). Finally, we estimated how well supported the site membership was relative to the five a priori drainage basins or a priori streams within a drainage basin using the compoplot function in adegenet (Figures S2 and S3). To demonstrate that the inclusion of sites with few gametophytes did not lead to a bias in the DAPC results, we conducted an additional DAPC excluding the two sites where the gametophyte count was $n < 6$ (refer to Figure S2).

We calculated pairwise genetic differentiation between all pairs of sites based on allele identity (F_{ST}) and geographic distance (km) using GPS coordinates with distance calculated along the surface of the earth in GenoDive ver. 3.06 (Meirmans, 2020; Appendix S1: Table S6). We performed Mantel tests using F_{ST} and geographic distance among all sites and within drainage basins that had more than two sites, using R version 4.3.1 (R Core Team, 2023) in order to test for isolation

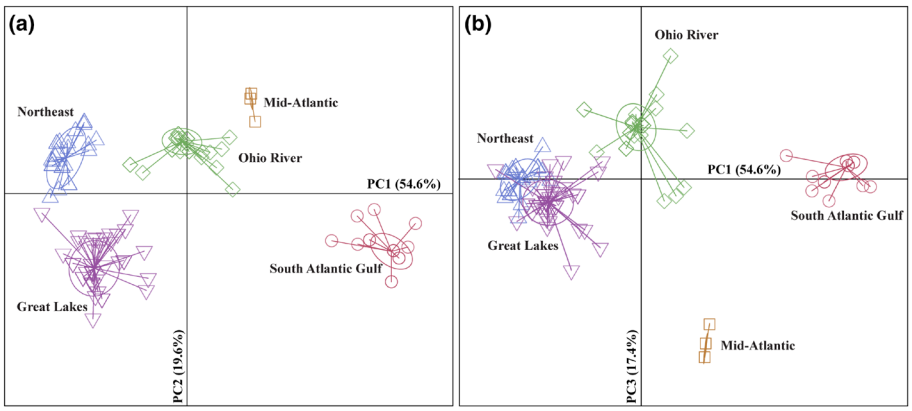

FIGURE 2 Discriminant analysis of principal components (DAPC) of *Batrachospermum gelatinosum* multilocus genotypes (MLGs) among the five drainage basins in eastern North America. Symbols represent gametophytes and correspond to the five a priori groups determined by drainage basins. (a) PC1 vs PC2 (b) PC1 vs PC3. [Color figure can be viewed at wileyonlinelibrary.com]

by distance (Figure 4). Using GenAlEx version 6.51b2 (Smouse et al., 2017), we calculated allele frequencies for each locus for each drainage basin (Figure 5).

To understand the distribution pattern and range expansion of the species, we ran simulations using the program DIYABC 2.1.0 (Cornuet et al., 2008, 2014) to compare a priori migration scenarios. Six scenarios hypothesized a lower latitude (i.e., southern) refugium with different expansion patterns after the Pleistocene glaciation (Figure 6). South Atlantic Gulf origin Scenarios 1–3 describe expansion from sites in the South Atlantic Gulf to higher latitudes. Mid-Atlantic origin Scenarios 1–3 describe expansion from sites in the Mid-Atlantic to higher latitudes. Based on the hypothesis that *Batrachospermum gelatinosum* dispersed to higher latitudes from lower latitude refugia following the Pleistocene glaciation, we first ran the DIYABC analysis using these six scenarios. After determining that the most genetically diverse populations were in the Great Lakes drainage basin (see Table 2), we decided to run an additional analysis that included a seventh scenario, Great Lakes: Great Lakes origin and expansion to lower latitudes. For both analyses, we generated a reference table of 1,000,000 generations per scenario. For each scenario, we allowed population sizes to change for each drainage basin throughout the population divergence events (Appendix S1: Table S7). We used the default mutation model parameters and then calculated one-sample—mean number of alleles, mean genetic diversity, and mean size variance—and two-sample—F_{ST}, classification index, and $(d\mu)^2$ distance—summary statistics. All priors and parameter ranges are presented in Table S7. To determine the scenario that most closely modeled dispersal, we used the logistic regression test implemented in DIYABC based on comparisons of 1% of simulated data sets closest to the observed data for the South Atlantic Gulf and Mid-Atlantic analysis (Scenarios 1–6) and South Atlantic Gulf and Mid-Atlantic with Great Lakes (Scenarios 1–7) analysis (Cornuet et al., 2008, 2014; Appendix S1: Table S8). To determine the goodness-of-fit of simulated scenarios, a PCA (Appendix S1: Figure S4) was conducted, using the posterior distributions of summary statistics in comparison to the observed data set using the perform model checking analysis (Appendix S1: Table S9).

RESULTS

Genotyping and null alleles

We genotyped a total of 311 gametophytes using 10 microsatellite loci from 18 sites spanning five major drainage basins in eastern North America (Figure 1, Table 1). Overall, null allele frequencies were low (0.0%–4.1%; Table S4). However, at locus Bgel_071, the null allele frequency was 22.6% when KY-CDK, OH-BOG, and OH-MCC were included but 0% when these sites were excluded. There appears to be a geographic pattern that is restricted to these three sites in the Ohio River Basin and most of their gametophytes ($n = 72/109$ without amplification). Due to the geographic pattern of nonamplification, locus Bgel_071 was only included to describe private allele frequency and was not used in other analyses.

Physiochemical data and genetic diversity

Although the five drainage basins represent a relatively broad geographic range (Figure 1), the in-stream characteristics (pH, conductivity, current velocity, stream depth, stream width, and water temperature) showed no significant differences among drainage basins (Table S5). The only significant variable was canopy cover (one-way ANOVA, $F(4,9)$ 19.09, $p < 0.002$): The South Atlantic Gulf has more canopy cover than the other drainage basins. In addition, a PCA showed no clustering of drainage basins based on the environmental characteristics measured (Figure S1).

FIGURE 3 Discriminant analysis of principal components (DAPC) of *Batrachospermum gelatinosum* multilocus genotypes (MLGs) within drainage basins showing the first three principal components. Symbols represent gametophytes and correspond to the a priori groups determined by streams within each drainage basin. Site numbers and abbreviations as in Table 1. (a,b) streams within the Ohio River basin. (c,d) streams within the Northeast basin. (e,f) streams within the Great Lakes basin. [Color figure can be viewed at wileyonlinelibrary.com]

There was a similar pattern of genetic diversity among the drainage basins. Unbiased diversity (*uh*) ranged from 0.000 to 0.422 (Table 2). The Great Lakes had three sites with *uh* values >0.353, but the other three were <0.059. Likewise in the Ohio River, one site (KY-CDK) was 0.341, whereas the other four sites were <0.053. Overall, the *uh* values were typically low with 10 of 16 sites having *uh* values <0.096. Mean allelic richness was calculated based on 96 observed genotypes and ranged from 1.0 to 2.0 but was <1.2 for most sites (Table 2). The sites with the higher allelic richness were in the Great Lakes, Ohio River, and Northeast. The number of eMLGs ranged from one to eight per site and were variable within drainage basins (Table 2). Three of the six Great Lakes sites had eMLG > 6, as did RI-CPR (6.6) in the Northeast. In KY-CDK from the Ohio River and CT-FLB from the Northeast, eMLG values were 5.74 and 5.73, respectively; however, 10 out of 16 sites had eMLG values <4.5.

