RESEARCH ARTICLE

Reproductive traits explain occupancy of predicted distributions in a genus of eastern North American understory herbs

Chelsea N. Miller[1,2] 🄳 | Monica Papeş[2] | Edward E. Schilling[2] | Charles Kwit[2,3]

[1]The Warnell School of Forestry and Natural Resources, University of Georgia Athens, GA, USA

[2]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA

[3]Department of Forestry, Wildlife and Fisheries, University of Tennessee, Knoxville, TN, USA

Correspondence
Chelsea N. Miller, The Warnell School of Forestry and Natural Resources, University of Georgia Athens, GA 30602, USA.
Email: Chelsea.Miller@uga.edu

Funding information
L. R. Hesler Herbarium Fund University of Tennessee Knoxville

Editor: Boris Leroy

Abstract

Aim: Abiotic, biotic and dispersal factors interact to shape species distributions. At broad geographic extents, abiotic factors are thought to exert the greatest influence on the distribution, while biotic and dispersal factors strongly influence the distribution regionally and locally. We test whether reproductive traits relating to biotic and dispersal factors explain differences between estimated potential and occupied geographic distributions for 21 species of *Trillium*.

Location: Eastern North America.

Methods: Fundamental niches and associated predicted suitable distributions were estimated using climate-calibrated ecological niche models. We defined occupied distributions as the intersection between known ranges and predicted suitable areas (PSAs). Proportional occupancy of the predicted suitable distribution (PO) was calculated by dividing the area of the occupied distribution by the PSA. We related reproductive traits (ovule number, seed set, number of seeds/plant, seed mass, adult biomass, flower type: sessile/pedicellate) to PO using beta regression models. AICc was used to assess model fit.

Results: There was considerable variation in PO across species (1.1%–96%, mean = 51%). Eighty-five percent of species with PO < 60% were sessile; 88% of species with PO > 60% were pedicellate. The best-fit beta regression (pseudo R^2 = .70) yielded significantly lower PO for sessile-flowered species; ovule number and seed mass were also significant predictors of PO.

Main conclusions: Variation in PO among study species can be explained by flower type, ovule number and seed mass—biotic traits related to dispersal ability. We posit that variation in dispersal potential stemming from primary and long-distance dispersers is related to occupancy of the predicted suitable distribution in *Trillium*. We exemplify a scenario in which life history traits explain why some species are range-restricted when their close relatives are widespread. Our methodology constitutes a powerful comparative framework that can be applied to diverse biological systems to inform conservation of rare species.

KEYWORDS
conservation, ecological niche model, endemism, habitat suitability, life history traits, seed dispersal, species distribution model

1 | INTRODUCTION

For more than a century, ecologists have attempted to understand the constraints on species' distributions (Dobzhansky, 1950; Grinnell, 1917; Hutchinson, 1957, 1978; Lomolino et al., 2005; MacArthur, 1984; Peterson, 2001). The concept of the ecological niche, which bears a long history (Chase & Leibold, 2003; Grinnell, 1917; Pocheville, 2015), is currently viewed as a multidimensional combination of environmental conditions that permit a species to survive and reproduce (Soberón & Arroyo-Peña, 2017) and is interwoven into explanations of patterns of biodiversity and species coexistence (Pocheville, 2015). This modern usage of the ecological niche is rooted in the niche concept as described by Grinnell (1917), which emphasized the influence of the environment on the physical distribution of biological populations and their evolution. Grinnell's niche concept encompassed abiotic factors such as temperature, precipitation and elevation, and biotic factors such as the presence of food, competitors or predators where the species in question existed (Grinnell, 1917; Pocheville, 2015). Hutchinson later advanced niche theory by introducing the concepts of the fundamental and realized niche, which respectively distinguish between the abiotic space that would permit a species to exist indefinitely in the absence of competitors, and the space actually occupied by a species (Hutchinson, 1957).

Today, biogeographers also recognize dispersal limitation as a critical factor constraining the distribution of species. Dispersal limitation, while not a formal component of the ecological niche, includes both extrinsic barriers preventing dispersal (e.g. fragmentation), as well as intrinsic dispersal limitations, which can be a function of traits that govern the ability of a species to reach areas that might otherwise be suitable (Soberón & Peterson, 2005). Dispersal limitation and differential migration have been postulated to influence current distribution and diversity patterns. This is particularly evident for forest plant species and communities in the northern hemisphere following glacial retreat (Jacquemyn et al., 2001; Svenning et al., 2008; Svenning & Skov, 2007; Verheyen & Hermy, 2001; Willner et al., 2009). To gain a better understanding of the constraints on a species' geographic distribution, the relative importance of abiotic, biotic and dispersal factors should be considered. One way to do this is by estimating the disparity between the fundamental and realized niche in geographic space. If a species is able to track its fundamental niche on the landscape, the primary constraints on the geographic distribution are presumed to be abiotic. Conversely, if the geographic distribution of a species' fundamental niche is larger than its occupied geographic distribution (i.e. its realized niche), it can be inferred that biotic factors, dispersal limitation or a combination of these affect the distribution of the species (Munguía et al., 2008; Peterson, 2006; Soberón & Peterson, 2005; Svenning & Skov, 2004).

An estimate of the fundamental niche can be obtained through ecological niche modelling (also referred to as species distribution modelling), which has been used extensively over the past two decades to relate species' distributions to abiotic factors (Araújo et al., 2019; Elith et al., 2006, 2011; Elith & Leathwick, 2009; Guisan & Zimmermann, 2000; Phillips et al., 2006). By combining known occurrences with a set of climate variables, ecological niche models

(ENMs) predict areas potentially suitable for the species in question. Much work has been conducted to determine whether ENMs estimate the realized or the fundamental niche; consensus suggests that ENM predictions are likely estimating an environmental space in between the two, depending on whether the species is close to or far from equilibrium with its environment (Araújo & Peterson, 2012; Guillera-Arroita et al., 2015; Peterson et al., 2011; Soberón & Nakamura, 2009). Nevertheless, the areas predicted to be suitable by climate-calibrated ENMs only take into consideration the impacts of abiotic factors, thus the expectation is that the model is more often closer to an estimate of the Hutchinsonian fundamental niche than the realized niche.

To that end, although most ENMs do not explicitly consider the effects of biotic factors and dispersal limitation on species distributions (Beale et al., 2014; Dormann et al., 2012; Elith & Leathwick, 2009; Leach et al., 2016), they can be used to address potential constraints on geographic distributions. Previous studies have used ENMs to assess the relationship between the fundamental and realized niche in a variety of ways. For example, Strubbe et al. (2013) and Villaverde et al. (2017) assessed niche conservatism in non-native bird species and niche shifts in bipolar sedges after long-distance dispersal events by estimating the overlap, equivalency, similarity, expansion and unfilling of climatic niches and potential distributions. Niche expansion in these studies was defined as a species moving into a new environment, whereas niche unfilling was defined as a species only partially filling its niche in the invaded range. Tingley et al. (2016) and Zhu et al. (2017) explored intraspecific variation in realized niche expansion and unfilling in an invasive skink and an invasive stink bug, respectively; their methods incorporated knowledge about native-range source populations and global introduction history into niche modelling approaches to explore the effects of intraspecific niche variation and different invaded-range environments on niche lability. Fewer studies have attempted to connect species traits that directly or indirectly relate to dispersal ability and/or species interactions to niche filling. For instance, Park et al. (2018) related mating system (self-pollinating versus outcrossing) in flowering plants to niche breadth using ENM and mixed-effects models, and determined that niche breadth was not greater for self-pollinating plant species than for their outcrossing relatives, despite larger geographic range sizes.

In this paper, we ask whether reproductive life history traits of herbaceous understory plants in the genus *Trillium* can explain differences in the magnitude of the mismatch between the predicted suitable distribution (geographic estimation of the fundamental niche) and the occupied distribution (geographic estimation of the realized niche). Reproductive life history traits of herbaceous understory plants are often associated with dispersal ability and biotic factors, such as competition, herbivory, mutualism or parasitism. Relating these to occupancy of suitable distributions among closely related species will elucidate whether geographic distributions are constrained primarily by abiotic factors, or whether biotic factors and dispersal limitations constrain the distribution.

Species of *Trillium* in eastern North America (ENA)—a *Trillium* biodiversity hotspot (Case & Case, 1997; Ohara, 1989)—can be divided

into two floral morphological types based on the presence ("pedicellate") or absence ("sessile") of a pedicel. "Sessile" refers to the attachment of the flower directly by its base, whereas "pedicellate" refers to taller and more noticeable flowers and fruits borne on pedicels. *Trillium* flower type is representative of many other biotic and dispersal factors that might impact extent of occupied distribution, relative to predicted suitable distribution. For example, although all trilliums are myrmecochorous (i.e., their seeds are dispersed short distances by ants), some of the seeds of pedicellate-flowered species are also known to be dispersed intermediate distances by yellow jackets ([*Vespula* spp.]; Jules, 1996; Zettler et al., 2001) and much longer distances via frugivory by white-tailed deer ([*Odocoileus virginianus*]; Vellend et al., 2003; Myers et al., 2004; Griffin & Barrett, 2004a, Griffin & Barrett, 2004b). Importantly, long-distance dispersal (LDD) by deer has not been recorded for any members of the sessile-flowered group. Several mechanisms might explain this phenomenon. Leaf mottling is present in most sessile-flowered species, which may camouflage them in the understory and reduce the probability of browsing by deer (Givnish, 1990). It is also possible that the greater plant height and biomass of many pedicellate-flowered species—in combination with the greater flower and fruit height conferred by the pedicel—might promote browsing and frugivory, and thus LDD, by deer by making pedicellate-flowered species more conspicuous in the understory. Because sessile flower positioning is a synapomorphy for the subgenus *Sessilium* (Farmer & Schilling, 2002), other reproductive traits that potentially impact the likelihood or efficacy of seed dispersal—such as ovule number, seed mass, seed setting rate and adult biomass—may also be linked to this phylogenetic distinction.

Given the notable differences in probability of LDD, other reproductive life history traits, and biotic interactions among pedicellate- and sessile-flowered trilliums, we hypothesize that the mismatch between the predicted suitable distributions and occupied distributions of 21 native *Trillium* species in ENA can be explained by flower type in combination with other relevant reproductive traits. To test this hypothesis, we first estimate proportional occupancies of predicted suitable distributions (also referred to as "range filling" and "niche filling" [Estrada et al., 2015; Fordham et al., 2012]) with ENMs. We then use a model-building framework to assess what set of reproductive life history traits (flower type, ovule number, seed setting rate, number of seeds per plant, seed weight and adult biomass) best predict proportional occupancy of predicted suitable distributions. Our study represents a comprehensive investigation of geographic distributions of eastern North American *Trillium* species as a function of traits related to reproduction and seed dispersal.

2 | MATERIALS AND METHODS

2.1 | Study system

Our study system is comprised of plants in the genus *Trillium*: perennial monocot rhizomatous herbs found in Northern Hemisphere temperate deciduous forests of eastern Asia and eastern and western North America (Freeman, 1975). Species in this genus are either pedicellate-flowered or sessile-flowered. Sessile-flowered trilliums are phylogenetically distinct, forming a monophyletic clade: subgenus *Sessilium* (formerly *Phyllantherum*; Farmer & Schilling, 2002; Case, 2002a, 2002b). Within the pedicellate species, subgenus *Delostylium* is monophyletic; all other pedicellate species form an informal group (subgenus *Trillium*; Farmer & Schilling, 2002; Millam, 2006). For the purposes of this study, flower type (sessile/pedicellate) will be used as a proxy for comparison of two groups: (i) the clade *Sessilium, and* (ii) an informal grouping of non-sessile species including the clade *Delostylium* and all other pedicellate species with currently unresolved taxonomy. Species in these two groups are characterized by reproductive differences such as the number of seeds produced per plant, seed setting rates, ovule number and seed weight (Ohara, 1989; Figure A1 in Appendix S1), as well as distinct floral scents and leaf mottling (Weakley, 2015). Pedicellate species can be further distinguished based on erect or declinate flower positioning. There are four pedicellate-declinate species which differ from pedicellate-erect species in having lower seed setting rates (Ohara, 1989).

