CASTANEA 89(2): 149–168 July

Molecular Barcoding the Rare Plants of Tennessee

Edward E. Schilling,[1*] Aaron Floden,[2] Jordan Reed,[1] Cory Hale,[1] Todd Crabtree,[3] and Caitlin Elam[3]

[1]Department of Ecology & Evolutionary Biology, University of Tennessee, Knoxville TN 37996

[2]Missouri Botanical Garden, St. Louis, MO 63121

[3]Natural Heritage Inventory Program, Tennessee Department of Environment & Conservation, Nashville TN 37243

ABSTRACT

The focus of this study was on obtaining DNA sequence information from selected rare plant species of Tennessee and their close relatives for the nuclear ribosomal ITS region, which is widely used for molecular barcoding. New sequence data was obtained for 71 species from 18 genera. The ITS region provided a good molecular barcode for nine of the genera. For the other genera, the results suggested that the ITS region would be uninformative, problematic, or require detailed analysis for application. For the orchid genus *Platanthera*, data from the ITS marker allowed unambiguous identification of vegetative plants of the federally listed *P. integrilabia*, which will help make efficient use of resources for conservation of the species. For the aquatic genus *Potamogeton*, the ITS data suggested the presence of unsuspected interspecific hybridization involving the rare species *P. tennesseensis* which showed the need for further investigation. The sequencing results provided validation of the separation of *Lobelia gattingeri* as distinct from *L. appendiculata*. The results of the study expanded the database of DNA sequences for rare plants of Tennessee and also highlighted the need for further study of the flora of the state and of the southeastern United States.

Key Words: conservation; molecular barcoding; nrDNA ITS; plant identification; rare plant species

INTRODUCTION

Even after two centuries of study, knowledge of the flora of the state of Tennessee is still incomplete. Tennessee includes a broad range of ecosystems and habitats, ranging from mountainous terrain in the eastern part of the state to flatter, lower-lying areas in the west and south which provide conditions for a broad range of plant species and communities (Tennessee Flora Committee 2015). Multiple species reach their southern range limit in the mountains or are disjunct from coastal plains or more western prairie habitats—in some cases when these have been isolated for long enough, speciation may have occurred to produce an endemic. Particularly in groups where morphological characters distinguishing species are obscure, the species status of disjunct populations may not have been assessed. Near the margins of some plant species ranges, individuals or populations tend to be rare, either globally or within the political boundary of the state. It is significant to provide protection for these rare plants, not just because of their occurrence within state boundaries, but because they help to characterize the genetic and ecophysiological limits of their lineages (Gaier and Resasco 2023). A key aspect of locating rare species, however, is being able to identify them accurately.

Identification of unknown plant material has long been a challenge, and the use of DNA sequence data to aid in the accurate identification of such material has been a major advance (Kress et al. 2005). There have been multiple uses of this approach both in taxonomy and ecology, reviewed by

*email address: eschilling@utk.edu

Received 17 September 2023; Accepted 30 July 2024

Gostel and Kress (2022). Not only can DNA sequence data be used to identify unknown samples, such data can also help to detect previously unrecognized taxa (Hebert et al. 2004; Schilling et al. 2019). In conservation biology, it can facilitate identification of material in the vegetative state, which can be helpful when rare plant surveys are done out of blooming season, or for groups such as the orchid genus *Platanthera* which don't flower every year. It also has wide applications in other fields as an aid for identifying, or narrowing the identification, of plant material from sources as disparate as human foods, medicinal preparations, and animal feces (Dugan et al. 2007; Nazar et al. 2023; Petrone et al. 2023).

A prerequisite for accurate and comprehensive identification using molecular data is possession of a database representing DNA sequences from all of the "known" taxa. A standard public database is GenBank, maintained by the NCBI. The coverage is still uneven, both by taxon and by geography, and there are many plant species that are not represented in the GenBank database, including several analyzed in the current study. For some geographical areas, efforts have been made to create comprehensive databases of species within their boundaries (Thornhill et al. 2017; Jones et al. 2021). Coverage is less thorough in many other geographical areas, particularly the southeastern United States. As a prelude to this study, we made a survey of species from Tennessee that are represented in the GenBank database for three widely used DNA markers, the nuclear ribosomal Internal Transcribed Spacer (ITS), and the plastid markers *rbcL* and *matK*. The result showed that 17% of 3,460 species of flowering plants found in Tennessee (and Kentucky) lacked data for any of the three markers (data not shown). For the ITS region, which has been particularly useful as a molecular barcode marker because of its combination of near universality in sequencing success, and appropriate level of variability in many genera to allow species-level identifications (Yao et al. 2010; Wang et al. 2014), the number of taxa lacking data was even higher, at 26%. For this study, we concentrated on obtaining and analyzing data for the ITS region, which because of its higher level of variability more frequently allows identification to the species level.

Although the ITS region is readily amplified and sequenced in the majority of flowering plants, a few aspects can prove problematic (Hollingsworth et al. 2011). One is the presence of fungal contamination—the primers used for routine amplification and sequencing were developed from fungi, and not surprisingly will co-amplify if there is mold or other fungal contamination to a specimen. This results in a direct sequence read in which there are multiple polymorphic positions to the point of being unusable. The fungal ITS region is typically about 50–100 bases shorter than the plant region, so fungal contamination can often be detected if there is "clean" sequence at the very end of an otherwise "messy" pherogram. Sequence can be obtained in many cases using angiosperm-specific primers located in the 5.8 coding region, but this requires additional sequence reactions making the process more expensive and time-consuming. Polymorphic sequences may also be obtained if the plant is a hybrid or hybrid-derived, owing to the presence of more than one copy, especially if there are indels that change the sequence length. This is notoriously the case in genera such as *Eupatorium*, where many species are hybrid-derived apomicts that are perpetuated asexually (Schilling and Grubbs 2016). The flip side is that detailed analysis involving additional sequencing reactions can determine the parentage in such cases, and thus identify hybrids or hybrid-derived species.

The goal of the current study was to improve coverage of flowering plant species from Tennessee for the molecular barcode marker ITS. We chose to focus on the state because that is a level at which significant conservation decisions are often made. Because of the large number of species still lacking data, the current study focused on genera that included rare taxa which lacked ITS data in Genbank, particularly those rated G1 (globally rare) and S1 (rare in the state), as reported in Crabtree (2021); this included *Astragalus bibullatus*, *Minuartia godfreyi* (now *Sabulina paludicola*), *Paysonia perforata*, *Paysonia stonensis, Stachys glandulosissima,* and *Stenanthium diffusum* (see Table 1 for authorities of taxonomic names). Sampling was extended to a few other genera containing state-listed species, most of which included multiple species lacking ITS data in Genbank, including *Ammoselinum*, *Desmodium, Galium*, *Lobelia*, *Phemeranthus*,

Table 1. List of genera and species (state-listed rare species shown in bold) and sources of nrDNA ITS sequences used for molecular barcoding; vouchers deposited at TENN herbarium; polymorphic, sample exhibited length polymorphisms that prevented generation of clean sequence.

Genus/Species	GenBank #	Location (voucher)
Ammoselinum Torr. & A. Gray		
A. butleri J.M.Coult. & Rose	MK087995	Missouri
A. butleri	OR392547 (this study)	TN, Loudon Co., *McNeilus 01-07*
A. popei Torr. & A.Gray	OR392546 (this study)	TN, Loudon Co., *Pyne 95-021*
A. popei	ERR5034947	Tamaulipas, Mexico
Arabis L.		
A. patens Sull.	FJ187964	North Carolina
A. patens	FJ187855	North Carolina
Borodinia N. Busch		
B.dentata (Raf.) P.J.Alexander & Windham	JX147004	Illinois
B. dentata	JX147003	Illinois
***B. dentata* (*B. shortii*)**	EU274846	Ontario, Canada
B. laevigata (Muhl. ex Willd.) P.J.Alexander & Windham	EU274858	TN, Marshall Co.
B. perstellata (E.L.Braun) P.J.Alexander & Windham	MK355973 (this study)	TN, Rutherford Co., *Bailey & Williams* s.n. 3/11/2000
B. perstellata	EU274862	TN, Davidson Co.
Astragalus L.		
A. bibullatus Barneby & E.L. Bridges	MK355968 (this study)	TN, Rutherford Co., *Wofford & Pyne 1984*
A. bibullatus	SRR26755391	no information
A. tennesseensis A.Gray ex Chapman	MK355969 (this study)	TN, Marshall Co., *McNeilus 98-135*
A. canadensis L.	MT610924	no information
A. canadensis	MF963993	Missouri
Desmodium Desv.		
D. canescens (L.) Poir.	KM098883	Ontario, Canada
D. ochroleucum M.A. Curtis ex Canby	MK355970 (this study)	TN, Perry Co., garden grown
D. ochroleucum	MK355971 (this study)	TN, Franklin Co., *Beck & Datillo* s.n., 2008
Diamorpha smallii Britton – see *Sedum smallii*, below		
Eupatorium L.		
E. leucolepis (DC.) Torr.& A.Gray	polymorphic (this study)	TN, Coffee Co., *Pyne #94-295*
E. leucolepis	DQ415738	Georgia
E. semiserratum DC.	FJ395156	South Carolina

