Digitized collections elucidate invasion history and patterns of awn polymorphism in *Microstegium vimineum*

Manuscript received; revision accepted

Running Head: Digitized records reveal invasion history of an exotic grass

Craig F. Barrett[1], Cynthia D. Huebner[1,2,3], Zoe A. Bender[1,4], Trezalka A. Budinsky[1,5], Cameron W. Corbett[1], Maribeth Latvis[6], Michael R. McKain[7], M'Kayla Motley[7], Samuel V. Skibicki[1], Hana L. Thixton[1], Mathilda V. Santee[1], Aubrey N. Cumberledge[1]

[1] Department of Biology, West Virginia University, 53 Campus Drive, Morgantown, West Virginia, USA 26506

[2] USDA Forest Service Northern Research Station, 180 Canfield Street, Morgantown, West Virginia, USA 26505

[3] Division of Plant and Soil Sciences, 4100 Agricultural Sciences Building, PO Box 6108 Morgantown, West Virginia, USA 26506

[4] Department of Biology, Gettysburg College, 300 North Washington Street, Gettysburg, Pennsylvania, USA 17325

[5] Department of Biological Sciences, University of Pittsburgh, 4249 Fifth Avenue, Pittsburgh, Pennsylvania, USA 15260

[6] Department of Natural Resource Management, South Dakota State University, 1390 College Avenue, South Dakota State University, Brookings, South Dakota, USA 57007

[7] Department of Biological Sciences, University of Alabama, 300 Hackberry Lane, Tuscaloosa, Alabama, USA 35487

* Corresponding author e-mail: craig.barrett@mail.wvu.edu, phone: 1 (304) 293-7506, OrcID: 0000-0001-8870-3672

ABSTRACT

Premise

Digitized collections can help illuminate the mechanisms behind establishment and spread of invasive plants. These databases provide a record of traits in space and time allowing the investigation of abiotic and biotic factors influencing invasive species.

Methods

Over 1,100 digitized herbarium records were examined to investigate invasion history and trait variation of *Microstegium vimineum*. Presence-absence of awns was investigated to quantify geographic patterns of this polymorphic trait, which serves several functions in grasses including diaspore burial and dispersal to germination sites. Floret traits were further quantified, and genomic analyses of contemporary samples conducted to investigate the history of introduction and spread into North America.

Key Results

Herbarium records revealed similar patterns of awn polymorphism in native and invaded ranges of *M. vimineum*, with awned forms predominating at higher latitudes and awnless forms at lower latitudes. Herbarium records and genomic data suggested initial introduction and spread of the awnless form in the southeastern US followed by a putative secondary invasion and spread of the awned form from eastern Pennsylvania. Awned forms have longer florets and floret size varies significantly with latitude. There is evidence of a 'transition zone' with short-awned specimens at mid-latitudes. Genomic analyses revealed two distinct clusters corresponding to awnless and awned forms, with evidence of admixture.

Conclusions

Our results demonstrate the power of herbarium data to elucidate the invasion history of a problematic weed in North America, and together with genomic data, reveal a possible key trait in introduction success: presence or absence of an awn.

KEYWORDS

Hygroscopic awns, habitat filtering, herbarium database, invasiveness, latitude, plant trait, pre-adaptation, seed burial syndrome, transition zone.

INTRODUCTION

Invasive species cause billions of dollars in damage to natural, agricultural, and urban systems globally, and together with habitat destruction, represent significant threats to biodiversity further intensified by globalization and climate change (Simberloff et al., 2013). Yet, the causes for invasion success in some species and not others remain an area of intense study (Jeschke and Heger, 2018; Enders and Jeschke, 2018; Enders et al., 2018; Dai et al. 2020; Enders et al., 2020). While recent introductions provide ideal

opportunities to evaluate the mechanisms behind invasion success *in situ*, established or naturalized invasive species present more of a challenge. Digitized museum collections provide evidence toward a solution, representing powerful tools to broadly study ecology, evolution, systematics, pathology, biogeography, climate change, and invasions (Page et al., 2015; Guralnick et al., 2016; Besnard et al., 2018; McAllister et al., 2018; Borges et al., 2020; Sutherland et al., 2021; Heberling, 2022). Such studies have been facilitated by recent efforts to integrate collections from thousands of institutions into centralized, searchable image databases (e.g. Beaman and Cellinese 2012; König et al., 2019; Heberling et al., 2020). Consolidated herbarium databases are particularly informative for studies of invasive plant species, allowing an approximation over time and space of introduction history, rate of spread, phenotypic plasticity, and rapid evolution of traits implicated in their invasion success (Crawford and Hoagland 2009; Marisco et al., 2010; Gallinat et al., 2018). In the absence of direct experimental evidence, digitized specimens allow broad-scale comparison of similarities or differences in patterns of trait variation among populations between the native and invasive ranges (e.g. morphology, phenology, ecological interactions), as well as the timing of invasion, changes in abundance over time, associations with topography and land use, and overall consequences for biodiversity (e.g. Buswell et al. 2011, Hodgkins and Rieseberg, 2011).

Grasses (family Poaceae) are among the most important plants on Earth from economic and ecological perspectives (Linder et al., 2018), and are substantially overrepresented among invasive plant species in both agricultural and natural habitats (Daehler, 1998; Kerns et al., 2020). Invasive grasses pose serious threats to natural, agricultural, and populated areas, and have been described as ecosystem engineers capable of altering ecosystem processes including fire regimes, moisture availability, and community dynamics (D'Antonio and Vitousek, 1992; D'Antonio et al., 2017; Fusco et al., 2019, Kerns et al., 2020).

One trait thought to be a key driver of ecological and evolutionary success in grasses is the awn (Humphreys et al., 2011; Linder et al., 2018; McAllister et al., 2018). These bristle-like extensions of the florets perform a range of functions, for example, in photosynthesis, water stress, dispersal, deterrence of herbivory, and 'burial syndrome' (Humphreys et al., 2011; Cavanaugh et al., 2019; Ntakirutimana and Xie, 2020;

Cavanaugh et al., 2020). 'Burial syndrome' relies on a suite of features that assists the diaspore into the soil and increases germination and survival, including a geniculate (bent or twisted) awn morphology (Fig. 1a). These geniculate awns demonstrate hygroscopic ability and are sensitive to fluctuations in moisture, changing shape, and moving due to the absorption of water. Increased humidity causes the awn to straighten (Fig. 1a-c), and as the awn dries and curls, it drills the diaspore downward, entrapping it in the substrate (Peart, 1979; 1981; Peart and Clifford, 1987; Chambers and MacMahon, 1994; Chambers, 1995; Garnier and Dajoz, 2001). This capability can also propel the diaspores across the surface of the substrate, providing a means of dispersal from dry to moist microsites (Cavanaugh et al. 2020).

Awn morphology has played an important role in grass taxonomy (Hitchcock, 1950; Watson and Dallwitz, 2012; Kellogg, 2015). The presence of awned florets is variable across Poaceae, and can be polymorphic within species. One such species that displays a presence-absence polymorphism of the awn is *Microstegium vimineum* (Trin.) A. Camus (stiltgrass; Fig. 1), which is native to eastern Asia and highly invasive in eastern North America (Fairbrothers and Gray, 1972; Chen et al., 2007). This annual, C4 grass has become one of the most problematic invaders in eastern North America and is thought to have been introduced as dried packing material in porcelain shipments from Asia (Tu, 2000). It has a mixed mating system, with terminal chasmogamous inflorescences (wind-pollinated) and axillary, cleistogamous inflorescences (self-pollinated), allowing production of up to 1,000 seeds per plant (Huebner, 2003; Cheplick, 2005). First collected in Knoxville, Tennessee, US, in 1919, it has spread over the last century to 30 US states and southern Canada (Fairbrothers and Gray, 1972; EDDMapS, 2021). Numerous studies have been conducted on the ecological impacts of *M. vimineum* in the past two decades (e.g. MacGrath et al., 2009; Baker and Dyer, 2010; Flory and Clay, 2010; Huebner, 2010a, b; Rauschert et al., 2010; Barfknecht et al., 2020), yet a comprehensive analysis of its invasion history in eastern North America is lacking. Currently, little is understood about geographic patterns of awn polymorphism in the native and invasive ranges and the eco-evolutionary implications of its presence or absence.

We examined >1,100 digitized herbarium records from Asian and US databases to investigate the invasion history and broad-scale geographic patterns of awn polymorphism in *M. vimineum*. Patterns derived from these data represent a crucial, initial step in determining what adaptive significance this trait may have and its potential role in facilitating the establishment and spread of this invasive grass in eastern North America. Further, we included field collections from 31 US populations across the invasive range for genomic analyses, in order to corroborate patterns from digitized records. Our objectives were to investigate the relationships between geography (specifically latitude) and awn polymorphism in *M. vimineum* to: 1) reconstruct the invasion history in the US, 2) compare broad-scale geographic patterns of the frequency of awned and awnless forms in Asia vs. the US, 3) determine if transition zones reveal intermediate awn types, 4) quantify relationships between awn and floret measurements (e.g. awn length, floret length/width, awn bending), 5) compare the abiotic climate conditions that may contribute to the maintenance of awned and awnless forms, and 6) characterize genomic variation and patterns of genetic distinctness among awnless and awned forms in the invasive range.

MATERIALS AND METHODS

Awn data from digitized herbarium sheets

Awn presence/absence data were obtained from several digitized collections databases. For the US, herbarium records with images were searched via the Southeastern Regional Network of Expertise and Collections (SERNEC; https://sernecportal.org/portal). SERNEC comprises digitized collections from 233 herbaria in 14 US states, and further links to several other such consortia. Other searches were conducted using the C. V. Starr Virtual Herbarium (New York Botanical Garden, NY; http://sweetgum.nybg.org), the Smithsonian US National Herbarium (US; https://collections.nmnh.si.edu), and the Ohio State University Herbarium (OS; https://herbarium.osu.edu). Any data that were unavailable in the databases were transcribed from the image of the specimen label. Many specimens without latitude, longitude, or elevation data were georeferenced from the locality information provided, using Google Maps (www.maps.google.com); vague localities that could not be recorded with high confidence (*viz*. within 1 km) were excluded. Awn presence/absence was determined by examining the high-resolution

image of each specimen. Non-flowering specimens, misidentified specimens, and specimens for which awn presence-absence could not be determined with confidence were excluded.

The Chinese Virtual Herbarium (CVH) was searched for the majority of Asian specimens (https://www.cvh.ac.cn); the CVH is a consortium of over 100 Chinese herbaria containing over five million digitized specimens. Additional digitized specimens were obtained via the Global Biodiversity Information Facility (GBIF, https://www.gbif.org; https://doi.org/10.15468/dl.caar7n), the Herbarium of National Taiwan University (TAI; https://tai2.ntu.edu.tw), the Kagoshima (Japan) University Museum Database (KAG; https://www.museum.kagoshima-u.ac.jp), and the Moscow State University (Russia) Herbarium (MW; https://plant.depo.msu.ru). GPS coordinates and locality information could not be obtained reliably from most Asian records, due to a lack of locality information or language barriers on hand-written specimen labels. Therefore, Provincial or Regional centroids were used, and awn presence-absence data were summarized as the proportion of specimens with awns present from a particular Province/Region/State. All Russian records were centered closely around Sochi (Caucasus) and Vladivostok (eastern Siberia), so coordinates for those respective cities were used. Centroids were also recovered for US states, to allow direct comparison of Asian and US specimen data. While centroids obviously provide lower spatial resolution than georeferenced specimen coordinates, they allow for broad-scale, grouped comparisons of awn presence-absence by region. Centroids were obtained via the R package 'rworldmap' v.1.3.6 (South, 2011). All subsequent data processing and analyses were conducted in R (R Core Team, 2020). Records were filtered for incomplete data, keeping only unique records (removal of duplicates), using 'tidyverse' v.1.3.1 (Wickham et al., 2019) and 'dplyr' v.1.0.7 (Wickham et al., 2020). Data were then grouped in 'tidyverse' by State (US, India), Province (China), Region (Japan), or Island (Philippines, Sumatra, Taiwan).

