Received: 22 April 2024 | Revised: 29 May 2024 | Accepted: 4 June 2024

DOI: 10.1002/pld3.615

RESEARCH ARTICLE

DNA barcoding of terrestrial invasive plant species in Southwest Michigan

Sneha Nath[1] | Joshua T. VanSlambrouck[1] | Janelle W. Yao[2,3] | Ashika Gullapalli[2,4] | Fayyaz Razi[2,3] | Yan Lu[1]

[1]Department of Biological Sciences, Western Michigan University, Kalamazoo, Michigan, USA

[2]Kalamazoo Area Math and Science Center, Kalamazoo, Michigan, USA

[3]Portage Central High School, Portage, Michigan, USA

[4]Portage Northern High School, Portage, Michigan, USA

Correspondence
Yan Lu, Department of Biological Sciences, Western Michigan University, 1903 W Michigan Ave, Kalamazoo, MI 49008-5410, USA.
Email: yan.1.lu@wmich.edu

Funding information
This work was supported in part by U.S. National Science Foundation grants (BIORETS grant DBI-2146882 and MI-LSAMP grant EES-2109942) and Western Michigan University.

Abstract

Because of the detrimental effects of terrestrial invasive plant species (TIPS) on native species, ecosystems, public health, and the economy, many countries have been actively looking for strategies to prevent the introduction and minimize the spread of TIPS. Fast and accurate detection of TIPS is essential to achieving these goals. Conventionally, invasive species monitoring has relied on morphological attributes. Recently, DNA-based species identification (i.e., DNA barcoding) has become more attractive. To investigate whether DNA barcoding can aid in the detection and management of TIPS, we visited multiple nature areas in Southwest Michigan and collected a small piece of leaf tissue from 91 representative terrestrial plant species, most of which are invasive. We extracted DNA from the leaf samples, amplified four genomic loci (ITS, *rbcL*, *matK*, and *trnH-psbA*) with PCR, and then purified and sequenced the PCR products. After careful examination of the sequencing data, we were able to identify reliable DNA barcode regions for most species and had an average PCR-and-sequencing success rate of 87.9%. We found that the species discrimination rate of a DNA barcode region is inversely related to the ease of PCR amplification and sequencing. Compared with *rbcL* and *matK*, ITS and *trnH-psbA* have better species discrimination rates (80.6% and 63.2%, respectively). When ITS and *trnH-psbA* are simultaneously used, the species discrimination rate increases to 97.1%. The high species/genus/family discrimination rates of DNA barcoding indicate that DNA barcoding can be successfully employed in TIPS identification. Further increases in the number of DNA barcode regions show little or no additional increases in the species discrimination rate, suggesting that dual-barcode approaches (e.g., ITS + *trnH-psbA*) might be the efficient and cost-effective method in DNA-based TIPS identification. Close inspection of nucleotide sequences at the four DNA barcode regions among related species demonstrates that DNA barcoding is especially useful in identifying TIPS that are morphologically similar to other species.

KEYWORDS
DNA barcoding, invasive plant species, PCR amplification and sequencing, species identification, terrestrial invasive plant species

1 | INTRODUCTION

Invasive species grow and reproduce quickly and have the potential to harm native species, the environment, public health, or the economy. We are particularly interested in terrestrial invasive plant species (TIPS) because many of them produce large amounts of seeds, which are easily distributed by the wind, birds, other animals, and unknowing humans (USDA, 2024b). TIPS thrive in natural and disturbed areas and outcompete native plant species (Rai et al., 2023). Some TIPS have a very aggressive root system capable of spreading a long distance and smothering the root systems of neighboring plants (Lawson et al., 2021; Rai et al., 2023). Some TIPS produce toxic chemicals that inhibit the growth and seed germination of other plants (Albouchi et al., 2013; Callaway & Aschehoug, 2000; Heisey, 1990; Oyeniyi et al., 2016; Rahman et al., 2017; Rai, 2015; Rai et al., 2023; Rai & Singh, 2024). Therefore, TIPS can rapidly replace a diverse ecosystem with a monoculture (Rai, 2015). Some TIPS produce toxic chemicals or allergens that adversely affect the health of humans and animals (Andonova et al., 2023; Boskabadi et al., 2021; Castaño-Quintana et al., 2013; Cipollini et al., 2008; Rai et al., 2023; Rai & Singh, 2020, 2024). The presence of TIPS on a property or agricultural land area will also decrease the property value and agricultural productivity and increase management costs (Fenichel et al., 2014; Holmes et al., 2009; Radosevich et al., 2007; USDA, 2024a).

Because of the detrimental effects of TIPS on native species, ecosystems, public health, and the economy, many countries and individual states or provinces have been actively looking for strategies to prevent the introduction and minimize the spread of TIPS (AIPAC, 2016; MDNR, 2009, 2015; USDA, 2024b; USDOI, 2021). The most efficient way to control invasive species is to prevent them from entering in the first place (Park & Potter, 2013). Thus, quick and accurate detection of invasive species is pivotal. Traditionally, invasive species detection and monitoring have heavily relied on morphological features (Xu et al., 2018). Since 2003, DNA-based species identification (i.e., DNA barcoding) has become more and more popular (Hebert, Cywinska, et al., 2003; Hebert, Ratnasingham, & deWaard, 2003). This technology refers to the use of DNA sequences from signature genomic loci to identify species. It includes four steps: (1) tissue harvesting and DNA extraction; (2) amplifying target genomic loci (i.e., DNA barcode regions) via PCR; (3) purification and sequencing of PCR products; (4) comparing the resulting sequences, that is, DNA barcode sequences, with those in reference sequence databases to identify matching species (IBLC, 2024).

Using appropriate genomic loci for PCR amplification is the key to the success of DNA-based species identification. The genomic loci for DNA barcoding must meet three criteria: (1) significant species-level sequence variation, (2) an appropriate length for PCR amplification, and (3) conserved bordering regions for universal primers to bind (Kress et al., 2005; Kress & Erickson, 2008). One potential plant DNA barcode region is the internal transcribed spacer (ITS) of the *18S-5.8S-26S* ribosomal genes in the nuclear genome (Álvarez & Wendel, 2003; Blaxter, 2004; Stoeckle, 2003). It has relatively high species-level discrimination rates (Chen et al., 2010; Hao et al., 2013; Kress et al., 2005; Xu et al., 2018). A second potential plant DNA barcode region is *trnH-psbA*, the spacer between *trnH* and *psbA* genes in the plastid genome, which encode transfer RNA-histidine and photosystem II reaction center protein D1, respectively (Kelly et al., 2010; Soltis et al., 2001). *trnH-psbA* also has relatively high species-level discrimination rates (Devey et al., 2009; Fazekas et al., 2008; Hao et al., 2013; Hollingsworth et al., 2009; Kress et al., 2005, 2009; Shaw et al., 2005). A third potential plant DNA barcode region is the plastid-encoded *ribulose bisphosphate carboxylase/oxygenase large subunit* (*rbcL*) gene, the most commonly sequenced plastid gene by plant systematists (Blaxter, 2004; de Vere et al., 2015; Fazekas et al., 2008; Gielly & Taberlet, 1994; Hollingsworth et al., 2009; Kress et al., 2009). A fourth potential plant DNA barcode region is the *maturase K* (*matK*) gene in the plastid genome (Hollingsworth et al., 2009; Kress et al., 2009; Kelly et al., 2010; de Vere et al., 2015; Lahaye et al., 2008; Shaw et al., 2005; Soltis et al., 2001; Yu et al., 2011). In addition to nuclear and plastid loci, mitochondrial genes (e.g., *cytochrome c oxidase subunit 1* [*CO1*]) have also been proposed for plant DNA barcoding (Chen et al., 2010; Chase et al., 2007; Fazekas et al., 2008; Kress et al., 2005; Pennisi, 2007). However, because plant mitochondrial genomes are highly variable in length, gene arrangement, and intergenic sequences, but highly conserved in protein-coding sequences, mitochondrial genes have limited usage in plant DNA barcoding (Chen et al., 2010; Chase et al., 2007; Fazekas et al., 2008; Kress et al., 2005; Pennisi, 2007).

Both morphology-based and DNA-based species identification methods have their own pros and cons (Hulley et al., 2018). The advantages of morphology-based species identification include objectivity and the ease of use (Kruk et al., 2010). However, many TIPS are morphologically similar to other species. For example, some plant materials, such as seeds, seedlings, vegetative tissues, and fruits, often lack diagnostic physical characteristics (Xu et al., 2018). This is especially challenging at points of entry, where TIPS are often transported in the form of seeds, seedlings, fruits, and fragmentary materials, not in the form of whole adult plants (Hollingsworth et al., 2009; Whitehurst et al., 2020). On the contrary, DNA-based species identification does not require access to whole adult plants (Hulley et al., 2018).

To investigate whether DNA-based species identification (i.e., DNA barcoding) can aid in the detection and management of TIPS, we visited multiple nature areas in Southwest Michigan to identify TIPS. We collected a small piece of leaf tissue from 91 representative terrestrial plant species, most of which are invasive. We extracted DNA from the leaf samples and amplified four genomic loci (i.e., DNA barcode regions) with PCR: ITS, *rbcL*, *matK*, and *trnH-psbA*. We inspected the PCR products with agarose gel electrophoresis, purified the PCR products with the gel extraction kit, and then submitted them for sequencing. After careful inspection of the sequencing data, we were able to identify reliable DNA barcode regions for most species. We compared the PCR-and-sequencing success rates of the four DNA barcode regions among different plant types. We also calculated and compared species/genus/family discrimination rates among the four DNA barcode regions. The average PCR-and-sequencing success

rate across the four DNA barcode regions and the 91 species (most of which are TIPS) is 87.9%. The high species/genus/family discrimination rates of DNA barcoding, especially when ITS and *trnH-psbA* are simultaneously used, indicate that DNA barcoding can be successfully employed in TIPS identification. DNA barcoding is especially useful to identify TIPS that are morphologically similar to other species as well as TIPS transported in forms that lack diagnostic physical attributes.

