Webbia: Journal of Plant Taxonomy and Geography, 2014
Vol. 69, No. 1, 1–12, http://dx.doi.org/10.1080/00837792.2014.927555

Taylor & Francis
Taylor & Francis Group

INVITED PAPER

A laboratory guide for generating DNA barcodes in grasses: a case study of *Leptochloa* s.l. (Poaceae: Chloridoideae)

Paul M. Peterson[a]*, Konstantin Romaschenko[a,b] and Robert J. Soreng[a]

[a]*Smithsonian Institution, Department of Botany, National Museum of Natural History, Washington, USA;* [b]*M.G. Kholodny Institute of Botany, National Academy of Sciences, Kiev, Ukraine*

(*Received 29 March 2014; final version received 20 May 2014*)

There is no easy way to identify to species, a small, vegetative leaf or culm sample of a grass and there are more than 12,000 species in this large, important family. The long-range aim of our study is to produce a standard DNA barcode library available to the public for all grasses (±1960 species) in North America (includes all Canada, Mexico and USA) that will facilitate the easy identification of these morphologically cryptic species. We provide a detailed protocol of the laboratory procedures for DNA extraction in grasses and the DNA-specific primers used for the polymerase chain reaction (PCR) enabling the laboratory work to be performed in any well-supplied molecular laboratory. In this paper we present a test of four barcodes [*matK*, *rbcL*, *rpl32-trnL* and internal transcribed spacer (ITS)] to discriminate among 50 taxa of grasses (55 samples), predominately in the subfamily Chloridoideae, and we used a tree-based method to identify relationships among species of *Leptochloa sensu lato*. The sequence divergence or discriminatory power based on uncorrected *p*-value, among the four DNA sequence markers was greatest in ITS (96%), followed by *rpl32-trnL* (25.6%), *matK* (3.0%) and *rbcL* (0.0%). *matK* was twice as effective in discriminating among the species compared with *rbcL*; *rpl32-trnL* was nearly 3.4 times better than *rbcL*; and nuclear rDNA ITS was 14 times better than *rbcL*. There are significant threshold levels of 0.0682 for ITS and 0.0732 for ITS + *rpl32-trnL* between intrageneric and intergeneric sequence divergences within the 16 species of *Dinebra* and between *Dinebra* and *Diplachne*, *Disakisperma* and *Leptochloa sensu stricto*. In our tree-based analyses of *Leptochloa* s.l. the following number of nodes with strong support (PP = 0.95−1.00) were successfully recovered (in descending order): combined ITS + *rpl32-trnL*, 43; ITS, 34; *rp32-trnL*, 27; *matK*, 19; and *rbcL*, 3.

Keywords: DNA sequences; identification; ITS; *matK*; *rbcL*; *rpl32-trnL*; taxonomy

Introduction

Despite efforts by many scientists, the standardization of DNA barcodes for all land plants has not yet been achieved. The barcode pair of *rbcL* and *matK* sequences is accepted as the core marker set, but supplementary markers, fine-tuned for each group of plants, are highly recommended (CBOL 2009; Hollingworth et al. 2011). Remarkably, members of the grass family (Poaceae) have yet to be thoroughly investigated to assess the applicability of the accepted barcodes, or for supplementary DNA barcode markers (Ward et al. 2009; Drumwright et al. 2011; Cai et al. 2012; Saarela et al. 2013; Zhang et al. 2013; Liu et al. 2014).

In light of their importance as grains for feeding the planet, forage for grazing animals, and as noxious weeds, and given their ubiquity in open spaces such as private lawns and parks, a complete barcode library for the identification of these cryptic plant species is urgently needed. Recognizing that these two loci discriminate only 70% of the plant species led us to test the use of two additional regions, *rpl32-trnL* and nuclear ribosomal internal transcribed spacer (ITS 1 and 2). Recent studies have demonstrated the suitability of both ITS and *rpl32-trnL* as additional barcode regions within the grasses (Romaschenko et al. 2008, 2010, 2011, 2012, 2014; Chen et al. 2010; Peterson et al. 2010a, 2010b, 2011, 2012, 2014; Cai et al. 2012).The primary goal of our long-range study is to produce a DNA barcode library for all grasses (±1960 species) in North America (includes all Canada, Mexico, and USA) with the addition of the 22 species that are Federally listed as Noxious weeds in the USA (Espejo Serna et al. 2000; Barkworth et al. 2003, 2007; Soreng et al. 2014).