Table 1. continued

Genus/Species	GenBank #	Location (voucher)
Galium L.		
G. lanceolatum Torr.	polymorphic (this study)	TN, Campbell Co., *Floden 1713*
G. latifolium Michx.	polymorphic (this study)	TN, Carter Co., *McCoy* sn. 6/3/2009
G. palustre L.	polymorphic (this study)	TN, Johnson Co., *Wofford 87-46*
G. uniflorum Michx.	polymorphic (this study)	TN, Marion Co., *Beck 2354*
Geum L.		
G. aleppicum Jacq.	KX645652	China
G. aleppicum	KX645653	China
G. aleppicum	SRR12361872 (ITS2 only)	South Dakota
G. (*Acomastylis*) *calthifolia* Scheutz	AJ302338	Sweden
G. canadense Jacq.	DQ006033	Maryland
G. canadense	SRR12361966	South Dakota
G. geniculatum Michx.	AJ302348	Sweden
G. geniculatum	MK355980 (this study)	TN, Carter Co., *Wofford & Patrick 81-17*
G. laciniatum Murray	MK355981 (this study)	TN, Carter Co., *McCoy* s.n. 6/3/2009
G. laciniatum	MG236982 (ITS2 only)	Ontario, Canada
G. laciniatum	SRR26753602 (partial)	no information
G. radiatum Michx.	OR392549 (this study)	NC, Mitchell Co., *Kral 60739*
G. vernum (Raf.) Torr.& A.Gray	AJ302355	Sweden
G. vernum	SRR26752852	no information
G. virginianum L.	MK355982 (this study)	TN, Marion Co., *Beck 6763*
Lobelia L.		
L. amoena Michx.	OR392550 (this study)	TN, Monroe Co., *Churchill 93133*
L. appendiculata A.DC.	OR392551 (this study)	AR, Sebastian Co., *Thomas 165533*
L. canbyi A. Gray	OR392552 (this study)	TN, Van Buren Co., *McNeilus 00-051*
L. cardinalis L.	AY350630	no information
L. cardinalis (*L. splendens*)	SRR12034797	cultivated
L. gattingeri A.Gray	OR392553 (this study)	TN, Wilson Co., McNeilus 00-501
L. inflata L.	OR392554 (this study)	TN, Campbell Co., *Floden 1080*
L. inflata	OR392555 (this study)	TN, Hamblen Co., *Floden & Schilling 2225*
L. nuttallii Roem.& Schult.	OR392556 (this study)	TN, Campbell Co., *Floden & Wofford 587*
L. puberula Michx.	OR392557 (this study)	TN, Knox Co., *Schilling* s.n. 9/4/2021
L. siphilitica L.	DQ006015	Maryland
L. siphilitica	ERR2040691	no information
L. spicata Lam.	OR392558 (this study)	TN, Campbell Co., *Floden & Wofford 566*

Table 1. continued

Genus/Species	GenBank #	Location (voucher)
Paxistima Raf.		
P. canbyi A.Gray	MK355972 (this study)	TN, Hawkins Co., *Somers & Smith* s.n. 1984
P. canbyi	EU328756	cultivated
P. myrsinites (Pursh) Raf.	EU328757	Arizona
Paysonia O'Kane & Al-Shehbaz		
P. densipila (Rollins) O'Kane & Al-Shehbaz	KU975783	Tennessee
P. lescurii (A.Gray) O'Kane & Al-Shehbaz	KU975754	Tennessee
P. perforata (Rollins) O'Kane & Al-Shehbaz	MK355974 (this study)	TN, Wilson Co., *McCoy & Bishop* s.n. 2011
P. perforata	MT175810	Tennessee
P. perforata	MT175812	Tennessee
P. stonensis (Rollins) O'Kane & Al-Shehbaz	MK355975 (this study)	TN, Rutherford Co., *Pyne & Shea* s.n. 1993
P. stonensis	KU975791	Tennessee
P. stonensis	MT175815	no information
Phemeranthus Raf.		
P. mengesii (W.Wolf) Kiger	OR388089 (this study)	TN, Marion Co., *Beck 6265*
P. brevifolius (Torr.) Hershk.	L78038	Arizona
P. confertifolius (Greene) Hershk.	L78039	New Mexico
P. spinescens (Torr.) Hershk.	L78040	Washington
Platanthera Rich.		
P. ciliaris (Willd.) Chapm.	MK356062 (this study)	TN, Campbell Co., *Floden 739*
P. ciliaris	OR392559 (this study)	TN, Campbell Co., Schilling photo voucher 2
P. ciliaris	SRR12576898	Virginia
P. clavellata (Michx.) Luer	MK356073 (this study)	TN, Van Buren Co., *Fleming & Bowman FCF-1535*
P. clavellata	SRR12576897	Virginia
P. clavellata	MG215705 (ITS2 only)	Ontario, Canada
P. clavellata	MG215730 (ITS2 only)	Newfoundland, Canada
P. clavellata	MG215907 (ITS2 only)	Newfoundland, Canada
P. cristata Lindl.	MK356072 (this study)	TN, McMinn Co., Schilling photo voucher 3
P. cristata	OQ474559	no information
P. cristata	OQ474562	no information
P. cristata	OQ474563	no information
P. cristata	OQ474564	no information
P. flava (L.) Lindl. var. *flava*	MK356064 (this study)	TN, Anderson Co., *Durr* s.n. 8/31/2016
P. flava* var. *herbiola (R.Br.) Luer	MK356065 (this study)	TN, Roane Co., *Patrick 4936*
P. grandiflora Lindl.	MG215797 (ITS2 only)	Newfoundland, Canada

Table 1. continued

Genus/Species	GenBank #	Location (voucher)
Platanthera Rich. continued		
P. grandiflora	MG216754 (ITS2 only)	Nova Scotia, Canada
P. grandiflora	MG216763 (ITS2 only)	Nova Scotia, Canada
P. integra (Nutt.) A.Gray	MK356066 (this study)	TN, Van Buren Co., *Jones 5353*
P. integrilabia (Correll) Luer	MK356061 (this study)	TN, Fentress Co., *McCoy, Kerr & Blount* s.n. 8/19/2009
P. integrilabia	MK356071 (this study)	TN, McMinn Co., Schilling photo voucher
P. lacera G.Don	MK356068 (this study)	TN, Coffee Co., *Crain et al. 30*
P. lacera	SRR12576894	cultivated
P. nivea (Nutt.) Luer	MK356067 (this study)	TN, Coffee Co., *Beatley & DeSelm 19464*
P. nivea	MF944379	no information
P. orbiculata (Pursh) Lindl.	OR388090 (this study)	TN, Unicoi Co., *Wofford 79-218*
P. orbiculata	MG216738	Yukon Territory, Canada
P. orbiculata	SRR12576890	Washington
P. peramoena (A.Gray) A.Gray	MK356069 (this study)	TN, McNairy Co., *Cooper* s.n. 7/14/1997
P. psycodes (L.) Lindl.	MK356070 (this study)	TN, Monroe Co., *Boom 557*
P. psycodes	MF944380	no information
Potamogeton L.		
P. amplifolius Tuck.	OR392560 (this study)	TN, Cumberland Co., *Bailey & Shea* s.n. 2000
P. amplifolius	KY695276	New Hampshire
P. crispus L.	MH171006	Türkiye
P. crispus	EF526372	Connecticut
P. crispus	GU814244	India
P. diversifolius Raf.	MK355987 (this study)	TN, Hamilton Co., *Rhinehart* s.n. 6/4/2000
P. diversifolius	KY695267	Louisiana
P. epihydrus Raf.	MK355988 (this study)	TN, Monroe Co., *Webb & Dennis* s.n. 1980
P. epihydrus	FJ151206	Connecticut
P. foliosus Raf.	MK355989 (this study)	TN, Dekalb Co., *Thompson 99-837*
P. foliosus	AY714292	Texas
P. foliosus	KF270907	USA
P. foliosus	GQ247410	Connecticut
P. nodosus Poir.	MK355990 (this study)	TN, Claiborne Co., *Floden 1279*
P. nodosus	HQ263519	Vermont
P. pulcher Tuck.	MF694350	Nova Scotia, Canada
P. pulcher	KY695275	Massachusetts
P. pusillus L.	GQ247421	Connecticut
P. tennesseensis Fernald	MK355991 (this study)	TN, Fentress Co., *Jones 6176*