Reproductive trilliums in ENA produce flowers from March to June, with flowering lasting 2–4 weeks, followed by the production of a single ovary containing seeds with elaiosomes (seed coat-borne appendages rich in lipids and other nutrients; Miller et al., 2020). Ripe ovaries dehisce and drop mature diaspores (the dispersal unit of the plant; the seed-elaiosome complex) in mid- to late summer. The seeds of *Trillium* are dispersed by ants (i.e. myrmecochory) primarily in the *Aphaenogaster fulva-rudis-texana* species group (DeMarco & Cognato, 2016; Ness et al., 2009; Umphrey, 1996). However, other seed dispersers have been noted for some species of *Trillium*. Yellow jackets (*Vespula* spp.) have been observed dispersing the seeds of both pedicellate and sessile species (Bale et al., 2003; Gonzales & Hamrick, 2005; Jules, 1996; Zettler et al., 2001), and white-tailed deer (*Odocoileus virginianus*) have been noted dispersing the seeds of two of the most ubiquitous pedicellate species in ENA, *T. grandiflorum* and *T. erectum* (Griffin & Barrett, 2004a; Myers et al., 2004; Vellend et al., 2003). This long-distance dispersal mechanism has not, to our knowledge, been recorded or observed for any sessile species.

At least 33 species of *Trillium* are native to ENA (NatureServe, 2020; Weakley, 2015). Fifteen of these are ranked as high conservation priority (e.g. G1, G2 or G3; NatureServe, 2020). Both the sessile and pedicellate groups include species that are characterized by narrow endemism (e.g. they inhabit restricted geographic regions or specific habitat types), and many of these endemic species co-occur locally with geographically widespread congeners. We included 21 of the 33 native ENA species of *Trillium* in our study, based on taxonomic certainty and number (>20), reliability and type of records. We excluded cryptic species and those whose taxonomy is poorly resolved, such as *T. lancifolium*, *T. simile* and the *T. pusillum* complex (Cabe, 1995; Cabe & Werth, 1995; Farmer, 2007; Osaloo et al., 1999; Timmerman-Erskine et al., 2002). Species with protected occurrence

records, such as *T. rugelii*, were not included. Species with too few occurrence data points included *T. delicatum*, *T. georgianum*, *T. gracile*, *T. tennesseense*, *T. oostingii*, *T. persistens* and *T. reliquum*; several of these are extremely range-restricted endemic species. Finally, we excluded *T. viride* due to concerns about occurrence data quality. Of the 21 species in our study (Table 1), nine are pedicellate (5 pedicellate-erect; 4 pedicellate-declinate) and 12 are sessile.

2.2 | Occurrence data

We sought to obtain every publicly available presence record for each *Trillium* species in ENA, these records dated back to 1900. The databases we searched included the Global Biodiversity Information Facility (GBIF; https://www.gbif.org/, accessed July 28—August 1, 2018), the SouthEast Regional Network of Expertise and Collections (SERNEC; http://sernecportal.org/portal/, accessed 1 August 2018, 26 August 2018, 30 August 2018 and 1 September 2018–4 September 2018; 12 July 2019), Tropicos (https://www.tropicos.org/, 28 March 2019), and online regional herbaria databases, such as the University of Tennessee Herbarium (TENN; UT Herbarium, 2020; https://herbarium.utk.edu/, accessed 10 August 2018–15 August 2018) and the Arnold Arboretum of Harvard University (https://www.arboretum.harvard.edu/, accessed 17 August 2018). Further details about the origin, sampling methods and sampling dates of

occurrence records are included in the Appendix S3. Approximately half of all records we obtained consisted of descriptive localities without latitude/longitude coordinates. To assign geographic coordinates to these localities, we used the GEOLocate software (Rios & Bart, 2010; https://www.geo-locate.org/, accessed from August 2018 to September 2019). A centroid of uncertainty with an area of 3 km^2 was automatically assigned to each locality by GEOLocate. Minimum uncertainty was adjusted manually based on specificity of record descriptions. Descriptive localities were georeferenced by one of three researchers, and all final coordinates and uncertainties were checked and confirmed by the first author.

2.3 | Study extent

We defined the study extent as ENA north of Mexico, including all U.S. states and Canadian provinces east of the western extent of the Mississippi River (–92.9 degrees longitude), because all 21 study species have known ranges that occur within the bounds of this extent and do not occur beyond this extent. We chose this extent because the Mississippi River represents the intuitive geographic boundary between eastern and western North America, and also because the Great Plains region, which begins west of the Mississippi River, is not part of the geographic range of *Trillium*. We used 1 km^2 spatial resolution (0.0083 decimal degrees; 30 arc-seconds) to match that of the

TABLE 1 List of *Trillium* species included in the study, subgenus, flower type and reproductive life history traits. Reproductive traits (biomass (g), ovule number, number of seeds per plant, seed setting rate (%) and seed mass (mg)) obtained from Ohara (1989)

Species	Subgenus	Flower type	Biomass (g)	Ovule No.	No. seeds/ plant	Seed setting rate (%)	Seed mass (mg)
T. catesbaei	*Delostylium*	Pedicellate-declinate	2.68	53.2	16.1	30.26	3.1
T. cernuum	*Trillium*	Pedicellate-declinate	NA	NA	NA	NA	NA
T. cuneatum	*Sessilium*	Sessile	6.85	160.3	47.7	29.76	4.02
T. decipiens	*Sessilium*	Sessile	3.95	94.6	47.9	50.63	8.86
T. decumbens	*Sessilium*	Sessile	2.56	49	19.5	39.8	6.83
T. discolor	*Sessilium*	Sessile	2.91	44	21	47.73	3.62
T. erectum	*Trillium*	Pedicellate-erect	6.95	105.1	80.3	76.4	5.04
T. flexipes	*Trillium*	Pedicellate-declinate	6.74	128.7	43.9	34.11	4.2
T. foetidissimum	*Sessilium*	Sessile	1.86	57	21	36.84	16
T. grandiflorum	*Trillium*	Pedicellate-erect	4.47	38.4	26	67.71	6.42
T. ludovicianum	*Sessilium*	Sessile	4.07	57.3	25.3	44.15	8
T. luteum	*Sessilium*	Sessile	5.77	47.3	32.2	68.08	3.58
T. maculatum	*Sessilium*	Sessile	3.15	155	24.2	15.61	8.17
T. nivale	*Trillium*	Pedicellate-erect	0.52	27.3	14.3	64.1	2.7
T. recurvatum	*Sessilium*	Sessile	2.67	100.4	7.8	7.77	10.5
T. sessile	*Sessilium*	Sessile	2.65	124.9	33.1	26.5	7.8
T. stamineum	*Sessilium*	Sessile	2.31	48.7	14.3	29.36	12.17
T. sulcatum	*Trillium*	Pedicellate-erect	NA	NA	71	NA	NA
T. underwoodii	*Sessilium*	Sessile	2.47	75.8	20.7	27.31	10.28
T. undulatum	*Trillium*	Pedicellate-erect	3.41	34.2	29.2	85.38	4.15
T. vaseyi	*Trillium*	Pedicellate-declinate	9.5	41.2	18.5	44.9	3

climate variable layers (see *Climate variables* below). Occurrence records for each species were spatially rarified to include one unique record per 1 km² using the spatial rarefication tool in SDMtoolbox 2.0 (Brown, 2014; Brown et al., 2017) in ESRI ArcMap (Version 10.7, Redlands, CA). Finally, we plotted all points (those that we georeferenced and those that had coordinates from databases) for each species on a map of the spatial extent, excluding any points falling outside the verified *Trillium* distributions generated by the Biota of North America Program (BONAP; Kartesz & BONAP [Biota of North America Program], 2015), and ensured the final set of occurrence records conformed to the known range of each species according to *Trillium* expert A. Floden (Missouri Botanical Garden). This resulted in a total of 10,068 records across the 21 species of *Trillium*.

2.4 | Extent of occurrence

We used the extent of occurrence (EOO) to delineate the known range of each species. EOO is defined as the area contained within the shortest continuous boundary which can be drawn to encompass all known, inferred or projected sites of occurrence of a taxon (IUCN, 2012). EOO was calculated by applying a convex hull technique in the package dismo in R (Hijmans et al., 2017) to all spatially rarified occurrence points. EOO for each species was confirmed by consulting maps of endemism generated by BONAP (Kartesz & BONAP [Biota of North America Program], 2015) and IUCN, and through personal communications with *Trillium* experts (A. Floden, Missouri Botanical Garden; T. Patrick, Georgia Department of Natural Resources).

2.5 | Climate variables

Climate variables were generated by Wang et al. (2016) at 1 km² resolution (ClimateNA v5.10 software package; available at http://tinyurl.com/ClimateNA), accessed September 2017. These data include 27 monthly, seasonal and annual climate variables calibrated at the extent of North America from 1961 to 1990 (Table A2 in Appendix S2). We estimate that ~75% of the 10,068 occurrence records originated within the past 100 years (see *Occurrence data* above and Appendix S3). Of this 75%, we estimate that ~75% of these originated within 1961–1990; ~12.5% originated prior to 1961; and ~12.5% after 1990. This yielded a final estimate of 56% (75% * 75%) of the 10,068 occurrences (~5,638 records) originating within the timeframe of the climate variables. Environmental variables used in our models were clipped to the extent of ENA (North America north of Mexico, bounded on the west by the Mississippi river) using ESRI ArcMap (v.10.7). Maps of the study extent and species occurrences are provided in Figure 1 and Appendix S4. We did not attempt to minimize correlation among the predictor variables prior to modelling, because removing correlated variables does not significantly alter model fit statistics when implementing machine learning algorithms, as these algorithms have built-in procedures to minimize multidimensionality and collinearity of predictor variables (Feng et al., 2019; Tanner et al., 2017). Instead, we relied on the modelling algorithm to rank the variables that contributed to model accuracy gain (Table 2).

2.6 | Training and testing of ecological niche models

We used the Maxent 3.4.1 program (Phillips et al., 2006) for modelling species ecological niches and projecting them as potential geographic distributions. Details about Maxent and of our methodology are provided in the Appendix S3. In brief, we ran two iterations of each ENM. The first iteration included all 27 climate variables and used the EOO to define the extent for model training and selection of background samples. From the initial model output, we obtained the environmental variables that contributed to the top 95% of gain in model fit; these were retained for use in a second and final iteration of each model (median number of variables used = 10 [*T. sulcatum*, *T. undulatum*]; minimum = 4 [*T. decumbens*, *T. discolor*]; maximum = 19 [*T. sessile*]; Table 2 and Appendix S3). From the second Maxent iteration, we used the most dissimilar variable (MoD) output of the multivariate environmental similarity surface analysis (MESS) to find the geographic extent in which none of the climatic variables were outside the range of values of the training region; in this way, we limited the extrapolation in the projections of models to ENA extent (for further details, see Appendix S4). The predicted suitable area (PSA; the geographic representation of estimated fundamental niche) obtained for each species with the second Maxent iteration was analysed within the geographic extent that limited model extrapolation. We modified certain model attributes (e.g. testing-training split of occurrences, number of iterations and regularization multipliers) to maximize model performance in the second Maxent iteration (see Appendix S4).