In this paper we present a test of four barcode marker regions (*matK*, *rbcL*, *rpl32-trnL* and ITS) to discriminate among 50 taxa of grasses, predominately in the subfamily Chloridoideae Kunth ex Beilschm. Within the data set, in addition to other grasses, we include 16 closely related species of *Dinebra* Jacq., one species of *Diplachne* P. Beauv., three species of *Disakisperma* Steud., two species of *Leptochloa* P. Beauv., and one species of *Trigonochloa* P.M. Peterson & N. Snow, all formally treated in *Leptochloa* s.l. (Peterson et al. 2012). Accurate species identification is of prime importance and, in addition to this use, we evaluate the use of barcode data to refine the classification of the grasses by constructing well-resolved phylogenies. As grass material offers special challenges for DNA study because it

*Corresponding author. Email: peterson@si.edu

contains large quantities of silica, we provide a detailed protocol of the laboratory procedures for DNA extraction and the DNA-specific primers used for the polymerase chain reaction (PCR), enabling the laboratory work to be performed in any well-supplied molecular laboratory.

Materials and Methods

Specimen collection and preservation

Dried leaf blades of recently collected (within the last 50 years) specimens and/or specimens that were dried rapidly (but not too hot, e.g. 30−40°C) and still green are optimal. If collecting fresh material for extraction it is best to place 5–10 leaf blades (depending on their size, 2−4 cm^2 of material) in silica gel (c.6−15 g) in an air-tight container (plastic ziplock bag) at room temperature for immediate desiccation. Drying of the plant samples usually occurs within 24 h. To store the dry material, place sample in a small paper envelope and place the envelope in a dry place. For long-term storage remove the silica desiccant.

DNA extraction

The DNA extraction from plant tissue follows the BioSprint 96 DNA Plant Kit (Qiagen, Valencia, CA; cat. no. 941557 or 941558) manufacturer's protocol. The Kit includes buffers RNeasy lysis tissue buffer (RLT), RNeasy pellet wash buffer (RPW), MagAttract Suspension for magnetic separation of DNA, a set of rod covers, and 96-well blocks. Before using the RPW buffer for the first time, 125 ml of isopropanol and one vial RNase a (1 × 200 µl) are added to each bottle of the buffer. In addition, distilled water or low-salt buffer and 96–100% ethanol are required.

Disruption of plant tissue

We used TissueLyser (Qiagen; cat. no. 85210), which includes the following procedures: homogenization, lysis and precipitation of the plant tissue.

(1) Place approximately 30–50 mg of each plant sample into a well of a polypropylene 96-well matrix screen-mate plate (Matrix Technologies cat. No. 4221), along with a 5-mm stainless steel bead (Qiagen; cat. No. 69,989) for disruption, or three tungsten 3-mm carbide beads (Qiagen; cat. No. 69,997). For best disruption of the tough grass tissue, use about 3 mg (a sealing cap) of fine (0.1 mm) crystal powder.
(2) Seal all the wells in the plate with cap-stripes and put entire plate in liquid nitrogen, then disrupt the plant material in a TissueLyser at 25–30 Hz for 1 min. It is best to use three cycles of flash-freezing and disruption or more, until all tissue is reduced to fine powder. The use of tungsten 3-mm carbide beads is particularly effective. In most cases they can be used without liquid nitrogen or even crystal powder.
(3) Remove the cap-stripes and pipette 300 µl of buffer RLT into each well. Recap the wells. Shake the entire plate more than 20 times and vortex it for 20 s. Alternatively, if working with old material it is best to shake the entire plate for several minutes and then store at room temperature for up to 6 h (if stored more than 6 h RLT buffer could corrupt the Tungsten beads).
(4) Centrifuge the entire plate at 3000–5000 ***g*** for 5 min.