Spatial distribution of awned and awnless records

Base maps for the eastern US and eastern Asia were generated in 'rworldmap.' The proportion of awned specimens per region and latitude/longitude centroids were specified as a data frame object in R, and pie graphs were placed on maps using the 'mapPies' function in 'rworldmap,' Plots of the proportion of awned specimens vs. latitude were

generated with 'ggplot2' v.3.3.5 (Wickham et al., 2016). Least-squares regression of the proportions of awned specimens vs. latitude (i.e. centroid latitude for each region) for Asian and US records was conducted in R. Asian and US specimens were treated separately to allow comparison of patterns between the native and invasive ranges.

Climate variables and awned forms in the US

In order to address the relationship between the presence-absence of awns and abiotic climate variation in the invasive range, we used publicly available data from WorldClim (https://www.worldclim.org/). The R libraries 'raster' v.3.5.2 and 'sp' v.1.4.5 (Pebesma and Bivand, 2005; Hijmans, 2020) were used to download data for the 19 BIOCLIM variables corresponding to GPS latitude/longitude data for all US specimens at 10 km resolution. In order to mitigate the potentially confounding effects of climate change over the past century, only specimens collected in the year 2000 onward were included, leaving a total of 207 records. The resulting data were subjected to Principal Components Analysis (PCA) in PAST v.4 (Hammer et al., 2001) employing a correlation matrix to account for different scales of the variables. A 'broken stick' analysis was applied to the Principal Components to determine the number of PC axes contributing significantly to the total variance in the dataset. The PCA was conducted using: 1) all BIOCLIM variables (n = 19), 2) temperature-related variables only (n = 11), and precipitation-related variables only (n = 8). Nonparametric Multivariate Analysis of Variance (NPMANOVA) was used to test for significant differences in BIOCLIM multivariate space among awned and awnless forms in PAST v.4. PCA loading scores were then exported from PAST and merged with the awn data frame in R. We used logistic regression to test the association between abiotic climate variables and awn presence-absence. Logistic regression was conducted in R among US records, regressing awn presence-absence as a binary response variable on the first principal component for temperature variables, precipitation variables, all BIOCLIM variables, and a model including temperature and precipitation with their interaction term. The model with the lowest Akaike Information Criterion score was used to predict the probability of being awned.

Mapping the invasion history of *M. vimineum*

Awn data from US specimens were used to create an animated.gif of invasion history using the R packages 'rgbif' v.3.6.0, 'ggplot2', 'tidyverse', 'gganimate' v.1.0.7, 'ggthemes' v.4.2.0, 'sf' v.0.9.5, 'tools' v.4.0.2, 'maps' v.3.3.0, 'rnaturalearth' v.0.1.0, and 'rnaturalearthdata' v.0.1.0 (Wickham et al., 2016; South, 2017a; 2017b; Pebesma, 2018; Chamberlain et al., 2020; Pedersen et al., 2020). The initial, filtered dataset was used for the animation, specifying the year collected as an integer to sequentially display the specimen data over time, beginning in 1919 and ending in 2019. An optional feature was added to query GBIF records in the R animation script using 'rgbif,' allowing reconstruction of the invasion history of any organism based on preserved records, observations, surveys, etc. The package 'gganimate' was used to produce the final animated figure. We also created a static representation of invasion history at six, roughly 15-year intervals: 1919-1935, 1936-1950, 1951-1965, 1966-1980, 1981-1995, and 1996-2020. Maps of awned and awnless records were plotted for each time interval with 'rworldmap,' 'ggplot2,' and summarized with 'ggpubr' v.0.4.0 (Kassambra, 2020). Lastly, invasion curves were generated for awnless and awned forms in 'ggplot2' using the cumulative summary function ('cumsum'), plotting unique localities over time, with the cumulative number of unique localities binned by year. R code is available at https://rpubs.com/cfb0001/771973.

Floret size, awn length, and intermediate forms

A subset of the US specimens was analyzed more closely to investigate patterns of floret size and awn length variation along a latitudinal 'transect' from Mississippi to New York (n = 73 total specimens examined). Specifically, forms were sought that displayed intermediate awn morphology between that of the awnless form and forms with long, exserted, geniculate awns. Only records with GPS coordinates and sufficient image resolution were chosen to minimize geographical uncertainty as well as error in measurements. Further, specimens were excluded in which multiple plants were present for a specimen and could not be differentiated. Only florets from a single plant were measured per record, with efforts made to measure florets from different inflorescences where possible, choosing the floret from the basal-most spikelet per raceme. Florets and awns were measured digitally with the ImageJ v.2 software (Rueden et al., 2017),

calibrating the number of pixels to 1.00 cm on the ruler provided for each specimen image. Measurements were taken from five florets per record, selected from racemes distributed along the length of the inflorescence, and only choosing florets that were clearly visible in their entirety. Four features were measured: 1) floret length (i.e. length of the upper lemma, excluding the awn); 2) total floret width at widest point; 3) length of the awn from point of emergence from the floret; and 4) angle of the geniculate awn, from 0-90°, where zero bending = 0° and full bending = 90°. Angle of awn bending was measured using the angle tool in ImageJ, approximated to the nearest 15° increment. All response variables as well as latitude were $\log_{10}$ transformed to mitigate the positive skew in the data. Relationships between latitude, awn type (awnless, short-awned, and long-awned; wherein 'long-awned' was determined by awns equal to or greater than the length of the floret) and floret measurements were calculated using a Generalized Linear Mixed-Model in R in the package 'lme4' ('glmer' function; Bates et al., 2015), accounting for non-independence among the five floret measurements per accession by treating these as a random effect. Tukey-corrected pairwise comparisons were conducted using the R package 'multcomp' (Hothorn et al., 2008). R code for all analyses is available at https://github.com/barrettlab/Awns-manuscript-R-code/wiki.

Genomic analyses of contemporary US collections

To assess broad-scale patterns of genetic relationships and population structure in *M. vimineum* across the invasive range, we sampled 51 individuals (Appendix S1, see the Supplementary Data with this article) from 31 localities. We used the CTAB method for DNA extraction (Doyle and Doyle, 1987) from 0.2 g silica-dried leaf tissue, one individual serving as a voucher from each locality. We used a recently developed protocol based on sequencing of inter-simple sequence repeat amplicons (ISSR-seq; Sinn, Simon et al., 2021). Briefly, gDNAs were quantified via Qubit Broad Range dsDNA assay (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and diluted to 20 ng/µl for amplification using four ISSR motifs in multiplex from UBC set #9 primers 848 [$(CA)_8RG$], 857 [$(AC)_8YG$], 868 [$(GAA)_6$], and 873 [$(GACA)_4$]. PCR reactions, library preparation, and Illumina NextSeq. 2000 sequencing (v3 chemistry, 2×100 bp paired end reads) were conducted as in Sinn, Simon, et al. (2021).

Barcoded read sets for each individual were run through a custom pipeline (https://github.com/btsinn/ISSRseq), which comprises a set of UNIX bash scripts to assemble a pseudoreference, clean and map reads to the reference, and call variants following Sinn, Simon, et al. (2021). The resulting filtered.vcf file was used in downstream analyses. Additionally, ANGSD (Korneliussen et al., 2014) was used to estimate genotype likelihoods from the 'HaplotypeCaller'.bam files produced by GATK (parameters: -GL 1 -nThreads 20 -doGlf 2 -doMajorMinor 1 -doMaf 2 -SNP_pval 2e-6). The R package 'phrynomics' was used to further filter the called variants, removing uninformative variants for downstream analysis (https://github.com/bbanbury/phrynomics-data). Relationships among individuals were estimated with SVDQuartets (Chifman and Kubatko, 2014; 2015), under a coalescent model and using all quartets, with 1,000 bootstrap replicates to assess branch support. The.vcf file was thinned to a single variant per locus with vcftools (-thin 3000 option, which is longer than the longest amplicon locus), to account for linkage among sites within loci, and converted to STRUCTURE format with PDGSpider (Lischer and Excoffier, 2012). Ancestry coefficients and group membership were estimated using parallelSTRUCTURE v.1.0 via the CIPRES Gateway (Miller et al., 2010), with K = 1-5, specifying 200,000 steps and the same number as burn-in (Besnier and Glover, 2013). Each run was replicated 10 times for each value of K, specifying the admixture model, without specifying population origin as a prior. StructureHarvester (Dent and vonHoldt, 2012) was used to estimate the optimal number of clusters using the 'delta K' method (Evanno et al. 2005). Clumpak (http://clumpak.tau.ac.il) was used to summarize results. Lastly, the ANGSD software v.0.935 was used to output pairwise distances between individuals based on genotype likelihoods (IBS matrix) and subjected to nonmetric multidimensional scaling (NMDS) to identify groupings among the accessions using the R package 'vegan' v.2.4-2 ('metaMDS' function; Oskasen et al., 2020).

RESULTS

Awn data from digitized herbarium records

A total of 1,145 specimen records remained after filtering records from a total of 2,913 available records, with 484 from Asia and 661 from the US (Appendix S2). In total, 717 were determined to be awnless, while 428 were awned, representing an overall proportion

of awned records to be 0.374. The proportion of awned records from Asia was 0.417, while the proportion from the US was 0.342.

Spatial distribution of awned and awnless proportions

In Asia, both awned and awnless records occur at low- and mid-latitudes (i.e. 0-40 degrees N), while awned records predominate above 40 degrees North latitude (Fig. 2). Records from northern China, northern Japan, Siberia, and the Caucasus are predominantly awned, representing the northernmost records in Asia. In the US (Fig. 3), there is a sharp latitudinal transition from predominantly awnless forms in the South (e.g. Alabama, Florida, Georgia, South Carolina, Tennessee) to awned forms in northern states (e.g. Massachusetts, New York, New Jersey, Pennsylvania). At mid-latitudes, e.g., Maryland, Ohio, Virginia, West Virginia, and North Carolina, there are various proportions of both forms.

Unweighted regressions revealed a significant, positive relationship between the proportion of awned specimens and latitude in both Asia (Table 1; Adjusted $R^2 = 0.135$, $p < 0.05$) and the US (Table 1; Adjusted $R^2 = 0.563$, $p < 0.01$). Further regressions of the proportion of awned specimens and latitude–weighted by sample sizes from each centroid–were also significant and positive for US (Table 1; Adjusted $R^2 = 0.673$, $p < 0.01$) and Asian records (Table 1; Adjusted $R^2 = 0.284$, $p < 0.01$). Overall there was a stronger positive relationship between the proportion of awned records and increasing latitude in the US, regardless of weighting, compared to Asian records.

Climate variables and awned forms in the US

Most BIOCLIM temperature variables had significant, negative correlations with latitude (Appendix S3), while Temperature Annual Range and Temperature Seasonality had significant, positive correlations (Pearson's $r = 0.74$ and 0.90, respectively; $p < 0.001$ for both). All precipitation variables were negatively correlated with latitude ($p < 0.01$ in all cases; Appendix S3). Principal Components Analysis of 19 BIOCLIM variables (US records only, post-2000) revealed that the first PC axis explained the majority of overall abiotic climate variation (Fig. 3; Appendix S4). This was also the case when analyzing temperature only (PC1 = 66.96%) and precipitation only (PC1 = 63.27%). One-way, nonparametric multivariate analysis of variance (NPMANOVA) revealed significant differences in multivariate climate space for awned and awnless US records ($F = 395.5$, p

< 0.0001). Logistic regression revealed a negative correlation between both temperature and precipitation and the proportion of awned individuals (p < 0.01 in both cases; Table 2; Fig. 3). Of the four models tested (temperature only, precipitation only, all BIOCLIM variables, and temperature × precipitation with interaction), the temperature × precipitation interaction model had the lowest Akaike Information Criterion score (119.953; Table 2). Both temperature and precipitation, and their interaction, were highly significant (p < 0.01). This model correctly classified the probability of being awned vs. awnless in 87.4% of cases.