Table 1. continued

Genus/Species	GenBank #	Location (voucher)
Potamogeton L. continued		
P. tennesseensis	polymorphic (this study)	TN, Scott Co., *Bailey* s.n. 2000
P. tennesseensis	polymorphic (this study)	TN, Cumberland Co., *Bailey & Shea* 6/21/2001
P. zosteriformis Fernald	EF526408	Connecticut
P. zosteriformis	KF270921	Canada
P. zosteriformis	GQ247438	Vermont
Pycnanthemum Michx.		
P. beadlei (Small) Fernald	MK355996 (this study)	TN, Carter Co., *Wofford et al. 79-233*
P. curvipes (Greene) E.Grant & Epling	polymorphic (this study)	TN, Polk Co., *Estes et al. 53*
P. incanum Michx.	AY506640	Indiana
P. loomisii Nutt.	polymorphic (this study)	TN, Campbell Co., *Floden 712*
P. montanum Michx.	polymorphic (this study)	TN, Sevier Co., *Phillippe & Cunningham 42644*
P. muticum (Michx.) Pers.	AY943494	Florida
P. pycnanthemoides (Leavenworth) Fernald	polymorphic (this study)	TN, Bledsoe Co., *Fleming et al FCF-1983*
P. tenuifolium Schrad.	GU381423	United States
P. torreyi Benth.	MK355997 (this study)	TN, Dickson Co., *Kral 35899*
P. verticillatum Pers.	MK355998 (this study)	TN, Polk Co., *Ocoee Survey Group* 8/9/2012
P. virginianum (L.) T.Durand & B.D.Jacks. ex B.L.Rob. & Fernald	DQ667319	no information
P. virginianum	JQ669131	no information
Scutellaria L.		
S. elliptica Muhl. ex Spreng. var. *elliptica*	MK356047 (this study)	TN, Bradley Co., *DeSelm* s.n. 1990
S. elliptica var. *hirsuta* (Short & R.Peter) Fernald	MK356048 (this study)	TN, Bradley Co., *DeSelm* s.n. 7/20/2002
S. incana Biehler var. *incana*	MK356049 (this study)	TN, Campbell Co., *Floden 1050*
S. incana var. *punctata* C.Mohr	MK356050 (this study)	TN, Hamilton Co., *Huskins 1271*
S. integrifolia L.	MK356051 (this study)	TN, McMinn Co., *DeSelm* s.n. 6/17/2006
S. integrifolia	ON890136	Florida
S. lateriflora L.	MK356052 (this study)	TN, Bradley Co., *DeSelm* s.n. 8/29/2006
S. lateriflora	ON890133	Florida
S. lateriflora	ON890134	Florida
S. lateriflora	SRR5602584	no information
S. leonardii Epling	MK356053 (this study)	TN, Marion Co., *Goodman & Larson 4640*
S. montana Chapm.	MK356054 (this study)	TN, Hamilton Co., *Bailey & Lincicome* s.n. 5/23/2002
S. nervosa Pursh	MK356055 (this study)	TN, Hawkins Co., *Floden & Schilling 2257*
S. ovata Hill	MK356056 (this study)	TN, Lawrence Co., *DeSelm* s.n. 6/23/84
S. parvula Michx. var. *parvula*	MK356046 (this study)	TN, Decatur Co., *McCoy & Martin* s.n. 5/25/2010

Table 1. continued

Genus/Species	GenBank #	Location (voucher)
Scutellaria L. continued		
S. parvula var. *australis* Fassett	MK356045 (this study)	TN, Roane Co., *Patrick 3418*
S. pseudoserrata Epling	MK356057 (this study)	TN, Roane Co., *DeSelm* s.n. 5/19/2008
S. saxatilis Riddell	MK356058 (this study)	TN, Johnson Co., *McCoy & Crabtree* s.n. 6/11/12
S. serrata Andrews	MK356059 (this study)	TN, Hawkins Co., *Floden & Schilling 2034*
Sedum L.		
S. acre L.	HE999635	no information
S. nevii A.Gray	OR392561 (this study)	TN, Polk Co., *Estes 2012-139*
S. pulchellum Michx.	OR392562 (this study)	TN, Hickman Co., *Estes 7164*
S. sarmentosum Bunge	EU592003	Shanghai, China
S. smallii (Britton) H.E. Ahles ***(Diamorpha smallii Britton)***	OR392548 (this study)	TN, Putnam Co., *McNeilus 375*
S. ternatum Michx.	OR392563 (this study)	TN, Fentress Co., *Estes & Beck 7390*
Stachys L.		
S. appalachiana D.B.Poind. & J.B.Nelson	polymorphic (this study)	TN, Johnson Co., *Carman & Rhinehart* s.n.
S. clingmannii Small	polymorphic (this study)	TN, Monroe Co., *Phillippe 38917*
S. floridana Shuttlew. ex Benth.	KF529590	Florida
S. glandulosissima Floden	OR392566 (this study)	TN, Polk Co., *Floden 904*
S. latidens Small	OR392564 (this study)	TN, Polk Co., *Floden & Hart 1658*
S. latidens	KF529608	North Carolina
S. tenuifolia Willd.	OR392565 (this study)	TN, Polk Co., *Floden & McFarland 1377*
Stenanthium (A.Gray) Kunth		
S. densum (Desr.) Zomlefer & Judd	AF303722	Florida
S. diffusum Wofford	MK355984 (this study)	TN, Morgan Co., *Wofford 20053*
S. gramineum (Ker-Gawl.) Morong	MK355986 (this study)	TN, White Co., *Estes 09125*
S. gramineum	AF303729	Florida
S. gramineum f. *robustum* (S.Watson) E.J.Palmer & Steyerm.	MK355985 (this study)	TN, *Floden & Datillo 2860*
S. tennesseense Sorrie & Weakley	MK355983 (this study)	TN, Coffee Co., *Crain et al.* s.n.
Thermopsis R.Br.		
T. divaricarpa A.Nelson	AY091575	North America
T. fraxinifolia Nutt. ex M.A. Curtis	OR392567 (this study)	TN, Polk Co., *Estes et al. 2012-150*
T. fraxinifolia	SRR26755951	no information
T. mollis (Michx.) M.A.Curtis ex A. Gray	OR392568 (this study)	TN, Sevier Co., *Pyne 94-071*
T. villosa (Walter) Fernald & B.G. Schub.	OR392569 (this study)	TN, Polk Co., *Estes et al. 247*
T. villosa	AF123444	no information
T. villosa	SRR26755950	no information

Platanthera, *Potamogeton*, *Sedum*, *Stachys*, and *Thermopsis*. The results showed that, in addition to lack of DNA sequence data, there remain gaps in the basic taxonomy and knowledge of distributions. Many of the rare species that lacked sequence data belonged to large, widespread genera that had been studied in other parts of their ranges, but not in the southeastern United States. An example is the orchid genus *Platanthera*, which has about 150 species, including 12 in Tennessee (POWO 2024; The Tennessee Flora Committee 2015). Of these, over half (seven) are considered rare in the state, including *P. integrilabia* which is on the Federally Endangered species list, and many of these lacked ITS sequence data in Genbank. Similarly, many species of *Scutellaria* from the southeastern United States, including the Federally Threatened *S. montana*, lack sequence data in GenBank. In still another example, the widespread genus *Lobelia* is represented in the Southeastern United States by an entire taxonomic section which is both endemic and also almost entirely confined to this region (Spaulding and Barger 2016), but most species including the rare *L. amoena* lacked DNA sequence data. In addition to filling gaps in basic knowledge, the results should stimulate further studies to resolve questions of systematics and evolution for plants of this region, as well as helping conservation efforts.