DNA isolation

We use Qiagen BioSprint 96 for DNA isolation, which includes the following procedures: (1) Mixing cleared lysate to cold isopropanol and MagAttract suspension; (2) binding DNA to magnetic particles of MagAttract suspension; (3) primary magnetic separation of DNA; (4) washing with ethanol; (5) eventual magnetic separation of DNA; and (6) DNA elusion.

Before starting the procedure several blocks (S-Blocks according to manufacturers' protocol) and microplates should be prepared: a 96-well microplate for rod-cover; another microplate with distilled (RNAse-free) water (200 µl); S-block with distilled water containing Tween-20 (50 µl Tween added to 250 µl H_2O) [500 µl]; two S-block with ethanol (96–100%) [500 µl], and one S-block with RPW buffer (200 µl).

(1) After centrifugation, transfer 200 µl cleared plant lysate into each well of an S-block.
(2) Add 200 µl cold isopropanol to each well (sometimes you can see the "web" of precipitated DNA at the bottom of the well when cold isopropanol is added).
(3) Add 20 µl of MagAttract suspension G to each well, seal the wells with cap-stripes (MagAttract suspension must be vortexed first for 1–3 min).
(4) Put all the previously prepared microplates and S-blocks into designated slots. The S-block with lysate goes to slot 1, and the microplate with water or low-salt buffer for final elusion of DNA goes to slot 6.
(5) Select the protocol "BS96 DNA Plant" and start the BioSprint 96. DNA elutes into the microplate with water or low-salt buffer, which yields 200 µl of the sample.
(6) Store DNA samples at –5°C until use, or at –20 to –80°C for long-term preservation.

Note: If processing fewer than 96 samples it is important that the number of wells be filled with buffers, ethanol and water. These should match the number and the location of the samples to be processed (e.g. A1–H1 to A6–H6 if processing 48 samples; A1–H1 to A3–H3 if processing 24 samples, etc.).

DNA concentration

In order to ensure the success of the PCR step, the optimal DNA concentration is estimated to be 50–150 ng/µl. The DNA concentration can be verified using a NanoDrop spectrophotometer. One of the indicators of the quality of the DNA samples is the 260/280 ratio, which should be reasonably close to 1.8 and the 260/230 ratio, which should be reasonably close to 2.0.

To verify that DNA is not degraded, standard electrophoresis can be run using a 1% agarose gel. The presence of the relatively large bands (>10 kb) would indicate the integrity of the DNA strands. Smearing and traces of RNA (small bands) normally indicate a poor quality DNA sample. If the DNA concentration is high, DNA samples should be diluted to 50–100 ng/µl. If the DNA concentration is low the elution volume could be reduced using a vacuum-centrifuge. The sample should be verified using a NanoDrop spectrophotometer. The process can be repeated if necessary until the desired DNA concentration is reached.

PCR primers and thermo-cycling parameters for barcode regions

The PCR primers for three plastid regions targeted for our study are based on the analysis of multiple available sequences from GenBank. The universal primers previously used to sequence barcode regions were redesigned to fit the conditions in grasses. The program FastPCR version 4.0.27 was employed to adjust the temperature and quality of the newly designed or modified primers and increase the PCR efficiency. The sequences, melting temperature, quality and references for the primers used are given in Table 1.

The ***matK*** core barcode region comprises 832 bp (excluding primer sequences) of the 1535 bp (54%) maturase K gene based on the *Oryza* plastid genome (Hiratsuka et al. 1989). The forward primer matK-454F is located at position 454 of the *matK* gene sequence (GPWGII 2011). It is slightly modified to meet the PCR conditions. The reverse primer matK-1315R is placed close to matK-5R (Kew; 5'- GTTCTAGCACAAGAAAGTCG-3') but contains fewer ambiguous positions and unmatched substitutions with tested *matK* sequences in grasses, and this helps with the PCR efficiency. Thermal cycling programme is: 95°C for 3 min; 40 cycles of 94°C for 30 s, 49°C for 1 min, 72°C for 1 min; final extension 72°C for 10 min.