2 | RESULTS

2.1 | Terrestrial plant species used in this study

During September 2022–September 2023, we identified and collected leaf tissues from 91 terrestrial plant species, including five grasses, 61 herbaceous plants, nine shrubs, 13 trees, and three vines (Table S1). Most of these species were identified from five nature areas in Southwest Michigan: Asylum Lake Preserve and WMU Goldsworth Valley Pond Area in Kalamazoo; Bishop's Bog Preserve, South Westnedge Park, and West Lake Nature Preserve in Portage. In terms of invasive status, this list includes 73 (terrestrial) invasive plant species (i.e., TIPS), five introduced (i.e., non-native) plant species, and 13 native plant species (Table S1). Most of these native plant species are included in this study because they are morphologically similar to some of the introduced or invasive plant species. The orders and families of the 91 terrestrial plant species belong to are shown in Figure S1.

As mentioned above, 73 TIPS were identified and analyzed in this study, including three grasses, 53 herbaceous plants, eight shrubs, seven trees, and two vines (Table S1). In terms of orders, 12 of them belong to Asterales (the daisy order), 10 belong to Lamiales (the mint order), nine belong to Caryophyllales (the pink or carnation order), another nine belong to Fabales, seven belong to Rosales, six belong to Brassicales, and the remaining 20 belong to 13 different orders (Figure 1a). In terms of families, 12 are in Asteraceae (the daisy family), nine are in Fabaceae (the legume family), six are in Brassicaceae (the mustard family), five are in Caryophyllaceae (the pink or carnation family), and the remaining 41 are in 23 different families (Figure 1b). We retrieved the total numbers of species for each relevant family from recent literature (Alves et al., 2014; Lai et al., 2018; Simpson, 2010) and plotted the numbers of TIPS/family identified in this study as a function of the total numbers of species in these families (Figure 2). We found that there is a positive correlation between the numbers of TIPS/family identified in this study and the total numbers of species in these families ($R^2 = .7263$). We also retrieved the numbers of invasive species/family listed in the Midwest Invasive Species Information Network (MISIN; https://www.misin.msu.edu/) database and plotted them against the total numbers of species in these families; a similar trend was observed (Figure S2).

We ground the frozen leaf tissues from representative TIPS plants into fine powder with a bead beater and extracted DNA from the powder with the DNeasy Plant Mini Kit. The average DNA yield for grasses, herbaceous plants, shrubs, trees, and vines is 33, 43, 48, 44, and 31 μg/g fresh weight, respectively (Figure 3). The average DNA yield among all the leaf samples is 43 μg/g fresh weight, and the average DNA purity ratio (OD260/280) is 1.81, which are considered suitable for downstream applications such as PCR amplification (Enan et al., 2017; Lucena-Aguilar et al., 2016).

2.2 | Selection of four DNA barcode regions

Using leaf DNA as the template, we amplified four genomic loci (i.e., DNA barcode regions) with PCR: ITS, *rbcL*, *matK*, and *trnH-psbA*. A diagram of these four DNA barcode regions is shown in Figure 4. We chose to amplify ITS and *trnH-psbA* because they have relatively high species-level discrimination rates (Chen et al., 2010; Fazekas et al., 2008; Hao et al., 2013; Kress et al., 2005; Shaw et al., 2005; Xu et al., 2018). We chose to amplify *rbcL* and *matK* as well because they are commonly sequenced by plant systematists and are easy to amplify, sequence, and analyze (Blaxter, 2004; de Vere et al., 2015; Fazekas et al., 2008; Gielly & Taberlet, 1994; Lahaye et al., 2008; Shaw et al., 2005; Soltis et al., 2001; Yu et al., 2011).

To amplify ITS, we used the locus-specific universal primers previously used for land plants by Stanford et al. (2000) and Kress et al. (2005). To amplify *rbcL* and *matK*, we adopted the locus-specific universal primers previously designed and optimized for land plants by Kress and Erickson (2007), Hollingsworth et al. (2009), and Setsuko et al. (2023). To amplify *trnH-psbA*, we used the locus-specific universal primers recently optimized for land plants by Setsuko et al. (2023). We mapped the four pairs of universal primers onto the genomes of three plant species: *Arabidopsis thaliana*, cultivated tobacco (*Nicotiana tabacum*), and bittersweet nightshade (*Solanum dulcamara*, an invasive perennial in the same family with tobacco). The locations where these primers bind to the four genomic loci in *A. thaliana* are shown in Figure 4. We also compared the length and divergence of the resulting PCR products between cultivated tobacco and bittersweet nightshade (Table 1 and Figure S3). Using cultivated tobacco as an example, the PCR product lengths of ITS, *rbcL*, *matK*, and *trnH-psbA* are 756, 749, 888, and 560 bp, respectively, well within the suggested length criterion for successful barcoding (Kress et al., 2005; Kress & Erickson, 2008). As shown in Table 1, ITS and *trnH-psbA* have much higher divergence values (14.5% and 10.7%, respectively) than *rbcL* and *matK* (1.4% and 2.0%, respectively).

2.3 | Sequence characteristics of the four DNA barcode regions

After amplifying the four barcode regions from leaf DNA with Phusion High-Fidelity DNA Polymerase, we inspected and purified the PCR products with agarose gels and the QIAquick Gel Extraction Kit. The purified PCR products were quantified with a NanoDrop spectrophotometer and then sequenced. Each sequence file was manually inspected to ensure accurate nucleotide calling and removal of the reverse primer sequence. Sequences that did not pass the initial

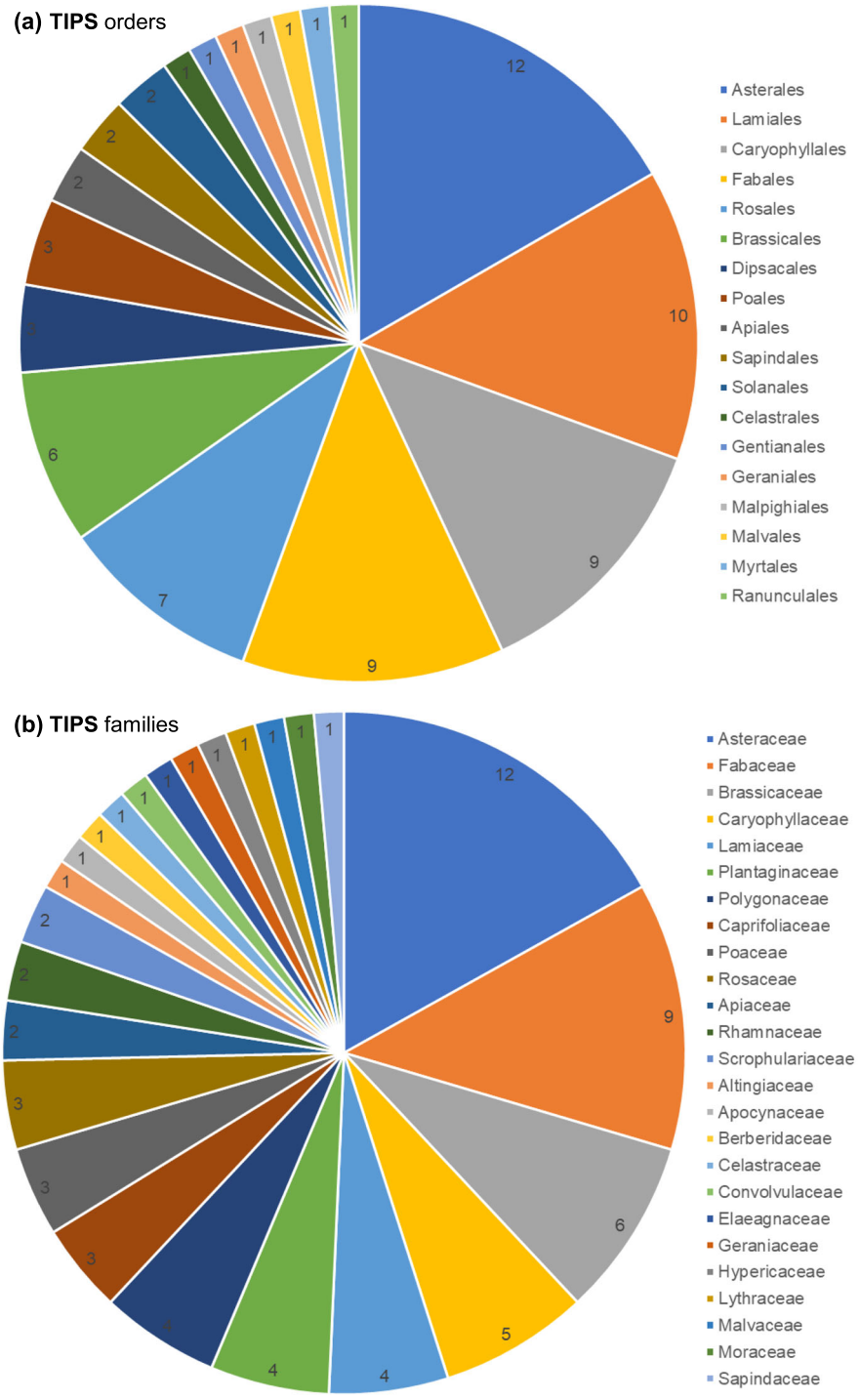

FIGURE 1 Orders (a) and families (b) the 73 TIPS identified and analyzed in this study belong to. The value in each pie slice indicates the number of analyzed TIPS in a specific order or family.

manual inspection were not used for downstream sequence analysis. The post-cleanup nucleotide sequences were submitted to the Nucleotide BLAST Portal of the National Center for Biotechnology Information (NCBI) to test whether the DNA-based species name matches with the morphology-based species name and to identify the reliable region of each nucleotide sequence (i.e., the reliable DNA barcode for each species at the four genomic loci). The reliable DNA barcode sequences were then submitted to the NCBI for public release (Table S2).