Allele frequency differed among drainage basins at each locus (Figure 5). At loci Bgel_021, Bgel_052, and Bgel_067, there were few alleles (≤5) overall, such that most alleles were shared among the drainage basins. However, the Ohio River basin had two unique alleles among these loci (Figure 5a–c). At loci Bgel_053, Bgel_056, Bgel_057, Bgel_059, and

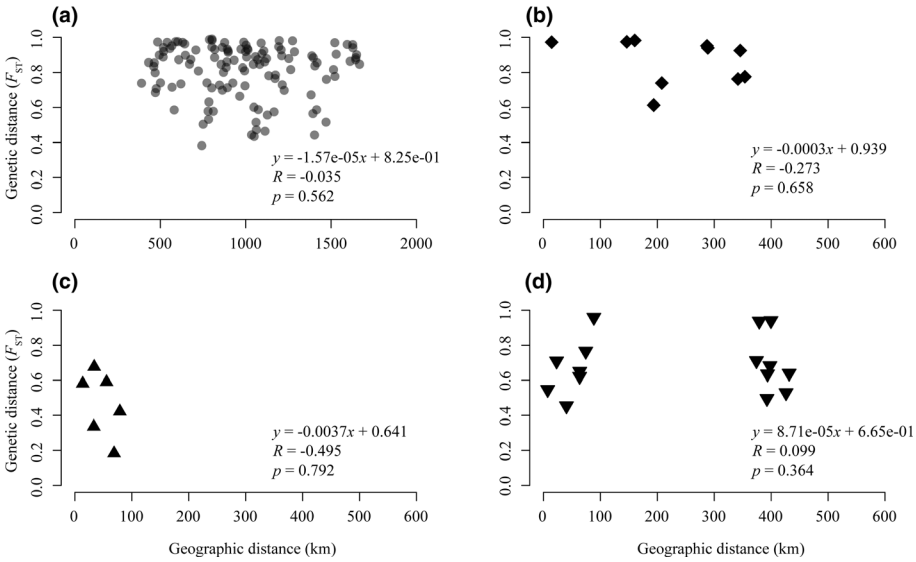

FIGURE 4 Tests for isolation by distance for *Batrachospermum gelatinosum* sites among and within the five major drainage basins in eastern North America. Pairwise genetic distances (F_{ST}) and geographic distances (km) are shown. Symbols as in Figure 1. All slopes were not significantly different from zero ($p > 0.05$). (a) all sites in eastern North America, (b) Ohio River sites, (c) Northeast sites, and (d) Great Lakes sites.

Bgel_073, we observed more alleles per locus, and at these loci, the South Atlantic Gulf, the Ohio River, the Northeast, and the Great Lakes had unique alleles (Figure 5d–h). Loci Bgel_070 and Bgel_071 had the greatest number of alleles per locus with 19 and 24, respectively (Figure 5i,j). At these two loci, all drainage basins had unique alleles, with the Great Lakes having 11 unique alleles.

The number of private alleles per stream varied between zero and five alleles within all drainage basins (Table 2). At eight sites, the number of private alleles was ≤1. Site MI-TRB had five private alleles at three loci (Bgel_053, Bgel_070, Bgel_073). Three sites each had four private alleles: sites AL-CRC (Bgel_053, Bgel_071), CT-FLB (Bgel_057, Bgel_071), and RI-CPR (Bgel_053, Bgel_059, Bgel_070, Bgel_071). At site CT-BLR, three private alleles at two loci (Bgel_053, Bgel_071) were recorded even though we only collected four gametophytes. Most of the private alleles were at Bgel_071 (33%), followed by Bgel_070 (22%) and Bgel_053 (17%).

Genetic structure

We observed genetic structure among and within drainage basins. Among the five drainage basins, the sites within each drainage basin clustered together, and each basin was separated along the first two PC axes (Figure 2a). The Great Lakes and Northeast sites clustered closer together, and the Mid-Atlantic and South Atlantic Gulf sites were further separated from the other drainage basins (Figure 2b). The three PCs explained 91.6% of the variation, with PC1 explaining 54.6%; PC2, 19.6%; and PC3, 17.4%. When we removed sites where $n < 6$ gametophytes, there was a similar pattern among sites, confirming that the inclusion of these sites with small sample sizes did not have an undue influence on the results (Figure S1c). The within-drainage basin plots for the Ohio River, Northeast, and Great Lakes showed all sites separated from each other using the first three PCs (Figure 3). Within the Ohio River, the five streams were genetically distinct, and PC1 explained 74.5% of the variation (Figure 3a). Within the Northeast, the four streams were genetically distinct, and PC1 explained 73.2% of the variation (Figure 3c). The six streams sampled in the Great Lakes were genetically distinct, and PC1 explained 59.4% of the variation (Figure 3e). Most notably, two sites on the same river, MI-TRM and MI-TRB, only ~8 km apart, were genetically distinct (Figure 3e,f).

Isolation by distance

Among the 18 sites, genetic differentiation (F_{ST}) ranged from 0.183–0.990, and geographic distance ranged from 7.5–1666.7 km (Table S6). We observed no evidence of isolation by distance among drainage basins (Mantel test, $R = -0.035$, $p = 0.562$, Figure 4a). Likewise, within a drainage basin, there was no evidence of genetic isolation (Figure 4b–d). Genetic differentiation varied among pairs of geographically close sites such as CT-IRM and RI-CPR (13.8 km) with $F_{ST} = 0.183$, AL-YEC and AL-CRC (17.3 km) with $F_{ST} = 0.841$, and MI-TRM and MI-TRB (7.5 km and same river) with $F_{ST} = 0.547$.

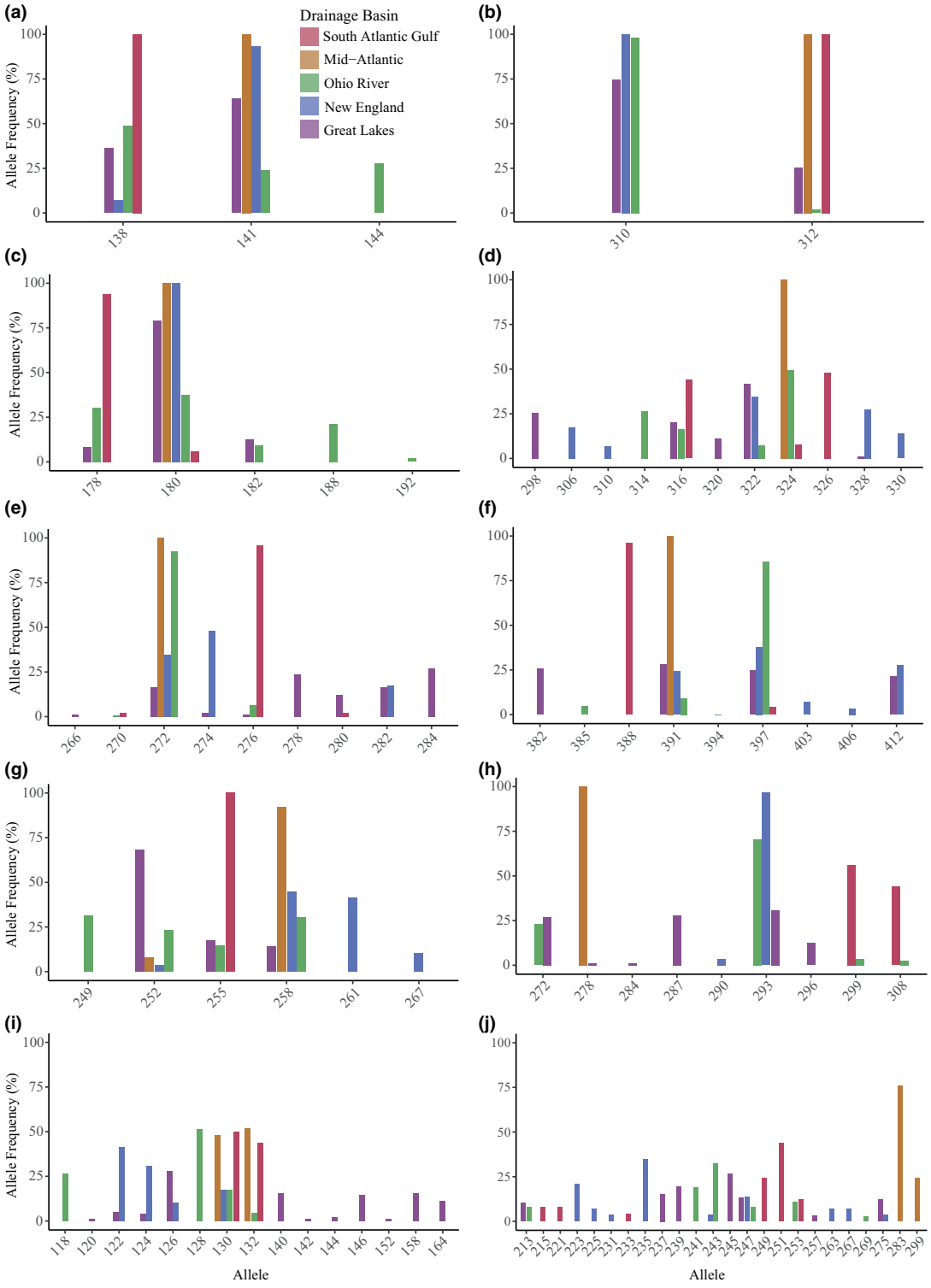

FIGURE 5 Bar plots showing allele frequency at 10 loci for each drainage basin. Colors represent drainage basins. (a) Locus Bgel_021, (b) Locus Bgel_052, (c) Locus Bgel_067, (d) Locus Bgel_053, (e) Locus Bgel_056, (f) Locus Bgel_057, (g) Locus Bgel_059, (h) Locus Bgel_073, (i) Locus Bgel 070, and (j) Locus Bgel 071.