Mapping the invasion history of *M. vimineum*

Animation 1 and Fig. 4 show the spread of *M. vimineum* in the eastern US, based on digital herbarium records (Animation 1, https://rpubs.com/cfb0001/771973). The first records of the invasion were of the awnless form, which appeared in 1919 (Knox County, Tennessee), followed by a gap in collections until 1931, when it was collected in Harlan County, Kentucky and Prince George's County, Virginia. By the late 1930s, the awnless form had spread to North Carolina, South Carolina, Georgia, northern Alabama, Kentucky, and southern Ohio. By 1950, it had spread throughout much of Kentucky, Virginia, Ohio, North Carolina, South Carolina, Georgia, and Alabama, while the awned form was restricted to eastern Pennsylvania, New Jersey, and with a single record in Virginia. By 1980, the awnless form had spread as far west as Arkansas and Louisiana, while the awned form had spread to Maryland, West Virginia, and North Carolina. The first awned records appeared in Berks County, eastern Pennsylvania, in 1938. Interestingly, awned records appeared to have remained concentrated in eastern Pennsylvania and New Jersey from the 1930s until the 1960s, with a few scattered records outside the region during this time period. By 1995, the awned form had spread to Connecticut, Massachusetts, eastern New York, and western Pennsylvania, and by 2020, it was present throughout much of Pennsylvania and southern New England, and as far west as Indiana. Invasion curve reconstructions (Fig. 4g) reveal apparent lag phases for both the awnless and awned forms, followed by linear expansion phases, representing spread from the points of origin. There is no apparent 'plateau phase' for either the awned or awnless form, suggesting continued invasive-range expansion to date.

Floret size, awn length, and intermediate forms

Investigation of 73 US records along a latitudinal gradient from Mississippi to New York revealed the presence of 'intermediate,' short-awned forms, between approximately 37° and 41° N latitude (Fig. 5). Short-awned records had a mean awn length of 0.169 cm (S.D. = 0.042 cm, range = 0.10-0.29 cm), whereas longer-awned records had a mean awn length of 0.464 cm (S.D. = 0.096, range = 0.27-0.89 cm). Table 3 summarizes the results of Generalized Linear Mixed models comparing floret and awn measurements. Awn lengths were significantly different among short-awned and long-awned records (Pseudo-R^2 = 0.88, $p < 0.001$). Notably, these intermediate-awned records appeared within a latitudinal 'transition zone' of overlap between awned and awnless records but were not observed outside of this zone of overlap (Fig. 5). The intermediate, short-awned records further displayed lower values of awn bending (i.e. awn angle) than those with longer awns (Fig. 5b; Pseudo-R^2 = 0.35, $p < 0.001$), but this did not vary significantly with latitude when awnless records were removed from the analysis. Floret length (i.e. lemma length, excluding the awn) displayed a positive relationship with latitude (Fig. 5c; Pseudo-R^2 = 0.65, $p < 0.05$), whereas floret width did not vary significantly with latitude. The length ratio of awns to florets (considering only awned records) further emphasized the presence of a transition zone between awned and awnless records (Fig. 5e). On average, awned records had significantly longer florets than awnless records, excluding the awn itself from measurements ($p < 0.001$). Specifically, the 'long' awned records had significantly longer florets than the short-awned or awnless records ($p < 0.001$; Tukey HSD $p < 0.05$ and $p < 0.001$, respectively), but short-awned records did not differ significantly in floret length from awnless records (Tukey HSD, $p = 0.494$).

Genomic relationships and population structure in the invasive range

Sequencing of ISSR amplicons yielded 4,714 final, called variants after conservative processing and filtering with GATK for 51 accessions from 31 localities (Appendix S5). After filtering non-parsimony-informative variants, 1,565 variants remained for phylogenetic analysis. Phylogenetic analysis in SVDQuartets under a coalescent model showed overall strong support (all branches had bootstrap values of 100), with two clearly defined clades corresponding to awned and awnless forms (Fig. 6a). One awnless accession from western North Carolina, US (NC SP 2-6) grouped closely with awned

accessions from Ohio and New York, US, but otherwise all awned and awnless accessions grouped together correspondingly. The NMDS plot (Fig. 6b) based on a distance matrix from genotype likelihoods for 39,198 variants yielded a similar overall pattern to the phylogenetic analysis, with awned and awnless accessions separated along NMDS axis 1; again, the exception was accession NC SP 2-6 from North Carolina. STRUCTURE analysis (Fig. 6c) further supported these findings, with K=2 as the optimal number of genomic clusters based on the 'delta K' method and the current sampling of accessions (K = 477.7; Appendix S6). Awned and awnless forms generally comprised distinct clusters, though there was some evidence for admixture, with the awnless NC SP 2-6 being a notable outlier. Specifically, there is limited evidence of admixture among the more northern, awned accessions, and the more southern, awnless accessions. A few awnless accessions from Kentucky, Illinois, Ohio, Virginia, and Tennessee showed some evidence of mixed ancestry among the two clusters, while some awned accessions from Pennsylvania, Virginia, New York, Ohio, Maryland, and West Virginia displayed similar patterns of mixed ancestry.

DISCUSSION

By analyzing >1,100 digital herbarium records of the invasive *Microstegium vimineum*, we were able to reconstruct the invasion history of this species in the US, detect evidence of at least two putative invasions, demonstrate strikingly similar geographical patterns of awn presence-absence polymorphism in the native and invasive ranges, and characterize patterns of floret and awn variation along a latitudinal gradient. The inclusion of genomic data from 51 contemporary samples demonstrates distinctness of awned and awnless forms, with some evidence of admixture, and supports the hypothesis of multiple introductions to the US. This study represents the most comprehensive use of digitized collections data for *M. vimineum* to date, underscoring the importance of herbaria and efforts to centralize digitized collections globally to begin to elucidate a plant's invasive success.

The potential role of awns in invasiveness of *M. vimineum*

Our data show a unique pattern of awn polymorphism, contributing to a novel hypothesis for its adaptive role in burial at higher latitudes. Würschum et al. (2020) found evidence of long and intermediate awn lengths in wheat at lower latitudes, and Li et al. (2015)

found that awn mass was negatively correlated with latitude, the opposite of our findings in *M. vimineum*, which they explain as either an adaptation to drier conditions or as having been artificially selected by growers. Grundbacher (1963) suggested that awns in more moist habitats may provide additional surface area for infection by soil-borne pathogens, thus potentially explaining the lack of awns at lower latitudes where precipitation is higher (Fig. 2). In the native grass *Sorghastrum nutans* (L.) Nash, shorter awns were observed in the southeastern US, while intermediate or longer awns were observed elsewhere, but the explanation for this pattern is unclear (Soper Gorden et al., 2016). There has been a significant amount of research on the adaptive value of awns in fire-prone regions, e.g. in African, Australian, and South American grasslands, with respect to increased survival and burial depth (e.g. Garnier and Dajoz, 2001; Johnson and Baruch, 2014). Individual diaspores with longer, hygroscopic awns typically have superior ability to find safe sites and to bury deeper in the soil, especially those with geniculate (twisted or bent) awns, thereby decreasing seed mortality (Cavanaugh et al., 2020), though the type of substrate is also important (smaller soil particle size results in greater burial depth; Molano-Flores 2012).

What then could explain the prevalence of awned forms in *M. vimineum* at higher latitudes in Asia and the US (Fig. 2)? Fire does not appear to be a logical explanation for the prevalence of awns in this species, as large fires are less frequent in the eastern US relative to drier regions in the western US, and actually have a higher incidence in the southeastern US (where awnless forms predominate) relative to the northeastern US and upper Midwest (e.g. Rudis and Skinner, 1991; Munn et al., 2003, LANDFIRE, 2012). In the US, precipitation differences between the Northeast and Southeast might explain this pattern, but these differences in precipitation are not extreme (i.e. the latitudinal pattern in precipitation is weaker than that for temperature; Appendix S3), and further, precipitation is heterogeneous throughout the eastern US as a result of longitude, elevation, and other factors (Peel et al., 2007).

One plausible hypothesis is that soil freezing may be a likely explanation for the north-south pattern in awn presence-absence, putatively providing a mechanism for habitat filtering post-introduction (*sensu* Weiher and Keddy, 1995). Though support in the literature for this geographic pattern is limited, there is evidence of awns being used for

seed burial in alpine systems and a greater likelihood of seeds with awns being found in alpine seedbanks (Chambers, 1995; Welling et al., 2004). In addition, buried seed of the awned grass *Bromus setifolius* var. *pictus* (Hook. f.) Skottsb. in a Pantagonian steppe (temperatures average around 3°C) had greater survival as seedlings than unburied seed (Rotundo and Aguiar 2004). Plants have evolved various mechanisms to survive freezing temperatures at higher latitudes and altitudes (Ambroise et al., 2020). As an annual grass, reliance upon mechanisms such as underground perennating organs is not applicable for *M. vimineum*, which relies completely on recruitment from the seed bank (Gibson et al., 2002; Redwood et al., 2018). Thus, to survive the more intense freezing cycles experienced at higher latitudes, seeds or seed dispersal units must be adapted to survive extreme conditions.

Alternatively, alleles associated with awn-formation may be linked to others that could better explain the north-south latitude pattern. For instance, silica deposition genes or their expression have been associated with awn development or expression (Peleg et al., 2010; Yamaji et al., 2012; Ntakirutimana and Xie, 2019). Silica in plants is important for structural support and defense against fungal pathogens and herbivory (Currie and Perry, 2007). Silica also slows plant tissue decomposition in some grasses (Miyake and Takahashi, 1983; Schaller et al., 2014). Theoretically, buried seed, especially over long periods, may need more protection from fungal pathogens, whereas some populations with extended growing seasons and large individuals may benefit more from increases in available nutrients in response to rapid decomposition. Plants with greater silica and awns may further lead to formation of a longer-term seed bank and lack of silica and awns may equate to a shorter-term seed bank and subsequent differences in seed longevity and/or dormancy.

To our knowledge, the geographic pattern of awn polymorphism in *M. vimineum* is the only case where such a north-south, awned-awnless pattern has been documented in grasses (e.g. Soper Gorden et al., 2016). Similar investigations of awns in other grass species spanning subtropical-temperate ranges, including invasive grasses, should be conducted to establish corroborating evidence for this pattern. Awns serve a variety of functions in grasses, and these roles are highly context-dependent (e.g. Humphreys et al., 2011; Schrager-Lavelle et al., 2017). Therefore, if awns are indeed adaptive for soil

burial at higher latitudes, this would represent a novel ecological context for the function of this trait and its role in invasion success. The fact that intermediate, short-awned forms were observed at relative low frequency at mid-latitudes suggests that awned and awnless forms intergrade where their ranges overlap (Fig. 5), that this trait is likely under genetic control, and further, that there may be selection against intermediate forms at both higher and lower latitudes, i.e. reinforcement (Dobzhansky, 1940; Harrison, 1993; Coyne and Orr, 1998; Curry, 2015). Clearly, additional study is needed to quantify the contribution of awns or other (possibly linked) traits to invasion success in this species.

Invasion history of *M. vimineum* via herbarium records

Our analysis of digitized herbarium records in the eastern US provides historical evidence of at least two separate invasions of *M. vimineum*. The initial invasion likely happened somewhere in the southeastern US (based on the initial 1919 Tennessee collection and subsequent collections in Tennessee, Kentucky, Virginia, North Carolina, and northern Alabama; Fig. 4). It is impossible to determine exactly when or where this initial introduction occurred, and it could have occurred much earlier than 1919, perhaps in the late 1800s. A secondary invasion in the 1930s in eastern Pennsylvania is evident, with only localized initial spread there and in neighboring New Jersey (Fig. 4). The gap between the initial collection by Ainslie (1919, Knox County, Tennessee) and both of the 1931 collections by Braun and Blake (Harlan County, Kentucky and Prince George County, Virginia, respectively) may be the result of low initial, post-introduction abundance, lack of familiarity with *M. vimineum* by collectors at that time, or separate, additional introductions (Fairbrothers and Gray, 1972).

The second putative introduction (i.e. awned form), and to a lesser degree the first introduction (awnless form), display characteristic 'lag phases,' in which they appear to establish locally or regionally prior to a more rapid expansion (Fig. 4a,b,g; Hengveld, 1988; 1989; Hobbs and Humphreys, 1995). The lag phase can have different causes, related to demography, changing environmental conditions, or genetic potential (Crooks and Soulé, 1999). While it is impossible to discern what the causes of the lag phase would have been for each of these putative introductions, investigation of genomic variation from herbarium samples, in combination with data on past environments and

the temporal dynamics of anthropogenic disturbance over the last century may shed light on this observation.