We then compared the PCR-and-sequencing success rates of the four DNA barcode regions (i.e., the DNA barcoding success rates) (Table 2). The overall success rate among the 91×4 PCR-and-sequencing samples is 87.9%. Of the four DNA barcode regions, *rbcL* has the highest PCR-and-sequencing success rate ($91/91 = 100.0\%$). The PCR-and-sequencing success rates of the other three DNA barcode regions are $82/91 = 90.1\%$ for *matK*, $76/91 = 83.5\%$ for *trnH-psbA*, and $71/91 = 78.0\%$ for ITS. Among the nine species that did produce a reliable *matK* barcode sequence in

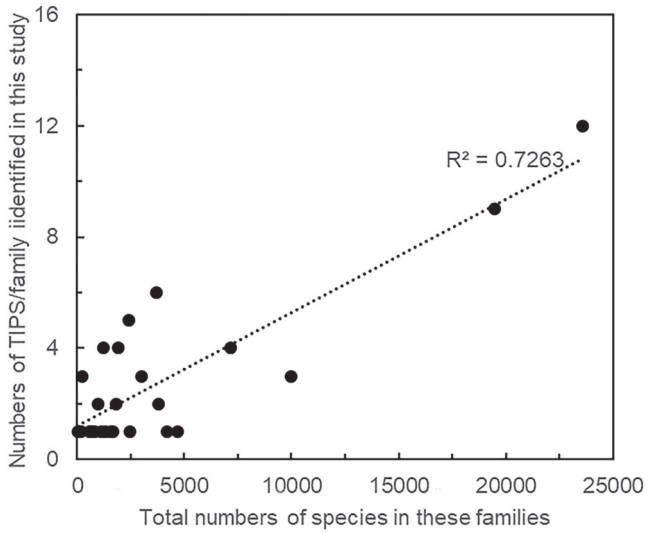

FIGURE 2 Relationship between the numbers of TIPS/family identified in this study and the total numbers of species in these families.

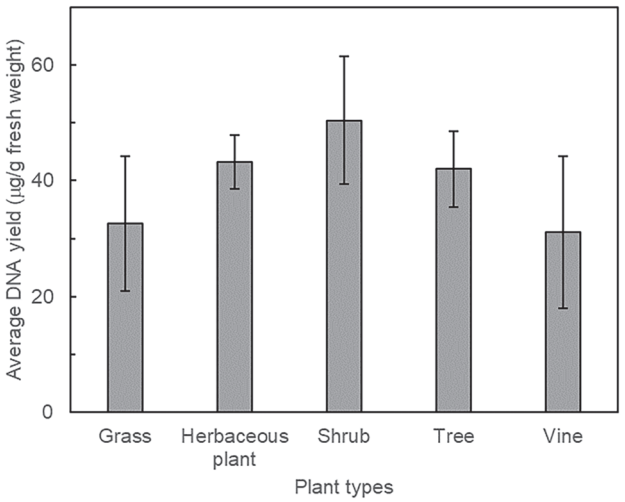

FIGURE 3 Average DNA yield from different TIPS plant types. Data are presented as means ± SE ($n = 5$ for grasses, 61 for herbaceous plants, 9 for shrubs, 13 for trees, and 3 for vines).

this study, seven have at least one primer with low homology (<80%) to the target sequence (Table S2). Among the 15 species that did generate a reliable *trnH-psbA* barcode sequence, 10 have at least one primer with low homology to the target sequence (Table S2). Among the 20 species that did not create a reliable ITS barcode sequence, 11 have at least one primer with low homology to the target sequence (Table S2). These observations indicate that the universality of ITS, *trnH-psbA*, and *matK* primers still needs improvements for some plant species.

We also compared the PCR-and-sequencing success rates among different plant types (Table 2). Of the five plant types, grasses have the highest PCR-and-sequencing success rate (95.0%). The PCR-and-sequencing success rates of the other four plant types are 91.7% for vines, 89.8% for herbaceous plants, 82.7% for trees, and 77.8% for shrubs.

2.4 | Species/genus/family discrimination rates of the four DNA barcode regions

We calculated and compared species discrimination rates among the four DNA barcode regions (Table 3). When a single genomic locus is used, ITS has the highest species discrimination rate (80.6%) and *trnH-psbA* has the second highest species discrimination rate (63.2%). The species discrimination rates of the other two DNA barcode regions are 58.5% for *matK* and 39.6% for *rbcL*. We plotted the species discrimination rate as a function of PCR-and-sequencing success rate and found that these two rates are inversely related ($R^2 = .9575$; Figure 5). When two DNA barcode regions are used, the combination of ITS + *trnH-psbA* has the highest species discrimination rate (97.1%) and the combination of ITS + *matK* has the second highest species discrimination rate (90.1%). The species discrimination rates for other combinations of two DNA barcode regions are 87.5% for ITS + *rbcL*, 80.5% for *matK* + *trnH-psbA*, 75.6% for *rbcL* + *trnH-psbA*, and 68.6% for *rbcL* + *matK*. When three DNA barcode regions are used, ITS + *rbcL* + *trnH-psbA* and ITS + *matK* + *trnH-psbA* have the same species discrimination rate (97.2%). The species discrimination rates for other two combinations are 93.1% for ITS + *rbcL* + *matK* and 84.3% for ITS + *rbcL* + *matK* + *trnH-psbA*. When all four DNA barcode regions are used, the species discrimination rate is still 97.2%. The reason this percentage is not 100% is because the ITS, *rbcL*, *matK*, and *trnH-psbA* DNA barcode sequences from mouseear hawkweed (*Hieracium pilosella*) are identical to a related species in the same genus: meadow hawkweed (*Hieracium caespitosum*).

We also calculated and compared genus discrimination rates among the four DNA barcode regions (Table 3). When a single DNA barcode region is used, ITS has the genus species discrimination rate (97.2%) and *trnH-psbA* has the second highest genus discrimination rate (96.1%). The species discrimination rates of the other two DNA barcode regions are 93.9% for *matK* and 91.2% for *rbcL*. When two DNA barcode regions are used, ITS + *rbcL*, ITS + *matK*, and ITS + *trnH-psbA* have the perfect genus discrimination rate (100.0%), *rbcL* + *matK* and *matK* + *trnH-psbA* have slightly lower genus discrimination rates (98.7%–98.8%), and *rbcL* + *trnH-psbA* has the lowest genus discrimination rate (97.6%). When three DNA barcode regions are used, ITS + *rbcL* + *matK*, ITS + *rbcL* + *trnH-psbA*, and ITS + *matK* + *trnH-psbA* have the perfect genus discrimination rate (100.0%), and ITS + *matK* + *trnH-psbA* has a slightly lower genus discrimination rate. The reason this percentage is not 100% is because the *rbcL*, *matK*, and *trnH-psbA* DNA barcode sequences from hairy cat's ear (*Hypochaeris radicata*) are identical to a species in a related genus: *Crepis rigescens*. As expected, when all four DNA barcode regions are used, the genus discrimination rate is 100.0%.