Dispersal following the Pleistocene glaciation

Results from the two DIYABC analyses (the South Atlantic Gulf and Mid-Atlantic analysis and the South Atlantic Gulf, Mid-Atlantic, and Great Lakes analysis) indicated that a Mid-Atlantic refugium origin was most likely (Mid-Atlantic Scenario 2, Figure 6). Dispersal likely occurred from the Mid-Atlantic to the Ohio River and the South-Atlantic Gulf and then from the Ohio River to both the Great Lakes and

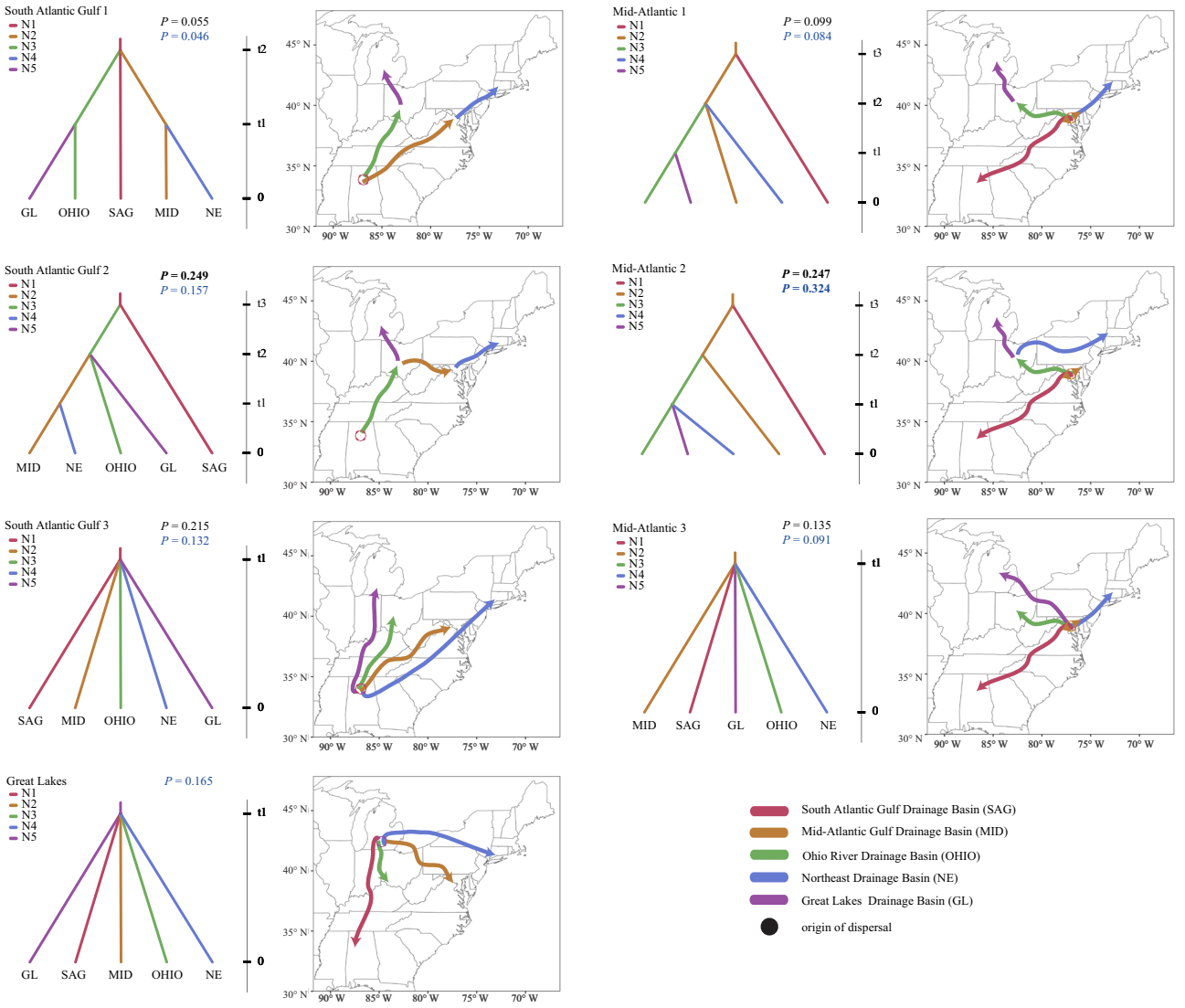

FIGURE 6 Graphical representations of DIYABC scenarios depicting potential colonization pathways of *Batrachospermum gelatinosum* with branching patterns on the left and a map on the right. The map was drawn using ggplot2 in R. The probability (*p*) values in black represent the analysis of the South Atlantic Gulf and Mid-Atlantic scenarios, while the *p* values in blue represent the analysis of South Atlantic Gulf, Mid-Atlantic, and Great Lakes scenarios. Bolded *p* values indicate most likely scenario. A time bar (not to scale) is shown to represent the population divergence events.

Northeast basins. This scenario (Mid-Atlantic Scenario 2, Figure 6) had the highest posterior probability in the South-Atlantic Gulf, Mid-Atlantic, and Great Lakes analysis (0.324) but had a similar probability (0.247) to the South Atlantic Gulf 2 scenario (0.249) in the South Atlantic and Mid-Atlantic analysis. In the analysis of posterior distributions for summary statistics, model-checking revealed that simulated data sets clustered closely around the observed data set (Figure S3). All estimates for demographic parameters are shown in Table S9.

DISCUSSION

In *Batrachospermum gelatinosum*, we observed strong genetic partitioning among the five major drainage basins. Contrary to our prediction, we detected no admixture between sites within a drainage basin, even between sites that were geographically close. Although we expected low genetic diversity in higher latitude basins, diversity statistics were highly variable among sites and did not show a clear pattern among drainage basins. Our analyses indicated that dispersal following the Pleistocene was likely from the Mid-Atlantic. We now discuss these results and their implications on the genetic diversity of *B. gelatinosum*.

Genetic diversity within and among drainage basins

We sampled across five geographically expansive drainage basins. At each location on the sampling date, we recorded as many physical and chemical parameters

as feasible, we could consider their potential influence on genetic diversity among drainage basins. However, the results showed that there were no significant differences among the drainage basins in the parameters measured except canopy cover, which was greatest in the South Atlantic Gulf. This difference may have been a result of latitudinal gradient in leaf-out as we sampled both low-latitude and high-latitude sites at roughly the same time of year. At a finer scale, Shainker-Connelly et al. (2024) also showed no correlation between population genetic statistics and stream environmental variables. Although environmental conditions undoubtedly determine the suitability of a site for colonization and subsequent proliferation of *Batrachospermum gelatinosum*, we did not find that the point estimates we measured influenced patterns of genetic diversity. As we were only able to take a single snapshot of physical and chemical parameters, continuous environmental data may provide insights on the influence of physical and chemical parameters on genetic diversity, as has been shown for marine algae with respect to physical parameters (Fouqueau et al., 2024).

We expected to observe greater genetic diversity at lower latitude sites as compared to sites at higher latitudes (House et al., 2010). However, genetic diversity varied among sites and did not appear to have a geographic pattern. For example, we observed streams with little to no diversity (e.g., OH-MCC, $uh = 0.015$) as well as streams with higher diversity (e.g., KY-CDK, $uh = 0.341$) within the same basin. In addition, the sites with the highest genetic diversity were from the Great Lakes and Northeast rather than the lower latitude sites in the South Atlantic Gulf and Mid-Atlantic. However, we note that we only sampled three sites in total from these two basins.