Genomic analysis of 51 accessions from contemporary field collections across the eastern US displays a strong pattern of differentiation among awnless and awned forms (Fig. 6). This finding provides further evidence for at least two separate invasions of *M. vimineum* in the US. There is also limited evidence of admixture between the two forms in some accessions, most notably in one awnless accession from Mitchell County, in northwestern North Carolina (NC SP 2-6), but also in several others. Novy et al. (2013) provided evidence of clinal variation in phenology and vegetative biomass at anthesis with latitude, with the main conclusion being that this invasive species has rapidly adapted within the introduced range over the past century. Given the annual life history of *M. vimineum*, this provides >100 generations over which evolution could have occurred. As an alternative explanation for this pattern, Novy et al. (2013) stated: "IBD [isolation by distance] could result from the northern phenotype (i.e. shorter time to flowering at lower biomass) being introduced in the northern United States and the southern phenotype (i.e. longer time to flowering at higher biomass) being introduced in the southern United States." While rapid evolution in flowering time could have evolved over a century resulting from a single introduction of an invasive annual plant (given there was sufficient genetic variation present), our findings suggest that the alternative scenario posited by Novy et al. (2013) is equally plausible, and possibly not mutually exclusive of their primary explanation of this pattern.

A regional genetic study of populations in northern and eastern West Virginia, US (Culley et al., 2016) suggested that while a significant amount of population structure exists for *M. vimineum*, there is also evidence for long- or at least intermediate-distance dispersal, likely mediated by human activity or movement via water (Huebner, 2007; Christen and Matlack, 2009; Huebner, 2010; Tekiela and Barney, 2013). The persistence of two distinct forms—awnless in the southeastern US, and awned in the northern US—suggests the possibility that both invasion history and selection for awned/awnless forms at higher/lower latitudes may have played roles in the current distribution and expansion of this species, though experimental evidence is needed to determine whether selection via habitat filtering (i.e. pre-adapted invasion) has played a role. The rationale here is that

even if, for example, long-distance seed dispersal is common in the invasive range, selection could counteract the predominantly homogenizing effects of gene flow, reinforcing the existence of awned forms at high latitudes and awnless forms at low latitudes, removing maladaptive alleles introduced by seed dispersal but maintaining new genotypic combinations that allow local adaptation (e.g. Edelaar and Bolnick, 2012). Admixture is now well known as a potential source of adaptive variation among multiple source introductions of invasive species (Dlugosch and Parker, 2007; Keller and Taylor, 2010; Keller et al., 2014; van Boheemen et al., 2017, 2018; Lachmuth et al., 2019), and has possibly played a role in local adaptation and continued range expansion of *M. vimineum* in North America. Following the rationale of Novy et al. (2013), if two successful invasions have indeed occurred (southern awnless and northern awned, respectively), then subsequent admixture and allelic shuffling could potentially allow the expression of novel phenotypes associated with phenology, physiology, and morphology, allowing 'fine tuning' of localized adaptation. Thus, if introductions of *M. vimineum* were frequent enough over time, then the invasion success of this species could at least partially be explained by a combination of habitat filtering and propagule pressure (Weiher and Keddy, 1995; Lockwood et al., 2005). In other words, there may have been many opportunities to "get it right" in terms of each initial invasion being preadapted to some degree—especially if low genetic diversity was associated with introduction bottlenecks—where subsequent spread, admixture, and selection have all contributed to invasion success.

CONCLUSIONS

We present the first comprehensive analysis of invasion history and intraspecific awn polymorphism in *M. vimineum*, relying on >1,100 digitized herbarium records from centralized databases. We demonstrate historical and contemporary-genomic evidence of at least two separate invasions in the US, a scenario that is increasingly becoming apparent in invasion biology (e.g. Keller and Taylor, 2010; Vallejo-Marín et al., 2021; Sutherland et al., 2021). A combination of habitat filtering, *M. vimineum* being an 'ideal weed' (in reference to awn traits), propagule pressure (i.e. the possibility of several introductions), and other factors may explain the success of this invasive species, and represent plausible hypotheses for future investigation. Though *M. vimineum* has several

plastic traits, these data support multiple introductions of the two awn-forms and possible subsequent habitat filtering instead of plasticity of the awn trait. While we hypothesize that habitat filtering may explain the success of the awned form in colder climates, more research is needed to explain the potential adaptive advantage of being awnless in warmer climates. Additional genomic analyses of material from herbarium specimens across space and time will allow more nuanced reconstruction of invasion history and determination of invasion mechanisms, further providing insights into any adaptive processes at the genomic level, i.e., by evaluating the relatedness of co-occurring awn types and estimating how long it takes awnless and awned forms to become extant in colder and warmer climates, respectively, thereby determining a rate of habitat filtering or competitive ability between the two forms. The current study further highlights the need for experimental studies, specifically exploring the role of awns and other potential trait polymorphisms in invasiveness. For example, seed bank germination studies or burial experiments using awned, awnless, and intermediate awned forms under different soil temperature/moisture regimes could be informative on fitness metrics such as germination rates, seed bank accumulation and viability over time, the effect of burial depth, differences in dormancy, and seedling survival to reproductive stage (e.g. Ramos et al., 2019). These data could also be used to better incorporate site invasibility with plant invasiveness by correlating local landscapes and physiography with invasion history, perhaps highlighting patterns along rivers and other corridors. Ultimately, what we have learned from the use of historical records of *M. vimimeum* can be applied to other established plant invaders and may help us predict and prevent future invasions.

ACKNOWLEDGEMENTS

Percifield, Donald Primerano, and Jun Fan. We thank WVU Genomics Core Facility for support provided to help make this publication possible and CTSI Grant #U54 GM104942 which in turn provides financial support to the WVU Core Facility. We further acknowledge WV-INBRE (P20GM103434), a COBRE ACCORD grant (1P20GM121299), and a West Virginia Clinical and Translational Science Institute (WV-CTSI) grant (2U54GM104942) in supporting the Marshall University Genomics Core (Research Citation: Marshall University Genomics Core Facility, RRID:SCR_018885).

CONFLICT OF INTEREST

There are no conflicts of interest to report.

AUTHORS' CONTRIBUTIONS

CFB and CDH conceived the manuscript; CFB conducted all data analyses and wrote the manuscript; CDH, ZAB, TAB, MVS, HMT, SVS, and ANC compiled data and helped write the manuscript; ML, MRM, and MM provided helpful discussion and helped write the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

Data extracted from herbarium image databases and genomic.vcf files (filtered and linkage disequilibrium-thinned variants) are provided at https://doi.org/10.5281/zenodo.6384405. Raw genomic data are provided in the Sequence Read Archive as FASTQ files (https://www.ncbi.nlm.nih.gov/sra) under BioProject PRJNA773774. R code used for this paper can be found at https://rpubs.com/cfb0001 and https://github.com/barrettlab/Awns-manuscript-R-code/wiki.

ORCID

Craig F. Barrett: https://orcid.org/0000-0001-8870-3672

REFERENCES

Ambroise, V., Legay, S., Guerriero, G., Hausman, J.-F., Cuypers, A., and Sergeant, K. 2020. The roots of plant frost hardiness and tolerance. *Plant and Cell Physiology* **61**: 3–20. https://doi.org/10.1093/pcp/pcz196

Barfknecht, D. F., Gibson, D. J., and K. M. Neubig. 2020. Plant community and phylogenetic shifts in acid seep springs over 49 years following *Microstegium vimineum* invasion. *Plant Ecology* **221**: 167–175. https://doi.org/10.1007/s11258-020-01002-7

Bates, D., Mächler, M., Bolker, B., and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**: 1–48. https://doi.org/10.18637/jss.v067.i01

Beaman R. S., and N. Cellinese. 2012. Mass digitization of scientific collections: New opportunities to transform the use of biological specimens and underwrite biodiversity science. *Zookeys* **209**: 7–17. https://doi.org/10.3897/zookeys.209.3313

Besnard, G., Gaudeul, M., Lavergne, S., Muller, S., Rouhan, G., Sukhorukov, A. P., Vanderpoorten, A., and F. Jabbour. 2018. Herbarium-based science in the twenty-first century. *Botany Letters* **165**: 323–327. https://doi.org/10.1080/23818107.2018.1482783

Besnier F., and K. A. Glover. 2013. ParallelStructure: A R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. *PLoS ONE* **8**: e70651. https://doi.org/10.1371/journal.pone.0070651

Borges, L. M., Reis, V. C., and R. Izbicki. 2020. Schrödinger's phenotypes: Herbarium specimens show two-dimensional images are both good and (not so) bad sources of morphological data. *Methods in Ecology and Evolution* **11**: 1296–1308. https://doi.org/10.1111/2041-210X.13450

Buswell, J. M., Moles, A. T., and S. Hartley. 2011. Is rapid evolution common in introduced plant species? *Journal of Ecology* **99**: 214–224. https://doi.org/10.1111/j.1365-2745.2010.01759.x

Carranza-Rojas, J., Goeau, H., Bonnet, P., Mata-Montero, E., and A. Joly. 2017. Going deeper in the automated identification of herbarium specimens. *BMC Evolutionary Biology* **17**: 181. https://doi.org/10.1186/s12862-017-1014-z

Cavanagh, A. M., Godfree, R. C., Morgan, J. W., Cavanagh, A. M., Godfree, R. C., and J. W. Morgan. 2019. An awn typology for Australian native grasses (Poaceae). *Australian Journal of Botany* **67**: 309–334. https://doi.org/10.1071/BT18216

Cavanagh, A. M., Morgan, J. W., and R. C. Godfree. 2020. Awn morphology influences dispersal, microsite selection and burial of Australian native grass diaspores. *Frontiers in Ecology and Evolution* **8**: 389. https://doi.org/10.3389/fevo.2020.581967

Chamberlain, S., Barve, V., McGlinn, D., Oldoni, D., Desmet, P., Geffert, L., and K. Ram. 2021. rgbif: Interface to the Global Biodiversity Information Facility API. R package version 3.3.0. URL: https://github.com/ropensci/rgbif

Chambers, J. C. 1995. Relationships between seed fates and seedling establishment in an alpine ecosystem. *Ecology* **76**: 2124–2133. https://doi.org/10.2307/1941687

Chambers, J. C., and J. A. MacMahon. 1994. A day in the life of a seed: Movements and fates of seeds and their implications for natural and managed systems. *Annual Review of Ecology and Systematics* **25**: 263–292. https://doi.org/10.1146/annurev.es.25.110194.001403

Cheplick, G. P. 2005. Biomass partitioning and reproductive allocation in the invasive, cleistogamous grass *Microstegium vimineum*: Influence of the light environment. The *Journal of the Torrey Botanical Society* **132**: 214–224. https://doi.org/10.3159/1095-5674(2005)132[214:BPARAI]2.0.CO;2

Chifman, J., and L. Kubatko. 2014. Quartet inference from SNP data under the coalescent. *Bioinformatics* **30**: 3317–3324. https://doi.org/10.1093/bioinformatics/btu530

Chifman, J., and L. Kubatko. 2015. Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of Theoretical Biology* **374***:* 35–47. https://doi.org/10.1016/j.jtbi.2015.03.006

Coyne, J. A., and H. A. Orr. 1998. The evolutionary genetics of speciation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **353**: 287–305. https://doi.org/10.1098/rstb.1998.0210

Crawford, P. H. C., and B. W. Hoagland. 2009. Can herbarium records be used to map alien species invasion and native species expansion over the past 100 years? *Journal of Biogeography* **36**: 651–661. https://doi.org/10.1111/j.1365-2699.2008.02043.x

Culley, T. M., Huebner, C. D., and A. Novy. 2016. Regional and local genetic variation in Japanese stiltgrass (*Microstegium vimineum*). *Invasive Plant Science and Management* **9**: 96–111. https://doi.org/10.1614/IPSM-D-15-00055.1

Currie, H.A., and C. C. Perry. 2007. Silica in plants: biological, biochemical and chemical studies. *Annals of Botany* **100**:, 1383–1389. https://doi.org/10.1093/aob/mcm247

Curry, C. M. 2015. An integrated framework for hybrid zone models. *Evolutionary Biology* **42**: 359–365. https://doi.org/10.1007/s11692-015-9332-9