MATERIALS AND METHODS

Samples were taken either from herbarium specimens deposited at TENN, or in some cases from fresh material (Table 1). Because of resource limitations, most species were sampled from a single specimen, so care was taken in utilizing specimens collected by experienced collectors who were most likely to correctly identify the species; use of deposited specimens also allows the identity to be re-checked by others. Molecular protocols generally followed those described in Schilling (2011). DNA was extracted from a small sample of leaf (ca. 0.2 g) ground in liquid N2, using the Qiagen Plant DNeasy kit protocol. ITS amplifications were performed in 20 µl reactions using 10–20 ng of genomic DNA, 10× PCR buffer (Promega), 1.8–2.25 mM MgCl2, 0.2 mM each dNTP, 1.25 units of Taq polymerase, and 0.2 µM each primer. Primers used were "ITS-4" (5'-TCCTCCGCTTATTGATATGC-3') and "ITS-5" (5'-GGAAGTAAAAGTCGTAACAAGG-3'; White et al. 1990). PCR was performed with the "ETS" protocol: 95°C for 2 min; 10 cycles of 95°C for 1 min, 55°C for 1 min, and 72°C initially for 1 min, with 4 sec added per cycle; 20 cycles of 95°C for 1 min, 50°C for 1 min, and 72°C initially for 1:40, with 4 sec added per cycle; and a final extension of 72°C for 7 min. A few samples were amplified and sequenced using the ''5.8S 79 for'' or ''ITS-5.8SR'' primers from the 5.8 nrDNA coding region (Schilling 2011), each coupled with the corresponding outer primer, ITS4 or ITS5. PCR products were checked on 1% agarose gels before being cleaned with ExoSAP-IT (USB, Cleveland, Ohio). All DNA sequencing was performed with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit, v. 3.1 (Perkin-Elmer/Applied Biosystems, Foster City, California, USA) and electrophoresed and detected on an ABI Prism 3100 automated sequencer (University of Tennessee Molecular Biology Resource Facility, Knoxville, Tennessee). The initial sequence data text files were edited following comparison with the same data displayed in four color electropherograms before they were analyzed further, using Sequencher v. 5.1 (Gene Codes Corporation, Ann Arbor, Michigan), with special attention paid to detecting evidence of positional and length polymorphisms. Sequence alignment was performed using MAFFT v7.308 (Katoh and Standley 2013) implemented in Geneious v. 9.1.7., and pairwise differences were calculated using the Geneious software. GenBank accession numbers are provided in Table 1. For some species, it was possible to add ITS sequences from additional samples using SRA data deposited in Genbank. For these samples, the SRA data were downloaded and the fragments assembled to the ITS region of the same (or closely related) species using the Geneious Map to Reference assembly algorithm, set to medium-low sensitivity and iterated five times. A total of 17 ITS sequences from nine genera were obtained from SRAs (Table 1).

RESULTS AND DISCUSSION

In general, the ITS region was successfully amplified and sequence was obtained from DNA extracts made from herbarium specimens of the study species. Although most specimens were collected after 2000, and thus less than 20 years old, successful results were obtained for seven specimens more than 35 years old. An exception was *Xyris*, for which attempts to extract DNA from herbarium specimens did not provide any PCR amplification results. A few samples with low DNA recovery were sequenced with primers from the internal coding 5.8S region, as noted in the methods section, producing a shorter amplification product which could be sequenced. The results of DNA sequencing of the ITS region are presented by genus in the following paragraphs, followed by an overall summary.

Ammoselinum (Apiaceae)

Ammoselinum includes 4 species of annual aquatic herbs (POWO 2024). It is represented in Tennessee by two species, the non-native *A. butleri* and the rare (S2/G4) *A. popei*. Prior to the start of this study, no DNA sequences had been deposited in Genbank for the genus, although sequences subsequently became available for *A. butleri*. One sample of each of the two Tennessee species was sequenced. The sequence for *A. butleri* was identical to one already in Genbank, whereas the sequence for *A. popei* was 1.5% (7 bp) different from that of *A. butleri*; a further ITS sequence for *A. popei* assembled from SRA ERR5034947 differed at 2 bp positions. Barcoding using the ITS region will thus distinguish samples of *A. popei* from *A. butleri*. A preliminary phylogenetic analysis was performed using ITS sequences downloaded from GenBank of other Apiaceae samples which suggested that *Ammoselinum*, *Spermolepis* Brongn. & Gris, and *Oligocladus* Chodat are not reciprocally monophyletic, so a change in classification that could affect the genus name for *A. popei* may be needed (data not shown).

Arenaria, ***Geocarpon***, ***Minuartia***, ***Mononeuria***, ***Sabulina***, ***Stellaria*** (Caryophyllaceae)

The stitchworts, in many treatments classified in either *Minuartia* Loefl. or *Mononeuria* Rchb., illustrate an application of molecular barcoding to basic classification. The initial barcoding results in the current study were obtained for a rare species listed formerly as *Minuartia* (or *Mononeuria*) *godfreyi* (Shinners) McNeill (S1/G1; the correct name is now proposed to be *Sabulina paludicola* [Fern. & B.G.Schub.] E.E.Schill.) and revealed the need for a major genus-level reclassification. This has been expanded and published as a separate study (Schilling et al. 2022), the results of which will be summarized here. *Mononeuria* comprised six species in Tennessee, with three listed as Endangered (including one that was Federally Listed), and another of Special Concern. The expanded barcoding data revealed that three of the species, including the type species of *Mononeuria*, *M. patula*, belonged to the *Sabulina* Rchb. clade. Because *Sabulina* is an older name than *Mononeuria*, these all were transferred to *Sabulina*. The remaining species were part of a second clade that was phylogenetically separate, for which the genus name *Geocarpon* Mack. applies and has priority in this clade. Two further results are of note. One is the demonstration that *Sabulina paludicola* (formerly *Minuartia godfreyi*) is diverse at the molecular level and may include more than a single species. The second is to cast doubt on whether material considered to be *Sabulina muscorum* (Fassett) E.E.Schill. (G1/S5) was correctly identified, and this species might not even occur in Tennessee. These results should stimulate further research on these tiny but interesting plants. First reports of ITS sequences for two other state-listed Caryophyllaceae species, *Arenaria lanuginosa* (Michx.) Rohrb. (S1/G5) and *Cerastium velutinum* Raf. (S1/G5T4), showed that each was distinct from other members of their respective genera. The level of variability makes the ITS marker a useful molecular barcode region for these genera of Caryophyllaceae in Tennessee.

Arabis, ***Boechera***, ***Borodinia*** (Brassicaceae)

Like Caryophyllaceae, parts of Brassicaceae have seen major rearrangements in genus-level classification, and these have affected species formerly placed in *Boechera*, in which three species considered rare in Tennessee are listed. Many of these are included in older floras as part of *Arabis*.

Two of these are now considered part of *Borodinia*, *B. perstellata* (S1/G2) and *B. dentata* (S1S2/G5; synonym is *Boechera shortii*), and a third has been returned to its earlier home in *Arabis*, *A. patens* (S1/G3). Prior to this study there were no sequences for *B. perstellata* in Genbank, and a newly obtained one matched exactly another that was subsequently deposited there. It differs by at least 5 bp from sequences in Genbank for *B. dentata* and both of these differ by >15% from *A. patens*. The ITS region is a good molecular barcode for the rare Tennessee species in this group.

Astragalus (Fabaceae)

Astragalus includes 3,066 species (POWO 2024), but only three of these occur in Tennessee, and two are found on limestone glades and considered rare, *A. bibullatus* (S1/G1) and *A. tennesseensis* (S3/G3). The ITS sequences of these were 1% (6 bp) different from one another. Other highly similar sequences in GenBank came from species from western North America, with *A. tennesseensis* differing by only two bp changes from *A. whitneyi* and *A. cusickii*. A shorter, ITS-2, sequence for *A. crassicarpus*, which may be a close relative (Barneby and Bridges 1987), was also two bp changes different from *A. tennesseensis*. In contrast, the sequence for the other species native to Tennessee, *A. canadensis*, differs from both by more than 7% (40–43 bp). Thus, the ITS barcode can uniquely identify each of the three Tennessee species of *Astragalus*, and may give information on the geographic origins of the rare species.

Desmodium (Fabaceae)

Desmodium includes 179 species (POWO 2024), with a complex taxonomy (Weakley et al. 2024). There are 17 species recorded for Tennessee, one of which, *D. ochroleucum*, is included in the state rare plant list (S1/G2G3). ITS sequences for two samples of this species were identical to one another, but they were also identical to the widespread *D. canescens*. Thus, the ITS marker will not uniquely identify an unknown sample to this species.

Eupatorium (Asteraceae)

Eupatorium includes two species listed as rare in the state. The ITS sequence of *E. godfreyanum* (S1/G4) is characterized by numerous polymorphisms that reflect its entirely hybrid origin from *E. rotundifolium* L. and *E. sessilifolium* L. (Siripun and Schilling 2006). Examination of material from Tennessee of the other rare species, identified as *E. leucolepis* (S1/G5), showed polymorphisms both for length and base pair substitution to suggest that the Tennessee populations are of hybrid origin, from *E. leucolepis* and *E. semiserratum*, even though their morphology appears to be identical to that of "pure" *E. leucolepis* from elsewhere (this is still under study). Genetically it may represent a distinct species, but its morphology appears to be identical to that of *E. leucolepis*, although this is also still under study. The situation for *Eupatorium* in Tennessee illustrates one of the limitations of using molecular barcodes based on ITS sequences—neither species gave unambiguous sequence on direct sequencing. However, with sequencing using multiple primers and careful interpretation of polymorphisms, the ITS region allows determination of the parentage of hybrids or hybrid-derived taxa of *Eupatorium*.