The variable pattern of genetic diversity among streams within a drainage basin may be explained by the prevailing reproductive mode. In a companion paper, we observed high levels of what we interpret to be intragametophytic selfing overall, although this varied by stream (Shainker-Connelly et al., 2024). We have not yet excluded the possibility of monospore production by the chantransia, as we were limited in our ability to distinguish between apomictic processes and intragametophytic selfing in driving these patterns (Shainker-Connelly et al., 2024). Nevertheless, patterns of the prevailing reproductive mode support our observations of genetic variation in this study. For example, the site with the greatest mean allelic richness and eMLGs (MI-CUT) was characterized by low intragametophytic selfing (pareto $\beta = 2.17$). A site in the same drainage basin, MI-CLI, had low genetic diversity and was primarily characterized by high intragametophytic selfing (pareto $\beta = 0.21$) (Shainker-Connelly et al., 2024). Any carpospores produced from intragametophytic selfing would be homozygous across the entire genome (Klekowski, 1973), which could result in the loss of any genetic diversity within a stream reach. This variable pattern could also be explained by founder events (as it is likely that few individuals would disperse long distances to new habitats) or bottleneck events, both of which would be compounded by intragametophytic selfing.

Genetic structure across large and fine spatial scales

Studies of the genetic structure of other freshwater red algal species have shown that genetic differentiation occurs across large spatial scales. Hall and Vis (2002), using inter simple sequence repeats, observed substantial genetic variation among gametophytes of *Virescentia viride-americana* (as *Batrachospermum helminthosum*) within 11 streams in the eastern United States. Genetic variation was partitioned as 45% within stream and 55% among streams, such that no conclusions about phylogeographic patterns were made. Hall and Vis (2002) concluded that the individuals in each stream segment were genetically unique. Chiasson et al. (2003) studied this same species and primarily the same stream sites, using *cox*2–3 spacer region sequence data and observed 13 haplotypes, but only four of the haplotypes were shared among two locations. Given the large number of haplotypes but few shared, the authors postulated that infrequent long-distance dispersal may explain the pattern of genetic variation observed across the large spatial scales. Hall and Vis (2002) did not find a correlation between genetic and geographic distance and Chiasson et al. (2003) observed that many streams had at least one haplotype that was not collected at other sites, showing the individuals from each stream may make an important contribution to genetic diversity within the species.

In *Batrachospermum gelatinosum*, we also observed partitioning of genetic structure among drainage basins across eastern North America. These streams varied in geographic distance, ranging from 7.5 km on the same river to 1666 km. Although strong genetic differentiation was observed among the drainage basins over a large latitudinal gradient, we observed no evidence of isolation by distance among the drainage basins. Our findings suggest that long-distance dispersal events may be rare for *B. gelatinosum* but that dispersal to drainage basins may have occurred infrequently from other sites.

Within each drainage basin, we expected admixture and gene flow over short distances. However, we observed strong genetic differentiation among sites. Sites within each drainage basin were genetically distinct, even when in proximity (e.g., 7.5 km). Comparable to our observation at large geographic scales, we observed no evidence of genetic isolation by distance within drainage basins. For example, MI-TRM and

MI-TRB (7.5 km apartment and in the same river)—the closest pair of sites—had greater genetic differentiation as measured by F_{ST} than CT-IRM and RI-CPR, which were 68 km apart and not in the same subdrainage basin. These within stream basin results suggest that even with further study, we would be unlikely to detect admixture, which supports the previous research on freshwater red algae that showed each site had a unique genetic composition (Hall & Vis, 2002). This pattern of isolation has also been observed in marine algae. In *Chondrus crispus*, high and low shore intertidal zones had distinct populations over short topographical distances and less than 2 m in tidal height (Krueger-Hadfield et al., 2013). Likewise, Reynes et al. (2021) and Buonomo et al. (2017) observed high levels of genetic differentiation over multiple spatial scales in the brown algae *Laminaria rodriguezii* and *Cystoseira amentacea*, respectively.

Pleistocene refugia and dispersal mechanisms

The phylogeography of eastern North American biota has, in large part, been influenced by the Pleistocene glaciation (Lomolino et al., 2015; Lyman & Edwards, 2022; Soltis et al., 2006). Postglacial colonization of newly opened habitats through migration from unglaciated refugia at lower latitudes has led to distinct genetic patterns in aquatic organisms, such as fish (Morgan et al., 2018; Strange & Burr, 1997) and mussels (Elderkin et al., 2008; Inoue et al., 2014), as well as terrestrial organisms, such as salamanders (Burkhart et al., 2019; Crespi et al., 2003) and angiosperms (Kim et al., 2018; Sewell et al., 1996). For some organisms, the northern part of their range is genetically homogeneous, such as for the ringed salamander (Burkhart et al., 2019). Other organisms, like the pygmy salamander, have genetically distinct populations throughout eastern North America (Crespi et al., 2003). Elderkin et al. (2008) observed one species of mussel, *Actinonaias ligamentina*, dispersed into northern areas from two genetically distinct glacial refugia, yet another species, *Elliptio dilatata*, likely dispersed from one refugium followed by low rates of gene flow in northern areas. Our results from *Batrachospermum gelatinosum* in this study as well as those from *Virescentia virideamericana* (Hall & Vis, 2002) showed that at least two freshwater red algae are more like the tiger salamander and the mussel *Elliptio dilatata* (Elderkin et al., 2008) in which one refugium contributed to the recolonization of higher latitudes. However, the reasons for these similar genetic patterns between freshwater red algae and other organisms are unknown.

Freshwater organisms disperse actively or passively, with passive dispersal being the most common in freshwater systems (Bohonak & Jenkins, 2003). For freshwater red algae, there are likely two potential modes of passive dispersal: water and adherence to motile organisms (Kristiansen, 1996; Vis, 2016). For water dispersal, there is evidence that dispersal of diatom taxa is likely constrained outside of individual hydrological networks, and flow direction plays a notable role (Liu et al., 2013). Most likely, freshwater red algae are similarly constrained, as they have a physically attached macroscopic gametophyte produced from the "chantransia." However, after carpospore or monospore dispersal via water flow, many factors may restrict the establishment of a downstream population, including stream depth and substrate availability. This constraint is evident when comparing two sites in Michigan that were collected on the same river, 7.5 km apart (MI-TRM and MI-TRB). These two populations are genetically distinct, and the downstream population, MI-TRB, had the greatest genetic diversity.

Motile organisms could and most likely do play roles in the dispersal of *Batrachospermum gelatinosum*. Within a drainage basin, fish, crayfish, and macroinvertebrates may disperse this alga over short and long distances (Fuelling et al., 2012; Hambrook & Sheath, 1987; Velasquez, 1940). It is known that some algae can pass through macroinvertebrate and fish guts and still be viable (Hambrook & Sheath, 1987; Velasquez, 1940). The freshwater red algal genus *Thorea* has been reported attached to the antennae of crayfish being carried up and downstream (Fuelling et al., 2012). Species in the order Batrachospermales have been documented as having gametophytes that grow attached to shells of snails (authors' personal observations). However, there is no known evidence of these potential dispersal mechanisms in *B. gelatinosum*.

Dispersal among drainage basins could potentially occur via waterfowl (Kristiansen, 1996), as studies have shown viable algae can be recovered from fecal pellets and washed off of feet (Proctor, 1963; Schlichting, 1960). Molecular studies on freshwater algae have detected evidence of long-distance dispersal and hypothesized waterfowl as a dispersal vector (Meiers et al., 1999; Oberholster et al., 2005), including for the freshwater red algal genus *Sirodotia* (Lam et al., 2012). However, *S. suecica* in eastern North America is located mostly in coastal plains, whereas *Batrachospermum gelatinosum* is much more widespread across many ecoregions (Vis et al., 1996; Vis & Necchi Jr., 2021). For *B. gelatinosum*, long-distance dispersal by waterfowl from a Mid-Atlantic refugium to higher latitudes is plausible. It is likely that these dispersal events did not occur more than once or a few times based on the general lack of genetic diversity within each drainage basin and within a stream, leading to genetically distinct sites. Nevertheless, it is just as likely that dispersal occurred more often than originally thought. This possibility arises from the variability in rates of intragametophytic selfing over time, as

suggested by Shainker-Connelly et al. (2024), which ultimately erodes genetic diversity.