Furthermore, we calculated family discrimination rates among the four DNA barcode regions (Table 3). When a single genomic locus is

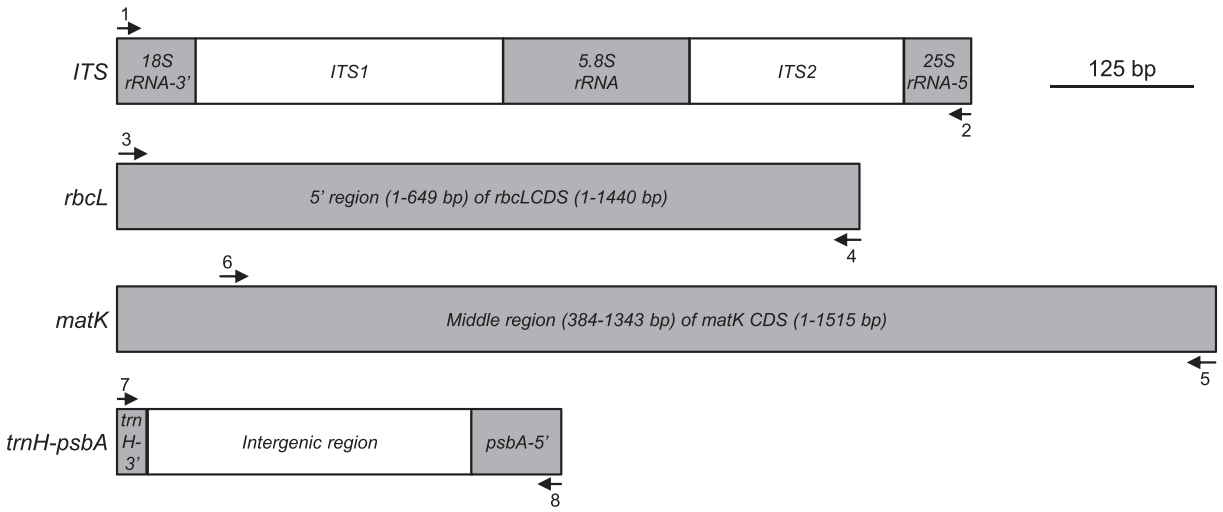

FIGURE 4 A diagram of the four genomic loci (i.e., DNA barcode regions) and four pairs of primers used in this study. Primer list: 1, ITS5a; 2, ITS4; 3, rbcLa-F; 4, NTrbcL-626L24; 5, matK_Kim_f; 6, HydmatK1041L; 7, NT55U; 8, PT119680L21. CDS, coding sequence. For simplicity reasons, *A. thaliana* genomic loci were used in this diagram. The lengths of the corresponding PCR products are 745 bp (ITS), 649 bp (*rbcL*), 872 bp (*matK*), and 387 bp (*trnH-psbA*), respectively.

TABLE 1 Length and divergence of the four DNA barcode regions in cultivated tobacco and bittersweet nightshade.

Locus	PCR product length (bp) in cultivated tobacco	PCR product length (bp) in bittersweet nightshade	Mismatches (bp) between two species	Indels (bp) between two species	Percent sequence divergence between two species
ITS	756	755	73	37	14.5%
rbcL	649	649	9	0	1.4%
matK	888	887	18	0	2.0%
trnH-psbA	560	550	42	18	10.7%

used, the family discrimination rate is 100% for all four DNA barcode regions. Consequently, when two, three, or four DNA barcode regions are used, the family discrimination rate is 100% for all 11 combinations. The high species/genus/family discrimination rates of DNA barcoding, especially when ITS and *trnH-psbA* are simultaneously used, indicate that DNA barcoding can be used alone in TIPS identification.

2.5 | DNA barcoding aids in morphology-based TIPS identification

Many TIPS are morphologically similar to other invasive or native species. One example is that of garlic mustard (*Alliaria petiolata*) and ground ivy (*Glechoma hederacea*) (Figure 6a). Garlic mustard is native to Europe and parts of Asia and was introduced into North America in the 1800s (Munger, 2001; Rodgers et al., 2008). Since then, it has become one of the most notorious TIPS in North America. It spreads aggressively by producing large amounts of seeds and allelopathic compounds that inhibit the seed germination and growth of other plant species. In the spring time, this species produces small, four-petal white flowers (Figure 6a). Ground ivy also is native to Eurasia and was brought into North America in the 1800s for ornamental or medicinal purposes (Waggy, 2009). Ground ivy is considered invasive in Michigan because it is toxic to many invertebrates and it spreads quickly by its above-ground stolons. This species produces small, tubular, lavender-colored flowers (Figure 6a), which are different from garlic mustard flowers. However, at the vegetative stage, the two species have similar round heart-shaped leaves with scalloped margins (Figure 6a), which are difficult to distinguish to the untrained eye. Fortunately, the alignments of ITS, *rbcL*, *matK*, and *trnH-psbA* DNA barcode sequences from garlic mustard and ground ivy displayed substantial nucleotide differences between the two species at all four barcode regions (Figure S4). This suggests that DNA barcoding can aid in the identification of garlic mustard and ground ivy, which are morphologically similar to each other.

A second example is that of common ragweed (*Ambrosia artemisiifolia*) and mugwort (*Artemisia vulgaris*) (Figure 6b). Common ragweed originated in North America but is considered invasive because of its competitive and allergenic nature (Rasmussen et al., 2017). This species spreads rapidly by producing lots of seeds. Mugwort is native to Eurasia and was introduced into North America in the 1600s for medicinal usages because it contains high amounts of medically active

TABLE 2 PCR-and-sequencing success rates of the four DNA barcode regions.

Plant type	ITS	*rbcL*	*matK*	*trnH-psbA*	Average among barcode regions
Grasses	80.0%	100.0%	100.0%	100.0%	95.0%
Herbaceous plants	83.6%	100.0%	91.8%	83.6%	89.8%
Shrubs	55.6%	100.0%	66.7%	88.9%	77.8%
Trees	69.2%	100.0%	92.3%	69.2%	82.7%
Vines	66.7%	100.0%	100.0%	100.0%	91.7%
Average among plant types	78.0%	100.0%	90.1%	83.5%	87.9%

TABLE 3 Species, genus, and family identification rates of the four DNA barcode regions.

Locus	Species discrimination rate	Genus discrimination rate	Family discrimination rate
ITS	80.6%	97.2%	100.0%
rbcL	39.6%	91.2%	100.0%
matK	58.5%	93.9%	100.0%
trnH-psbA	63.2%	96.1%	100.0%
ITS + *rbcL*	87.5%	100.0%	100.0%
ITS + *matK*	90.1%	100.0%	100.0%
ITS + *trnH-psbA*	97.1%[a]	100.0%	100.0%
rbcL + *matK*	68.6%	98.8%[b]	100.0%
rbcL + *trnH-psbA*	75.6%	97.6%	100.0%
matK + *trnH-psbA*	80.5%	98.7%[b]	100.0%
ITS + *rbcL* + *matK*	93.1%	100.0%	100.0%
ITS + *rbcL* + *trnH-psbA*	97.2%[a]	100.0%	100.0%
ITS + *matK* + *trnH-psbA*	97.2%[a]	100.0%	100.0%
rbcL + *matK* + *trnH-psbA*	84.3%	98.8%[b]	100.0%
ITS + *rbcL* + *matK* + *trnH-psbA*	97.2%[a]	100.0%	100.0%

[a]The minor difference in the species discrimination rate among ITS + *trnH-psbA*, ITS + *rbcL* + *trnH-psbA*, ITS + *matK* + *trnH-psbA*, and ITS + *rbcL* + *matK* + *trnH-psbA* is because some terrestrial invasive species do not have all four reliable DNA barcode regions identified in this study.
[b]The minor difference in the genus discrimination rate among *rbcL* + *matK*, *matK* + *trnH-psbA*, and *rbcL* + *matK* + *trnH-psbA* is also because some terrestrial invasive species do not have all four reliable DNA barcode regions identified in this study.

ketones and alkaloids (Ekiert et al., 2020). Mugwort in North America rarely produces viable seeds; instead, it spreads aggressively via the persistent rhizome system. Both common ragweed and mugwort produce deeply lobed leaves and inconspicuous flowers (Figure 6b), which makes it challenging to morphologically distinguish one species from the other (Ustyuzhanin et al., 2017). This problem can be solved by DNA barcoding. The alignments of ITS, *rbcL*, *matK*, and *trnH-psbA* DNA barcode sequences from these two species demonstrated significant nucleotide differences at all four barcode regions (Figure S5).

A third example is that of Canada thistle (*Cirsium arvense*) and bull thistle (*Cirsium vulgare*) (Figure 6c). Canada thistle is native to Europe and western Asia and was accidentally introduced to North America in the 1600s (Moore, 1975; Zouhar, 2001). Bull thistle is native to Europe, western Asia, and northern Africa and was brought into North America as a seed contaminant (Zouhar, 2002). These two TIPS are listed as prohibited noxious weeds by the Michigan Department of Agriculture and Rural Development (MDARD, 1994). Canada thistle propagates by laterally spreading rhizomes and by producing seeds; bull thistle reproduces solely by seed. Both Canada thistle and bull thistle produce lobed, lance-shaped leaves with spiny margins and lavender pink-colored flowers (Figure 6c). This makes it difficult to distinguish the two species solely based on morphology. Such challenge may be overcome by DNA barcoding. As shown in Figure S6, the alignments of the four DNA barcode sequences from Canada thistle and bull thistle showed sufficient nucleotide differences at two DNA barcode regions: ITS and *trnH-psbA*.

A fourth example is that of rabbitfoot clover (*Trifolium arvense*), alsike clover (*Trifolium hybridum*), red clover (*Trifolium pratense*), and white clover (*Trifolium repens*) (Figure 6d). Rabbitfoot clover, alsike clover, and white clover are native to Eurasia, whereas red clover is native to both Eurasia and Africa (Coladonato, 1993). These clover species were introduced to North America during the 1600s–1800s for cultivation of hay, pasture, and fodder; stabilization of soil; or medicinal purposes (Coladonato, 1993). They have now been widespread across North America but are considered invasive or posing threats in some areas, such as Hawaii, Ontario, Minnesota, and

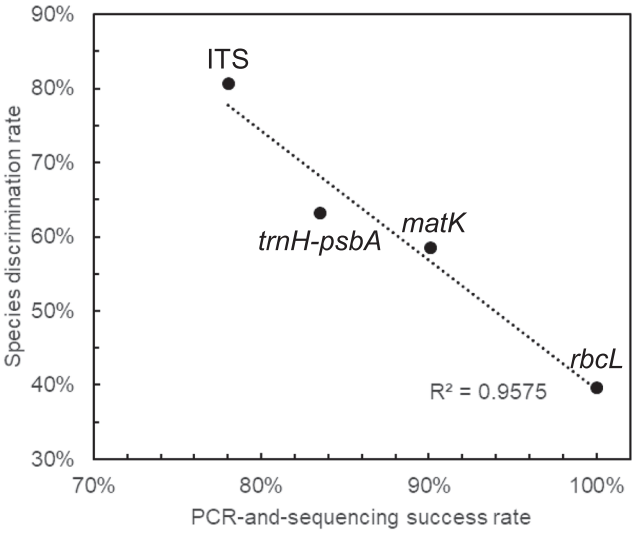

FIGURE 5 Relationship between the species discrimination rate and the PCR-and-sequencing success rate of the four DNA barcode regions.