Galium (Rubiaceae)

Galium includes 646 species (POWO 2024), and is represented in Tennessee by 18 species (five nonnative), and two, *G. asprellum* and *G. palustre*, are state-listed (both S1/G5). Four native species were sampled for ITS sequencing, including *G. palustre*, and all returned sequence that was polymorphic for at least one length variant, which renders subsequent sequence unreadable and would require further sequencing to interpret. The ITS region does not appear to be a suitable molecular barcode marker for rapid identification of unknowns within *Galium* in Tennessee.

Geum (Rosaceae)

Generic level classification of *Geum* has been problematic. It has variously been divided into smaller genera, but molecular and morphological data do not present a clear picture, and for

now they all continue to be considered part of *Geum*, which includes 56 accepted species (POWO 2024). Four of the seven species listed for Tennessee are considered rare in the state, including *G. radiatum* (S1/G2), for which the current study presents the first ITS sequence report. Most of the species had distinct ITS sequences, although *G. laciniatum* and *G. virginianum* were identical; these differed from *G. geniculatum* by 1% (5 bp). The newly obtained sequence of *G. radiatum* was highly distinct, 28–33 bp different from *G. laciniatum*/*G. virginianum* and *G. geniculatum*; this species has been included in segregate genera (*Acomastylis* Greene, *Sieversia* Willd.), and the most similar ITS sequence in Genbank is for the boreal *G. calthifolium* Scheutz, which is 2% (13 bp) different. ITS sequence data can be used to identify the rarest Tennessee species (*G. geniculatum* and *G. radiatum*, both S1/G2), but not *G. laciniatum* (S1/G5). Genbank data for the geographically widespread *G. aleppicum* (S1/G5) appeared to indicate that it is distinct, however, this is complicated by a lack of consistency in the sequences deposited in Genbank, which suggested it needs further investigation. Sequence from various sites in Asia differed by as much as 13–37 bp from one another. The single sample available from North America (South Dakota, assembled from an SRA) was at least 6 bp different for just the ITS-2 region from all of these, and closest to samples of *G. geniculatum* (3 differences in ITS-2). It was not possible to include samples of *G. aleppicum* from Tennessee in the current study, so it is unclear if ITS sequences would clarify its identity. As a side note, a BLAST search of the ITS2 region of *G. geniculatum* returned strong matches to sequences labeled as *Agrimonia striata* (99%; MG236054) and *Prunus triloba* (99%; JF421470), both produced by mass molecular barcoding projects; the lack of careful scrutiny in such projects shows the need for caution in using GenBank as a definitive source of reference material.

Lobelia (Campanulaceae)

Lobelia is a widespread genus with 441 accepted species (POWO 2024), and is another example of a widespread, species-rich genus that has been well studied in some geographic regions, but not in the southeastern United States. Of the nine Tennessee species, only *L. amoena* (S1S2/G4) is state-listed as rare. A newly obtained ITS sequence for *L. amoena* differed by at least 23 bp (>3%) from all other sequences for the genus deposited in Genbank as well as five other new ones from the current study (Table 2). Another taxon, classified by some as *L. appendiculata* var. *gattingeri* (but accepted elsewhere as the distinct species *L. gattingeri*) gave ITS sequences that were 29 bp different from *L. appendiculata* s.s., and thus provides support that the two taxa should be recognized as distinct species. The Tennessee species of *Lobelia* differed from one another by at least 23 bp differences, whereas samples within a species differed by 0–2 bp (Table 2). Thus, the ITS region appears to be suitable as a molecular barcode for species of *Lobelia* from Tennessee.

Paxistima (Celastraceae)

Paxistima includes two North American species (POWO 2024), one of which, *P. canbyi*, occurs in eastern North America and is rare in Tennessee (S1/G2). An ITS sequence was obtained for a sample originating from Tennessee, and it was identical to a sequence in Genbank that had been obtained from a cultivated plant of unknown origin. The ITS sequence is highly distinctive, differing by 14 bp (2%) from the western North American *P. myrsinites*, and by >10% from all other Genbank sequences.

Paysonia (Brassicaceae)

Paysonia includes eight accepted species (POWO 2024), and is represented in Tennessee by four species, three of which are state-listed. ITS sequences were obtained for the two species listed as S1/G1, *P. perforata* and *P. stonensis*, and these differed from each other at two positions. Previous work using shotgun sequencing approaches had documented slight variability within individuals for ITS sequence (Mazie and Baum 2016), and *P. stonensis*, *P. lescurii*, and *P. densipila* were not unambiguously separated from one another based on ITS sequence data; inclusion of clones reported

Table 2. Pairwise differences between species of *Lobelia* for nrDNA ITS sequences. The main diagonal shows pairwise differences among samples within the species (–, only single sample available). Species abbreviations: amo, *amoena*; app, *appendiculata*; can, *canbyi*; car, *cardinalis*; gat, *gattingeri*; inf, *inflata*; nut, *nuttallii*; pub, *puberula*; sip, *siphilitica*.

	amo	app	can	car	gat	inf	nut	pub	sip	spi
amo	–									
app	45	*0-2*								
can	49	43	–							
car	44	53	54	*0*						
gat	48	29	52	61	*0-2*					
inf	68	70	56	78	77	*2*				
nut	50	52	37	52	62	69	–			
pub	48	44	24	60	53	57	45	–		
sip	23	34	34	39	37	60	38	35	*2*	
spi	49	27	43	54	31	73	55	44	38	–

in GenBank extended this result to *P. perforata*. The ITS region does not provide a suitable marker for confirming species-level identification of Tennessee species of this genus.

Phemeranthus (Portulacaceae)

Phemeranthus includes 25 New World species (POWO, 2024). Each of the three Tennessee species of *Phemeranthus* is state-listed, and there are currently no ITS sequences for any of them, although the genus has been studied from other parts of its range. Because of technical difficulties with obtaining sufficient DNA from herbarium samples, it was only possible to obtain sequences for the ITS-2 region. A sample of *P. mengesii* was sequenced successfully, and was >6% different from samples of other species of the genus on Genbank, but samples of *P. calcaricus* and *P. teretifolius* exhibited length polymorphisms making it impossible to obtain clean sequence. Both *P. calcaricus* and *P. teretifolius* have been reported to be tetraploids, whereas *P. mengesii* is reported to be diploid. Thus, ITS sequences may be informative for phylogenetic study of *Phemeranthus* in Tennessee, but may not be ideal for molecular barcodes because of difficulty in obtaining clean sequence.

Platanthera (Orchidaceae)

Platanthera includes 150 accepted species (POWO 2024) distributed broadly across the northern hemisphere. There are 13 taxa of *Platanthera* recorded for Tennessee, of which seven are listed as rare, including the Federally Endangered *P. integrilabia*. Although *Platanthera* has been subject to molecular phylogenetic investigation (Bateman et al. 2009), most of the Tennessee species including *P. integrilabia* lacked an entry for the ITS region in GenBank prior to this study. A combination of newly generated ITS sequences and ones downloaded from GenBank accounted for all Tennessee species, and almost all had distinctive sequences, mostly differing by 10–76 bp (Table 3); species where it was possible to obtain multiple samples revealed infraspecific variation of 0–4 bp. The two varieties of *P. flava* differed by a single bp change. The species *P. ciliaris* and *P. cristata*, as well as *P. blepharoglottis* (which does not occur in Tennessee) have been documented to lack differentiation for the ITS marker, and to hybridize extensively at some sites (Evans et al. 2023). *Platanthera grandiflora* and *P. psycodes* differed by only 3 bp, although only ITS-2 sequence data were available; these have also been documented to hybridize. In contrast, the ITS region differentiates *P. integrilabia* from all other species in Tennessee, or with records in Genbank, by 10–73 bp (Table 3). The distinctiveness of the ITS sequence for *P. integrilabia* allows positive identification of unknowns of this species based on vegetative material, from which DNA sequence is readily obtained. During this project, we confirmed identification of five unknowns as *P. integrilabia*, and conversely determined that an additional eight samples were not this species. Thus, the ITS molecular barcode functions well for *Platanthera* in Tennessee and has already had practical application.