In the widespread freshwater mussel *Cumberlandia monodota*, analyses using DIYABC pointed to the presence of two glacial refugia followed by the simultaneous dispersal into the extant range following the Pleistocene glaciation (Inoue et al., 2014). Our DIYABC analyses indicated that the Mid-Atlantic served as a refugium for *Batrachospermum gelatinosum*. We had originally predicted the South Atlantic Gulf as the location of the refugial sites. One potential explanation may be that *B. gelatinosum* persisted in refugia that were just south of the ice margins, like the American beech did (see Soltis et al., 2006, figure 6). To verify persistence in refugia near the ice margins, future studies could concentrate sampling in the Tennessee River basin and other higher latitude locations at the ice margins in the South Atlantic Gulf and Mid-Atlantic. There may be more than one refugial area from which *B. gelatinosum* dispersed into higher latitudes, such as has been reported for mussel and fish species (April et al., 2013; Elderkin et al., 2008; Inoue et al., 2014).

For this study, we were only able to sample two locations from the South Atlantic Gulf and one in the Mid-Atlantic. Historical records, literature reports, and herbarium specimens suggest that *Batrachospermum gelatinosum* is more common in these drainage basins (Vis et al., 1996; Vis & Sheath, 1997). Nevertheless, we visited several previously reported stream locations without finding macroscopic gametophytes to sample. One of the locations in Alabama from Chiasson et al. (2003) had been altered by housing development and is no longer suitable for this taxon. Other locations at the time of visit had no flowing water, and given the lack of a desiccation-resistance *B. gelatinosum* may be extirpated. Additionally, gametophytes are seasonal and may not have been present at the time of sampling, but the chantransia (not sampled) is believed to be perennial. Future work should assess the patterns of genetic variation in the chantransia (sporophyte) alongside gametophytes, as has been done in marine red algae (tetrasporophytes and gametophytes; see review in Krueger-Hadfield et al., 2021).

The results from our DIYABC analyses revealed that *Batrachospermum gelatinosum* first dispersed from the Mid-Atlantic basin to the South Atlantic Gulf, then to the Ohio River basin. From there, it may have simultaneously dispersed to the Great Lakes and Northeast basins. However, the posterior probability of this best-fit scenario only had moderate support in the South Atlantic Gulf, Mid-Atlantic, and Great Lakes analysis. In the South Atlantic Gulf and Mid-Atlantic analysis, dispersal from the South Atlantic Gulf had similar probability. In this scenario (South Atlantic Gulf 2, Figure 6), *B. gelatinosum* dispersed to the Ohio River from the South Atlantic Gulf. From the Ohio River, dispersal simultaneously occurred to the Great Lakes and Mid-Atlantic, followed by dispersal to the Northeast. The low probability for this analysis could be due to the small number of sites among the southernmost drainage basins. Adding more sites would increase our power to assess patterns of genetic diversity within these basins. Due to intragametophytic selfing, any remaining genetic variation following dispersal events has a high probability of erosion and may limit predictions of demographic history analyses (Shainker-Connelly et al., 2024). Rates of intragametophytic selfing varied by stream; therefore, sampling more sites could uncover more diversity and aid in the understanding of *B. gelatinosum* dispersal following the Pleistocene glaciation.

Future directions

Previous biogeographic studies of *Batrachospermum gelatinosum* relied exclusively on DNA sequence data of the ITS rRNA region or a portion of the *cox*1 gene (House et al., 2010; Vis et al., 1996; Vis & Sheath, 1997). Anywhere from one to three haplotypes per stream segment were recorded, with low haplotypic diversity over a large geographic range (House et al., 2010; Vis et al., 1996; Vis & Sheath, 1997). In the present study, we used microsatellite data that have more discriminating power and observed strong genetic partitioning among the five major drainage basins with no admixture between geographically proximate sites. Further sampling is necessary to disentangle the patterns of dispersal to higher latitudes following the Pleistocene glaciation. As the macroscopic gametophyte stage is seasonal, future studies could take advantage of sampling the microscopic diploid chantransia. The chantransia is hypothesized to persist as long as streams are not dry. In our sampling, we were only able to collect a total of three sites from the unglaciated eastern North America. Collecting more samples within this unglaciated region could further our understanding of the dispersal of this alga following glacial retreat. Approaches that include more sampling across the genome (e.g., single-nucleotide polymorphisms, SNPs; Delord et al., 2018) could also validate demographic history analyses and provide additional support for the results. Given that numerous freshwater red algae (e.g., *Tuomeya americana*, *Virescentia viride-americana*, *Sheathia* spp.) have similar distributions in eastern North America, developing population-level genetic tools for other taxa would allow for comparative studies to determine if the results from *B. gelatinosum* are representative of all freshwater red algae. *Batrachospermum gelatinosum* may be another example of a taxon that did not come from a glacial refugium in the most southern

locations in the range but rather dispersed to higher latitudes from refugia close to the ice margins.

AUTHOR CONTRIBUTIONS

Roseanna M. Crowell: Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Sarah J. Shainker-Connelly:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); visualization (equal); writing – original draft (supporting); writing – review and editing (equal). **Stacy A. Krueger-Hadfield:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); visualization (equal); writing – original draft (supporting); writing – review and editing (equal). **Morgan L. Vis:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

We thank G. Lindsey, B. Thornton, A. Oetterer, M. Amsler, B. Anderson, C. Schneider, R. Standaert, and W. Chiasson for help collecting *B. gelatinosum*; A. Cao and M. Crowley for use of the capillary sequencer at the Helfin Center for Genomic Sciences at the University of Alabama at Birmingham (UAB); D. Alvarado-Serrano for his assistance with analyses; and J. Schenk, H. Ballard, and S. Kuchta for serving on the dissertation committee of RMC. Algae from MD-HOU were sampled under a permit from the Maryland Department of Natural Resources, and from MI-CLI, MI-CUT, MI-KCL, and MI-HLK under a permit from the Michigan Department of Natural Resources. This project was supported by a Phycological Society of America Grant-in-Aid of Research (to RMC), an Ohio University Student Enhancement Award (to RMC), the Ohio Center for Ecology and Evolutionary Studies Fellowship (to RMC), the Ohio University Roach Fund (to RMC), the Ohio University Graduate Student Senate Original Work Grant (to RMC), Ohio University research incentive funds (MLV), the UAB Department of Biology's Harold Martin Outstanding Student Development Award (to SJSC), start-up funds from the College of Arts and Sciences at UAB (to SAKH), a National Science Foundation (NSF) CAREER Award (DEB-2141971 to SAKH). SJSC was supported by the UAB Blazer Fellowship and the NSF Graduate Research Fellowship (2020295779). This manuscript represents the partial fulfillment of the PhD dissertation of RMC.

FUNDING INFORMATION

Phycological Society of America Grant-in-Aid of Research (to RMC); Ohio University Student Enhancement Award (to RMC); Ohio Center for Ecology and Evolutionary Studies Fellowship (to RMC); Ohio University Roach Fund (to RMC); Ohio University Graduate Student Senate Original Work Grant (to RMC); Ohio University research incentive funds (to MLV); UAB Department of Biology's Harold Martin Outstanding Student Development Award (to SJSC); UAB College of Arts and Sciences start-up funds (to SAKH); National Science Foundation CAREER Award (DEB-2141971 to SAKH); NSF GRFP(2020295779 to SJSC); UAB Blazer Fellowship (to SJSC).

ORCID

Roseanna M. Crowell https://orcid.org/0000-0003-3117-2094
Stacy A. Krueger-Hadfield https://orcid.org/0000-0002-7324-7448
Morgan L. Vis https://orcid.org/0000-0003-3087-1563

REFERENCES

Adamack, A. T., & Gruber, B. (2014). PopGenReport: Simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*, *5*(4), 384–387. https://doi.org/10.1111/2041-210X.12158

April, J., Hanner, R. H., Dion-Côté, A.-M., & Bernatchez, L. (2013). Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. *Molecular Ecology*, *22*(2), 409–422. https://doi.org/10.1111/mec.12116

Becker, R. A., Chambers, J. M., & Wilks, A. R. (1988). *The new S language*. Wadsworth & Brooks/Cole.