We evaluated final model performance with three metrics: (i) the area under the curve (AUC) of the receiver operating characteristic (ROC), which reflects the ability of the model to correctly predict presences relative to the proportion of the area predicted present (Phillips et al., 2006); (ii) omission error, or the false-negative rate that calculates the proportion of presences incorrectly predicted as absences (Fielding & Bell, 1997); and (iii) the Boyce Index, an evaluation method for presence-only models that measures how much model predictions differ from a random distribution of the known presences across prediction gradients (Boyce et al., 2002; Hirzel et al., 2006).

AUC is a widely used metric for model performance, although it is known to have shortcomings. For instance, it is highly dependent on the prevalence of the species, which is generally not known (Smith, 2013). The inclusion of omission error as a second measure of model performance mitigates the risk of relying solely on AUC; because AUC is threshold-independent, and omission error relies on a user-chosen threshold (in our case, 0.1 training omission error), this metric provides a more balanced assessment of model performance. We used the Maxent logistic value corresponding to
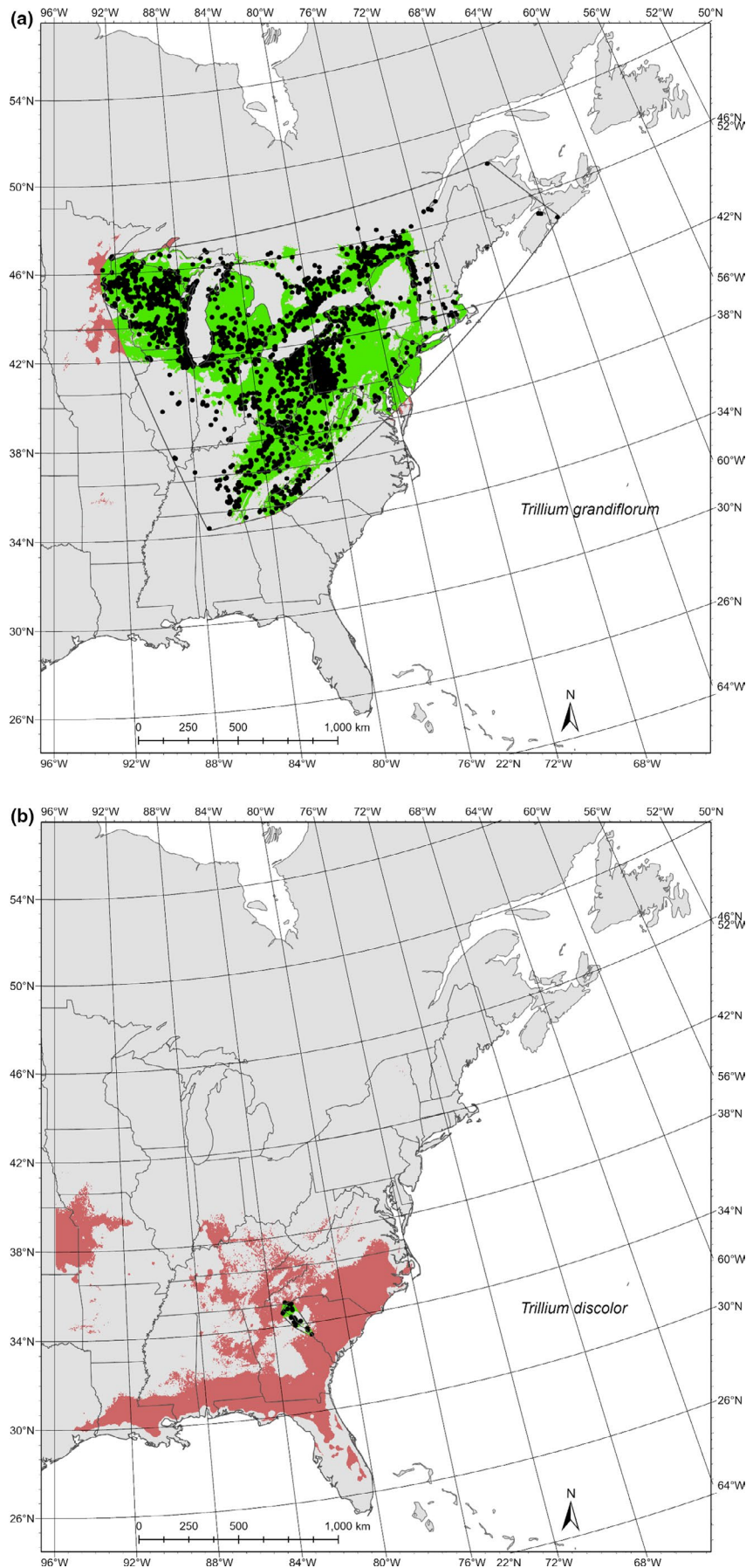
FIGURE 1 Maps illustrating the potentially suitable area (PSA; i.e. an estimate of the fundamental niche) in pink, produced by Maxent algorithm for (a) *Trillium grandiflorum* (pedicellate-erect) and (b) *T. discolor* (sessile) in eastern North America. The area in grey represents the study extent (ENA north of Mexico bounded by the western extent of the Mississippi river). Training and testing occurrences are depicted as black dots. Polygons depicting the extent of occurrence (EOO) are overlaid. Areas in green represent the intersection between the PSA and the EOO (i.e. our estimate of the realized niche). The area of the realized niche divided by the area of the fundamental niche yields the proportional occupancy (PO) of the fundamental niche. For *T. grandiflorum*, PO = 96%; for *T. discolor*, PO = 1.1%

TABLE 2 Percent contribution of climatic variables[a] to ecological niche models for each species (ordered alphabetically)

Species (number of variables)	ahm	bffp	cmd	dd_0	dd_18	dd5	dd18	effp	emt	eref	ext	Ffp	map
T. catesbaei (8)	0.044	0	0.318	0.102	0	0	0.045	0	0	0	0.075	0	0
T. cernuum (12)	0	0.042	0.048	0	0	0	0	0	0.006	0	0.016	0	0.005
T. cuneatum (15)	0.043	0.009	0.358	0	0	0	0	0.020	0.024	0	0.022	0.010	0.019
T. decipiens (7)	0	0.099	0.026	0	0	0	0	0	0	0.014	0	0	0
T. decumbens (4)	0	0	0	0	0	0	0	0	0	0	0	0	0.266
T. discolor (4)	0	0	0.016	0	0	0	0	0	0	0	0	0	0
T. erectum (9)	0	0	0.034	0	0	0	0	0.026	0	0.052	0.243	0	0.019
T. flexipes (11)	0.088	0	0	0	0	0	0.300	0.007	0	0.026	0.088	0	0
T. foetidissimum (12)	0	0.313	0.118	0	0	0	0	0	0	0.033	0	0	0
T. grandiflorum (13)	0.083	0	0.024	0	0	0	0.450	0	0	0	0.196	0	0.038
T. ludocivianum (6)	0	0	0	0	0	0	0	0	0	0	0.170	0	0
T. luteum (8)	0.035	0	0	0.192	0	0	0	0	0	0	0	0	0
T. maculatum (13)	0.008	0	0.011	0	0.008	0.074	0.021	0.030	0	0	0.039	0	0.094
T. nivale (14)	0	0	0.008	0.238	0	0	0	0	0.003	0.032	0.047	0	0.089
T. recurvatum (14)	0	0	0.010	0.005	0	0	0.240	0.005	0	0.027	0.021	0	0
T. sessile (19)	0.008	0	0.004	0.236	0	0	0.114	0.027	0.009	0.077	0.109	0	0.009
T. stamineum (11)	0.085	0	0.131	0	0	0.225	0	0	0	0	0.129	0	0
T. sulcatum (10)	0	0.020	0.180	0	0	0	0	0.102	0.030	0	0.071	0.206	0
T. underwoodii (6)	0.015	0	0.201	0	0	0	0	0	0	0	0	0	0.048
T. undulatum (10)	0.021	0	0.100	0	0	0	0	0	0.328	0.063	0.100	0	0.137
T. vaseyi (6)	0	0	0.609	0	0	0	0	0	0	0	0	0	0

Species (number of variables)	mar	mat	mcmt	msp	mwmt	nffd	pas	ppt_sm	ppt_wt	rh	shm	tave_sm	tave_wt	td
T. catesbaei (8)	0	0	0	0	0	0	0.125	0.064	0.229	0	0	0	0	0
T. cernuum (12)	0	0.425	0	0	0.019	0	0.019	0.005	0.300	0	0.106	0	0	0.008
T. cuneatum (15)	0.034	0	0.244	0.012	0.044	0	0.035	0	0.014	0	0	0	0	0.113
T. decipiens (7)	0	0	0	0	0.147	0.604	0	0.056	0	0.054	0	0	0	0
T. decumbens (4)	0	0	0	0	0.437	0	0.116	0	0	0	0	0	0	0.274
T. discolor (4)	0.317	0	0	0	0	0	0	0.647	0	0.020	0	0	0	0
T. erectum (9)	0	0	0.292	0	0	0	0	0.126	0.177	0	0	0	0	0.032
T. flexipes (11)	0.012	0	0	0	0	0.116	0	0.004	0.063	0.133	0	0	0	0.164
T. foetidissimum (12)	0	0.002	0	0	0.029	0.347	0.033	0.014	0.017	0	0.004	0.024	0	0.065
T. grandiflorum (13)	0.015	0	0	0.013	0.066	0	0.033	0	0.009	0.024	0.027	0	0	0.023
T. ludocivianum (6)	0.027	0	0	0	0.484	0	0	0.210	0.046	0	0	0	0	0.063
T. luteum (8)	0	0	0	0	0	0.312	0.072	0.095	0	0.101	0.108	0	0	0.084
T. maculatum (13)	0.323	0	0	0.203	0.061	0	0	0	0	0	0	0.020	0	0.108
T. nivale (14)	0.048	0	0	0	0.001	0.238	0.058	0	0.047	0.032	0.132	0	0	0.027
T. recurvatum (14)	0.016	0	0.044	0	0.016	0	0.275	0.048	0.047	0.208	0.038	0	0	0
T. sessile (19)	0	0	0.026	0.006	0	0	0.005	0.020	0.007	0.011	0.042	0.039	0.032	0.218
T. stamineum (11)	0.054	0	0	0.024	0	0	0	0.111	0.008	0.089	0.038	0.106	0	0
T. sulcatum (10)	0	0	0.035	0.298	0	0	0.033	0	0	0.026	0	0	0	0
T. underwoodii (6)	0.220	0	0	0	0	0.009	0	0	0	0	0.508	0	0	0

TABLE 2 (Continued)

Species (number of variables)	mar	mat	mcmt	msp	mwmt	nffd	pas	ppt_sm	ppt_wt	rh	shm	tave_sm	tave_wt	td
T. undulatum (10)	0	0	0.008	0	0.143	0	0	0.099	0	0.001	0	0	0	0
T. vaseyi (6)	0	0	0.002	0.109	0	0	0	0.092	0	0	0.159	0	0	0.028

Note: The total number of variables used in the final model for each species is included in parentheses after the species name in column 1. Number of variables ranges from 4 to 19, with a median value of 10.