Table 3. Pairwise differences between taxa of *Platanthera* for nrDNA ITS sequences. The main diagonal shows pairwise differences among samples within the species (-, only single sample available). Note that data for *P. grandiflora* included only the ITS-2 region. Taxon abbreviations: cil, *ciliaris*; cla, *clavellata*; cri, *cristata*; f/f, *flava* var. *flava*; f/h, *flava* var. *herbiola*; gra, *grandiflora*; int, *integra*; ila, *integrilabia*; lac, *lacera*; niv, *nivea*; orb, *orbiculata*; per, *peramoena*; psy, *psycodes*.

	cil	cla	cri	f/f	f/h	gra	int	ila	lac	niv	orb	per	psy
cil	*0-2*												
cla	19	*0-4*											
cri	2	18	*0*										
f/f	74	74	73	–									
f/h	73	73	72	1	–								
gra	43	38	42	33	33	*0*							
int	23	22	22	76	75	41	–						
ila	11	16	10	73	72	41	20	*0*					
lac	69	64	68	60	59	20	68	65	*0*				
niv	22	21	21	76	75	40	25	19	65	*3*			
orb	52	49	51	46	46	53	52	50	47	50	*0*		
per	68	63	67	58	57	20	67	64	17	64	48	–	
psy	76	73	75	63	62	3	76	72	30	74	56	25	*3*

Potamogeton (Potamogetonaceae)

Potamogeton is a cosmopolitan genus of 90 accepted species (POWO 2024) represented by ten taxa in Tennessee, three of which are considered rare: *P. amplifolius* (S1/G5), *P. epihydrus* (S1S2/G5), and *P. tennesseensis* (S2/G2G3). As is often the case with aquatic plants, identification can be difficult, and there is frequent hybridization, making molecular barcoding a useful tool (Kaplan et al. 2018). Sequence data of the ITS region for all of the Tennessee species except for *P. tennesseensis* were available prior to this study. ITS sequence was obtained for nine samples, including three samples of *P. tennesseensis*. Most of the sequences obtained for newly sequenced samples matched or were at most 1 bp different from Genbank listings for the same species. An exception was *P. amplifolius*, for which the ITS sequence was superficially identical to the Genbank sample for *P. pulcher*, and seven positions different from the Genbank record of *P. amplifolius*. On close inspection there was evidence from both length and positional polymorphisms in the sequencing pherograms that the specimen chosen for analysis was of hybrid origin, from *P. amplifolius*/*P. pulcher*. Note that there is also a sequence labelled as *P. amplifolius* in GenBank (EF526338) that is 100% identical to *Persicaria hydropiperoides*, a likely error. Newly obtained ITS sequence for one sample of *P. tennesseensis* differed by seven bp and three gap positions from all other sequences available in Genbank. The other two samples labeled as *P. tennesseensis*, however, gave sequence with multiple polymorphisms, both length and positional. The pattern of one of these, when compared to the sample with clean sequence, suggested that it was a hybrid of *P. tennesseensis* × *P. epihydrus*, and of the other a hybrid with *P. diversifolius*. Thus, the ITS region might not allow unambiguous identification for all samples of *Potamogeton*, but could be useful in detection of hybrids. It could also be used to track hybridization, which could be a potential threat to species integrity.

Pycnanthemum (Lamiaceae)

Pycnanthemum includes 19 accepted species native to North America (POWO 2024). The taxonomy and species boundaries of many are still not fully resolved, and the situation is complicated by polyploidy and hybridization (Chambers and Chambers 1971). Three species, *P. beadlei* (S1S2/G2G3), *P. torreyi* (S1/G2), and *P. verticillatum* (S1/G5) are listed as rare in Tennessee. ITS sequencing was done for samples representing all of the Tennessee species except for *P. incanum*, which was available in GenBank. Clean sequence was obtained for only a few species. Most samples exhibited a polyC/polyG stretch in the ITS2 region which proved impossible to resolve easily, and there were

no sequence differences for the approximately 400 bp leading up it. The exceptions included *P. beadlei* and *P. verticillatum*, both of which were 9 bp (>1%) different from the nearest sample in Genbank. For *P. torreyi*, which was 15 bp (>2%) different from the nearest sequence in Genbank, the morphology of putative available specimens differed significantly from the type of the species, raising doubt about whether this species actually occurs in Tennessee. Thus, the ITS region is not a useful molecular barcode in general for identifying unknown samples of *Pycnanthemum*, although detailed analysis of it could provide clues to the origins of polyploids in the genus as well as resolving taxonomic issues.

Scutellaria (Lamiaceae)

Scutellaria is a cosmopolitan genus of 476 accepted species (POWO 2024) represented by 16 taxa in Tennessee. Currently there are two species listed as rare, *S. arguta*, which is placed by Weakley et al. (2024) as a variety of *S. ovata*, and the Federally Listed *S. montana* (G4/S4). Material of *S. arguta* (*S. ovata* var. *bracteata*) was not available for analysis, but ITS sequences were obtained for all of the other species; most were not in Genbank prior to this study. Each of the Tennessee species had distinctive ITS sequence differing at 2–55 bp, (Table 4), but *S. montana*, *S. incana* var. *punctata*, *S. saxatilis*, and *S. leonardii* all had one or more length polymorphisms that would complicate obtaining sequence using a single Sanger sequence. In the case of *S. montana*, there were two length polymorphisms as well as two polymorphic bp positions, which would be consistent with an allopolyploid origin for this species. There were eight polymorphic bp positions in *S. leonardii*, suggesting a hybrid origin for this taxon, which has also been classified as *S. parvula* var. *missouriensis*; the distinctive ITS sequence compared to *S. parvula* supports its recognition as a distinct species. Thus, the ITS region would be informative for study of species relationships of *Scutellaria* in Tennessee, but the polymorphisms suggest care will be required for its use as a barcoding marker.

Sedum, ***Diamorpha*** (Crassulaceae)

Sedum is a cosmopolitan genus of 482 accepted species (POWO 2024), although the generic boundaries are still not certain. There are two introduced and three native species of the genus in Tennessee, including the state-listed *S. nevii* (S1/G3). *Diamorpha* is a monotypic genus that is somewhat questionably segregated from *Sedum* (Weakley et al. 2024), and *D. smallii* (=*Sedum smallii*) is state-listed (S1S2/G4). The ITS sequences for both *Diamorpha* and *Sedum nevii* were >10% different (63–81 bp) than the highest matches in Genbank as well as from each other, and from newly obtained sequences for the other two native Tennessee species of *Sedum*, *S. pulchellum* and *S. ternatum*. The introduced species (*S. acre*, *S. sarmentosum*) were >25% different from any native species for the ITS region. Thus, the ITS region should be a useful molecular barcode for *Sedum* in Tennessee.

Stachys (Lamiaceae)

Stachys includes 372 accepted species (POWO, 2024) with a widespread distribution. Although the genus has been the subject of multiple studies, the species of the southeastern United States remain understudied, and new species are still being described (Floden 2016; Keener and Davenport 2016). Seven species occur in Tennessee, including three that are considered rare: *S. appalachiana* (S1/G1G2), *S. clingmanii* (S1S2/G2), and *S. glandulosissima* (S1/G1). ITS records were available in GenBank for only *S. latidens* and the non-native *S. floridana*. Sequences were attempted for all of the native Tennessee species, but most samples had length polymorphisms that prevented unambiguous reads. A notable exception was the recently described *S. glandulosissima* and it was 16 bp (>2%) different from any ITS sequence in Genbank. A sample of *S. latidens* also gave clean sequence, but it was about 8% different from a sequence in Genbank for this species, raising questions of identification. As with *Pycnanthemum*, the ITS region appears to be of limited use for molecular barcoding of *Stachys* in Tennessee.

Table 4. Pairwise differences between taxa of *Scutellaria* for nrDNA ITS sequences. The main diagonal shows pairwise differences among samples within the species (-, only single sample available). Note that sequences for *S. elliptica* var. *elliptica* and *S. elliptica* var. *hirsuta* were identical, and these are not shown separately. Taxon abbreviations: ell, *elliptica* var. *elliptica*; i/i, *incana* var. *incana*; pun, *incana* var. *punctata*; int, *integrifolia*; lat, *lateriflora*; leo, *leonardii*; mon, *montana*; ner, *nervosa*; ova, *ovata*; aus, *parvula* var. *australis*; par, *parvula* var. *parvula*; pse, *pseudoserrata*; sax, *saxatilis*; ser, *serrata*.

	ell	i/i	i/p	int	lat	leo	mon	ner	ova	p/a	p/p	pse	sax	ser
ell	*1*													
i/i	2	–												
i/p	11	9	–											
int	5	3	11	*0*										
lat	41	43	51	46	*0-1*									
leo	45	47	55	50	39	–								
mon	6	4	11	7	45	49	–							
ner	2	2	11	5	43	47	6	–						
ova	41	43	51	46	42	28	45	43	–					
p/a	43	45	53	48	37	12	47	45	25	–				
p/p	42	44	52	47	36	11	46	44	24	1	–			
pse	5	3	12	6	45	49	7	4	45	47	46	–		
sax	41	43	49	45	24	39	45	43	40	39	38	45	–	
ser	10	8	17	11	44	49	12	10	45	47	46	11	46	–

Stenanthium (Melanthiaceae)

The classification of *Stenanthium* has been adjusted multiple times, and as currently circumscribed it includes six species (POWO 2024), four of which occur in Tennessee, including two which are state-listed as rare: *S. diffusum* (S1/G1) and *S. tennesseense* (S2/G2). ITS sequences were obtained from a sample of each, and they are all nearly identical, with a single bp difference recorded in *S. tennesseense*. Thus, the ITS region does not provide a suitable marker for confirming species-level identification of Tennessee species of this genus.