Bohonak, A. J., & Jenkins, D. G. (2003). Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters*, *6*(8), 783–796. https://doi.org/10.1046/j.1461-0248.2003.00486.x

Buonomo, R., Assis, J., Fernandes, F., Engelen, A. H., Airoldi, L., & Serrão, E. A. (2017). Habitat continuity and stepping-stone oceanographic distances explain population genetic connectivity of the brown alga *Cystoseira amentacea*. *Molecular Ecology*, *26*(3), 766–780. https://doi.org/10.1111/mec.13960

Burkhart, J. J., Puckett, E. E., Beringer, C. J., Sholy, C. N., Semlitsch, R. D., & Eggert, L. S. (2019). Post-Pleistocene differentiation in a central interior highlands endemic salamander. *Ecology and Evolution*, *9*(19), 11171–11184. https://doi.org/10.1002/ece3.5619

Cánovas, F. G., Mota, C. F., Serrão, E. A., & Pearson, G. A. (2011). Driving south: A multi-gene phylogeny of the brown algal family Fucaceae reveals relationships and recent drivers of a marine radiation. *BMC Evolutionary Biology*, *11*(1), 371. https://doi.org/10.1186/1471-2148-11-371

Chiasson, W. B., Machesky, N. J., & Vis, M. L. (2003). Phylogeography of a freshwater red alga, *Batrachospermum helminthosum*, in North America. *Phycologia*, *42*(6), 654–660. https://doi.org/10.2216/i0031-8884-42-6-654.1

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., & Estoup, A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data.

Bioinformatics, *30*(8), 1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Cornuet, J.-M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J.-M., Balding, D. J., Guillemaud, T., & Estoup, A. (2008). Inferring population history with DIY ABC: A user-friendly approach to approximate Bayesian computation. *Bioinformatics*, *24*(23), 2713–2719. https://doi.org/10.1093/bioinformatics/btn514

Crespi, E. J., Rissler, L. J., & Browne, R. A. (2003). Testing Pleistocene refugia theory: Phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. *Molecular Ecology*, *12*(4), 969–984. https://doi.org/10.1046/j.1365-294X.2003.01797.x

Crowell, R. M., Shainker-Connelly, S. J., Vis, M. L., & Krueger-Hadfield, S. A. (2024). Microsatellite development in the freshwater red alga *Batrachospermum gelatinosum* (L.) De Candolle (Batrachospermales, Rhodophyta). *Cryptogamie Algologie*, *45*(5), 53–62. https://doi.org/10.5252/cryptogamie-algologie2024v45a5

de Meeûs, T., McCoy, K. D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-Boussès, S., & Renaud, F. (2007). Population genetics and molecular epidemiology or how to "débusquer la bête." *Infection, Genetics and Evolution*, *7*(2), 308–332. https://doi.org/10.1016/j.meegid.2006.07.003

Delord, C., Lassalle, G., Oger, A., Barloy, D., Coutellec, M.-A., Delcamp, A., Evanno, G., Genthon, C., Guichoux, E., Le Bail, P.-Y., Le Quilliec, P., Longin, G., Lorvelec, O., Massot, M., Reveillac, E., Rinaldo, R., Roussel, J.-M., Vigouroux, R., Launey, S., & Petit, E. J. (2018). A cost-and-time effective procedure to develop SNP markers for multiple species: A support for community genetics. *Methods in Ecology and Evolution*, *9*(9), 1959–1974. https://doi.org/10.1111/2041-210X.13034

Díaz-Tapia, P., Maggs, C. A., Macaya, E. C., & Verbruggen, H. (2018). Widely distributed red algae often represent hidden introductions, complexes of cryptic species or species with strong phylogeographic structure. *Journal of Phycology*, *54*(6), 829–839. https://doi.org/10.1111/jpy.12778

El Mousadik, A., & Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, *92*(7), 832–839. https://doi.org/10.1007/BF00221895

Elderkin, C. L., Christian, A. D., Metcalfe-Smith, J. L., & Berg, D. J. (2008). Population genetics and phylogeography of freshwater mussels in North America, *Elliptio dilatata* and *Actinonaias ligamentina* (Bivalvia: Unionidae). *Molecular Ecology*, *17*(9), 2149–2163. https://doi.org/10.1111/j.1365-294X.2008.03745.x

Engel, C. R., Destombe, C., & Valero, M. (2004). Mating system and gene flow in the red seaweed *Gracilaria gracilis*: Effect of haploid–diploid life history and intertidal rocky shore landscape on fine-scale genetic structure. *Heredity*, *92*(4), 289–298. https://doi.org/10.1038/sj.hdy.6800407

Fouqueau, L., Reynes, L., Tempera, F., Bajjouk, T., Blanfuné, A., Chevalier, C., Laurans, M., Mauger, S., Sourisseau, M., Assis, J., Lévêque, L., & Valero, M. (2024). *Influence of oceanography and geographic distance on genetic structure: How varying the sampled domain influences conclusions in* Laminaria digitata. bioRxiv https://doi.org/10.1101/2023.05.11.540379

Fuelling, L., Adams, J., Badik, K., Bixby, R., Caprette, C., Caprette, H., Chiasson, W., Davies, C., DeColibus, D., Glascock, K., Hall, M., Perry, W., Schultz, E., Taylor, D., Vis, M., & Verb, R. (2012). An unusual occurrence of *Thorea hispida* (Thore) Desvaux chantransia on rusty crayfish in West Central Ohio. *Nova Hedwigia*, *94*(3–4), 355–366. https://doi.org/10.1127/0029-5035/2012/0018

Gruber, B., & Adamack, A. T. (2015). Landgenreport: A new R function to simplify landscape genetic analysis using resistance surface layers. *Molecular Ecology Resources*, *15*(5), 1172–1178. https://doi.org/10.1111/1755-0998.12381

Guillemin, M.-L., Faugeron, S., Destombe, C., Viard, F., Correa, J. A., & Valero, M. (2008). Genetic variation in wild and cultivated populations of the haploid–diploid red alga *Gracilaria chilensis*: How farming practices favor asexual reproduction and heterozygosity. *Evolution*, *62*(6), 1500–1519. https://doi.org/10.1111/j.1558-5646.2008.00373.x

Hall, M. M., & Vis, M. L. (2002). Genetic variation in *Batrachospermum helminthosum* (Batrachospermales, Rhodophyta) among and within stream reaches using intersimple sequence repeat molecular markers. *Phycological Research*, *50*(2), 155–162. https://doi.org/10.1046/j.1440-1835.2002.00269.x

Hambrook, J. A., & Sheath, R. G. (1987). Grazing of freshwater Rhodophyta. *Journal of Phycology*, *23*(4), 656–662. https://doi.org/10.1111/j.1529-8817.1987.tb04220.x

Hambrook, J. A., & Sheath, R. G. (1991). Reproductive ecology of the freshwater red alga *Batrachospermum boryanum* Sirodot in a temperature headwater stream. *Hydrobiologia*, *218*(3), 233–246. https://doi.org/10.1007/BF00038837

House, D. L., Vandenbroek, A. M., & Vis, M. L. (2010). Intraspecific genetic variation of *Batrachospermum gelatinosum* (Batrachospermales, Rhodophyta) in eastern North America. *Phycologia*, *49*(5), 501–507. https://doi.org/10.2216/09-104.1

Inoue, K., Monroe, E. M., Elderkin, C. L., & Berg, D. J. (2014). Phylogeographic and population genetic analyses reveal Pleistocene isolation followed by high gene flow in a wide ranging, but endangered, freshwater mussel. *Heredity*, *112*(3), 282–290. https://doi.org/10.1038/hdy.2013.104

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

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*(1), 94. https://doi.org/10.1186/1471-2156-11-94

Jombart, T., Pontier, D., & Dufour, A.-B. (2009). Genetic markers in the playground of multivariate analysis. *Heredity*, *102*(4), 4. https://doi.org/10.1038/hdy.2008.130

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, *6*, e00208. https://doi.org/10.3389/fgene.2015.00208

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. https://doi.org/10.7717/peerj.281

Kim, S.-H., Cho, M.-S., Li, P., & Kim, S.-C. (2018). Phylogeography and ecological niche modeling reveal reduced genetic diversity and colonization patterns of skunk cabbage (*Symplocarpus foetidus*; Araceae) from glacial refugia in eastern North America. *Frontiers in Plant Science*, *9*, e00648. https://doi.org/10.3389/fpls.2018.00648

Klekowski, E. J. (1969). Reproductive biology of the Pteridophyta. II. Theoretical considerations. *Botanical Journal of the Linnean Society*, *62*(3), 347–359. https://doi.org/10.1111/j.1095-8339.1969.tb01972.x

Klekowski, E. J. (1973). Genetic endemism of Galapagos *Pteridium*. *Botanical Journal of the Linnean Society*, *66*(2), 181–188. https://doi.org/10.1111/j.1095-8339.1973.tb02168.x