The ***rbcL*** core barcode region comprises 824 bp (excluding primer sequences) of the 1434 bp (57%) ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit gene (CBOL 2009). The primers rbcLa_SIrev (5'-GTAAAATCAAGTCCACCRCG-3') and rbcLa_SIfor (5'-ATGTCACCACAAACAGAGACTAAAGC-3') usually provide the highest levels of amplification success among plants (Kress et al. 2009). The retrievable region with these primers consists of 554 bp of the 5'-portion of the gene. However, the duplication and mitochondrial *rbcL* gene transfer is found in grasses (Cummings et al. 2003). A study of Stipeae grasses found multiple double peaks in the *rbcL* trace chromatogram (Syme et al. 2012), and this condition was found in our study of chloridoid grasses. The mitochondrial copies in grasses are pseudogenes containing many indels (Syme et al. 2012). The copies of the mitochondrial pseudogene of *rbcL* available from GenBank were used to develop a specific set of primers to amplify only the chloroplast copies. The new region also encompasses the 5' portion of the *rbcL* gene, it is 270 bp longer than the region used previously (Kress et al. 2009) and it is compatible in size with the *matK* barcode region. Thermal cycling programme is: 95°C for 3 min; five cycles of 94°C for 30 s, 56°C for 1 min, 72°C for 1 min; 30 cycles of 94°C for 30 s, 54°C for 1 min, 72°C for 1 min; final extension 72°C for 10 min.

The ***rpl32-trnL*** barcode region comprises 162 bp (excluding primer sequences) of 180 bp (90%) of ribosomal protein L32 (rpl32) encoding gene small single copy (SSC) region of the plastid genome and 524 bp of rpl32-trnL intergenic spacer of entire length according to the *Oryza* plastid genome. The primers trnL(UAG) as reverse and rpL32-F as forward provide stable amplification of the region with high success in angiosperms (Shaw et al. 2007). The region is easily amplified within the grasses following previous PCR parameters in Peterson et al. (2010a). The thermal cycling programme is: 95°C for 3 min; 35 cycles of 94°C for 40 s, 54°C for 40 s, 72°C for 1.40 min; final extension 72°C for 10 min.

Table 1. Regions studied, sequences, melting temperature (°C), and quality of primers used for polymerase chain reaction and sequencing.

Region	Primers	Sequence (5'–3')	Temp.	Quality	Reference
matK	matK-454F	CATATAGARATACCYTAYCCTATC	51.9	68	GPWG (2001)*
	matK-1315R	GCTAAAGTTCTAGCRCATGAAAG	53.1	88	This study
	matK-1326R	TCTAGCRCATGAAAGYCGAAGT	56		Cuénoud et al. (2002)*
rbcL	rbcL-109F	TGGCAGCATTCCGAGTAASTCCT	60.3	73	This study
	rbcL-926R	CATACGCAATGCTTTAGCTAATACACG	56.4	80	This study
rpl32-trnL	trnL(UAG)	CTGCTTCCTAAGAGCAGCGT	60	120	Shaw et al. (2007)
	rpl32-F	CAGTTCCAAAAAAACGTACTTC	53.7	103	Shaw et al. (2007)
ITS	ITS5a	CCTTATCATTTAGAGGAAGGAG	53.7	82	Stanford et al. (2000)
	ITS4	TCCTCCGCTTATTGATATGC	55	38	White et al. (1990)

*modified.

The **ITS** barcode region comprises on average 669 bp of nuclear ribosomal DNA. It includes complete sequences of internal transcribed spacer 1 (ITS1), 5.8S ribosomal RNA gene, and internal transcribed spacer 2 (ITS2). The suggested pair of primers includes ITS4 as a reverse, and ITS5a as forward. The ITS5a primer has a sequence that disables the amplification of copies not of plant origin. It helps to avoid sequencing unrelated copies of endophytic fungi, which are often present in the tissue of cool-temperate grasses (Crawford et al. 2010; Jani et al. 2010). Thermal cycling programme is: 95°C for 4 min; 35 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 1.20 min; final extension 72°C for 10 min.