Thermopsis (Fabaceae)

Thermopsis is a genus of 29 accepted species found in temperate North America and east Asia (POWO 2024). One of the three species that occur in Tennessee, *T. fraxinifolia*, is rare in the state and throughout its range (S3/G3). Although there are multiple ITS sequences in GenBank for *Thermopsis*, only one of the Tennessee species, *T. villosa*, is represented. Sequences were obtained from a sample of each species, and those for *T. fraxinifolia* and *T. mollis* were identical to one another, and 2 bp different from that of *T. villosa*. Thus, *Thermopsis* is another widespread genus which is poorly sampled in the southeast and provides another case where the ITS molecular barcode will not uniquely identify a rare species.

General Discussion

A primary goal of this study was to obtain sequence data for the widely used molecular barcode marker ITS for some of the rarest plant species in Tennessee, and this was accomplished for 33 state-listed species (Table 1), and data were also obtained for 38 species or varieties that were congeneric with rare ones. For an additional four rare species, it was revealed that there were length polymorphisms that would prevent accurate sequence determination from direct sequence. This has expanded significantly the data available in Genbank for flowering plant species in Tennessee, and helped to highlight some taxonomic problems that need resolution.

A secondary outcome of the study was to assess whether identification of an unknown sample would be possible using the ITS region as a molecular barcode, particularly for sterile samples collected out of season that might be state-listed or rare species. Overall, the ITS region produced

mixed results as a molecular barcode for identification of the rare Tennessee plant species that were analyzed. Although a metric such as the "barcoding gap", which shows whether interspecific variance exceeds intraspecies variance, has been suggested as a way to quantify identification success (Meyer and Paulay 2005), this appears to be of limited application for the current study. Unlike a case cited for grasses in a local region of Tennessee (Drumwright et al. 2011), we envision that in most cases it would be possible to determine the plant genus from morphology, and it would be a question of distinguishing the rare species from other more common ones. For nine genera (*Ammoselinum*, *Borodinia*, *Astragalus*, *Diamorpha*, *Lobelia*, *Paxistima*, *Platanthera*, *Scutellaria*, and *Sedum*) the rare species that were the target of the study exhibited at least five bp differences from the next closest species in the genus, so that molecular barcoding with ITS would likely lead to a correct identification. For nine other genera, ITS would not likely be successful in identifying an unknown sample. For four of these genera (*Desmodium*, *Paysonia*, *Stenanthium*, and *Thermopsis*), the sequence of the rare species was not different from at least one other species of the genus. For another five genera (*Eupatorium*, *Galium*, *Phemeranthus*, *Pycnanthemum*, and *Stachys*), the presence of length polymorphisms rendered direct sequencing results unusable. For two genera (*Geum*, *Potamogeton*), identification of one rare species would be possible, but one or more other rare species were not distinct.

Besides adding to the sequence database, the results of the current study have helped clarify issues of classification at various levels. Most strikingly, in Caryophyllaceae, the generic identity of multiple rare species was changed (Schilling et al. 2022), including *Geocarpon cumberlandensis* (S2/G3; *Minuartia cumberlandensis*), which was recently removed from the Federally Endangered Species List; *Sabulina paludicola* (S1/G1; *Minuartia godfreyanum*); *Geocarpum groenlandicum* (S1/G5; *Minuartia groenlandica*), *Sabulina muscorum* (SU/GNU, *Minuartia muscorum*, the presence of which in Tennessee is still unconfirmed); and *Sabulina fontinalis* (*Stellaria fontinalis*; S3/G3). The data suggested that a change in generic level classification is also needed for *Ammoselinum popei*. Confirmation of the distinctiveness as species previously considered to be varieties was provided for *Lobelia gattingeri* (*L. appendiculata* var. *gattingeri*) and *Scutellaria leonardii* (*S. parvula* var. *missouriensis*). In a similar manner, molecular barcoding was used to help justify retention of *Eriogonum harperi* (S1/G4) as distinct from *E. longifolium* (Floden 2022).

The case of *Platanthera* showed the distinctive potential for the molecular barcoding approach. With the exception of *P. ciliaris* and *P. cristata*, which appear to be actively hybridizing (Evans et al. 2023), the other species that occur in Tennessee had distinctive ITS sequences. Thus, it is possible to identify vegetative material to species. This received a practical application in this study in connection to the Federally Endangered *P. integrilabia*. Multiple samples taken from basal leaf rosettes of vegetative plants that were clearly *Platanthera*, and in locations where this species might occur, were analyzed using the molecular barcode approach. Of 13 unknown samples, five were demonstrated to be *P. integrilabia*, whereas the remaining were shown to be either *P. ciliaris* (two) or *P. clavellata* (six). This allows protection of sites that have been identified even before the plants have reach flowering maturity. This has a direct impact on site restoration and species recovery efforts. It allows for the prioritization of limited resources to restore habitat at sites which have undergone succession to the point of nearing extinction. Some sites that were artificially maintained in an open condition which benefited *P. integrilabia* are no longer maintained and woody plants are dominant. One such site was chosen as a donor site for material to be used in establishing populations on public lands where the plants could be managed and monitored thus preserving that genetic lineage. When the plants were cleaned from the soil it became apparent that there were more plants than were visible above ground. For every plant with an associated leaf above ground there were approximately 10 more plants represented by only the tuberous roots (Matt Richards formerly of Atlanta Botanical Gardens, pers. comm.) The endophytic fungi associated with this species allow it to persist without photosynthesizing for a period of time. When conditions degrade, as woody plants shade the habitat, vegetative plants dominate on the surface along with some subsurface plants. The donor site still had flowering plants but at many sites flowering

has not been observed or rarely occurs. The ability to identify *P. integrilabia* at these at-risk sites allows state and federal agencies to make informed decisions about restoration and recovery efforts. Without the ability to identify vegetative plants no resources would be directed towards sites with only vegetative *Platanthera* plants. The same approach can be extended to other rare species of *Platanthera*.

Another genus for which the results showed the value of molecular barcoding using ITS sequencing was *Lobelia*. Each of the seven species that was sampled had a distinctive ITS sequence. Beyond simple identification of unknown samples, the ITS marker appears to have potential to shed further light on the systematics and evolution of the genus in the southeastern United States, where it has undergone a distinct radiation to produce sect. *Lobelia* (Lammers 2011). It also provided evidence that the calcareous glade endemic *L. gattingeri* should be recognized as a species distinct from the more western and southern *L. appendiculata*. *Lobelia gattingeri* is restricted to limestone glades and found in Tennessee in only eight counties; its conservation status should be reassessed.

Aquatic plants often present taxonomic challenges, and the results for *Potamogeton* suggested that further investigation of the rare Tennessee species is warranted. Previous studies have documented the extensive occurrence of hybridization in the genus (Kaplan and Fehrer 2013), and a more recent study used molecular markers including the ITS region to document an unexpected parentage for the rare *P. floridanus* (Kaplan et al. 2018). The presence of numerous polymorphisms in samples of the genus from Tennessee suggested that hybridization occurs in two rare species, *P. amplifolius* and *P. tennesseensis*. In the case of *P. tennesseensis*, which is rare both within the state and globally (S2/G2G3), the contrast between clean, distinctive sequence from one sample and highly polymorphic sequence from two others raises the possibility that hybridization may be a threat. Management plans for *P. tennesseensis* should be made to ensure that non-hybrid populations of the species continue to exist.

Results of this study highlight the need for further study of the flora of Tennessee, and of the southeastern United States in general. As noted in Schilling et al. (2022), new plant species continue to be named from this area at a rate of over 8 per year, and numerous distinctions between species remain to be clarified. Use of DNA sequence results in conjunction with careful assessment of morphology can help to bring clarity, and the data reported in the current study adds to the overall database of information available for such studies.

ACKNOWLEDGMENTS

The authors thank J. May of the University of Tennessee Molecular Biology Resource Facility for help with generating DNA sequence data, and staff at the University of Tennessee herbarium (TENN) for access to specimens. The authors also thank editors and reviewers at Castanea for helpful feedback. Funding from the Hesler Fund of the University of Tennessee Herbarium and from the Tennessee Department of Environment and Conservation is gratefully acknowledged.