D'Antonio, C. M., and P. M. Vitousek. 1992. Biological invasions by exotic grasses, the grass/fire cycle, and global change. *Annual Review of Ecology and Systematics* **23**: 63–87. https://doi.org/10.1146/annurev.es.23.110192.000431

D'Antonio, C. M., Yelenik, S. G., and M. C. Mack. 2017. Ecosystem vs. community recovery 25 years after grass invasions and fire in a subtropical woodland. *Journal of Ecology*, **105**: 1462–1474. https://doi.org/10.1111/1365-2745.12855

Daehler, C. C. 1998. The taxonomic distribution of invasive angiosperm plants: Ecological insights and comparison to agricultural weeds. *Biological Conservation* **84**: 167–180. https://doi.org/10.1016/S0006-3207(97)00096-7

Dlugosch, K. M., and I. M. Parker. 2008. Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* **17**: 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *The American Naturalist* **74**: 312–321. https://doi.org/10.1086/280899

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Earl, D. A. and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361 https://doi.org/10.1007/s12686-011-9548-7

Edelaar, P., and D. I. Bolnick. 2012. Non-random gene flow: An underappreciated force in evolution and ecology. *Trends in Ecology and Evolution* **27**: 659–665. https://doi.org/10.1016/j.tree.2012.07.009

Enders, M., Hütt, M.-T., and J. M. Jeschke. 2018. Drawing a map of invasion biology based on a network of hypotheses. *Ecosphere* **9**: e02146. https://doi.org/10.1002/ecs2.2146

Enders, M., Havemann, F., Ruland, F., Bernard-Verdier, M., Catford, J. A., Gómez-Aparicio, L., Haider, et al. 2020. A conceptual map of invasion biology: Integrating hypotheses into a consensus network. *Global Ecology and Biogeography* **29**: 978–991. https://doi.org/10.1111/geb.13082

Evanno, G., Regnaut, S. and J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**: 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fairbrothers, D. E., and J. R. Gray. 1972. *Microstegium vimineum* (Trin.) A. Camus (Gramineae) in the United States. *Bulletin of the Torrey Botanical Club*, **99**: 97. https://doi.org/10.2307/2484205

Flory, S. L., and K. Clay. 2010. Non-native grass invasion suppresses forest succession. *Oecologia* **164**: 1029–1038. https://doi.org/10.1007/s00442-010-1697-y

Fusco, E. J., Finn, J. T., Balch, J. K., Nagy, R. C., and B. A. Bradley. 2019. Invasive grasses increase fire occurrence and frequency across US ecoregions. *Proceedings of the National Academy of Sciences of the United States of America* **116**: 23594-23599. https://dx.doi.org/10.1073%2Fpnas.1908253116

Gallinat, A. S., Russo, L., Melaas, E. K., Willis, C. G., and R. B. Primack. 2018. Herbarium specimens show patterns of fruiting phenology in native and invasive plant species across New England. *American Journal of Botany* **105**: 31–41. https://doi.org/10.1002/ajb2.1005

Garnier, L. K. M., and I. Dajoz. 2001. Evolutionary significance of awn length variation in a clonal grass of fire-prone savannas. *Ecology* **82**: 1720–1733. https://doi.org/10.1890/0012-9658(2001)082[1720:ESOALV]2.0.CO;2

Gibson, D. J., Spyreas, G., and J. Benedict. 2002. Life history of *Microstegium vimineum* (Poaceae), an invasive grass in southern Illinois. *Journal of the Torrey Botanical Society*, **129**: 207. https://doi.org/10.2307/3088771

Grundbacher, F. 1963. The physiological function of the cereal awn. *The Botanical Review* **29***:* 366–381.

Guralnick, R. P., Zermoglio, P. F., Wieczorek, J., LaFrance, R., Bloom, D., and L. Russell. 2016. The importance of digitized biocollections as a source of trait data and a new VertNet resource. https://doi.org/10.1093/database/baw158

Hammer, Ø., Harper, D. A. T., and P. D. Ryan. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4**: 9. http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Harrison, R. G. 1993. Hybrid zones and the Evolutionary Process. 364 pp. Oxford University Press, Oxford, U. K.

Heberling, J. M. 2022. Herbaria as big data sources of plant traits. *International Journal of Plant Sciences* **183**: 87-118. https://doi.org/10.1086/717623

Heberling, J. M., Miller, J. T., Noesgaard, D., Weingart, S. B., and D. Schigel. 2021. Data integration enables global biodiversity synthesis. *Proceedings of the National Academy of Sciences* **118**: e2018093118. https://doi.org/10.1073/pnas.2018093118

Hedrick, B. P., Heberling, J. M., Meineke, E. K., Turner, K. G., Grassa, C. J., Park, D. S., Kennedy, J., et al. 2020. Digitization and the future of natural history collections. *BioScience* **70**: 243–251. https://doi.org/10.1093/biosci/biz163

Hijmans, R. J. 2020. raster: Geographic data analysis and modeling. R package version 3.0-12. https://CRAN.R-project.org/package=raster

Hodgins, K. A., and L. Rieseberg. 2011. Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *Journal of Evolutionary Biology* **24**: 2731–2749. https://doi.org/10.1111/j.1420-9101.2011.02404.x

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

Huebner, C. D. 2003. Vulnerability of oak-dominated forests in West Virginia to invasive exotic plants: temporal and spatial patterns of nine exotic species using herbarium records and land classification data. *Castanea* **68**: 1-14.

Huebner, C. D. 2007. Strategic management of five deciduous forest invaders using *Microstegium vimineum* as a model species. In*:* Cavender, N. D. (Ed), *Proceedings of the 2007 Ohio Invasive Plants Research Conference: Continuing Partnerships for Invasive Plant Management; 2007 January 18; Delaware, OH, Ohio Biological Survey:* 19-28. https://www.fs.usda.gov/treesearch/pubs/14215

Huebner, C. D. 2010a. Establishment of an invasive grass in closed-canopy deciduous forests across local and regional environmental gradients. *Biological Invasions* **12**: 2069–2080. https://doi.org/10.1007/s10530-009-9609-y

Huebner, C. D. 2010b. Spread of an invasive grass in closed-canopy deciduous forests across local and regional environmental gradients. *Biological Invasions* **12**: 2081–2089. https://doi.org/10.1007/s10530-009-9610-5

Humphreys, A. M., Antonelli, A., Pirie, M. D., and H. P. Linder. 2011. Ecology and evolution of the diaspore "burial syndrome." *Evolution* **65**: 1163–1180. https://doi.org/10.1111/j.1558-5646.2010.01184.x

Jeschke, J. M., and T. Heger. 2018. Invasion Biology: Hypotheses and Evidence. 188 pp. CABI.

Johnson, E. E., and Z. Baruch. 2014. Awn length variation and its effect on dispersal unit burial of *Trachypogon spicatus* (Poaceae). *Revista De Biologia Tropical* **62**: 321–326.

Kassambara, A. 2020. ggpubr: 'ggplot2' based publication ready plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr

Keller, S. R., and D. R. Taylor. 2010. Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology* **23**: 1720–1731. https://doi.org/10.1111/j.1420-9101.2010.02037.x

Keller, S. R., Fields, P. D., Berardi, A. E., and D. R. Taylor. 2014. Recent admixture generates heterozygosity-fitness correlations during the range expansion of an invading species. *Journal of Evolutionary Biology* **27**: 616–627. https://doi.org/10.1111/jeb.12330

Kerns, B. K., Tortorelli, C., Day, M. A., Nietupski, T., Barros, A. M. G., Kim, J. B., and M. A. Krawchuk. 2020. Invasive grasses: A new perfect storm for forested ecosystems? *Forest Ecology and Management* **463**: 117985. https://doi.org/10.1016/j.foreco.2020.117985

König, C., Weigelt, P., Schrader, J., Taylor, A., Kattge, J., and H. Kreft. 2019. Biodiversity data integration—The significance of data resolution and domain. *PLOS Biology*, **17**: e3000183. https://doi.org/10.1371/journal.pbio.3000183

Korneliussen, T.S., Albrechtsen, A. and R. Nielsen. 2014. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**: 356. https://doi.org/10.1186/s12859-014-0356-4

Lachmuth, S., Molofsky, J., Milbrath, L., Suda, J., and S. R. Keller. 2019. Associations between genomic ancestry, genome size and capitula morphology in the invasive meadow knapweed hybrid complex (*Centaurea × moncktonii*) in eastern North America. *AoB PLANTS* **11**: https://doi.org/10.1093/aobpla/plz055

Lan, Z., Su, Z., Guo, M. C., Alvarado, E., Guo, F., Hu, H., and G. Wang. 2021. Are climate factors driving the contemporary wildfire occurrence in China? *Forests* **12**: 392. https://doi.org/10.3390/f12040392

LANDFIRE: Fire Regime Group, U.S. Department of Agriculture and U.S. Department of the Interior. Accessed 21 Oct 2021 at https://landfire.gov/geoareasmaps/2012/CONUS_FRG_c12.pdf

Li, X., Yin, X., Yang, S., Yang, Y., Qian, M., Zhou, Y., Zhang, C., et al. 2015. Variations in seed characteristics among and within *Stipa purpurea* populations on the Qinghai-Tibet Plateau. *Botany* **93**: 651–662. https://doi.org/10.1139/cjb-2015-0011

Linder, H. P., Lehmann, C. E. R., Archibald, S., Osborne, C. P., and D. M. Richardson. 2018. Global grass (Poaceae) success underpinned by traits facilitating colonization, persistence and habitat transformation: Grass success. *Biological Reviews*, **93**: 1125–1144. https://doi.org/10.1111/brv.12388

Lischer, H. E. L., and L. Excoffier. 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**: 298–299. https://doi.org/10.1093/bioinformatics/btr642

Little, D. P., Tulig, M., Tan, K. C., Liu, Y., Belongie, S., Kaeser-Chen, C., Michelangeli, F. A., et al. 2020. An algorithm competition for automatic species identification from herbarium specimens. *Applications in Plant Sciences* **8**:. https://doi.org/10.1002/aps3.11365

Lockwood, J. L., Cassey, P., and T. Blackburn. 2005. The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution* **20**: 223–228. https://doi.org/10.1016/j.tree.2005.02.004

Marsico, T. D., Burt, J. W., Espeland, E. K., Gilchrist, G. W., Jamieson, M. A., Lindström, L., Roderick, G.K., et al. 2010. Underutilized resources for studying the evolution of invasive species during their introduction, establishment, and lag phases. *Evolutionary Applications* **3**: 203–219. https://doi.org/10.1111/j.1752-4571.2009.00101.x

McAllister, C. A., McKain, M. R., Li, M., Bookout, B., and E. A. Kellogg. 2019. Specimen-based analysis of morphology and the environment in ecologically dominant grasses: The power of the herbarium. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**: 20170403. https://doi.org/10.1098/rstb.2017.0403

McGrath, D. A., and M. A. Binkley. 2009. *Microstegium vimineum* invasion changes soil chemistry and microarthropod communities in Cumberland plateau forests. *Southeastern Naturalist* **8**: 141–156. https://doi.org/10.1656/058.008.0113

Miller, M.A., Pfeiffer, W., and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010*, 1–8.