[a]Monthly variables: Tmin (minimum temperature for a given month, °C), Tmax (maximum temperature for a given month, °C), Tave (mean temperature for a given month, °C), Ppt (total precipitation for a given month, mm). Bioclimatic variables: MAT (mean annual temperature, °C), MWMT (mean temperature of the warmest month, °C), MCMT (mean temperature of the coldest month, °C), TD (difference between MCMT and MWMT, as a measure of continentality, °C), MAP (mean annual precipitation, mm), MSP (mean summer [May to Sep] precipitation (mm), AHM (annual heat moisture index, calculated as [MAT+10]/[MAP/1000]), SHM (summer heat moisture index, calculated as MWMT/[MSP/1000], DD_0 (degree days below 0°C, chilling degree days), DD5 (degree days above 5°C [growing degree days]), DD_18 (degree days below 18°C), DD18 (degree days above 18°C), NFFD (the number of frost-free days), bFFP (the julian date on which the frost-free period begins), eFFP (the julian date on which the frost-free period ends), FFP (frost-free period), PAS (precipitation as snow [mm]), EMT (extreme minimum temperature over 30 years), EXT (extreme maximum temperature over 30 years), Eref (Hargreave's reference evaporation), CMD (Hargreave's climatic moisture index), MAR (mean annual solar radiation [MJ/m² d-1] [excludes areas south of US]), RH (mean annual relative humidity, %), Tave_wt (winter [Dec to Feb] mean temperature, °C), Tave_sm (summer [Jun to Aug] mean temperature, °C), PPT_wt (winter [Dec to Feb] precipitation, mm) and PPT_sm (summer [Jun to Aug] precipitation, mm) (AdaptWest Project, 2015; Andreas Hamann's website, 2013).

0.1 omission error of training data as a threshold to convert the model suitability values to binary form (suitable above threshold, unsuitable below threshold) and calculate omission error of the testing data points. Models with ≤0.3 testing omission error were considered reliable.

Finally, the use of the Boyce Index, which is threshold-independent like AUC, further alleviates concerns regarding the reliance on AUC to evaluate models. The Boyce Index partitions the predicted distribution into a series of suitability classes; by calculating both the predicted and expected frequencies of evaluation points, a ratio (predicted: expected, F_i) is generated. Areas of low predicted suitability will have F_i <1, whereas areas of high suitability will have F_i >1. Good models display a monotonically increasing predicted: expected curve, where F_i values increase with increasing predicted suitability. The Boyce Index measures this monotonic increase using the Spearman rank correlation coefficient, and thus, the quality of the model is indicated by the value of the Spearman coefficient. Positive values indicate models with predictions consistent with the distribution of presences in the test dataset, values close to zero indicate models not different from a random model, and negative values indicate models which predict low suitability where presences are actually more frequent (i.e. incorrect models) (Hirzel et al., 2006). We used the package "ecospat" in R to calculate Boyce Index values (Broennimann et al., 2021).

2.7 | Determining the relationship between predicted suitable area and occupied distribution

To calculate the proportional occupancy of the predicted suitable area (corresponding to estimated fundamental niche) for each species, we projected all rasters to Albers equal-area map projection, then overlaid the EOO and the PSA rasters in ArcMap (v.10.7) and used the Raster Calculator feature to identify the intersection. We

then divided the area of the intersection (km²) by the total area of the PSA to yield the proportion of occupancy of the predicted suitable area for each species, PO (Equation 1).

$$PO = \frac{[EOO \cap PSA]}{PSA} \tag{1}$$

2.8 | Incorporating reproductive life history traits

We used beta regression to relate reproductive life history traits to PO. Beta regression, a technique for modelling data limited to the open interval (0,1) (Douma & Weedon, 2019), is ideal for modelling PO, which is a proportion derived from continuous numbers (e.g. square kilometres). Alternative methods for modelling data bounded between 0 and 1, such as generalized linear models with a binomial error structure or logistic regression, are best used for proportions consisting of two categories derived from discrete counts. Beta regression consists of the same components as generalized linear mixed models (the random component, the systematic component and the link function), and we employed the mean-precision parametrization, which consists of μ as the expected value and Φ as a measure of precision (i.e. the inverse of dispersion). The use of these two parameters lends the beta distribution a great deal of flexibility in terms of the range of shapes the distribution can take (Douma & Weedon, 2019).

We ran beta regression models with several life history predictor variables and used Akaike's Information Criterion (AICc; calculated using AICcmodavg [Mazerolle, 2019]) to determine best-fit models. The following continuous reproductive life history traits were considered as predictors: biomass (e.g. fruiting plants were separated into component organs and dried in an oven for 48 hr [Ohara, 1989]), number of ovules per flower, seed setting rate (%), number of seeds per plant and seed weight (mg). The nominal predictor "flower type" (three categories: pedicellate-erect,

pedicellate-declinate and sessile) was also considered. Values of all reproductive life history predictor variables were obtained from Ohara (1989; Table 1). Because the Ohara (1989) dataset only contained information for 19 of our 21 study species, excluding *T. cernuum* and *T. sulcatum*, we did not include either of these species in beta regression models. Prior to implementing beta regressions, we evaluated the complete reproductive life history dataset for multicollinearity using the mctest package in R (Imdad & Aslam, 2018; Imdad et al., 2019; Imdadullah et al., 2016) and the variance inflation factor (VIF) of each predictor using the car package in R (Fox & Weisberg, 2011). Details of multicollinearity and VIF tests are provided in Appendix S3. Because the tests did

not detect collinearity among biomass, ovule number, seed weight or flower type (Table A1 in Appendix S2), we modelled PO using beta regressions with these four predictors individually and in combination.

We created multi-factor mean beta regression models by sequentially adding the best-fit factors and comparing model fit with single-factor models, and we generated mean models that considered interaction effects by sequentially adding best-fit factors as pairwise interactions. We also assessed the effects of different link functions (probit, loglog and cloglog, as compared to the default logit) on model fit for the best combination of factors, and explored the effects of specifying factors for the precision

TABLE 3 Beta regression model formulas, k (number of estimated parameters including intercept and Φ, the coefficient for the precision model), AICc scores, changes in AICc scores in relation to the lowest AICc score (ΔAICc), loglikelihood and pseudo R^2 values associated with 30 candidate models testing the effects of reproductive life history traits ("FT"—flower type [three categories], "ovule"—ovule number, biomass and seed mass) on the proportional occupancy of the fundamental niche for 21 species of *Trillium*

Model	k	AICc	ΔAICc	LogLikelihood	Pseudo R^2
FT + ovule + seedmass	**6**	**−4.83**	**0.00**	**11.92**	**.70**
FT + ovule + seedmass, link =probit	**6**	**−4.81**	**0.02**	**11.90**	**.71**
FT + ovule + seedmass, link =loglog	**6**	**−4.18**	**0.65**	**11.59**	**.69**
FT + ovule + seedmass, link =cloglog	**6**	**−4.11**	**0.72**	**11.55**	**.65**
FT + ovule	**5**	**−1.89**	**2.94**	**8.25**	**.55**
FT + biomass + ovule + seedmass	7	0.18	5.01	12.00	.70
FT	4	3.28	8.11	3.77	.31
FT +seedmass	5	3.28	8.11	5.67	.44
1	2	4.57	9.40	0.09	NA
biomass * seedmass +FT + ovule	8	5.49	10.32	12.46	.72
ovule * seedmass +FT + biomass	8	5.78	10.61	12.31	.71
FT +biomass + seedmass	6	5.81	10.64	6.60	.51
ovule	3	5.85	10.68	0.88	.08
biomass * ovule +FT + seedmass	8	6.36	11.19	12.02	.70
biomass	3	6.57	11.40	0.51	.04
FT +biomass	5	6.62	11.45	4.00	.33
FT * ovule	7	7.14	11.97	8.52	.56
seedmass	3	7.41	12.24	0.10	.00
FT * seedmass	7	8.32	13.15	7.90	.56
biomass +ovule	4	8.69	13.52	1.08	.10
ovule +seedmass	4	9.10	13.93	0.88	.08
biomass +seedmass	4	9.47	14.30	0.70	.06
biomass * seedmass	5	10.16	14.99	2.23	.19
biomass * ovule	5	12.05	16.88	1.28	.11
ovule * seedmass	5	12.24	17.07	1.19	.10
FT * seedmass +biomass + ovule	9	12.26	17.09	12.87	.73
biomass +ovule + seedmass	5	12.31	17.14	1.15	.10
FT * biomass +ovule + seedmass	9	13.03	17.86	12.48	.74
FT * ovule +biomass + seedmass	9	13.34	18.17	12.33	.72
FT * biomass	7	15.21	20.04	4.48	.37

The table is sorted by ΔAICc, which places the most likely models at the top. Bold indicates models for which change in AICc score <5. All models without specified link functions employed the default logit link.

model. To assess overall model fit, we noted loglikelihood values (generated via maximum likelihood estimation) and calculated AICc scores. We considered a total of 30 models (Table 3). To evaluate significant differences among categories of the nominal factor "flower type" in the best-fit model, we performed a pairwise contrasts post hoc test in the R package multcomp (Hothorn et al., 2008).

3 | RESULTS

3.1 | Ecological niche models

Overall, ENMs had reliable performance metrics. Testing AUC was ≥0.7 for 17 species and ranged from 0.62 to 0.69 for 4 species. Testing omission error was low for most ENMs, ranging from zero (*T. discolor*) to 0.31 (*T. ludovicianum*), with an average of 0.17 (Table 4). Boyce Index Spearman rank correlation coefficients were all positive and close to 1 (Table 4), indicating that our models generated predictions consistent with the distribution of presences in the test datasets.

The climate variable most frequently used was CMD (Hargreave's climatic moisture index), in the models of 17 species. The variables that were used in at least half (11) of the species' models and contributed on average ≥5% were RH (mean annual relative humidity, %), EXT (extreme maximum temperature over 30 years, °C), MWMT (mean temperature of the warmest month, °C), PAS (precipitation as snow, mm), PPT_sm (summer [Jun to Aug] precipitation, mm), PPT_wt (winter [Dec to Feb] precipitation, mm) and TD (difference between mean temperature of the coldest month and mean temperature of the warmest month [measure of continentality], °C) (Table 2).

The intersection between the PSA (binary maps of the potentially suitable area, based on modelled fundamental niche), and an approximation of the known range, EOO, represents our estimate of the occupied distribution (corresponding to the realized niche; Figure 1 and Appendix S4). PSA ranged from 23,114 km² (*T. ludovicianum*) to 1,352,496 km² (*T. cernuum*), with an average of 456,791.4 km² (Table 4). EOO ranged from 12,464 km² (*T. discolor*) to 2,483,803 km² (*T. cernuum*), with an average of 649,624.8 km². The intersection between the PSA and the EOO ranged from 7,235 km² (*T. discolor*) to 963,994 km² (*T. grandiflorum*), with an average of 282,743.8 km² (Table 4).

Simple linear models, evaluated using ANOVA, yielded significant, positive relationships between PSA and EOO ($F_{1,19}$ = 92.16, p <.001, R^2 = 0.82; Figure A3a in Appendix S1), between PSA and the number of occurrences used in ENMs ($F_{1,19}$ = 17.86, p <.001, R^2 = 0.45; Figure A3b in Appendix S1), and between EOO and the number of occurrences used in ENMs ($F_{1,19}$ = 28.68, p <.001, R^2 = 0.58; Figure A3c in Appendix S1). A Pearson's product moment correlation test found that PO (the proportional occupancy of PSA), was significantly, positively correlated with EOO (t_{19} = 4.79, p =.001, ρ = 0.74).

3.2 | Relationship between the predicted suitable area and occupied distribution

There was considerable variation in PO across the 21 species of *Trillium*. PO ranged from 1.1% (*T. discolor*) to 96% (*T. grandiflorum*), with an average of 51% (Figure 2). Sessile-flowered species comprised the majority of species with PO <60% (11/13), and only one sessile species (*T. sessile*) had PO >60%. Pedicellate-flowered species as a group showed broad variation in PO, ranging from 29% to 96%; however, 7 of 8 species with PO >60% were pedicellate (Figure 2).