LITERATURE CITED

Barneby, R.C. and E.L. Bridges. 1987. A new species of *Astragalus* (Fabaceae) from Tennessee's Central Basin. Brittonia 39:358–363.

Bateman, R.M., K.E. James, Y.B. Luo, R.K. Lauri, T. Fulcher, P.J. Cribb, and M.W. Chase. 2009. Molecular phylogenetics and morphological reappraisal of the *Platanthera* clade (Orchidaceae: Orchidinae) prompts expansion of the generic limits of *Galearis* and *Platanthera*. Ann. Bot. 104:431–445.

Chambers, H.L. and K.L. Chambers. 1971. Artificial and Natural Hybrids in *Pycnanthemum* (Labiatae). Brittonia. 23:71–88.

Crabtree, T. 2021. Tennessee Heritage Program Rare Plant List. Tennessee Department of Environment and Conservation, Division of Natural Areas, Nashville, Tennessee. (**https://www.tn.gov/content/dam/tn/environment/natural-areas/documents/Rare-Plant-list.pdf**, 23 July 2024).

Drumwright, A.M., B.W. Allen, K.A. Huff, P.A. Ritchey, and A.B. Cahoon. 2011. Survey and DNA barcoding of Poaceae in flat rock cedar glades and Barrens State Natural Area, Murfreesboro, Tennessee. Castanea 76:300–310.

Dugan, L.E., M.F. Wojciechowski, and L.R. Landrum. 2007. A large scale plant survey: Efficient vouchering with identification through morphology and DNA analysis. Taxon 56:1238–1244.

Evans, S.A., D.F. Whigham, I. Hartvig, and M.K. McCormick. 2023. Hybridization in the fringed orchids: An analysis of species boundaries in the face of gene flow. Diversity 15:384. (doi: 10.3390/d15030384).

Floden, A. 2016. A new endemic hedgenettle (*Stachys*: Lamiaceae) from Tennessee. Phytoneuron. 2016–53:1–6.

Floden, A. 2022. Recognition of *Eriogonum harperi* (Polygonaceae) based on molecular barcoding, achene morphology, and life history. Phytoneuron 2022-64:1–7.

Gaier A.G. and J. Resasco. 2023. Does adding community science observations to museum records improve distribution modeling of a rare endemic plant? Ecosphere. 2023 Mar 23; 14e4419 (doi: 10.1002/ecs2.4419).

Gostel, M.R. and W.J. Kress. 2022. The expanding role of DNA barcodes: Indispensable tools for ecology. evolution, and conservation. Diversity, 14:213 (doi: 10.3390/d14030213).

Hebert, P.D.N., E.H. Penton, J.M. Burns, D.H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgenator*. Proc. Nat. Acad. Sci. USA 101:14812–14817.

Hollingsworth P.M., S.W. Graham, and D.P. Little. 2011. Choosing and Using a Plant DNA Barcode. PLoS ONE 6(5): e19254 (doi:10.1371/journal.pone.0019254).

Jones, L., A.D. Twyford, C.R. Ford, T.C.G. Rich, H. Davies, L.L. Forrest, M.L. Hart, H. McHaffie, M.R. Brown, and P.M. Hollingsworth. 2021. Barcode UK: A complete DNA barcoding resource for the flowering plants and conifers of the United Kingdom. Mol. Ecol. Resour. 21:2050–2062.

Kaplan, Z. and J. Fehrer. 2013. Molecular identification of hybrids from a former hot spot of *Potamogeton* hybrid diversity. Aquatic Bot. 105:34–40.

Kaplan, Z., J. Fehrer, V. Bambasová, and C.B. Hellquist. 2018. The endangered Florida pondweed (*Potamogeton floridanus*) is a hybrid. PloS One 13:e0195241 (doi: 10.1371/journal.pone.0195241).

Katoh, K. and D.M. Standley. 2013. MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30: 772–780.

Keener B. R. and L.J. Davenport. 2016. Two new hedge-nettles (*Stachys*: Lamiaceae) from the Blue Ridge outliers of East Central Alabama, U.S.A. J. Bot. Res. Inst. Texas 10:315–23.

Kress, W.J., K.J. Wurdack, E.A. Zimmer, L.A. Weigt, and D.H. Janzen. 2005. Use of DNA barcodes to identify flowering plants. Proc. Nat. Acad. Sci. USA 102:8369– 8374.

Lammers, T.G. 2011. Revision of the infrageneric classification of *Lobelia* L. (Campanulaceae: Lobelioideae). Ann. Missouri Bot. Gard. 98:37–62.

Mazie, A.R. and D.A. Baum. 2016. Clade-specific positive selection on a developmental gene: branchless TRICHOME and the evolution of stellate trichomes in *Physaria* (Brassicaceae). Mol. Phylogenet. Evol. 100:31–40.

Meyer C.P. and G. Paulay. 2005. DNA barcoding: error rates based on comprehensive sampling. Plos Biology, 3:2229–2238.

Nazar, N., C. Howard, A. Slater, and T. Sgamma, T. 2023. Challenges in medicinal and aromatic plants DNA barcoding–lessons from the Lamiaceae. Plants 11:137 (doi: 10.3390/plants11010137).

Petrone, B.L, A. Aqeel, S. Jiang, H.K. Durand, E.P. Dallow, J.R. McCann, H.K. Dressman, Z. Hu, C.B. Tenekjian, W.S. Yancy Jr., P.-H. Lin, J.J. Scialla, P.C. Seed, J.F. Rawls, S.C. Armstrong, J. Stevens, and L.A. David. 2023. Diversity of plant DNA in stool is linked to dietary quality, age, and household income. Proc. Nat. Acad. Sci. USA 120 e2304441120 (doi: 10.1073/pnas.2304441120).

POWO. 2024. Plants of the World Online. Royal Botanic Gardens, Kew, United Kingdom. (**http://www.plantsoftheworldonline.org/**, 17 May 2024).

Schilling, E.E. 2011. Hybrid genera in Liatrinae (Asteraceae: Eupatorieae). Mol. Phylogenet. Evol. 59:158–167.

Schilling, E.E. and K.C. Grubbs. 2016. Systematics of the *Eupatorium mohrii* complex (Asteraceae). Syst. Bot. 41:787–795.

Schilling E.E., A. Floden, J. Lampley, T.S. Patrick, and S.B. Farmer. 2019. A new species of *Trillium* (Melanthiaceae) from central Georgia and its phylogenetic position in subgenus *Sessilium*. Syst. Bot. 44:107–14.

Schilling, E.E., A.J. Floden, A.S. Weakley, C. Winder, and R.L. Small. 2022. Molecular barcoding reveals unexpected diversity in eastern North American stitchworts (Caryophyllaceae). Bot. J. Linn. Soc. 200:75–84.

Siripun K.C. and E.E. Schilling. 2006. Molecular confirmation of the hybrid origin of *Eupatorium godfreyanum* (Asteraceae). Amer. J. Bot. 93:319–325.

Spaulding, D.D. and T.W. Barger 2016. Keys, distribution, and taxonomic notes for the Lobelias (*Lobelia*, Campanulaceae) of Alabama and adjacent states. Phytoneuron 2016-76:1–60.

The Tennessee Flora Committee. 2015. Guide to the vascular flora of Tennessee. The University of Tennessee Press, Knoxville, Tennessee.

Thornhill A.H., B.G. Baldwin, W.A. Freyman, S. Nosratinia, M.M. Kling, N. Morueta-Holme, T.P. Madsen, D.D. Ackerly, and B.D. Mishler. 2017. Spatial phylogenetics of the native California flora. BMC Biology 15:96 (doi: 10.1186/s12915-017-0435-x).

Wang, X.C., C. Liu, L. Huang, J. Bengtsson-Palme, H. Chen, J.H. Zhang, D. Cai, and J. N-Q. Li. 2014. ITS1: A DNA barcode better than ITS2 in eukaryotes? Mol. Ecol. Resour. 15:573–586 (doi: 10.1111/1755-0998.12325).

Weakley, A.S., and Southeastern Flora Team. 2024. Flora of the southeastern United States Web App. University of North Carolina Herbarium, North Carolina Botanical Garden, Chapel Hill.

White, T.J., T.D. Bruns, S.B. Lee, and J.W. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. p. 315–322 *In*: Innis, M.A., D.H. Gelfand, J.I. Sninsky, and T.J. White (eds.). PCR protocols: A guide to methods and applications. Academic Press, San Diego, California.

Yao, H., J. Song, C. Liu, K. Luo, J. Han, Y. Li, X. Pang, H. Xu, Y. Zhu, P. Xiao, and S. Chen. 2010. Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS ONE, 5:e13102 (doi: 10.1371/journal.pone.0013102).