Wisconsin (HISC, 2012; MIPN, 2024). Rabbitfoot clover, alsike clover, and red clover do not produce stolons; therefore, they have an erect habit. White clover produces stolons; consequently, this species has an erect to creeping habit. The leaflets of rabbitfoot clover are narrowly oblong-shaped, and the flowerhead of rabbitfoot clover is fuzzy and soft like a rabbit's foot (Figure 6d). Therefore, it is easy to distinguish rabbitfoot clover from other clover species. However, alsike clover, red clover, and white clover have similar trifoliate leaves and flowerheads (Figure 6d). Although red clover and white clover have V-shaped watermarks on their leaflets, the watermarks on white clover leaflets are not always obvious (Figure 6d). These morphological characteristics make it challenging to distinguish between some clover species, especially alsike clover and white clover, which have similarly colored flowerheads (Figure 6d). DNA barcoding may solve this problem. As shown in Figure S7, these four clover species have an identical *rbcL* sequence; however, sufficient nucleotide differences exist at ITS and *matK*. It should be noted that we did not perform alignments with *trnH-psbA*, because we were unable to obtain reliable *trnH-psbA* sequences from rabbitfoot clover or alsike clover (Table S2).

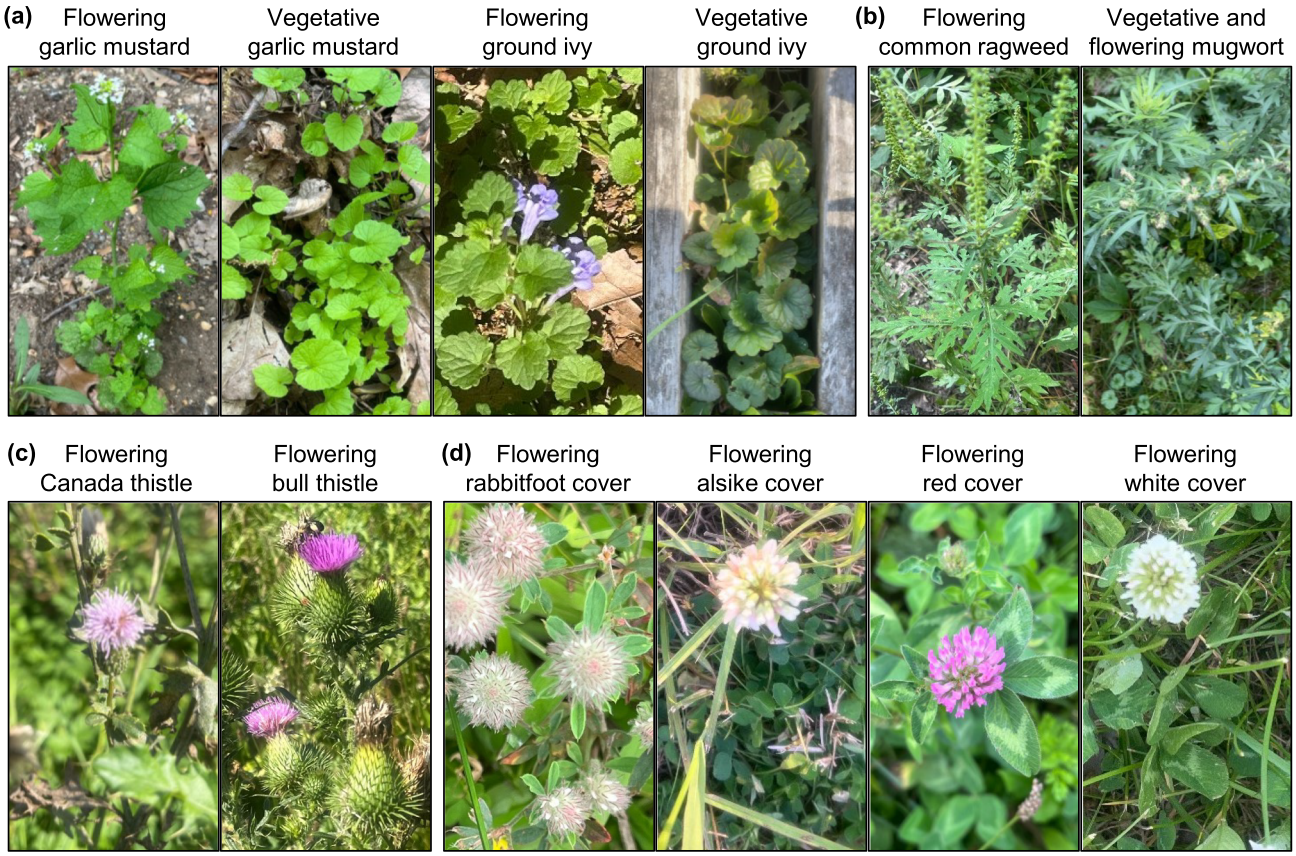

FIGURE 6 DNA barcoding to aid in the identification of TIPS that are morphologically similar to other species. (a) Garlic mustard (*Alliaria petiolata*) and ground ivy (*Glechoma hederacea*). (b) Common ragweed (*Ambrosia artemisiifolia*) and mugwort (*Artemisia vulgaris*). (c) Canada thistle (*Cirsium arvense*) and bull thistle (*Cirsium vulgare*). (d) Rabbitfoot clover (*Trifolium arvense*), alsike clover (*Trifolium hybridum*), red clover (*Trifolium pratense*), and white clover (*Trifolium repens*).

3 | DISCUSSION

Whether DNA barcoding is successful heavily depends on the amplification and sequencing of specific DNA barcode regions from target species (Hollingsworth et al., 2009; Kress et al., 2005; Wang et al., 2017; Xu et al., 2018). In this study, the overall success rate of amplifying and sequencing four DNA barcode regions from 91 terrestrial plant species, most of which are invasive, is 87.9% (Table 2). This suggests that DNA barcoding can be used in the identification of invasive plant species, especially TIPS.

The species discrimination rate of a DNA barcode region seems to be inversely related to the ease of PCR amplification and sequencing (Figure 5). In this study, ITS has the highest species discrimination rate (80.6%; Table 3) but the lowest PCR-and-sequencing success rate (78.0%; Table 2). The relatively high species discrimination rate and the relatively low PCR-and-sequencing success rate of ITS were also reported by other studies (Kress & Erickson, 2007; Wang et al., 2017; Xu et al., 2018). *trnH-psbA* has the second highest species discrimination rate (63.2%) and the second lowest PCR-and-sequencing success rate (83.5%). *rbcL* has the lowest species discrimination rate (39.6%) but the highest PCR-and-sequencing success rate (100.0%). This is not surprising because *rbcL* has very low sequence divergence (Kress et al., 2005, 2009). Our comparison of mismatches and indels between cultivated tobacco and bittersweet nightshade also suggests that *rbcL* has the lowest sequence divergence among the four DNA barcode regions (Table 1).

The inverse relationship between the species discrimination rate and the PCR-and-sequencing success rate of DNA barcode regions indicates that the selection of DNA barcode regions is complicated by the trade-off between high sequence divergence and the need for universal application (Kress et al., 2005; Kress & Erickson, 2007). Apparently, using one single DNA barcode region is not enough to discriminate all the terrestrial plant species included in this study (Table 3). Among the four single DNA barcode regions tested, the average species discrimination rate is 6.5%. Fortunately, using two DNA barcode regions simultaneously may resolve this issue: the average species discrimination rate of the six two-barcode combinations is 83.2%. Among them, the combination of ITS + *trnH-psbA* has the highest species discrimination rate (97.1%). Further increases in the number of DNA barcode regions show little or no additional increases in the species discrimination rate (Table 3). Therefore, dual-barcode approaches (e.g., ITS + *trnH-psbA*) seem to be the efficient and cost-effective methods in DNA-based TIPS identification. This finding is in line with the proposal that the combination of ITS and *trnH-psbA* is the most appropriate dual-DNA barcode combo for invasive plants in China (Xu et al., 2018). Similar approaches (e.g., *rbcL* + *trnH-psbA* and *rbcL* + *matK*) were proposed for land plants in general by Kress and Erickson (2007) and Hollingsworth et al. (2009), respectively. However, in our study, the species discrimination rates of *rbcL* + *trnH-psbA* and *rbcL* + *matK* (i.e., 75.6% and 68.6%) are much lower than that of ITS + *trnH-psbA* (97.1%). Another aspect to look at when choosing which two DNA barcode regions to use is the simultaneous failure rate of PCR and sequencing of the two DNA barcode regions. Among the 91 terrestrial plant species included in this study, only two failed at both ITS and *trnH-psbA* loci (Table S2). Therefore, using two DNA barcode regions ITS and *trnH-psbA* together is a promising approach for DNA-based terrestrial species identification, especially TIPS identification.