3.3 | Incorporating life history traits

The best-fit beta regression model included the non-interactive effects of flower type, ovule number, and seed mass for the mean model and employed a logit link (AICc = −4.83; Loglik =11.92, df =6, pseudo R^2 = 0.70; X^2 = 23.65, p <.001; Table 3). The mean model yielded significant effects of all three predictors on PO. The median PO value of pedicellate-erect species was twice that of the median PO value of sessile species; this difference was determined to be significant by a pairwise contrasts post hoc test (p <.001; Figure 3). Ovule number and seed mass, both continuous variables, had small, positive effects on PO (ovule coeff. = 0.02, p <.001 [Figure 4a]; seed mass coeff. = 0.16, p =.003 [Figure 4b].). The unspecified precision model yielded an estimate of 8.12 for Φ (z = 3.18, p =.001). Models with low ΔAICc (>5) included the versions of this model with different link functions (loglog, cloglog and probit; Table 3), as well as the model containing flower type and ovule number (Table 3). None of the interaction effects in any of the models considered were significant.

Simple linear models, evaluated using ANOVA, found significant differences in seed mass and seed setting rate among flower types for the 19 species included in the Ohara (1989) dataset, with sessile species having significantly greater seed mass than pedicellate species (mean =8.32 ∓ 3.67 mg; mean =4.09 ∓ 1.32 mg, respectively; $F_{1,17}$ = 10.34, p =.005), and pedicellate species having significantly higher seed setting rates than sessile species (mean =57.55 ∓ 21.33; mean =35.30 ∓ 16.29, respectively; $F_{1,17}$ = 15.70, p =.001; Figure A1 in Appendix S1).

4 | DISCUSSION

Our results support the hypothesis that variation in proportional occupancies (*POs*) of *Trillium* species' predicted suitable areas (based on models of fundamental niches) can be explained by flower type—a component of trillium life history that relates to animal-mediated seed dispersal ability (Griffin & Barrett, 2004a; Jules, 1996; Myers et al., 2004; Vellend et al., 2003; Zettler et al., 2001., b.) and conservation status (NatureServe, 2020; Figure A2 in Appendix S1). Flower type was a significant predictor of PO, with sessile-flowered species

TABLE 4 A list of the final ecological niche model characteristics (calibration and validation sample sizes, accuracy metrics, extent of occurrence and predicted suitable area)[a] for each species (ordered alphabetically)

Species	Total occ.	Train occ.	Test occ.	% Train	% Test	Boyce Index (Spearman correlation coefficient)	Sensitivity/specificity training data (AUC)	Sensitivity/specificity testing data (AUC)	10% training presence test omission rate	PSA (km²)	EOO (km²)	EOO ∩ PSA (km²)	PO
T. catesbaei	274	57	217	0.21*	0.79*	0.97	0.87	0.72	0.3	258,788	308,649	112,546	0.43
T. cernuum	609	305	304	0.5	0.5	0.99	0.85	0.83	0.13	1,352,496	2,483,803	931,153	0.69
T. cuneatum	365	183	182	0.5	0.5	0.95	0.73	0.7	0.15	545,655	455,496	291,839	0.53
T. decipiens	29	21	8	0.7	0.3	0.91	0.84	0.73	0.25	91,742	68,360	27,224	0.30
T. decumbens	50	16	34	0.32*	0.68*	0.85	0.8	0.69	0.15	89,395	33,433	13,261	0.15
T. discolor	20	15	5	0.75+	0.25+	0.69	0.73	0.68	0	613,004	12,464	7,235	0.01
T. erectum	1,991	996	995	0.5	0.5	0.99	0.78	0.79	0.1	856,466	1,873,455	795,181	0.93
T. flexipes	439	220	219	0.5	0.5	0.94	0.83	0.78	0.19	644,018	1,339,115	565,325	0.88
T. foetidissimum	112	39	73	0.35*	0.65*	0.90	0.87	0.76	0.18	28,611	40,246	16,301	0.57
T. grandiflorum	2,399	1,200	1,199	0.5	0.5	0.97	0.77	0.76	0.11	1,001,650	1,943,268	963,994	0.96
T. ludovicianum	47	31	16	0.66*	0.34*	0.95	0.88	0.79	0.31	23,114	60,896	13,340	0.58
T. luteum	147	67	80	0.45*	0.55*	0.90	0.71	0.72	0.13	283,464	93,179	62,665	0.22
T. maculatum	75	38	37	0.5	0.5	0.69	0.72	0.62	0.14	315,289	232,963	159,818	0.51
T. nivale	195	137	58	0.7	0.3	0.94	0.81	0.77	0.17	552,251	654,384	339,972	0.62
T. recurvatum	1,118	378	740	0.34*	0.66*	0.99	0.76	0.71	0.19	728,779	764,182	388,129	0.53
T. sessile	597	299	298	0.5	0.5	0.98	0.8	0.75	0.17	600,646	1,081,210	502,722	0.84
T. stamineum	121	97	24	0.8+	0.2+	0.93	0.71	0.69	0.08	218,826	126,598	87,675	0.40
T. sulcatum	97	68	29	0.7*	0.3*	0.94	0.81	0.7	0.3	55,685	171,902	33,829	0.61
T. underwoodii	85	31	54	0.36*	0.64*	0.93	0.82	0.74	0.24	323,303	107,794	43,242	0.13
T. undulatum	1,187	594	593	0.5	0.5	0.99	0.85	0.85	0.11	930,424	1,713,392	559,239	0.60
T. vaseyi	111	48	63	0.39*	0.61*	0.97	0.87	0.82	0.19	79,003	77,320	22,919	0.29

[a]Total occ. is the total number of occurrences used per species. In total, 10,068 occurrences were used in this study. Training occ. is the number of occurrences used for training the model. Testing occ. is the number of occurrences used to test the model. % Training and % Testing indicate the percentage of the total number of occurrences used for training and testing the model, respectively. Asterisks indicate species for which the first method of data splitting was used (i.e. georeferenced occurrences with ≤3 km² uncertainty were used as training data; all other occurrences were used as testing data). Plus signs indicate species (2) for which crossvalidation with 5 replicates was used and the best replicate (highest AUC and lowest test omission error) was retained for analysis. Boyce Index (Spearman correlation coefficient), sensitivity/specificity for training data (AUC), sensitivity/specificity for testing data (AUC), and 10 per. training presence test omission rate are all model fit statistics. PSA (i.e. estimate of the fundamental niche) is the predicted suitable area (km²) produced by ENM. EOO is an estimate of the area (km²) of the known range. EOO ∩ PSA is the area of intersection between the PSA and EOO. PO is the proportional occupancy of the fundamental niche, calculated as in Equation 1.
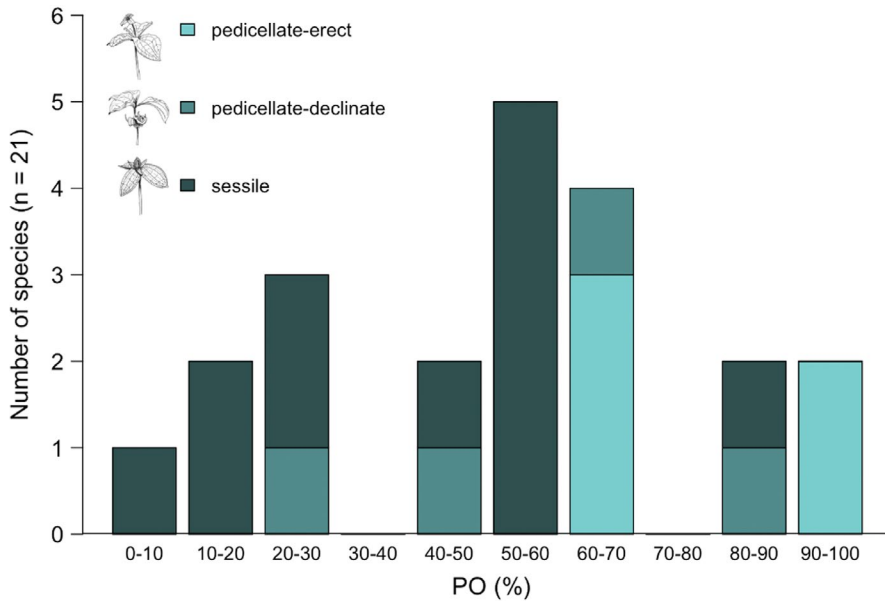
FIGURE 2 Frequency histogram illustrating the proportional occupancy of the fundamental niche (PO, expressed as a percentage) for 21 species of *Trillium* based on flower type. Light turquois indicates species with pedicellate-erect flowers; medium turquois indicates species with pedicellate-declinate flowers; and dark turquois indicates species with sessile flowers. All but two of the 12 species with PO ≤ 60% are sessile, whereas only one of the 7 species with PO ≥ 60% is sessile
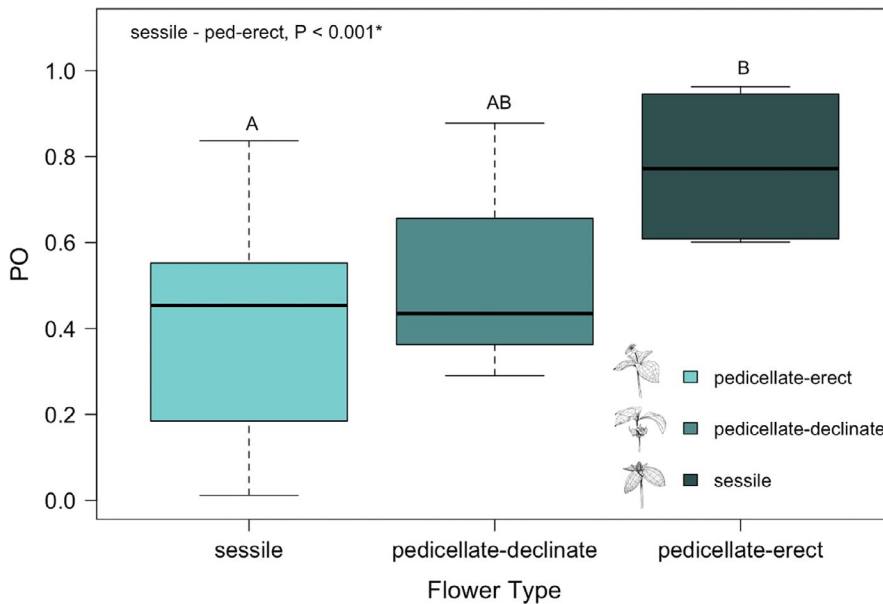

FIGURE 3 Boxplots illustrating median, interquartile range and outliers of the proportional occupancy of the fundamental niche (PO) for pedicellate-erect (light turquois), pedicellate-declinate (medium turquois) and sessile (dark turquois) species of *Trillium*. Letters indicate significant pairwise differences among flower types determined by post hoc tests. Sessile-flowered species have significantly lower PO than pedicellate-erect species ($p < .001$)

having significantly lower proportional occupancy of their predicted suitable areas than pedicellate-erect species. This finding may be explained by a variety of factors associated with flower type, including vegetative and reproductive differences, microhabitat preferences and resulting demographic consequences, or differences in the likelihood and frequency of short- and long-distance seed dispersal. Furthermore, because sessile flowers are a synapomorphy for subgenus *Sessilium*, variation in PO may be linked to phylogenetic differences between this clade and other species in genus *Trillium*. Ovule number and seed mass were also significant predictors of PO, although the coefficient estimates for both of these factors were small. This suggests that species with higher ovule numbers and larger seed masses tended to have larger PO. The pseudo R^2 values for the five most likely beta regression models (those with ΔAICc <5) were all greater than 0.5 (Table 3), indicating that reproductive

life history traits are a useful measure of potential range filling in this genus.