Kristiansen, J. (1996). Dispersal of freshwater algae—A review. *Hydrobiologia*, *336*(1), 151–157. https://doi.org/10.1007/BF00010829

Krueger-Hadfield, S. A., Byers, J. E., Bonthond, G., Terada, R., Weinberger, F., & Sotka, E. E. (2021). Intraspecific diversity and

genetic structure in the widespread macroalga *Agarophyton vermiculophyllum*. *Journal of Phycology*, *57*(5), 1403–1410. https://doi.org/10.1111/jpy.13195

Krueger-Hadfield, S. A., Collén, J., Daguin-Thiébaut, C., & Valero, M. (2011). Genetic population structure and mating system in *Chondrus crispus* (Rhodophyta). *Journal of Phycology*, *47*(3), 440–450. https://doi.org/10.1111/j.1529-8817.2011.00995.x

Krueger-Hadfield, S. A., & Hoban, S. M. (2016). The importance of effective sampling for exploring the population dynamics of haploid–diploid seaweeds. *Journal of Phycology*, *52*(1), 1–9. https://doi.org/10.1111/jpy.12366

Krueger-Hadfield, S. A., Kollars, N. M., Strand, A. E., Byers, J. E., Shainker, S. J., Terada, R., Greig, T. W., Hammann, M., Murray, D. C., Weinberger, F., & Sotka, E. E. (2017). Genetic identification of source and likely vector of a widespread marine invader. *Ecology and Evolution*, *7*(12), 4432–4447. https://doi.org/10.1002/ece3.3001

Krueger-Hadfield, S. A., Roze, D., Mauger, S., & Valero, M. (2013). Intergametophytic selfing and microgeographic genetic structure shape populations of the intertidal red seaweed *Chondrus crispus*. *Molecular Ecology*, *22*(12), 3242–3260. https://doi.org/10.1111/mec.12191

Krueger-Hadfield, S. A., Shainker-Connelly, S. J., Crowell, R. M., & Vis, M. L. (2024). The eco-evolutionary importance of reproductive system variation in the macroalgae: Freshwater reds as a case study. *Journal of Phycology*, *60*(1), 15–25. https://doi.org/10.1111/jpy.13407

Lam, D. W., Entwisle, T. J., Eloranta, P., Kwandrans, J., & Vis, M. L. (2012). Circumscription of species in the genus *Sirodotia* (Batrachospermales, Rhodophyta) based on molecular and morphological data. *European Journal of Phycology*, *47*(1), 42–50. https://doi.org/10.1080/09670262.2011.645885

Lehrian, S., Bálint, M., Haase, P., & Pauls, S. U. (2010). Genetic population structure of an autumn-emerging caddisfly with inherently low dispersal capacity and insights into its phylogeography. *Journal of the North American Benthological Society*, *29*(3), 1100–1118. https://doi.org/10.1899/09-100.1

Lemmon, P. E. (1956). A spherical densiometer for estimating forest overstory density. *Forest Science*, *2*(4), 314–320.

Lemmon, P. E. (1957). A new instrument for measuring forest overstory density. *Journal of Forestry*, *55*(9), 667–668.

Liu, J., Soininen, J., Han, B.-P., & Declerck, S. (2013). Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *Journal of Biogeography*, *40*(12), 2238–2248. https://doi.org/10.1111/jbi.12160

Lomolino, M., Pijanowski, B., & Gasc, A. (2015). The silence of biogeography. *Journal of Biogeography*, *42*(7), 1187–1196. https://doi.org/10.1111/jbi.12525

Lyman, R. A., & Edwards, C. E. (2022). Revisiting the comparative phylogeography of unglaciated eastern North America: 15 years of patterns and progress. *Ecology and Evolution*, *12*(4), e8827. https://doi.org/10.1002/ece3.8827

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press.

Matschiner, M., & Salzburger, W. (2009). TANDEM: Integrating automated allele binning into genetics and genomics workflows. *Bioinformatics*, *25*(15), 1982–1983. https://doi.org/10.1093/bioinformatics/btp303

Meiers, S. T., Proctor, V. W., & Chapman, R. L. (1999). Phylogeny and biogeography of *Chara* (Charophyta) inferred from 18S rDNA sequences. *Australian Journal of Botany*, *47*(3), 347–360. https://doi.org/10.1071/bt97089

Meirmans, P. G. (2020). Genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources*, *20*(4), 1126–1131. https://doi.org/10.1111/1755-0998.13145

Morgan, T. D., Graham, C. F., McArthur, A. G., Raphenya, A. R., Boreham, D. R., Manzon, R. G., Wilson, J. Y., Lance, S. L., Howland, K. L., Patrick, P. H., & Somers, C. M. (2018). Genetic population structure of the round whitefish (*Prosopium cylindraceum*) in North America: Multiple markers reveal glacial refugia and regional subdivision. *Canadian Journal of Fisheries and Aquatic Sciences*, *75*(6), 836–849. https://doi.org/10.1139/cjfas-2016-0528

Oberholster, P. J., Botha, A. M., Muller, K., & Cloete, T. E. (2005). Assessment of the genetic diversity of geographical unrelated *Microcystis aeruginosa* strains using amplified fragment length polymorphisms (AFLPs). *African Journal of Biotechnology*, *4*(5), 389–399.

Proctor, V. (1963). Viability of *Chara* oospores taken from migratory water birds. *Ecology*, *43*, 528–529.

Provan, J., Murphy, S., & Maggs, C. A. (2005). Tracking the invasive history of the green alga *Codium fragile* ssp. *tomentosoides*. *Molecular Ecology*, *14*(1), 189–194. https://doi.org/10.1111/j.1365-294X.2004.02384.x

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reynes, L., Thibaut, T., Mauger, S., Blanfuné, A., Holon, F., Cruaud, C., Couloux, A., Valero, M., & Aurelle, D. (2021). Genomic signatures of clonality in the deep-water kelp *Laminaria rodriguezii*. *Molecular Ecology*, *30*(8), 1806–1822. https://doi.org/10.1111/mec.15860

Schlichting, H. E. (1960). The role of waterfowl in the dispersal of algae. *Transactions of the American Microscopical Society*, *79*(2), 160–166. https://doi.org/10.2307/3224082

Sewell, M. M., Parks, C. R., & Chase, M. W. (1996). Intraspecific chloroplast DNA variation and biogeography of North American *liriodendron* L. (Magnoliaceae). *Evolution*, *50*(3), 1147–1154. https://doi.org/10.1111/j.1558-5646.1996.tb02355.x

Shainker-Connelly, S. A., Crowell, R. M., Stoeckel, S., Vis, M. L., & Krueger-Hadfield, S. A. (2024). Population genetics of the freshwater red alga *Batrachospermum gelatinosum* (Rhodophyta) I: Frequent intragametophytic selfing in a monoicous, haploid-diploid species. *Journal of Phycology*, *60*(6).

Sheath, R. G. (1984). The biology of freshwater red algae. In F. E. Round & D. J. Chapman (Eds.), *Progress in Phycological research* (Vol. *3*, pp. 89–156). Biopress Ltd.

Sheath, R. G., & Cole, K. M. (1992). Biogeography of stream macroalgae in North America. *Journal of Phycology*, *28*(4), 448–460. https://doi.org/10.1111/j.0022-3646.1992.00448.x

Sheath, R. G., & Hymes, B. J. (1980). A preliminary investigation of the freshwater red algae in streams of southern Ontario, Canada. *Canadian Journal of Botany*, *58*(11), 1295–1318. https://doi.org/10.1139/b80-161

Smouse, P. E., Banks, S. C., & Peakall, R. (2017). Converting quadratic entropy to diversity: Both animals and alleles are diverse, but some are more diverse than others. *PLoS ONE*, *12*(10), e0185499. https://doi.org/10.1371/journal.pone.0185499

Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, *15*(14), 4261–4293. https://doi.org/10.1111/j.1365-294X.2006.03061.x

Stoeckel, S., Arnaud-Haond, S., & Krueger-Hadfield, S. A. (2021). The combined effect of haplodiplonty and partial clonality on genotypic and genetic diversity in a finite mutation population. *Journal of Heredity*, *112*(1), 78–91. https://doi.org/10.1093/jhered/esaa062

Strange, R. M., & Burr, B. M. (1997). Intraspecific phylogeography of north American highland fishes: A test of the Pleistocene vicariance hypothesis. *Evolution*, *51*(3), 885–897. https://doi.org/10.1111/j.1558-5646.1997.tb03670.x