The high species/genus/family discrimination rates of DNA barcoding (Table 3), especially when ITS and *trnH-psbA* are used together, indicate that DNA barcoding can be used alone in TIPS identification. This is especially important at points of entry, where TIPS are often transported in the form of seeds, seedlings, vegetative tissues, fruits, or fragmentary materials (Hollingsworth et al., 2009; Whitehurst et al., 2020), which lack diagnostic physical characteristics (Xu et al., 2018). DNA barcoding is also very useful in discriminating TIPS that are morphologically similar to other invasive or native species. As shown in Figure 6, many TIPS have similar physical features as other species, which makes it challenging to discriminate one species from another solely based on morphology. Luckily, the ITS, *rbcL*, *matK*, and *trnH-psbA* genomic loci from different terrestrial plant species often displayed substantial nucleotide differences (Figures S4, S5, S6, and S7), allowing us to use DNA barcoding to distinguish TIPS from their native or invasive look-alikes.

In summary, TIPS pose detrimental effects on native species, ecosystems, public health, and the economy. DNA barcoding can be successfully used in TIPS identification, either by itself or in combination with the traditional, morphology-based approach. Among the four DNA barcode regions tested in this work, ITS and *trnH-psbA* have better species discrimination rates than the other two. Furthermore, the species discrimination rate increases substantially when both ITS and *trnH-psbA* are used. Therefore, the ITS + *trnH-psbA* dual-barcode approach seems to be the most efficient and cost-effective method in DNA-based TIPS identification.

4 | MATERIALS AND METHODS

4.1 | Leaf sample collection

We harvested a small piece of leaf tissue (<.1 g) from representative plants with surgical scissors and surgical forceps and then placed it into a 2-mL plastic tube, which contained a 3.2-mm Chrome steel disruption bead (Research Products International). The harvested leaf tissues were stored on dry ice for temporary preservation during field work and were immediately transferred to a −80°C freezer for long-term storage after field work. To avoid cross contamination between samples, the scissors and forceps were washed and wiped with an excess amount of distilled water, disinfected with 70% ethanol, rinsed again with an excess amount of distilled water, and blotted dry with disposable paper wipes.

The metadata of each tissue sample (e.g., geographic coordinates, site locations, and plant images) were originally recorded in the Midwest Invasive Species Information Network (MISIN; https://www.misin.msu.edu/) database in real time. Each record contains the following information: common name, scientific name, observation date, site location, geographic location (latitude and longitude), area size (individual/several, <1000 square feet, 1000 square feet to

.5 acres, .5–1.0 acres, or >1.0 acres), density (sparse, patchy, dense, or monoculture), and images of the plant(s).

4.2 | DNA extraction

The frozen leaf tissue samples were ground into fine powder with the TissueLyser II bead beater (Qiagen) at the oscillation frequency of 30 Hz for 90 s. To make sure all the samples were disrupted evenly, the adapter sets for 2-mL tubes were rotated 180° after grinding for 45 s. To ensure that leaf tissues stay frozen during and after grinding, the adaptor sets were stored at −80°C overnight before use and the samples were stored on dry ice after grinding. DNA was extracted from the frozen tissue powder using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. The DNA concentrations were determined with the NanoDrop 2000 spectrophotometer (Thermo Scientific).

4.3 | PCR amplification

The ITS, *rbcL*, *matK*, and *trnH-psbA* genomic loci were amplified via PCR on the C1000 Touch Thermal Cycler (Bio-Rad). All PCR reactions were carried out in a total volume of 50 μL, including 10 μL of template DNA (100 ng or less), 2.5 μL of 10-μM forward primer, 2.5 μL of 10-μM reverse primer, 1 μL of 10-mM dNTP mix, .5 μL of Phusion High-Fidelity DNA Polymerase (Thermo Scientific; 2.5 units/μL), 10 μL of 5X Phusion High-Fidelity Buffer, and 23.5 μL of nuclease-free water. The primers used in this study are listed in Table S3.

According to the manufacturer's instructions, the following PCR program was used: initial denaturation at 98°C for 30 s; 36 cycles of denaturation at 98°C for 10 s, annealing at 55°C for 30 s, and extension at 72°C for 30 s; final extension at 72°C for 5 min.

4.4 | Inspection and purification of PCR products

After being mixed with 10 μL of 6X DNA Gel Loading Dye (Thermo Scientific), each PCR reaction was inspected by 1% agarose gel electrophoresis and stained with SYBR Safe DNA Gel Stain (Thermo Scientific). Each successful PCR product was excised from the agarose gel with an x-tracta disposable gel band extraction tool and purified with the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. The concentrations of gel-purified PCR products were determined with the NanoDrop 2000 spectrophotometer (Thermo Scientific).

4.5 | Sanger sequencing

Sanger sequencing of gel-purified PCR products was performed with an Applied Biosystems ABI 3730XL Genetic Analyzer (Thermo Fisher Scientific) at the Genomics Core of the Michigan State University Research Technology Support Facility. Each 12-μL sequencing reaction contains 40 ng of a gel-purified PCR product and 30 picomoles (i.e., 3 μL of 10 μM) of the primer.

4.6 | Data analysis

Each Sanger DNA sequencing chromatogram (i.e., trace file in the ab1 format) was manually inspected with the Chromas chromatogram viewing and editing software to make sure that the nucleotide calls are correct. The post-inspection nucleotide sequences were exported in the fasta format and then archived in the Jellyfish sequence analysis and alignment software. Each nucleotide sequence was further edited in Jellyfish to remove the reverse primer sequence at the 3′ end. Sequences (e.g., noisy sequences) that did not pass the initial manual inspection with the Chromas software were not exported or used for downstream sequence analysis.

The post-cleanup nucleotide sequences were submitted to the Nucleotide BLAST Portal (https://blast.ncbi.nlm.nih.gov/Blast.cgi) of the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/) to (a) test whether the DNA-based species name matches with the morphology-based species name and (b) identify the reliable region of each nucleotide sequence (i.e., reliable DNA barcode for each species at the four genomic loci). The reliable DNA barcode sequences were archived in Jellyfish and subsequently submitted to the NCBI for public release. Reliable ITS DNA barcode sequences were submitted via the Eukaryotic Nuclear rRNA/ITS Submission Portal (https://submit.ncbi.nlm.nih.gov/subs/genbank/); reliable *rbcL*, *matK*, and *trnH-psbA* DNA barcode sequences were submitted via the BankIt Submission Portal (https://www.ncbi.nlm.nih.gov/WebSub/index.cgi).

To calculate species/genus/family discrimination rates for the four DNA barcode regions, the numbers of matching species/genera/families for each reliable DNA barcode were recorded. For example, if an ITS DNA barcode had only one matching species in the NCBI, this barcode would be successful at species discrimination and would be assigned the numeric code 1. If an ITS DNA barcode had two or more matching species, this barcode would be unsuccessful at species discrimination and would be assigned the numeric code 0. The mean value of such numeric codes among all the ITS DNA barcode sequences was the species discrimination rate of the ITS DNA barcode region.

The Pairwise Sequence Alignment tool (https://www.ebi.ac.uk/jdispatcher/psa/emboss_needle), which employs the Needleman-Wunsch alignment algorithm, was used to align two DNA barcode sequences from the same locus (Madeira et al., 2022). The Multiple Sequence Alignment tool (i.e., Clustal Omega; https://www.ebi.ac.uk/jdispatcher/msa/clustalo) was used to align three or more DNA barcode sequences from the same locus (Madeira et al., 2022). Default parameters were used in both cases.

4.7 | Accession numbers

The nucleotide sequences of the four DNA barcode regions amplified from various terrestrial invasive plant species in this study can be found in the GenBank/EMBL databases. The accession numbers are listed in Table S2.

AUTHOR CONTRIBUTIONS

Yan Lu conceived and designed the project. Sneha Nath, Joshua T. VanSlambrouck, Janelle W. Yao, Ashika Gullapalli, Fayyaz Razi, and Yan Lu performed the experiments and analyzed the data. Yan Lu wrote the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge the Western Michigan University Asylum Lake Preserve Management Council, Dr. Steven Kohler (Emeritus Professor of Biological Sciences at WMU), Ms. Kathleen Hoyle (Director of Portage Parks and Recreation), and Ms. Annie Pryor (Recreation Program Manager of Portage Parks and Recreation) for their strong support of this project. The authors thank Ms. Claire Peterson at Michigan State University for her assistance in downloading the relevant metadata stored in the MISIN database. The authors also thank Mr. Alex J. Kolstoe, Ms. Charlotte Resek, Mr. Thomas Wheatley, Ms. Danielle Ahn, Mr. Theophilus Addo, and Ms. Glory James for being wonderful co-workers.

CONFLICT OF INTEREST STATEMENT

The Authors did not report any conflict of interest.

DATA AVAILABILITY STATEMENT

All research data are included in the main text and supporting information. Sequencing data from this work have been deposited to the NCBI. The accession numbers can be found in Table S2.