That PO is relatively low for the majority of species in this study (mean =51%) suggests that factors not captured by traditional, climate-calibrated ENMs play a substantial role in determining the geographic distributions of most species in this genus. Whether those factors can be classified as species interactions, biotic characteristics of the environment, dispersal limitations or historical biogeographical factors, we cannot determine here. However, our study establishes that reproductive life history traits capture components of some of the above, and therefore exploring the potential effects of these traits on dispersal ability is a useful starting point for further investigation. Sessile-flowered species belong to the subgenus *Sessilium*, comprising a monophyletic clade found in eastern North America (Farmer & Schilling, 2002; Millam, 2006).
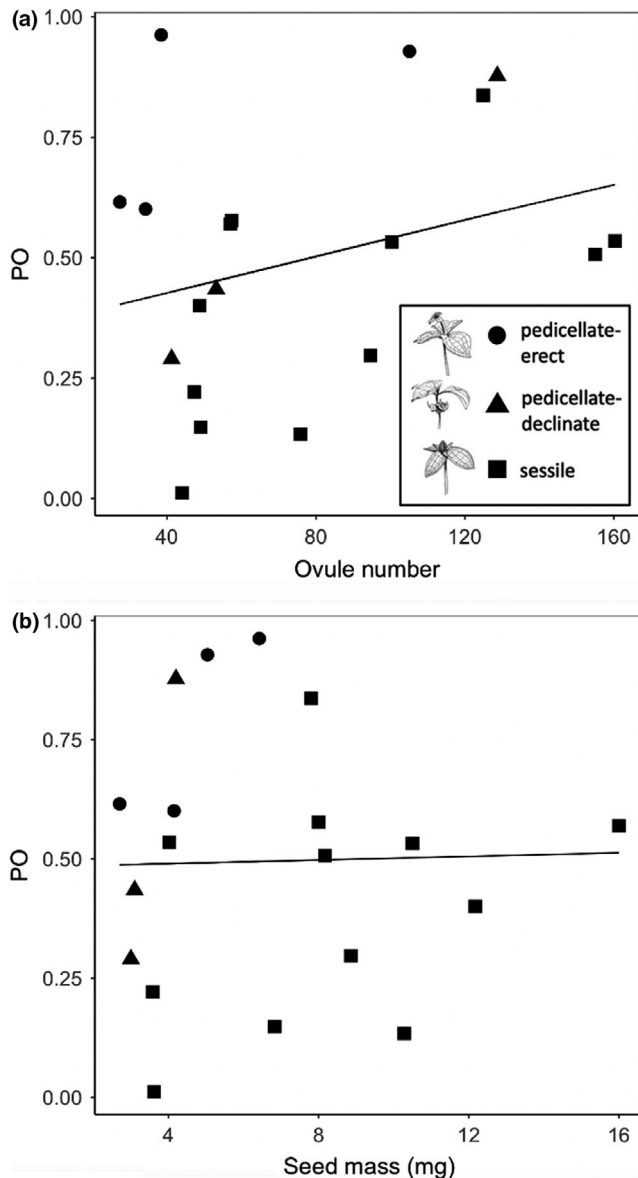
FIGURE 4 Scatterplots with trendlines illustrating the relationships between (a) ovule number and PO and (b) seed mass (mg) and PO, which, in combination with flower type, were both significant predictors of PO according to the best-fit beta regression (Table 3). Shape of points denotes flower type (circles = pedicellate-erect; triangles = pedicellate-declinate; squares = sessile). The trendlines and scatter of the points in each panel depict the isolated relationships between ovule number and PO, and seed mass and PO, respectively, and are included for illustrative purposes; these trendlines do not reflect the pseudo R^2 value for the overall best-fit model

Within the pedicellate species, subgenus *Delostylium* (the "delostylis group") is a monophyletic clade endemic to the southeastern U.S. (Farmer, 2007). The remaining pedicellate species are informally grouped into the subgenus *Trillium*, which spans North America and Asia (Farmer & Schilling, 2002; Millam, 2006). The conservation status of trilliums in ENA differs as a function of flower type. Both the sessile and pedicellate groups contain species ranked by NatureServe (2020) as highest concern (G1: "Critically Imperiled")

and lowest concern (G5: "Secure"). For example, *Trillium delicatum* (G1), *T. gracile* (G2), *T. oostingii* (G1) *and T. tennesseense* (G1) are all sessile-flowered species, while *T. georgianum* (G1) and *T. persistens* (G1) are pedicellate-flowered species. However, 64% of ENA pedicellate species have rankings of G5 or G4 (low concern), compared to only 47% of sessile species (NatureServe, 2020), suggesting that sessile species, as a clade, are of greater conservation concern than non-sessile species in ENA.

The set of 21 study species was represented by ranks G3, G4 and G5. Twenty-five percent of sessile study species comprised the highest conservation risk category (G3), compared to only 11% of pedicellate study species. Likewise, only 16% of sessile study species were ranked as lowest concern (G5), compared to 55% of pedicellate study species (Figure A2 in Appendix S1). That more sessile trilliums are of greater conservation concern than pedicellate species in ENA may be a reflection of the higher prevalence of range-restriction and narrow endemism in sessile species. We note a few caveats to this assertion. First, while most pedicellate species in ENA are geographically widespread and of low conservation concern, there are two range-restricted pedicellate species in ENA ranked as "Critically Imperiled" (e.g. *T. georgianum* and *T. persistens*). Additionally, while most sessile species are indeed restricted to the southeastern United States, and many are endemic to specific regions (e.g. *T. discolor* and *T. decumbens* are endemic to the southern Appalachians; *T. underwoodii*, *T. ludovicianum* and *T. foetidissimum* are endemic to the southern coastal plains), two sessile species, *T. sessile* and *T. recurvatum*, are geographically widespread, with ranges extending as far north as Wisconsin. These species are exceptions to the general trends we highlight concerning differences in range size, endemism, conservation status, and PO among sessile and pedicellate species of *Trillium*. Second, range size and habitat specificity are not the only factors considered by NatureServe when assigning conservation ranks (Master et al., 2003; Regan et al., 2004), so range size is not wholly synonymous with conservation status. Finally, we acknowledge that it is possible that a species could persist in the environmental space outside of the region represented by the training data in our models; therefore, our estimates of potential occupancy could be negatively biased (i.e. estimates could be smaller than they are in reality.) Despite these caveats, a connection can be drawn between higher frequency of narrow endemism and greater conservation threat associated with sessile trilliums, and our finding that species in this group have significantly lower PO than pedicellate species. Thus, at-risk, endemic species may be less likely to occupy their entire predicted suitable areas (based on models of fundamental niches) than geographically widespread species. This is consistent with the notion that range-restricted, endemic plant species are constrained to a greater extent by biotic factors, dispersal limitations or a combination of these non-abiotic factors compared to geographically widespread species. There is evidence for this trend in endemic flora in Australia (Rossetto & Kooyman, 2005; Rossetto et al., 2008), central Europe (Essl et al., 2011) and the Mediterranean region (Youssef et al., 2011).

In his comprehensive study of life history evolution in the genus *Trillium*, Ohara (1989) compared life history characteristics for 27 species of *Trillium* across Japan and North America. He noted "remarkable differences" in traits among sessile and pedicellate species, finding that pedicellate species had higher individual adult biomass, larger reproductive outputs and smaller seeds than sessile-flowered species (Ohara, 1989). Among our investigated species, seed mass and seed setting rate differed significantly among flower types, but biomass and ovule number did not differ among flower types. Somewhat paradoxically, seed mass, which we found to be significantly larger for sessile study species, had a significant, positive effect on PO; this apparent contradiction to our finding that sessile study species had significantly lower PO than pedicellate study species may be explained by the small value of the positive coefficient estimate for seed mass. Alternatively, pedicellate-declinate species, which did not have significantly different values of PO from either pedicellate-erect or sessile species (Figure 3), may explain this result; pedicellate-declinate species may have larger seed masses than pedicellate-erect species as a whole. Ovule number, which was not significantly different among pedicellate and sessile species, had a very weak positive effect on PO in combination with the effects of flower type and seed mass; as the sole factor in a beta regression model, ovule number was not a significant predictor of PO. As such, we will focus the remainder of our discussion on the stronger effects of flower type on the proportional occupancy of the predicted suitable area.

In ENA, Ohara (1989) observed that pedicellate species tended to occupy the northeast, whereas most sessile species were found in the southeast. This is generally consistent with our estimates of the known ranges of sessile and pedicellate species. Ohara also noted habitat differences, wherein pedicellate species occupied beech and sugar maple forests, and sessile species occupied alluvial flood plain terraces, river bottoms and river bluffs (Ohara, 1989). Patrick (1984) and Millam (2006) confirmed this observation, reporting that pedicellate and sessile species are characterized by microhabitat differences such as elevation and edaphic factors. The results of these studies suggest that microhabitat preferences, habitat specificity and resulting demographic effects may also contribute to our finding that pedicellate species are more likely to occupy their entire predicted suitable areas than sessile species. Given that sessile species are generally of greater conservation concern than pedicellate species, future studies should focus on risks to sessile trilliums associated with microhabitat preferences and habitat specificity.

To our knowledge, the only other study that connected life history traits with the proportional occupancy of predicted suitable area using our same methods found that dispersal mechanism was a significant predictor of PO for 89 Mexican mammal species (Munguía et al., 2008). In our study, all trillium species share a primary dispersal mechanism: myrmecochory. However, differences in seed dispersal rates and probabilities may nevertheless explain our finding that flower type and other seed-related traits are related to PO. There is empirical evidence that the rate of ant-mediated seed dispersal is significantly lower for the sessile

species *T. discolor* than its sympatric, pedicellate congener *T. catesbaei* in the southern Appalachians (Miller & Kwit, 2018). In a follow-up study, Miller et al. (2020) found that, of five southeastern *Trillium* species, *T. catesbaei* (the only pedicellate species in the group) had elaiosomes with greater concentrations of important signalling compounds and nutrients; this species also had the highest probability of seed dispersal by ants in the field. Although these studies only considered one pedicellate species in comparison with several sessile species, they provide evidence that pedicellate trilliums may produce seeds that are more attractive to ant dispersers and suggest that when sympatric, pedicellate trilliums may outcompete their sessile congeners for dispersal services. Consistently lower rates of seed dispersal by ants could contribute to the increased instances of range-restricted endemism observed in sessile trilliums.

It is also likely that differences in the probability of long-distance dispersal (LDD) events among flower types explain the observed variation in PO. In addition to myrmecochory, two geographically widespread pedicellate species in ENA, *T. grandiflorum* and *T. erectum*, are prone to frugivory and seed dispersal via herbivorous white-tailed deer (Griffin & Barrett, 2004a; Myers et al., 2004; Vellend et al., 2003). Whether other species of *Trillium* similarly obtain LDD from white-tailed deer is unknown. Notably, we are not aware of any evidence that sessile species of *Trillium* are adapted to dispersal by deer. Empirically calibrated diffusion models have illustrated that, given the relationship between current distributions of North American ant-dispersed woodland herbs and the extent of the glaciers at the Last Glacial Maximum (LGM), the migration of plants such as those in the genus *Trillium* must rely on occasional, unknown form(s) of LDD (Cain et al., 1998). Few comprehensive phytogeographical studies have been conducted for trilliums in ENA, but an interesting comparison can be drawn between two studies that used similar methods to reconstruct the glacial history and post-glacial colonization of two trillium species. Gonzales et al. (2008) found that *T. cuneatum*—a sessile species with a moderately sized range in the southeastern United States (BONAP; Kartesz & BONAP [Biota of North America Program], 2015)—survived the LGM in multiple refugia across the southeastern United States, but that the Appalachian Mountains functioned as a barrier to northward dispersal. In contrast, Griffin and Barrett (2004a) demonstrated that *T. grandiflorum*—a ubiquitous pedicellate species with a large, northerly range (BONAP; Kartesz & BONAP [Biota of North America Program], 2015)—also survived the LGM in two southern refugia, but its dispersal was not subsequently impeded by the Appalachian Mountains. Griffin and Barrett (2004a) concluded that occasional LDD events, such as those provided by white-tailed deer, must be responsible for the post-glacial recolonization of northern areas by *T. grandiflorum*. By that same logic, it is possible that the lack of post-glacial recolonization of northern areas by *T. cuneatum* may be due to the lack of LDD.