PCR mix

Genomic DNA (50–100 ng/reaction) is to be combined with 1× reaction buffer (200 mM Tris–HCl, 500 mM NH_4) (Bioline Biolase, Taunton, Madison, USA) without Mg^{2+}, 2 mM $MgCl_2$, 200 mM dNTP's, 1.5 µl of *Taq* polymerase (Bioline Biolase, Taunton, MA, USA), 40 pmol/l each of forward and reverse primers.

PCR setup

Components:	Volume
Nuclease-free ddH_2O	11.75 µl
Reaction buffer (×10)	2 µl
$MgCl_2$ (50 mM)	1 µl
dNTPs (5 mM)	1 µl
Primer "Forward" (10 mM)	0.5 µl
Primer "Reverse" (10 mM)	0.5 µl
Taq polymerase (5 U/µl)	0.25 µl
DNA template (70 ng)	2 µl
Dimethylsulphoxide	1 µl
Total volume:	20.0 µl

Gel electrophoresis of PCR products

To check the results of the amplification of barcode genomic regions the products of the PCR can be separated in 1% agarose gel using electrophoresis:

(1) Mix 2 µl of DNA loading buffer (5×), 3 µl of H_2O and 2 µl of amplified PCR product.
(2) Load the mix into well of agarose gel containing ethidium bromide made with Tris acetic acid EDTA (TAE) buffer (0.5 µl) with 10 mg/ml ethidium bromide per 100 ml 1% agarose.
(3) Load a ladder marker to identify the bands of the proper size.
(4) Run PCR products at 100–120 V until clear separation of the DNA bands.
(5) Inspect the resulting bands under UV light.

DNA purification and sequencing reaction

Use ExoSAP-IT (USB Corp. Cat No 78201 1 ML), to clean up PCR products less than 100 bp.

(1) Add nuclease-free ddH_2O to the ExoSAP–IT solution in proportion 9 to 1.
(2) Mix 9 µl of delute ExoSAP–IT to 18 µl of PCR product (at general ratio 1:2).
(3) Use the thermal cycler to incubate the mix at 37°C for 30 min and then inactivate ExoSAP-IT by heating to 80°C for 20 min.
(4) Proceed to automated DNA sequencing procedure.

DNA sequencing setup

Components:	Volume
BigDye Terminator sequence mix	0.8 µl
Reaction buffer (5×) (400 mM Tris; 10 mM $MgCl_2$)	2.0 µl
1 mM primer	2.0 µl
Purified DNA	3.0 µl
Nuclease-free ddH_2O	4.2 µl
Total volume	12.0 µl

Thermal cycling program for DNA sequencing is: 94 °C for 20 s; 33 cycles of 94 °C for 20 s, 50 °C for 15 s, 60 °C for 4 min.

DNA sequencing is best performed with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit, v. 3.1 (Perkin-Elmer/Applied Biosystems, Foster City, CA, USA) and electrophoresed and detected on an ABI Prism 3100 automated sequencer (University of Tennessee Molecular Biology Resource Facility, Knoxville, TN, USA).

Analysis of barcode performance

Taxon sampling

The species chosen to test the usefulness of the four barcodes were taken from an existing study of *Leptochloa*, a morphologically heterogeneous grass complex that was recently re-examined and placed in five genera (Peterson et al. 2012; Snow and Peterson 2012; Snow et al. 2013). The list of 50 taxa (55 samples) used in this study is given in Appendix 1. Thirty-one samples of *Dinebra*, *Diplachne*, *Disakisperma*, *Leptochloa* and *Trigonochloa,* all formally treated in *Leptochloa* s.l., were included in our study to test the discriminatory power among closely related species.