ORCID

Joshua T. VanSlambrouck https://orcid.org/0009-0006-0672-2129
Yan Lu https://orcid.org/0000-0002-3374-7376

REFERENCES

AIPAC. (2016). *Australian Weeds Strategy 2017 to 2027*. Australian Invasive Plants and Animals Committee. https://www.agriculture.gov.au/sites/default/files/sitecollectiondocuments/pests-diseases-weeds/consultation/aws-final.pdf

Albouchi, F., Hassen, I., Casabianca, H., & Hosni, K. (2013). Phytochemicals, antioxidant, antimicrobial and phytotoxic activities of *Ailanthus altissima* (Mill.) Swingle leaves. *South African Journal of Botany*, *87*, 164–174. https://doi.org/10.1016/j.sajb.2013.04.003

Álvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, *29*, 417–434. https://doi.org/10.1016/S1055-7903(03)00208-2

Alves, I. A. B. S., Miranda, H. M., Soares, L. A. L., & Randau, K. P. (2014). Simaroubaceae family: Botany, chemical composition and biological activities. *Revista Brasileira de Farmacognosia*, *24*, 481–501. https://doi.org/10.1016/j.bjp.2014.07.021

Andonova, T., Muhovski, Y., Slavov, I., Vrancheva, R., Georgiev, V., Apostolova, E., Naimov, S., Mladenov, R., Pavlov, A., & Dimitrova-Dyulgerova, I. (2023). Phenolic profile, antioxidant and DNA-protective capacity, and microscopic characters of *Ailanthus altissima* aerial substances. *Plants*, *12*, 920. https://doi.org/10.3390/plants12040920

Blaxter, M. L. (2004). The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences*, *359*, 669–679. https://doi.org/10.1098/rstb.2003.1447

Boskabadi, J., Askari, Z., Zakariaei, Z., Fakhar, M., & Tabaripour, R. (2021). Mild-to-severe poisoning due to Conium maculatum as toxic herb: A case series. *Clinical Case Reports*, *9*, e04509. https://doi.org/10.1002/ccr3.4509

Callaway, R. M., & Aschehoug, E. T. (2000). Invasive plants versus their new and old neighbors: A mechanism for exotic invasion. *Science*, *290*, 521–523. https://doi.org/10.1126/science.290.5491.521

Castaño-Quintana, K., Montoya-Lerma, J., & Giraldo-Echeverri, C. (2013). Toxicity of foliage extracts of *Tithonia diversifolia* (Asteraceae) on *Atta cephalotes* (Hymenoptera: Myrmicinae) workers. *Industrial Crops and Products*, *44*, 391–395. https://doi.org/10.1016/j.indcrop.2012.11.039

Chase, M. W., Cowan, R. S., Hollingsworth, P. M., van den Berg, C., Madriñán, S., Petersen, G., Seberg, O., Jørgsensen, T., Cameron, K. M., Carine, M., Pedersen, N., Hedderson, T. A. J., Conrad, F., Salazar, G. A., Richardson, J. E., Hollingsworth, M. L., Barraclough, T. G., Kelly, L., & Wilkinson, M. (2007). A proposal for a standardised protocol to barcode all land plants. *Taxon*, *56*, 295–299. https://doi.org/10.1002/tax.562004

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., & Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*, *5*, e8613. https://doi.org/10.1371/journal.pone.0008613

Cipollini, D., Stevenson, R., Enright, S., Eyles, A., & Bonello, P. (2008). Phenolic metabolites in leaves of the invasive shrub, *Lonicera maackii*, and their potential phytotoxic and anti-herbivore effects. *Journal of Chemical Ecology*, *34*, 144–152. https://doi.org/10.1007/s10886-008-9426-2

Coladonato, M. (1993). Trifolium repens. In *In: Fire effects information system, [online]*. United States Department of Agriculture. https://www.fs.usda.gov/database/feis/plants/forb/trirep/all.html

de Vere, N., Rich, T. C. G., Trinder, S. A., & Long, C. (2015). DNA barcoding for plants. In J. Batley (Ed.), *Plant genotyping: Methods and protocols* (pp. 101–118). Springer. https://doi.org/10.1007/978-1-4939-1966-6_8

Devey, D. S., Chase, M. W., & Clarkson, J. J. (2009). A stuttering start to plant DNA barcoding: Microsatellites present a previously overlooked problem in non-coding plastid regions. *Taxon*, *58*, 7–15. https://doi.org/10.1002/tax.581003

Ekiert, H., Pajor, J., Klin, P., Rzepiela, A., Ślesak, H., & Szopa, A. (2020). Significance of *Artemisia vulgaris* L. (common mugwort) in the history of medicine and its possible contemporary applications substantiated by phytochemical and pharmacological studies. *Molecules*, *25*, 4415. https://doi.org/10.3390/molecules25194415

Enan, M. R., Palakkott, A. R., & Ksiksi, T. S. (2017). DNA barcoding of selected UAE medicinal plant species: A comparative assessment of herbarium and fresh samples. *Physiology and Molecular Biology of Plants*, *23*, 221–227. https://doi.org/10.1007/s12298-016-0412-9

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D. M., Hajibabaei, M., & Barrett, S. C. H. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE*, *3*, e2802. https://doi.org/10.1371/journal.pone.0002802

Fenichel, E. P., Richards, T. J., & Shanafelt, D. W. (2014). The control of invasive species on private property with neighbor-to-neighbor spillovers. *Environmental and Resource Economics*, *59*, 231–255. https://doi.org/10.1007/s10640-013-9726-z

Gielly, L., & Taberlet, P. (1994). The use of chloroplast DNA to resolve plant phylogenies: Noncoding versus rbcL sequences. *Molecular Biology and Evolution*, *11*, 769–777. https://doi.org/10.1093/oxfordjournals.molbev.a040157

Hao, D., Gu, X.-j., Xiao, P., & Peng, Y. (2013). Chemical and biological research of clematis medicinal resources. *Chinese Science Bulletin*, *58*, 1120–1129. https://doi.org/10.1007/s11434-012-5628-7

Hebert, P. D., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, *270*, 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D., Ratnasingham, S., & deWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, *270*(Suppl 1), S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Heisey, R. M. (1990). Allelopathic and herbicidal effects of extracts from tree of heaven (*Ailanthus altissima*). *American Journal of Botany*, *77*, 662–670. https://doi.org/10.1002/j.1537-2197.1990.tb14451.x

HISC. (2012). *Maui - Haleakala National Park - invasive plant field guide*. Hawaii Invasive Species Council. https://www.643pest.org/fieldguides/NPS_CARDS_HALE_12112012_final.pdf

Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M. W., Cowan, R. S., Erickson, D. L., Fazekas, A. J., Graham, S. W., James, K. E., Kim, K.-J., Kress, W. J., Schneider, H., van AlphenStahl, J., Barrett, S. C. H., van den Berg, C., Bogarin, D., … Little, D. P. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, *106*, 12794–12797. https://doi.org/10.1073/pnas.0905845106

Holmes, T. P., Aukema, J. E., Von Holle, B., Liebhold, A., & Sills, E. (2009). Economic impacts of invasive species in forests. *Annals of the New York Academy of Sciences*, *1162*, 18–38. https://doi.org/10.1111/j.1749-6632.2009.04446.x

Hulley, E. N., Taylor, N. D. J., Zarnke, A. M., Somers, C. M., Manzon, R. G., Wilson, J. Y., & Boreham, D. R. (2018). DNA barcoding vs. morphological identification of larval fish and embryos in Lake Huron: Advantages to a molecular approach. *Journal of Great Lakes Research*, *44*, 1110–1116. https://doi.org/10.1016/j.jglr.2018.07.013

IBLC. (2024). *DNA barcoding: A tool for specimen identification and species discovery*. International Barcode of Life Consortium. https://ibol.org/about/dna-barcoding/

Kelly, L. J., Ameka, G. K., & Chase, M. W. (2010). DNA barcoding of African Podostemaceae (river-weeds): A test of proposed barcode regions. *Taxon*, *59*, 251–260. https://doi.org/10.1002/tax.591023

Kress, W. J., & Erickson, D. L. (2007). A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*, *2*, e508. https://doi.org/10.1371/journal.pone.0000508

Kress, W. J., & Erickson, D. L. (2008). DNA barcodes: Genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences*, *105*, 2761–2762. https://doi.org/10.1073/pnas.0800476105

Kress, W. J., Erickson, D. L., Jones, F. A., Swenson, N. G., Perez, R., Sanjur, O., & Bermingham, E. (2009). Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences*, *106*, 18621–18626. https://doi.org/10.1073/pnas.0909820106

Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., & Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences*, *102*, 8369–8374. https://doi.org/10.1073/pnas.0503123102

Kruk, C., Huszar, V. L. M., Peeters, E. T. H. M., Bonilla, S., Costa, L., Lürling, M., Reynolds, C. S., & Scheffer, M. (2010). A morphological classification capturing functional variation in phytoplankton. *Freshwater Biology*, *55*, 614–627. https://doi.org/10.1111/j.1365-2427.2009.02298.x

Lahaye, R., van der Beank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T. G., & Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences*, *105*, 2923–2928. https://doi.org/10.1073/pnas.0709936105

Lai, Y. J., Li, S. J., & Wang, W. M. (2018). Evolutionary trends in leaf morphology and biogeography of Altingiaceae based on fossil evidence. *Palaeoworld*, *27*, 415–422. https://doi.org/10.1016/j.palwor.2018.06.002

Lawson, J. W., Fennell, M., Smith, M. W., & Bacon, K. L. (2021). Regeneration and growth in crowns and rhizome fragments of Japanese knotweed (*Reynoutria japonica*) and desiccation as a potential control strategy. *PeerJ*, *9*, e11783. https://doi.org/10.7717/peerj.11783