Suetsugu, K., Nozaki, T., Hirota, S. K., Funaki, S., Ito, K., Isagi, Y., Suyama, Y., & Kaneko, S. (2023). Phylogeographical evidence for historical long-distance dispersal in the flightless stick insect *Ramulus mikado*. *Proceedings of the Royal Society B:*

Biological Sciences, *290*, 20231708. https://doi.org/10.1098/rspb.2023.1708
Tibayrenc, M., & Ayala, F. J. (2002). The clonal theory of parasitic protozoa: 12 years on. *Trends in Parasitology*, *18*(9), 405–410. https://doi.org/10.1016/S1471-4922(02)02357-7
van der Strate, H. J., Van De Zande, L., Stam, W. T., & Olsen, J. L. (2002). The contribution of haploids, diploids and clones to fine-scale population structure in the seaweed *Cladophoropsis membranacea* (Chlorophyta). *Molecular Ecology*, *11*(3), 329–345. https://doi.org/10.1046/j.1365-294X.2002.01448.x
Velasquez, G. T. (1940). On the viability of algae obtained from the digestive tract of the gizzard shad, *Dorosoma cepedianum* (Le Sueur). *The American Midland Naturalist*, *22*(2), 376–412. https://doi.org/10.2307/2420316
Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer. https://doi.org/10.1007/978-0-387-21706-2
Vis, M. L. (2016). Biogeography of river algae. In O. Necchi, Jr. (Ed.), *River Algae* (pp. 219–243). Springer. https://doi.org/10.1007/978-3-319-31984-1_10
Vis, M. L., & Necchi, O., Jr. (2021). *Freshwater red algae: Phylogeny*. Springer. https://doi.org/10.1007/978-3-030-83970-3
Vis, M. L., & Sheath, R. G. (1997). Biogeography of *Batrachospermum gelatinosum* (Batrachospermales, Rhodophyta) in North America based on molecular and morphological data. *Journal of Phycology*, *33*(3), 520–526. https://doi.org/10.1111/j.0022-3646.1997.00520.x
Vis, M. L., Sheath, R. G., & Cole, K. M. (1996). Distribution and systematics of *Batrachospermum* (Batrachospermales, Rhodophyta) in North America. 8a. Section *Batrachospermum*: *Batrachospermum gelatinosum*. *European Journal of Phycology*, *31*(1), 31–40. https://doi.org/10.1080/09670269600651161
Voisin, M., Engel, C. R., & Viard, F. (2005). Differential shuffling of native genetic diversity across introduced regions in a brown alga: Aquaculture vs. maritime traffic effects. *Proceedings of the National Academy of Sciences*, *102*(15), 5432–5437. https://doi.org/10.1073/pnas.0501754102

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1.

Figure S1. Principal components analysis showing cumulative effects of environmental parameters for each drainage basin. Colors represent each drainage basin. PC1 explains 29.5% and PC2 explains 23.4% of the variation. Width is the stream width, depth is stream depth, cond is conductivity, temp is water temperature, canopy is % canopy cover and log_H.ions.nM is pH.

Figure S2. DAPC (discrete analysis of principal components) depicting relationships of multilocus genotypes (MLGs) of *Batrachospermum gelatinosum* from five drainage basins in the eastern United States (a, b) Individuals are color-coded corresponding to the five *a priori* groups based on drainage basin. (a) Using the *compoplot* function as implemented in *adegenet*, we determined the *a priori* assignment of five drainage basins. (b) Cross-validation using the *xvalDapc* function as implemented in *adegenet*. The optimal number of principal components to retain was 23. DAPC excluding two sites in the Northeast basin with low sample size ($n < 6$) (c–f). Symbols represent thalli and the symbols correspond to the five *a priori* groups (drainage basins). The first three principal components are shown: (c) PC1: 53.4% and PC2: 22.1% and (d) PC1: 53.4% and PC3: 17.5%. (e) Cross-validation using the *xvalDapc* function as implemented in *adegenet*. The optimal number of principal components to retain was 23. (f) Using the *compoplot* function as implemented in *adegenet*, we determined the *a priori* assignment of five drainage basins.

Figure S3. Discrete analysis of principal components (DAPC) depicting relationships of multilocus genotypes (MLGs) of *Batrachospermum gelatinosum* within drainage basins in the eastern United States. Ohio River Basin (a, b): (a) Using the *compoplot* function as implemented in *adegenet*, we determined the *a priori* assignment of five streams. (b) Cross-validation using the *xvalDapc* function as implemented in *adegenet*. The optimal number of principal components to retain was four. Northeast Basin (c, d): (c) Using the *compoplot* function as implemented in *adegenet*, we determined the *a priori* assignment of four streams. (d) Cross-validation using the *xvalDapc* function as implemented in *adegenet*. The optimal number of principal components to retain was five. Great Lakes Basin (e, f): (e) Using the *compoplot* function as implemented in *adegenet*, we determined the *a priori* assignment of six streams. (f) Cross-validation using the *xvalDapc* function as implemented in *adegenet*. The optimal number of principal components to retain was 20. South Atlantic-Gulf (g): Using the *compoplot* function as implemented in *adegenet*, we determined the *a priori* assignment of two streams.

Figure S4. Principal component analysis (PCA) representing the goodness of fit of posterior estimates compared to prior estimates (open circles) for the best fit DIYABC scenarios from the South Atlantic Gulf and Mid-Atlantic analysis in relationship to the observed datasets (yellow circle).

Table S1. Locations and physiochemical data measured at each site where *Batrachospermum gelatinosum* gametophytes were sampled. Site numbers as not taken in figure 1 and table 1. Dashes indicated that the measurement was not taken.

Table S2. Information for multiplex PCRs of 13 microsatellite loci developed for genotyping of Batrachospermum gelatinosum (Crowell et al., 2024). Simplex PCRs were only used for reruns and locus Bgel_056 (see methods). Primer concentrations are given for F* (labeled forward), F (unlabeled forward), and R (unlabeled reverse).

Table S3. Raw size ranges for each binned allele. All loci are shown along with the assigned dye and multiplex or simplex assignment. For each allele, the minimum and maximum raw allele sizes observed are shown, along with the binned call for the raw range. If there was only one raw size for an allele, it was placed between the raw_min and raw_max columns.

Table S4. Null allele frequencies for each locus.
Table S5. ANOVA or Kruskal-Wallis tables for each drainage basin and physiochemical measurements. Significant *p*-values are shown in bold.
Table S6. Genetic and geographic distances between 18 *Batrachospermum gelatinosum* sites. Geographic distance between sites in km (above) and genetic distance represented by F_{ST} (below). Site abbreviation as in s1.
Table S7. Parameter priors for all DIYABC scenarios simulating origins of *Batrachospermum gelatinosum* sites following the Pleistocene glaciation. N1-N5 are effective population sizes for the corresponding populations. t1–t3 are divergence times for sites. Mutation rates: μmic is the mean mutation rate; pmic is the mean distribution of the number of repeats of microsatellites; snimic is the mean rate of single nucleotide insertions/deletions.
Table S8. Logistic regression results for the scenarios estimating the origins of *Batrachospermum gelatinosum* sites following the Pleistocene glaciation based on DIYABC simulations. Results are presented as the mean value with 95% confidence interval in brackets for the inclusion of 50 summary statistics from different numbers of datasets (*n*) closest to those from the observed dataset.
Table S9. Parameter posterior estimates for the best supported DIYABC scenario that predicted a refugium for *Batrachospermum gelatinosum* in the Mid-Atlantic with dispersal to the South Atlantic Gulf followed by dispersal to the Ohio River and finally simultaneous dispersal to the Great Lakes and Northeast (Mid-Atlantic Scenario 2). The first number is the result from the South Atlantic Gulf and Mid-Atlantic analysis, the second number is the result for the South Atlantic Gulf, Mid-Atlantic and Great Lakes analysis.

How to cite this article: Crowell, R. M., Shainker-Connelly, S. J., Krueger-Hadfield, S. A., & Vis, M. L. (2024). Population genetics of the freshwater red alga *Batrachospermum gelatinosum* (Rhodophyta) II: Phylogeographic analyses reveal spatial genetic structure among and within five major drainage basins in eastern North America. *Journal of Phycology*, *60*, 1437–1455. https://doi.org/10.1111/jpy.13512