Comparison of sequence divergence

DNA barcoding primarily relies on sequence data to identify a specimen (Hebert et al. 2003). Nucleotide similarity is inferred through pairwise comparison of the sequences and is expressed as a genetic distance or the number of substitutions per site. Sequence similarity is derived from uncorrected *p*-values and these are preferred when working with closely related species (Srivathsan and Meier 2011; Collins et al. 2012). Uncorrected *p*-values are a measure of the likelihood that the

association between sequences is due to random chance. In general, *p*-values < 0.05 indicate a statistically significant association, similarity or possible homology. If the *p*-value for an alignment of two sequences is >0.05, we conclude that the alignment score is not statistically significant and the two sequences are not closely related. To test the discriminatory power of the four barcode regions we compared the average mean of *p*-value for each region as well as combined ITS+*rpl32-trnL*, and analysed the rate at which each region discriminates between sequences/taxa using the statistical threshold at $p < 0.05$ (Meier et al. 2006; Hawlitschek et al. 2013). The program PAUP* 4.0b10 was run to produce a matrix before performing the distance analyses (Swofford 2000).

We analysed intrageneric distances within *Dinebra*, including the United States Department of Agriculture Federally listed noxious weed, *Dinebra chinensis* (L.) P.M. Peterson & N. Snow (Federal Noxious Weed List 2012). We assessed the average distance values among the 16 species of *Dinebra* (intrageneric) and compared these with other species in *Leptochloa* s.l (intergeneric). In addition, we provided a comparison of the intergeneric and intrageneric distances for *Dinebra* among the four barcode regions. This allows us to determine if a DNA marker is more variable within (intra) a genus or between (inter) genera.

Tree-based discrimination among Leptochloa s.l.

A DNA Blast search is a common procedure for identifying unknown taxa (Altschul et al. 1997). However, Blast provides no guidance to the researcher for choosing among multiple best matches (Munch et al. 2008a, 2008b). We also use a tree-based method to identify taxa based on their placement within clades. This approach is based on neighbour joining (Hebert et al. 2003; Barrett and Hebert 2005; Munch et al. 2008b; Saarela et al. 2013) or Bayesian inference (Munch et al. 2008a; Hawlitschek et al. 2013). We used a tree-based method to identify relationships among species of *Leptochloa* s.l. using four separate barcode markers (*rbcL*, *matK*, *rpl32-trnL* and ITS) and combined ITS + *rpL32-trnL*. The Bayesian posterior probabilities (PP) were estimated using MrBayes 3.01 (Huelsenbeck and Ronquist 2001). All parameters were left at default settings. Each Bayesian analysis was initiated with random starting trees and was run for one million generations, sampling once per 100 generations. The analysis was continued until the value of standard deviation of split sequences dropped below 0.01 and the potential scale reduction factor was close to or equal to 1.0 (Huelsenbeck and Ronquist 2001). The fraction of the sampled values discarded as burn-in was set at 0.25. Posterior probabilities of 0.95–1.00 were considered strong support. The number of recovered strongly supported clades was counted for each barcoding set.

Results

A total of 100 sequences representing all taxa in our study are newly reported in GenBank and, of these, six are new for *rpl32-trnL* (11%), seven are new for ITS (13%), 33 are new for *matK* (89%) and 54 are new for *rbcL* (100%).

The amount of sequence divergence (average mean uncorrected *p*-value ± standard deviation) in the PAUP matrix for *rbcL* was 0.0115 ± 0.0064, for *matK* was 0.0255 ± 0.0113, for *rpl32-trnL* was 0.0380 ± 0.0195, and for ITS was 0.1636 ± 0.0483, and for ITS + *rpl32-trnL* it was 0.1987 ± 0.0685 (Figure 1). Therefore, *matK* was twice as effective in discriminating among the species compared with *rbcL*; *rpl32-trnL* was nearly 3.4 times better than *rbcL*; and nuclear rDNA ITS was 14 times better than *rbcL*. The ITS region had the highest level of discriminatory power and is the most variable among the barcode regions (Figure 1). The largest amount of sequence divergence among plastid markers is found in the *rpl32-trnL*, which exceeds the fast evolving protein-encoding *matK* region by 1.6 times. The sequence divergence between the nuclear rDNA ITS and *rpl32-trnL* is 4.1-fold, and this is roughly the same value as between *rpl32-trnL* and *rbcL*.