Miyake, Y. and E. Takahashi. 1983. Effect of silicon on the growth of solution-cultured cucumber plant. *Soil Science and Plant Nutrition* **29**: 71-83. https://doi.org/10.1080/00380768.1983.10432407

Munn, I. A., Zhai, Y., and D. L. Evans. 2003. Modeling forest fire probabilities in the South-Central United States using FIA Data. *Southern Journal of Applied Forestry* **27**: 11–17. https://doi.org/10.1093/sjaf/27.1.11

Novy, A., Flory, S. L., and J. M. Hartman. 2013. Evidence for rapid evolution of phenology in an invasive grass. *Journal of Evolutionary Biology* **26**: 443–450. https://doi.org/10.1111/jeb.12047

Ntakirutimana, F., and W. Xie. 2019. Morphological and genetic mechanisms underlying awn development in monocotyledonous grasses. *Genes* **10**: 573. https://doi.org/10.3390/genes10080573

Ntakirutimana, F., and W. Xie. 2020. Unveiling the actual functions of awns in grasses: from yield potential to quality traits. *International Journal of Molecular Sciences* **21**: 7593. https://doi.org/10.3390/ijms21207593

Oksanen, J., Blanchet F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., and H. Wagner. 2020. vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Page, L. M., MacFadden, B. J., Fortes, J. A., Soltis, P. S., and G. Riccardi. 2015. Digitization of biodiversity collections reveals the biggest data on biodiversity. *BioScience* **65**: 841–842. https://doi.org/10.1093/biosci/biv104

Pearson, K. D., Nelson, G., Aronson, M. F. J., Bonnet, P., Brenskelle, L., Davis, C. C., Denny, E. G., et al. 2020. Machine learning using digitized herbarium specimens to advance phenological research. *BioScience* **70**: 610–620. https://doi.org/10.1093/biosci/biaa044

Peart, M. H. 1979. Experiments on the biological significance of the morphology of seed-dispersal units in grasses. *Journal of Ecology* **67**: 843. https://doi.org/10.2307/2259218

Peart, M. H. 1981. Further experiments on the biological significance of the morphology of seed-dispersal units in grasses. *Journal of Ecology* **69**: 425. https://doi.org/10.2307/2259677

Peart, M. H., and H. T. Clifford. 1987. The influence of diaspore morphology and soil-surface properties on the distribution of grasses. *Journal of Ecology* **75**: 569. https://doi.org/10.2307/2260435

Pebesma, E. 2018. Simple features for R: standardized support for spatial vector data. *The R Journal* **10**: 439–446. https://doi.org/10.32614/RJ-2018-009

Pebesma, E. J., and R. S. Bivand. 2005. Classes and methods for spatial data in R. *R News* **5**: https://cran.r-project.org/doc/Rnews/

Pedersen, T. L., and D. Robinson. 2020. Gganimate: A grammar of animated graphics. https://CRAN.R-project.org/package=gganimate

Peel, M. C., Finlayson, B. L., and T. A. McMahon. 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* **11**: 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Peleg, Z., Saranga, Y., Fahima, T., Aharoni, A. and R. Elbaum. 2010. Genetic control over silica deposition in wheat awns. *Physiologia plantarum* **140**: 10–20. https://doi.org/10.1111/j.1399-3054.2010.01376.x

R Core Team. 2020. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL: https://www.R-project.org/

Ramos, D. M., Valls, J. F. M., Borghetti, F., and M. K. J. Ooi. 2019. Fire cues trigger germination and stimulate seedling growth of grass species from Brazilian savannas. *American Journal of Botany* **106**: 1190-1201. https://doi.org/10.1002/ajb2.1345

Rauschert, E. S. J., Mortensen, D. A., Bjørnstad, O. N., Nord, A. N., and N. Peskin. 2010. Slow spread of the aggressive invader, *Microstegium vimineum* (Japanese stiltgrass). *Biological Invasions* **12**: 563–579. https://doi.org/10.1007/s10530-009-9463-y

Redwood, M. E., Matlack, G. R., and C. D. Huebner. 2018. Seed longevity and dormancy state suggest management strategies for garlic mustard (*Alliaria petiolata*) and Japanese stiltgrass (*Microstegium vimineum*) in deciduous forest sites. *Weed Science* **66**: 190–198. https://doi.org/10.1017/wsc.2017.74

Richards, C. L., Bossdorf, O., Muth, N. Z., Gurevitch, J., and M. Pigliucci. 2006. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* **9**: 981–993. https://doi.org/10.1111/j.1461-0248.2006.00950.x

Rotundo, J. L. and M. R. Aguiar. 2004. Vertical seed distribution in the soil constrains regeneration of *Bromus pictus* in a Patagonian steppe. *Journal of Vegetation Science* **15**: 515-522. https://doi.org/10.1111/j.1654-1103.2004.tb02290.x

Rudis, V. A., and T. V. Skinner. 1991. Fire's importance in South Central U.S. forests: Distribution of fire evidence. *Fire and the Environment: Ecological and Cultural*

Perspectives; Proceedings of an International Symposium; 1990 March 20-24; Knoxville, TN. (In GTR-SE69). https://www.fs.usda.gov/treesearch/pubs/506

Rueden, C. T., Schindelin, J., and Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., and K. W. Eliceiri. 2017. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18**: 529. https://doi.org/10.1186/s12859-017-1934-z

Schaller, J., Hines, J., Brackhage, C., Bäucker, E., and M. O. Gessner. 2014. Silica decouples fungal growth and litter decomposition without changing responses to climate warming and N enrichment. *Ecology* **95**: 3181-3189. https://doi.org/10.1890/13-2104.1

Schrager-Lavelle, A., Klein, H., Fisher, A., and M. Bartlett. 2017. Grass flowers: An untapped resource for floral evo-devo. *Journal of Systematics and Evolution* **55**: 525–541. https://doi.org/10.1111/jse.12251

Simberloff, D., Martin, J.-L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., Courchamp, F., et al. 2013. Impacts of biological invasions: What's what and the way forward. *Trends in Ecology and Evolution* **28**: 58–66. https://doi.org/10.1016/j.tree.2012.07.013

Sinn, B. T., Simon, S. J., Santee, M. V., DiFazio, S. P., Fama, N. M., and C. F. Barrett. 2021. ISSRseq: An extensible, low-cost, and efficient method for reduced representation sequencing. *Methods in Ecology and Evolution* **00**: 1–14. https://doi.org/10.1111/2041-210X.13784

Soper Gorden, N. L., Winkler, K. J., Jahnke, M. R., Marshall, E., Horky, J., Hudelson, C., and J. R. Etterson. 2016. Geographic patterns of seed mass are associated with climate factors, but relationships vary between species. *American Journal of Botany* **103**: 60–72. https://doi.org/10.3732/ajb.1500295

South, A. 2011. rworldmap: A new R package for mapping global data. *The R Journal* **3**: 35-43. http://journal.r-project.org/archive/2011-1/RJournal_2011-1_South.pdf

South A. 2017. rnaturalearth: World map data from natural Earth. R package version 0.1.0. https://CRAN.R-project.org/package=rnaturalearth

Sutherland, B. L., Barrett, C. F., Beck, J. B., Latvis, M., McKain, M. R., Sigel, E. M., and N. J. Kooyers. 2021. Botany is the root and the future of invasion biology. *American Journal of Botany* **108**: 549–552. https://doi.org/10.1002/ajb2.1642

Tekiela, D. R., and J. N. Barney. 2013. Quantifying *Microstegium vimineum* seed movement by non-riparian water dispersal using an ultraviolet-marking based recapture method. *PLOS ONE* **8**: e63811. https://doi.org/10.1371/journal.pone.0063811

Tu, M. 2000. Element stewardship abstract for *Microstegium vimineum*. The Nature Conservancy. URL: https://www.invasive.org/gist/esadocs/documnts/micrvim.pdf

Vallejo-Marín, M., Friedman, J., Twyford, A. D., Lepais, O., Ickert-Bond, S. M., Streisfeld, M. A., Yant, L., et al. 2021. Population genomic and historical analysis suggests a global invasion by bridgehead processes in *Mimulus guttatus*. *Communications Biology*, **4**: 327. https://doi.org/10.1038/s42003-021-01795-x

van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., and K. A. Hodgins. 2017. Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Molecular Ecology*, **26**: 5421–5434. https://doi.org/10.1111/mec.14293

van Boheemen, L. A., Atwater, D. Z., and Hodgins, K. A. 2019. Rapid and repeated local adaptation to climate in an invasive plant. *New Phytologist* **222**: 614–627. https://doi.org/10.1111/nph.15564

Weiher, E., and P. A. Keddy. (1995). Assembly rules, null models, and trait dispersion: new questions from old patterns. *Oikos* **74**: 159. https://doi.org/10.2307/3545686

Wickham, H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag. https://ggplot2.tidyverse.org

Wickham H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., and G. Grolemund. 2019. Welcome to the tidyverse. *Journal of Open Source Software,* **4**: 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., and K. Müller. 2020. dplyr: A grammar of data manipulation. R package version 1.0.2. https://CRAN.R-project.org/package=dplyr

Würschum, T., Jähne, F., Phillips, A. L., Langer, S. M., Longin, C. F. H., Tucker, M. R., and W. L. Leiser. 2020. Misexpression of a transcriptional repressor candidate

provides a molecular mechanism for the suppression of awns by Tipped 1 in wheat. *Journal of Experimental Botany* **71**: 3428–3436. https://doi.org/10.1093/jxb/eraa106

Yamaji, N., Chiba, Y., Mitani-Ueno, N. and J. F. Ma. 2012. Functional characterization of a silicon transporter gene implicated in silicon distribution in barley. *Plant Physiology* **160**: 1491-1497. https://doi.org/10.1104/pp.112.204578

Table 1. Least-squares regressions of latitude (centroids) vs. the proportion of awned specimens for Asian and US records. Unweighted and weighted regressions (by the total number of records per State, Province, or Region) are shown. 'df' = degrees of freedom, 'Std. Error = standard error.'

	Proportion of awned records			
	US	US (weighted)	Asia	Asia (weighted)
Latitude Coefficient	0.080[***]	0.105[***]	0.015[**]	0.026[***]
Std. Error	(0.015)	(0.016)	(0.006)	(0.007)
Intercept Coefficient	-2.570[***]	-3.534[***]	0.063	-0.405[*]
Std. Error	(0.574)	(0.600)	(0.205)	(0.226)
Observations	21	21	33	33
R^2	0.585	0.689	0.162	0.306
Adjusted R^2	0.563	0.673	0.135	0.284
Residual Std. Error	0.291 (df = 19)	1.322 (df = 19)	0.307 (df = 31)	0.902 (df = 31)
F-Statistic	26.772[***](df = 1; 19)	42.102[***](df = 1; 19)	5.999[**](df = 1; 31)	13.682[***](df = 1; 31)

Note: *$p<0.1$; **$p<0.05$; ***$p<0.01$.

Table 2. Logistic regression models for US specimens collected from the year 2000 onward showing the probability of being awned vs. Principal Component 1 of BIOCLIM Temperature and Precipitation variables. 'Std. Error' = standard error.

|	Temp. model	Precip. model	All BIOCLIM model	Temp. × Precip. model
Temperature Coefficient (PC1)	-1.115***			-1.226***
	-0.147			-0.23
Precipitation Coefficient (PC1)		-0.965***		-0.760***
		-0.138		-0.253
All BIOCLIM Coefficient (PC1)			-1.026***	
			-0.14	
Temperature (PC1) × Precipitation (PC1) Interaction Coefficient				-0.343***
				-0.104
Intercept Coefficient	-0.411*	-0.340*	-0.705**	-0.664*
	-0.231	-0.191	-0.274	-0.353
Observations	207	207	207	207
Log Likelihood	-65.364	-92.878	-60.237	-55.977
Akaike Information Criterion	134.727	189.757	124.473	119.953

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 3. Mixed-effects Generalized Linear Regression models for US specimens showing $\log_{10}$-transformed awn and floret measurements vs. $\log_{10}$-latitude and awn type (long, short, or awnless). Fixed-effect independent variables are shown (left), while random effects were used to account for non-independent variation in five floret measurements for each accession. Awnless records were removed from the awn angle and awn length regressions to avoid zero-values. 'Std. Error' = standard error.

	Floret length	Floret width	Awn angle	Awn length
Intercept coefficient	-0.99[***]	-1.11[*]	-1.55	0.26
Std. error	(0.29)	(0.46)	(6.41)	(1.00)
Awn type (long, short, awnless)	-0.05[***]	-0.03		
Std. error	(0.01)	(0.02)		
Awn type (short vs. long)			-0.71[***]	-0.45[***]
Std. error			(0.16)	(0.02)
$\log_{10}$ latitude	0.41[*]	0.12	1.79	-0.37
Std. error	(0.18)	(0.29)	(4.01)	(0.62)
Pseudo-R^2 (fixed effects/total)	0.36/0.65	0.04/0.39	0.19/0.35	0.82/0.88
AIC	-1232.78	-830.19	379.02	-360.32
BIC	-1209.46	-806.87	395.12	-344.22
Log likelihood	622.39	421.09	-184.51	185.16
Number of observations	360	360	185	185
Number of groups: accession	72	72	37	37
Variance: accession (Intercept)	0.00	0.00	0.09	0.00
Variance: residual	0.00	0.00	0.37	0.01

Note: *p<0.1; **p<0.05; ***p<0.01.