These studies, in conjunction with our findings here, provide compelling support for the notion that sessile-flowered trilliums in ENA may be characterized by higher frequencies of narrow

endemism, ranges restricted to the southeastern United States, and lower occupancy of predicted suitable areas because they do not achieve LDD to the same extent as pedicellate-flowered species. As such, we posit that variation in dispersal potential in plants of the genus *Trillium* stemming from all sources, including ants as primary dispersers and any potential LDD vectors, is strongly related to both range size and a species' ability to occupy its suitable area (as predicted by models of fundamental niche). These insights, gained from studying the relationship between proportion of occupancy and reproductive life history traits in species with sample sizes large enough to calibrate models, might be extrapolatable to the rarest and most-threatened members of the genus. For instance, one attribute typically associated with the pedicellate flower type—relatively tall plant stature—is lacking for the G1 pedicellate species *T. georgianum*, known commonly as the Georgia Dwarf Trillium (Schilling et al., 2017). The short stature of this extremely range-restricted species may preclude it from obtaining LDD by white-tailed deer. However, we stress the need for caution in extrapolating insights drawn from more common species to species with extremely small distributions, given that the latter could be range-restricted due to fine-scale parameters such as habitat specificity and/or local adaptation that would not be detected via coarse-scale distributional comparisons. In this study, we exemplify a scenario in which reproductive life history traits may shed light on why some species are geographically range-restricted when their close relatives are geographically widespread. Our methodology, particularly if combined with estimates of niche breadth, overlap, equivalency, similarity and expansion, constitutes a powerful comparative framework that can be widely applied to diverse biological systems around the globe to yield important insights into the conservation of rare species.

ACKNOWLEDGEMENTS

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13297.

DATA AVAILABILITY STATEMENT

All data generated for use in this manuscript are archived in the Dryad Digital Repository and are accessible at the following: Miller, Chelsea (2021), Data associated with ecological niche models and post-ENM statistical analyses for *Trillium* species distributions, Dryad, Dataset, https://doi.org/10.5061/dryad.6m905qg03.

ORCID

Chelsea N. Miller [iD] https://orcid.org/0000-0002-8214-1565

REFERENCES

Andreas Hamann's website. (2013). https://sites.ualberta.ca/~ahamann/data.html. Accessed September, 2017

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93(7), 1527–1539. https://doi.org/10.1890/11-1930.1

Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R. B., Zimmermann, N. E., & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, 5(1), eaat4858. https://doi.org/10.1126/sciadv.aat4858

The Arnold Arboretum of Harvard University. (2020). *The President and Fellows of Harvard College*. Retrieved from https://www.arboretum.harvard.edu/

Bale, M. T., Zettler, J. A., Robinson, B. A., Spira, T. P., & Allen, C. R. (2003). Yellow jackets may be an underestimated component of an ant-seed mutualism. *Southeastern Naturalist*, 2(4), 609–614.

Beale, C. M., Brewer, M. J., & Lennon, J. J. (2014). A new statistical framework for the quantification of covariate associations with species distributions. *Methods in Ecology and Evolution*, 5, 421–432. https://doi.org/10.1111/2041-210X.12174

Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. (2002). Evaluating resource selection functions. *Ecological Modelling*, 157(2–3), 281–300. https://doi.org/10.1016/S0304-3800(02)00200-4

Broennimann, O., Di Cola, V., & Guisan, A. (2021). *ecospat: Spatial ecology miscellaneous methods*. R Package Version 3.2. Retrieved from https://CRAN.R-project.org/package=ecospat

Brown, J. L. (2014). SDMtoolbox: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, 5(7), 694–700. https://doi.org/10.1111/2041-210X.12200

Brown, J. L., Bennett, J. R., & French, C. M. (2017). SDMtoolbox 2.0: The next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ*, 5, e4095.

Cabe, P. R. (1995). The *Trillium pusillum* Michaux (Liliaceae) complex in Virginia. I. Morphological Investigations. *Castanea*, 1–14.

Cabe, P. R., & Werth, C. (1995). The *Trillium pusillum* Michaux (Liliaceae) complex in Virginia. *II. Isozyme Evidence. Castanea*, 15–29.

Cain, M. L., Damman, H., & Muir, A. (1998). Seed dispersal and the Holocene migration of woodland herbs. *Ecological Monographs*, 68, 325–347.

Case, F. W., & Case, R. B. (1997). *Trillium*. Timber Press.

Case, F. W. (2002a). "*Trillium subg. Trillium*". In Flora of North America Editorial Committee (Ed.), *Flora of North America North of Mexico (FNA)*. 26. New York and Oxford –via eFloras.org, Missouri Botanical Garden, St. Louis, MO & Harvard University Herbaria. Cambridge, MA.

Case, F. W. (2002b). *"Trillium subg.Phyllantherum"*. In Flora of North America Editorial Committee (Ed.), *Flora of North America North of Mexico (FNA)*. 26. New York and Oxford – via eFloras.org, Missouri Botanical Garden, St. Louis, MO & Harvard University Herbaria, Cambridge, MA.

Chase, J. M., & Leibold, M. A. (2003). *Ecological niches: Linking classical and contemporary approaches*. University of Chicago Press.

DeMarco, B. B., & Cognato, A. I. (2016). A multiple-gene phylogeny reveals polyphyly among eastern North American *Aphaenogaster* species (Hymenoptera: Formicidae). *Zoologica Scripta*, 45, 512–520.

Dobzhansky, T. (1950). Mendelian populations and their evolution. *The American Naturalist*, 84, 401–418. https://doi.org/10.1086/281638

Dormann, C. F., Schymanski, S. J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin, X., Römermann, C., Schröder, B., &

Singer, A. (2012). Correlation and process in species distribution models: Bridging a dichotomy. *Journal of Biogeography*, *39*, 2119–2131. https://doi.org/10.1111/j.1365-2699.2011.02659.x

Douma, J. C., & Weedon, J. T. (2019). Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution*, *10*(9), 1412–1430. https://doi.org/10.1111/2041-210X.13234

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., ... E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, *17*(1), 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Essl, F., Dullinger, S., Plutzar, C., Willner, W., & Rabitsch, W. (2011). Imprints of glacial history and current environment on correlations between endemic plant and invertebrate species richness. *Journal of Biogeography*, *38*(3), 604–614. https://doi.org/10.1111/j.1365-2699.2010.02425.x

Estrada, A., Meireles, C., Morales-Castilla, I., Poschlod, P., Vietes, D., Araújo, M. B., & Early, R. (2015). Species' intrinsic traits inform their range limitations and vulnerability under environmental change. *Global Ecology and Biogeography*, *24*, 849–858.

Farmer, S. B., & Schilling, E. E. (2002). Phylogenetic analyses of Trilliaceae based on morphological and molecular data. *Systematic Botany*, *27*(4), 674–692.

Farmer, S. B. (2007). *Systematics of* Trillium *subgenus* Delostylis. University of Tennessee. PhD Dissertation.

Feng, X., Park, D. S., Walker, C., Peterson, A. T., Merow, C., & Papeş, M. (2019). A checklist for maximizing reproducibility of ecological niche models. *Nature Ecology and Evolution*, *3*, 1382–1395. https://doi.org/10.1038/s41559-019-0972-5

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, 38–49.

Fordham, D. A., Akçakaya, H. R., Araújo, M. B., Elith, J., Keith, D. A., Pearson, R., Auld, T. D., Mellin, C., Morgan, J. W., Regan, T. J., Tozer, M., Watts, M. J., White, M. W., Wintle, B. A., Yates, C., & Watts, B. W. (2012). Plant extinction risk under climate change: Are forecast range shifts alone a good indicator of species vulnerability to global warming? *Global Change Biology*, *18*, 1357–1371. https://doi.org/10.1111/j.1365-2486.2011.02614.x

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*, 2nd ed. Sage. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Freeman, J. D. (1975). Revision of *Trillium* subgenus *Phyllantherum* (Liliaceae). *Brittonia*, *27*(1), 1–62. https://doi.org/10.2307/2805646

GBIF: The Global Biodiversity Information Facility (2020). *What is GBIF?*. Retrieved from https://www.gbif.org/what-is-gbif

Givnish, T. J. (1990). Leaf mottling: Relation to growth form and leaf phenology and possible role as camouflage. *Functional Ecology*, 463–474. https://doi.org/10.2307/2389314

Gonzales, E., & Hamrick, J. L. (2005). Distribution of genetic diversity among disjunct populations of the rare forest understory herb, Trillium Reliquum. *Heredity*, *95*(4), 306–314. https://doi.org/10.1038/sj.hdy.6800719

Gonzales, E., Hamrick, J. L., & Chang, S. M. (2008). Identification of glacial refugia in south-eastern North America by phylogeographical analyses of a forest understorey plant, *Trillium cuneatum*. *Journal of Biogeography*, *35*(5), 844–852. https://doi.org/10.1111/j.1365-2699.2007.01834.x

Griffin, S. R., & Barrett, S. C. H. (2004). Post-glacial history of *Trillium grandiflorum* (Melanthiaceae) in eastern North America: Inferences from phylogeography. *American Journal of Botany*, *91*, 465–473.

Griffin, S. R., & Barrett, S. C. (2004). Genetic variation in *Trillium erectum* (Melanthiaceae), a widespread forest herb in eastern North America. *Canadian Journal of Botany*, *82*(3), 316–321.

Grinnell, J. (1917). The niche-relationships of the California Thrasher. *The Auk*, *34*(4), 427–433. https://doi.org/10.2307/4072271

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., McCarthy, M. A., Tingley, R., & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, *24*(3), 276–292. https://doi.org/10.1111/geb.12268

Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., & Hijmans, M. R. J. (2017). Package 'dismo'. *Circles*, *9*(1), 1–68.

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, *199*(2), 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427. https://doi.org/10.1101/SQB.1957.022.01.039

Hutchinson, G. E. (1978). *An introduction to population biology*. Yale Univ Press.

Imdad, M. U., & Aslam, M. (2018). *mctest: Multicollinearity Diagnostic Measures*. Retrieved from https://CRAN.R-project.org/package=mctest, R package version 1.2.5

Imdad, M. U., Aslam, M., Altaf, S., & Ahmed, M. (2019). Some new diagnostics of multicollinearity in linear regression model. *Sains Malaysiana*, *48*(9), 2051–2060. https://doi.org/10.17576/jsm-2019-4809-26

Imdadullah, M., Aslam, M., & Altaf, S. (2016). mctest: An R package for detection of collinearity among regressors. *The R Journal*, *8*(2), 499–509.

IUCN (2012). *IUCN Red list categories and criteria: Version 3.1*, 2nd ed. IUCN. iv + 32 pp.