Lucena-Aguilar, G., Sánchez-López, A. M., Barberán-Aceituno, C., Carrillo-Ávila, J. A., López-Guerrero, J. A., & Aguilar-Quesada, R. (2016). DNA source selection for downstream applications based on DNA quality indicators analysis. *Biopreservation and Biobanking*, *14*, 264–270. https://doi.org/10.1089/bio.2015.0064

Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, *50*, W276–W279. https://doi.org/10.1093/nar/gkac240

MDARD. (1994). *Prohibited and restricted weeds*. Michigan Department of Agriculture and Rural Development. https://www.michigan.gov/-/media/Project/Websites/mdard/documents/pesticide-plant-pest/planthealth/michigan_prohibited_and_restricted_weeds.pdf?rev=044acf897d5b432ca34f4b5b2e575657

MDNR. (2009). *Meeting the challenge of invasive plants: A framework for action*. Michigan Department of Natural-Resources. https://www.michigan.gov/invasives/-/media/Project/Websites/invasives/Documents/Response/Invasives_strategy_final.pdf?rev=24c0ade5cd64469b9aefffeda3b2e3b6

MDNR. (2015). *Michigan's terrestrial invasive species state management plan*. Michigan Department of Natural-Resources. https://www.michigan.gov/-/media/Project/Websites/dnr/Documents/PublicLands/Terrestrial_invasivesp_plan.pdf

MIPN. (2024). *Midwest invasive plant list*. Midwest Invasive Plant Network. https://www.mipn.org/plantlist/

Moore, R. J. (1975). The biology of Canadian weeds.: 13. *Cirsium arvense* (L.) Scop. *Canadian Journal of Plant Science*, *55*, 1033–1048. https://doi.org/10.4141/cjps75-163

Munger, G. T. (2001). Alliaria petiolata. In *In: Fire effects information system, [online]*. United States Department of Agriculture. https://www.fs.usda.gov/database/feis/plants/forb/allpet/all.html

Oyeniyi, T. A., Odekanyin, O. O., Kuku, A., & Otusanya, O. O. (2016). Allelopathic effects of *Tithonia diversifolia* extracts on biochemical parameters and growth of *Vigna unguiculata*. *International Journal of Biology*, *8*, 46–57. https://doi.org/10.5539/ijb.v8n3p45

Park, D. S., & Potter, D. (2013). A test of Darwin's naturalization hypothesis in the thistle tribe shows that close relatives make bad neighbors. *Proceedings of the National Academy of Sciences*, *110*, 17915–17920. https://doi.org/10.1073/pnas.1309948110

Pennisi, E. (2007). Wanted: A barcode for plants. *Science*, *318*, 190–191. https://doi.org/10.1126/science.318.5848.190

Radosevich, S. R., Holt, J. S., & Ghersa, C. M. (2007). *Ecology of weeds and invasive plants: Relationship to agriculture and natural resource management*. John Wiley & Sons. https://doi.org/10.1002/9780470168943

Rahman, R., Qudah, T., Shakoor, A., Nisar, S., & Al-Qudah, T. (2017). Spiny sowthistle facts and health benefits: A review. *International Journal of Chemical and Biochemical Sciences*, *11*, 71–74. https://www.iscientific.org/wp-content/uploads/2019/09/9-IJCBS-17-11-9.pdf

Rai, P. K. (2015). Paradigm of plant invasion: Multifaceted review on sustainable management. *Environmental Monitoring and Assessment*, *187*, 759. https://doi.org/10.1007/s10661-015-4934-3

Rai, P. K., Lee, S. S., Bhardwaj, N., & Kim, K.-H. (2023). The environmental, socio-economic, and health effects of invasive alien plants: Review on Tithonia diversifolia (Hemsl.) A. Gray in Asteraceae. *South African Journal of Botany*, *162*, 461–480. https://doi.org/10.1016/j.sajb.2023.09.038

Rai, P. K., & Singh, J. S. (2020). Invasive alien plant species: Their impact on environment, ecosystem services and human health. *Ecological Indicators*, *111*, 106020. https://doi.org/10.1016/j.ecolind.2019.106020

Rai, P. K., & Singh, J. S. (2024). Ecological insights and environmental threats of invasive alien plant Chromolaena odorata: Prospects for sustainable management. *Weed Biology and Management*, *24*, 15–37. https://doi.org/10.1111/wbm.12286

Rasmussen, K., Thyrring, J., Muscarella, R., & Borchsenius, F. (2017). Climate-change-induced range shifts of three allergenic ragweeds (*Ambrosia* L.) in Europe and their potential impact on human health. *PeerJ*, *5*, e3104. https://doi.org/10.7717/peerj.3104

Rodgers, V. L., Stinson, K. A., & Finzi, A. C. (2008). Ready or not, garlic mustard is moving in: *Alliaria petiolata* as a member of Eastern North American forests. *Bioscience*, *58*, 426–436. https://doi.org/10.1641/B580510

Setsuko, S., Yoshimura, K., Ueno, S., Worth, J. R. P., Ujino-Ihara, T., Katsuki, T., Noshiro, S., Fujii, T., Arai, T., & Yoshimaru, H. (2023). A DNA barcode reference library for the native woody seed plants of Japan. *Molecular Ecology Resources*, *23*, 855–871. https://doi.org/10.1111/1755-0998.13748

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., Siripun, K. C., Winder, C. T., Schilling, E. E., & Small, R. L. (2005). The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, *92*, 142–166. https://doi.org/10.3732/ajb.92.1.142

Simpson, M. G. (2010). *Plant systematics*. Academic Press. https://doi.org/10.1016/C2009-0-02260-0

Soltis, D. E., Kuzoff, R. K., Mort, M. E., Zanis, M., Fishbein, M., Hufford, L., Koontz, J., & Arroyo, M. K. (2001). Elucidating deep-level phylogenetic relationships in Saxifragaceae using sequences for six chloroplastic and nuclear DNA regions. *Annals of the Missouri Botanical Garden*, *88*, 669–693. https://doi.org/10.2307/3298639

Stanford, A. M., Harden, R., & Parks, C. R. (2000). Phylogeny and biogeography of Juglans (Juglandaceae) based on matK and ITS sequence data. *American Journal of Botany*, *87*, 872–882. https://doi.org/10.2307/2656895

Stoeckle, M. (2003). Taxonomy, DNA, and the bar code of life. *Bioscience*, *53*, 796–797. https://doi.org/10.1641/0006-3568(2003)053%5B0796:TDATBC%5D2.0.CO;2

USDA. (2024a). *Economic and social impacts of invasive species*. United States Department of Agriculture. https://www.invasivespeciesinfo.gov/subject/economic-and-social-impacts

USDA. (2024b). *Invasive plants*. United States Department of Agriculture. https://www.fs.usda.gov/wildflowers/invasives/index.shtml

USDOI. (2021). *U.S. Department of the Interior Invasive Species Strategic Plan Fiscal Years 2021-2025*. United States Department-of-the Interior. https://www.doi.gov/sites/doi.gov/files/doi-invasive-species-strategic-plan-2021-2025-508.pdf

Ustyuzhanin, A., Dammer, K.-H., Giebel, A., Weltzien, C., & Schirrmann, M. (2017). Discrimination of common ragweed (*Ambrosia artemisiifolia*) and mugwort (*Artemisia vulgaris*) based on bag of visual words model. *Weed Technology*, *31*, 310–319. https://doi.org/10.1614/WT-D-16-00068.1

Waggy, M. A. (2009). Glechoma hederacea. In *In: Fire effects information system, [online]*. United States Department of Agriculture. https://www.fs.usda.gov/database/feis/plants/forb/glehed/all.html

Wang, A., Gopurenko, D., Wu, H., & Lepschi, B. (2017). Evaluation of six candidate DNA barcode loci for identification of five important invasive grasses in eastern Australia. *PLoS ONE*, *12*, e0175338. https://doi.org/10.1371/journal.pone.0175338

Whitehurst, L. E., Cunard, C. E., Reed, J. N., Worthy, S. J., Marsico, T. D., Lucardi, R. D., & Burgess, K. S. (2020). Preliminary application of DNA barcoding toward the detection of viable plant propagules at an initial, international point-of-entry in Georgia, USA. *Biological Invasions*, *22*, 1585–1606. https://doi.org/10.1007/s10530-020-02204-w

Xu, S. Z., Li, Z. Y., & Jin, X. H. (2018). DNA barcoding of invasive plants in China: A resource for identifying invasive plants. *Molecular Ecology Resources*, *18*, 128–136. https://doi.org/10.1111/1755-0998.12715

Yu, J., Xue, J.-H., & Zhou, S.-L. (2011). New universal matK primers for DNA barcoding angiosperms. *Journal of Systematics and Evolution*, *49*, 176–181. https://doi.org/10.1111/j.1759-6831.2011.00134.x

Zouhar, K. (2001). *Cirsium arvense. In: Fire effects information system, [online]*. United States Department of Agriculture. https://www.fs.usda.gov/database/feis/plants/forb/cirarv/all.html

Zouhar, K. (2002). Cirsium vulgare. In *In: Fire effects information system, [online]*. United States Department of Agriculture. https://www.fs.usda.gov/database/feis/plants/forb/cirvul/all.html

SUPPORTING INFORMATION

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

How to cite this article: Nath, S., VanSlambrouck, J. T., Yao, J. W., Gullapalli, A., Razi, F., & Lu, Y. (2024). DNA barcoding of terrestrial invasive plant species in Southwest Michigan. *Plant Direct*, *8*(6), e615. https://doi.org/10.1002/pld3.615