**FIGURE LEGENDS**
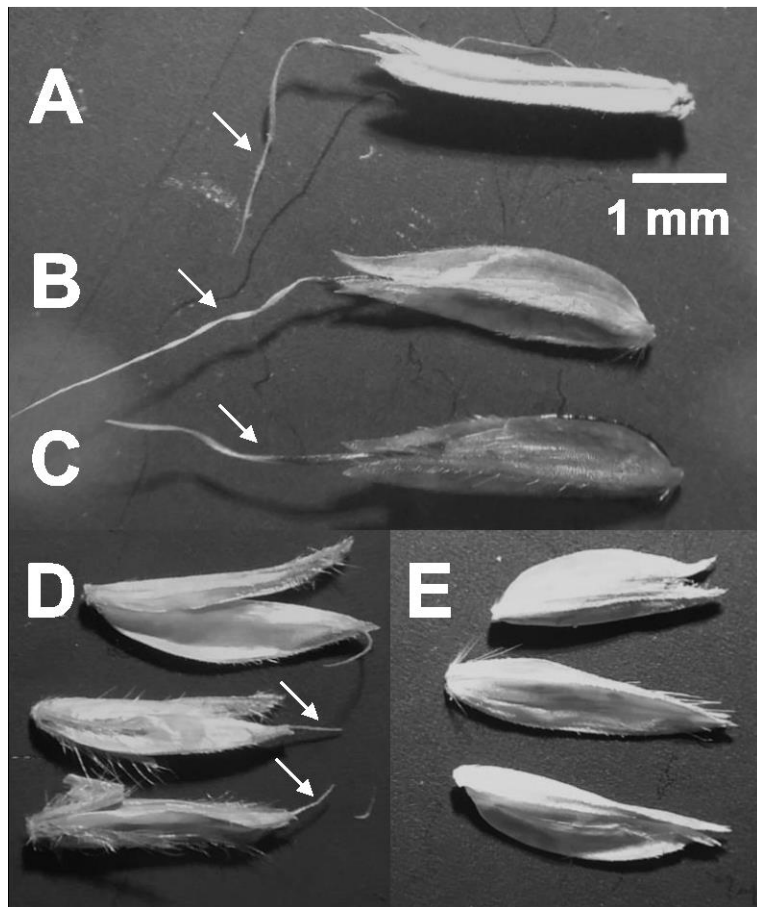
Fig. 1. Florets of *Microstegium vimineum*. **a**. Awned floret, with an arrow illustrating the geniculate awn at a 90° angle (fully dried). **b**. Awned floret, with an arrow illustrating the geniculate awn at a 45° angle, one minute after soaking 5 minutes in water. **c**. Awned floret with an arrow showing straightened, geniculate awn immediately after soaking 5 minutes in water. Florets from Tompkins County, New York, USA. **d**. Florets with arrows illustrating intermediate, short-awned phenotypes (Sussex County, Delaware, USA). **e**. Awnless florets (Knox County, Tennessee, USA).
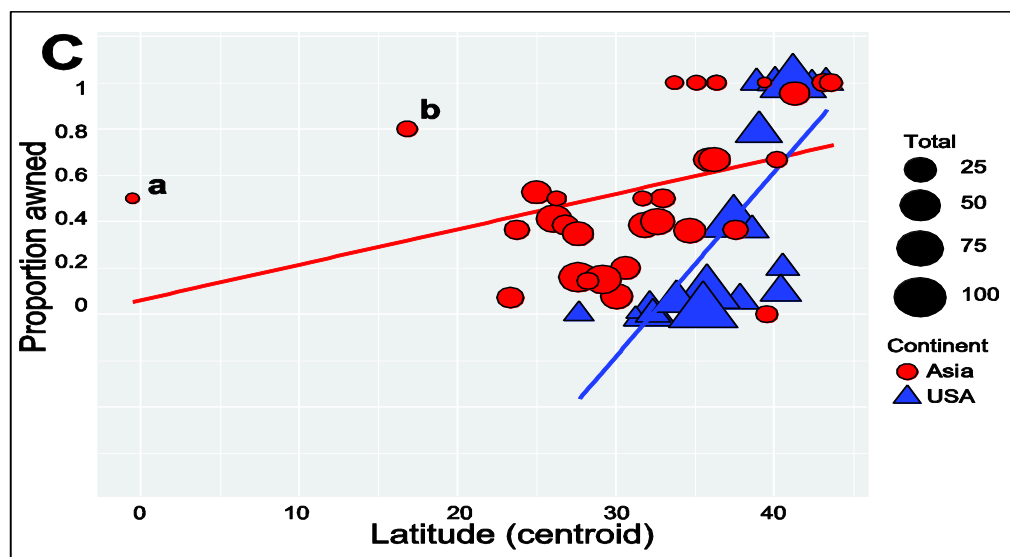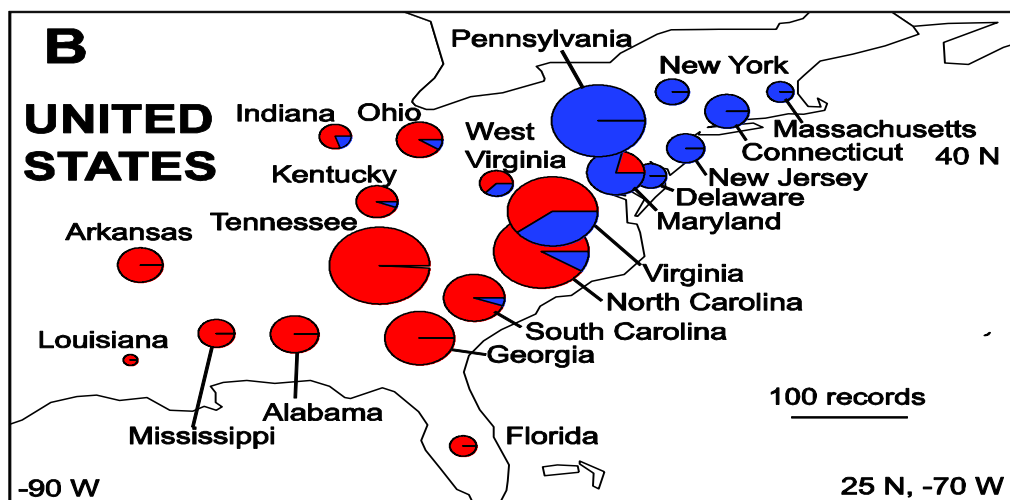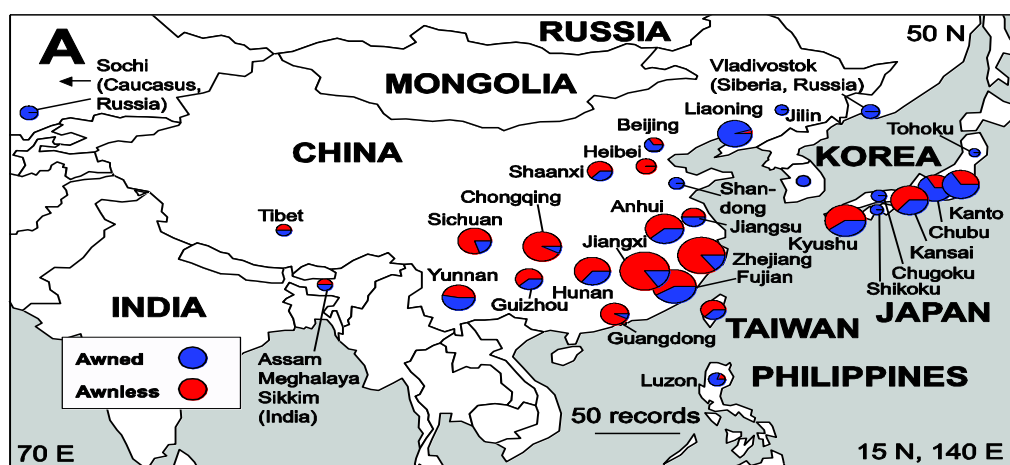
Fig. 2. Maps of the proportions of awned and awnless forms of *Microstegium vimineum* in: **a.** Asia (native range) and **b.** the US (introduced range). Pie graphs are scaled proportionally within each map (not between the two maps) to the number of records examined for each centroid. Red = awnless form, blue = awned form. Coordinates given at the corners of each map are degrees latitude and longitude. Arrow in 'a.' (Sochi, Caucasus, Russia) is west of the map boundary. **c.** Relationship between latitude (degrees North, x-axis) and the proportion of awned records (y-axis) for Asian (red circles) and US records (blue triangles). The total number of records examined for each point is given by the scale to the right. Outliers are indicated by 'a' (Sumatra) and 'b' (Philippines).
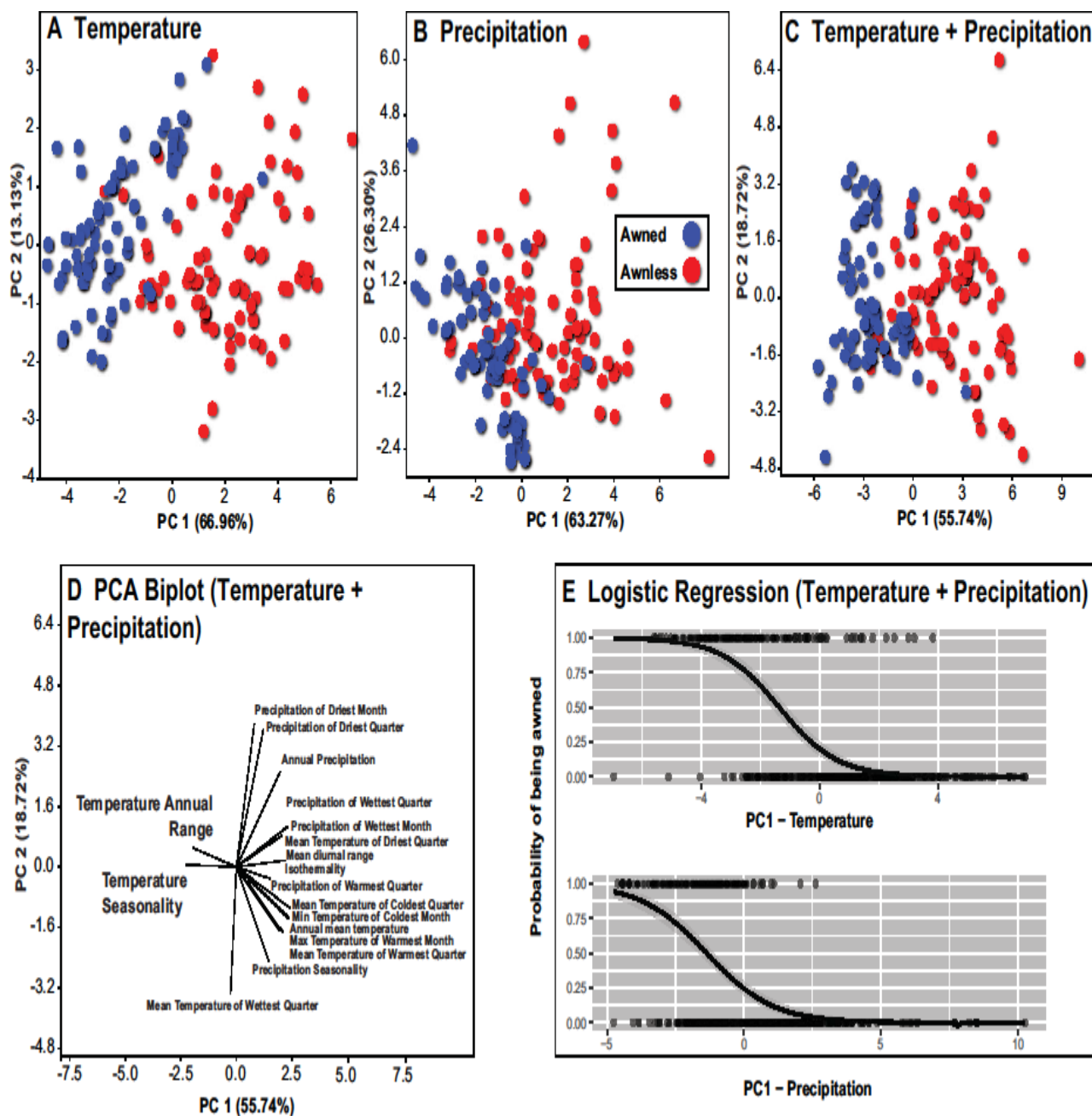
Fig. 3. Principal Components Analysis of 19 BIOCLIM variables for all US records collected after the year 2000 of *Microstegium vimineum*, based on the correlation matrix. Red = awnless forms, blue = awned forms. Percentages in parentheses on each PC axis represent the total percent variance explained by each axis. **a**. BIOCLIM temperature variables. **b**. BIOCLIM precipitation variables. **c**. All BIOCLIM variables. **d**. PCA biplot of temperature and precipitation variables. **e**. Logistic regression plots of temperature and