Jacquemyn, H., Butaye, J., & Hermy, M. (2001). Forest plant species richness in small, fragmented mixed deciduous forest patches: The role of area, time and dispersal limitation. *Journal of Biogeography*, *28*(6), 801–812. https://doi.org/10.1046/j.1365-2699.2001.00590.x

Jules, E. S. (1996). Yellow jackets (*Vespula vulgaris*) as a second seed disperser for the myrmecochorous plant, *Trillium Ovatum*. *American Midland Naturalist*, 367–369. https://doi.org/10.2307/2426720

Kartesz, J., & BONAP [Biota of North America Program]. (2015). *North American Plant Atlas*. Website: http://bonap.net/napa. Chapel Hill, North Carolina [maps generated from J. Kartesz. (2015). Floristic synthesis of North America, version 1.0 (in press). Biota of North America Program (BONAP).]

Leach, K., Montgomery, W. I., & Reid, N. (2016). Modelling the influence of biotic factors on species distribution patterns. *Ecological Modelling*, *337*, 96–106. https://doi.org/10.1016/j.ecolmodel.2016.06.008

Lomolino, M. V., Riddle, B. R., & Brown, J. H. (2005). *Biogeography*, 3rd ed. Sinauer Associates.

MacArthur, R. H. (1984). *Geographical ecology: Patterns in the distribution of species*. Princeton University Press.

Master, L. L., Morse, L. E., Weakley, A. S., Hammerson, G. A., & Faber-Langendoen, D. (2003). *NatureServe conservation status criteria*. NatureServe Arlington.

Mazerolle, M. J. (2019). *AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c)*. R package version 2.2-1.

Millam, K. C. (2006). *The Trillium erectum species-complex (Melanthiaceae): Insights from molecular systematics and biogeography*. The University of Wisconsin-Madison.

Miller, C. N., & Kwit, C. (2018). Overall seed dispersal effectiveness is lower in endemic *Trillium* species than in their widespread congeners. *American Journal of Botany*, *105*(11), 1847–1857.

Miller, C. N., Whitehead, S. R., & Kwit, C. (2020). Effects of seed morphology and elaiosome chemical composition on attractiveness of five *Trillium* species to seed-dispersing ants. *Ecology and Evolution*, *00*, 1–14. https://doi.org/10.1002/ece3.6101

Munguía, M., Townsend Peterson, A., & Sánchez-Cordero, V. (2008). Dispersal limitation and geographical distributions of mammal species. *Journal of Biogeography*, *35*(10), 1879–1887. https://doi.org/10.1111/j.1365-2699.2008.01921.x

Myers, J. A., Vellend, M., Gardescu, S., & Marks, P. L. (2004). Seed dispersal by white-tailed deer: Implications for long-distance dispersal, invasion, and migration of plants in eastern North America. *Oecologia*, *139*(1), 35–44.

NatureServe (2020). *NatureServe Explorer: An online encyclopedia of life*. NatureServe. http://explorer.natureserve.org/

Ness, J. H., Morin, D., & Giladi, I. (2009). Uncommon specialization in a mutualism between a temperate herbaceous plant guild and an ant: Are *Aphaenogaster* ants keystone mutualists? *Oikos*, *118*, 1793–1804.

Ohara, M. (1989). Life history evolution in the genus *Trillium. Plant Species Biology*, *4*(1), 1–28. https://doi.org/10.1111/j.1442-1984.1989.tb00044.x

Osaloo, S. K., Utech, F. H., Ohara, M., & Kawano, S. (1999). Molecular systematics of Trilliaceae I. Phylogenetic analyses of *Trillium* using matK gene sequences. *Journal of Plant Research*, *112*(1), 35–49.

Park, D. S., Ellison, A. M., & Davis, C. C. (2018). Mating system does not predict niche breath. *Global Ecology and Biogeography*, *27*(7), 804–813. https://doi.org/10.1111/geb.12740

Patrick, T. S. (1984). *Trillium sulcatum* (Liliaceae), a new species of the southern Appalachians. *Brittonia*, *36*(1), 26–36. https://doi.org/10.2307/2806287

Peterson, A. T. (2001). Predicting species' geographic distributions based on ecological niche modeling. *The Condor*, *103*(3), 599–605. https://doi.org/10.1093/condor/103.3.599

Peterson, A. T. (2006). Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics*, *3*, 59–72. https://doi.org/10.17161/bi.v3i0.29

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). *Ecological niches and geographic distributions (MPB-49)*, Vol. *49*. Princeton University Press.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3–4), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pocheville, A. (2015). The ecological niche: History and recent controversies. In T. H. Heams, P. Huneman, G. Lecointre, & M. Silberstein (Eds.), *Handbook of evolutionary thinking in the sciences* (pp. 547–586). Springer.

Regan, T. J., Master, L. L., & Hammerson, G. A. (2004). Capturing expert knowledge for threatened species assessments: A case study using NatureServe conservation status ranks. *Acta Oecologica*, *26*(2), 95–107. https://doi.org/10.1016/j.actao.2004.03.013

Rios, N. E., & Bart, H. L. (2010). *GEOLocate (Version 3.22)* [Computer software]. Tulane University Museum of Natural History. https://www.geo-locate.org/default.html

Rossetto, M., & Kooyman, R. M. (2005). The tension between dispersal and persistence regulates the current distribution of rare palaeo-endemic rain forest flora: A case study. *Journal of Ecology*, *93*(5), 906–917. https://doi.org/10.1111/j.1365-2745.2005.01046.x

Rossetto, M., Kooyman, R., Sherwin, W., & Jones, R. (2008). Dispersal limitations, rather than bottlenecks or habitat specificity, can restrict the distribution of rare and endemic rainforest trees. *American Journal of Botany*, *95*(3), 321–329. https://doi.org/10.3732/ajb.95.3.321

Schilling, E. E., Floden, A., Lampley, J., Patrick, T. S., & Farmer, S. B. (2017). A new species in *Trillium* subgen. Delostylium (Melanthiaceae, Parideae). *Phytotaxa*, *296*(3), 287–291.

SERNEC Data Portal. (2020). Retrieved from http://:sernecportal.org/portal/index.php [Aug 1, 26, 30, and Sept. 1 - 4, 2018; July 12, 2019]

Smith, A. B. (2013). On evaluating species distribution models with random background sites in place of absences when test presences disproportionately sample suitable habitat. *Diversity and Distributions*, *19*(7), 867–872. https://doi.org/10.1111/ddi.12031

Soberón, J., & Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, *2*, 1–10. https://doi.org/10.17161/bi.v2i0.4

Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, *106*(Supplement 2), 19644–19650. https://doi.org/10.1073/pnas.0901637106

Soberón, J., & Arroyo-Peña, B. (2017). Are fundamental niches larger than the realized? Testing a 50-year-old prediction by Hutchinson. *PLoS One*, *12*(4), e0175138. https://doi.org/10.1371/journal.pone.0175138

Strubbe, D., Broennimann, O., Chiron, F., & Matthysen, E. (2013). Niche conservatism in non-native birds in Europe: Niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, *22*(8), 962–970. https://doi.org/10.1111/geb.12050

Svenning, J. C., & Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters*, *7*(7), 565–573. https://doi.org/10.1111/j.1461-0248.2004.00614.x

Svenning, J. C., & Skov, F. (2007). Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters*, *10*(6), 453–460. https://doi.org/10.1111/j.1461-0248.2007.01038.x

Svenning, J. C., Normand, S., & Skov, F. (2008). Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, *31*(3), 316–326. https://doi.org/10.1111/j.0906-7590.2008.05206.x

Tanner, E. P., Papeş, M., Elmore, R. D., Fuhlendorf, S. D., & Davis, C. A. (2017). Incorporating abundance information and guiding variable selection for climate-based ensemble forecasting of species' distributional shifts. *PLoS One*, *12*, e0184316. https://doi.org/10.1371/journal.pone.0184316

Timmerman-Erskine, M., Dute, R. R., & Boyd, R. S. (2002). Morphometric analysis of the *Trillium pusillum* Michaux complex (Trilliaceae) of the southeastern United States. *Castanea*, 109–119.

Tingley, R., Thompson, M. B., Hartley, S., & Chapple, D. G. (2016). Patterns of niche filling and expansion across the invaded ranges of an Australian lizard. *Ecography*, *39*(3), 270–280. https://doi.org/10.1111/ecog.01576

Tropicos.org. (2020). *Missouri Botanical Garden*. Retrieved from http://www.tropicos.org

Umphrey, G. J. (1996). Morphometric discrimination among sibling species in the *fulva–rudis–texana* complex of the ant genus *Aphaenogaster* (Hymenoptera: Formicidae). *Canadian Journal of Zoology*, *74*(3), 528–559.

UT Herbarium – TENN. (2020). *The University of Tennessee, Knoxville*. Retrieved from https://herbarium.utk.edu/

Vellend, M., Myers, J. A., Gardescu, S., & Marks, P. L. (2003). Dispersal of *Trillium* seeds by deer: Implications for long-distance migration of forest herbs. *Ecology*, *84*(4), 1067–1072.

Verheyen, K., & Hermy, M. (2001). The relative importance of dispersal limitation of vascular plants in secondary forest succession in Muizen Forest, Belgium. *Journal of Ecology*, 829–840. https://doi.org/10.1046/j.0022-0477.2001.00596.x

Villaverde, T., González-Moreno, P., Rodríguez-Sánchez, F., & Escudero, M. (2017). Niche shifts after long-distance dispersal events in bipolar sedges (Carex, Cyperaceae). *American Journal of Botany*, *104*(11), 1765–1774.

Wang, T., Hamann, A., Spittlehouse, D. L., & Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One*, *11*, e0156720.

Weakley, A. S. (2015). *Flora of the southern and mid-Atlantic states. Working draft. University of North Carolina Herbarium (NCU)*. North Carolina Botanical Garden. University of North Carolina at Chapel Hill.

Willner, W., Di Pietro, R., & Bergmeier, E. (2009). Phytogeographical evidence for post-glacial dispersal limitation of European beech forest species. *Ecography*, *32*(6), 1011–1018. https://doi.org/10.1111/j.1600-0587.2009.05957.x

Youssef, S., Baumel, A., Véla, E., Juin, M., Dumas, E., Affre, L., & Tatoni, T. (2011). Factors underlying the narrow distribution of the Mediterranean annual plant *Arenaria provincialis* (Caryophyllaceae). *Folia Geobotanica*, *46*(4), 327–350. https://doi.org/10.1007/s12224-011-9101-1

Zettler, J. A., Spira, T. P., & Allen, C. R. (2001). Yellow jackets (*Vespula* spp.) disperse *Trillium* (spp.) seeds in eastern North America. *The American Midland Naturalist*, *146*(2), 444–446.

Zhu, G., Gariepy, T. D., Haye, T., & Bu, W. (2017). Patterns of niche filling and expansion across the invaded ranges of Halyomorpha halys in North America and Europe. *Journal of Pest Science*, *90*(4), 1045–1057. https://doi.org/10.1007/s10340-016-0786-z

BIOSKETCH

Chelsea N. Miller is a postdoctoral researcher in the Forest Entomology Lab at the Warnell School of Forestry and Natural Resources (University of Georgia, Athens) interested in the effects of climate change and anthropogenic and natural disturbances on insect-plant interactions. Using observational and experimental laboratory and field techniques, computational analyses, and ecological niche modelling, Miller addresses questions concerning the impacts of weather disturbances on the interactions between host trees and phytophagous insects and explores how fine-scale variation in plant-insect interactions scale up to impact regional geographic distributions of species. Miller's website, which contains details on her past and current research and outreach activities, can be accessed at http://www.chelseanicolemiller.com/.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Miller CN, Papeş M, Schilling EE, Kwit C. Reproductive traits explain occupancy of predicted distributions in a genus of eastern North American understory herbs. *Divers Distrib*. 2021;00:1–18. https://doi.org/10.1111/ddi.13297