precipitation axes (x-axes) vs. the probability of being awned (y-axes) for US records. Each x-axis represents the first Principal Component (PC) for 11 BIOCLIM temperature variables (above) and eight precipitation variables (below) based on a correlation matrix. Gray shading around the logistic regression line indicates the 95% confidence interval.
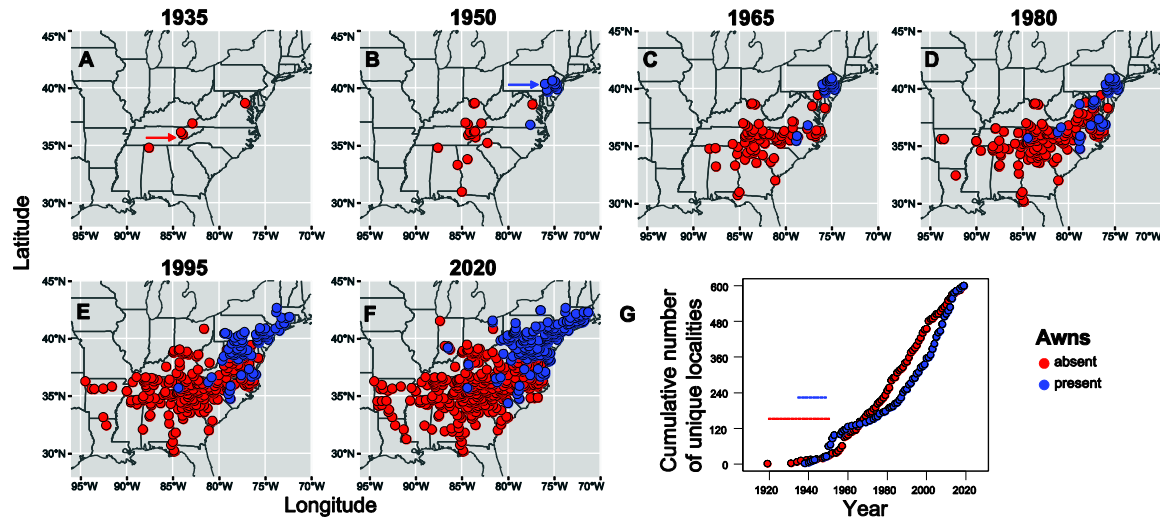
Fig. 4. Mapping the invasion history of *Microstegium vimineum* in the eastern US. Each slice represents 15 years (**a-f**), except for the first (a, 1919-1935, top left) and last (f, 1995-2020, bottom right). Red and blue circles indicate awnless and awned forms, respectively. Red arrow in a = 1919, Knox County, Tennessee (awnless); blue arrow in b = 1938, Berks County, Pennsylvania (awned). **g**. Plot of the number of unique localities colonized over time for awnless (red) and awned (blue) forms, binned by year. Dashed line = putative 'lag phase' for each.
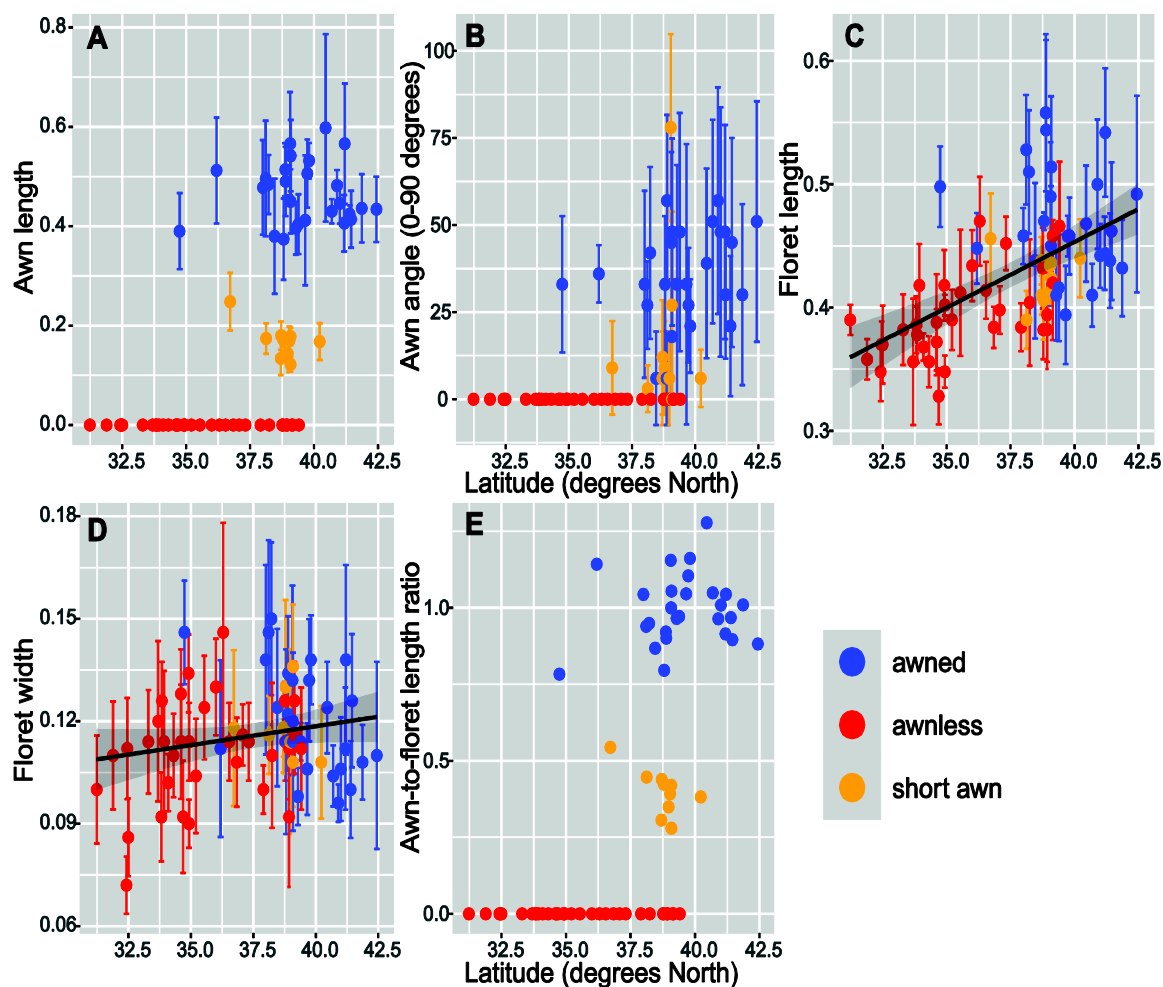
Fig. 5. Relationships between awn and floret measurements and latitude for 73 US records of *Microstegium vimineum*. All measurements are in centimeters except for awn angle and awn:floret length ratio. Values are means ± standard deviation based on five replicate florets per record. **a**. Awn length. **b**. Degree of awn bending (awn angle). **c**. Floret length. **d**. Floret width. **e**. Awn:floret length ratio. Colors indicate awned forms (blue), awnless forms (red), and intermediate, short-awned forms (orange).
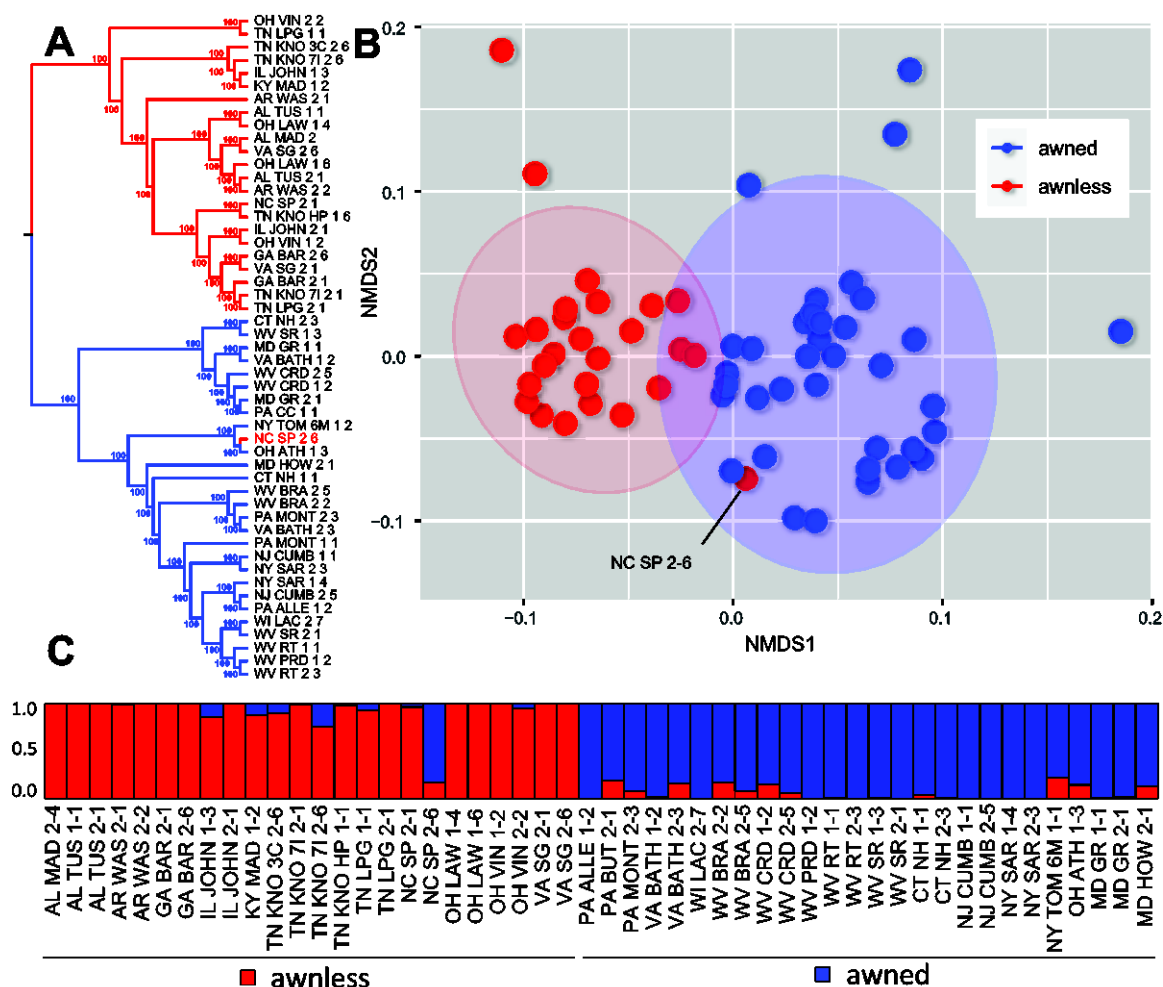
Fig. 6. Genomic analyses of field-collected US *Microstegium vimineum* accessions. **a.** Coalescent-based phylogenetic analysis of 1,565 SNPs with SVDQuartets, with parsimony-uninformative SNPs removed. Numbers adjacent to branches indicate bootstrap support values (1,000 pseudoreplicates). Red = awnless accessions and blue = awned accessions. **b.** Nonmetric multidimensional scaling ordination plot based on a pairwise distance matrix from ANGSD genotype likelihoods of 39,198 SNPs (K = 3 dimensions, stress = 0.176). **c.** Ancestry plot based on 3,148 SNPs after linkage disequilibrium thinning in parallelSTRUCTURE (highest deltaK was 477.7 for K = 2).