The percent discrimination rate (at $p < 0.05$) further demonstrates complete predominance of the non-coding regions (ITS and *rpl32-trnL*) over the coding regions (*rbcL* and *matK*) (Figure 2). The most disparate results are for *rbcL* with 0.0% of the pairwise alignments having a *p*-value > 0.05 (= low discrimination). Whereas ITS and combined ITS + *rpL32-trnL* have a discrimination rate of 96%. The difference between the discrimination rate of *matK* paired with *rpl32-trnL* is greater than between their average *p*-values, and the discrimination rate for *rpl32-trnL* (25.6%) exceeds that of *matK* (3.0%) by 8.5 times.

Analysis of the sequence divergence (uncorrelated *p*-value) of the four barcode regions among 16 species

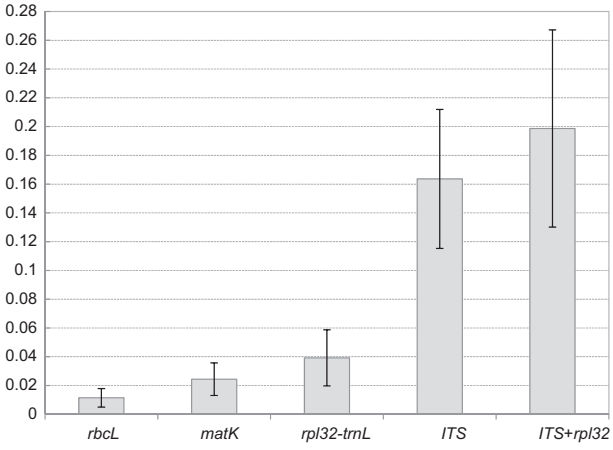

Figure 1. Average mean sequence divergence (uncorrected *p*-value) and standard deviation among DNA markers for 55 taxa of grasses, including 23 species of *Leptochloa* s.l.

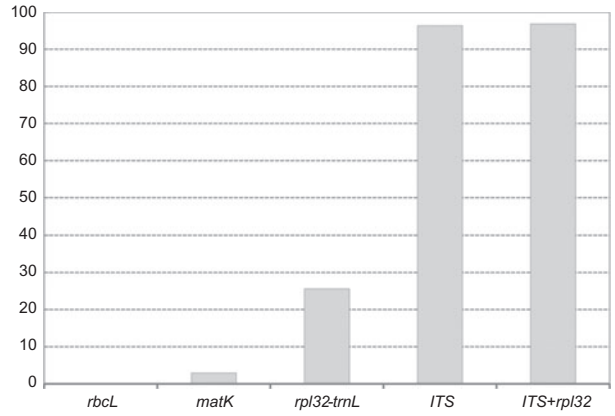

Figure 2. Percent discrimination rate of *rbcL*, *matK*, *rpl32-trnL* and internal transcribed spacer (ITS), and ITS + *rpl32-trnL* barcode regions (statistical threshold $p < 0.05$).

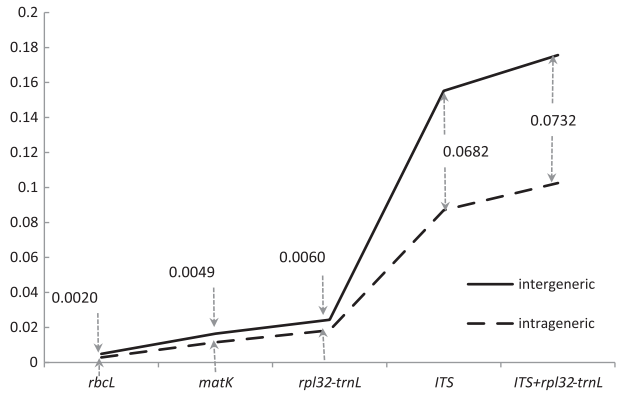

Figure 4. Threshold variation between intrageneric and intergeneric sequence divergence (uncorrected *p*-value) of the four barcode markers among 16 species of *Dinebra* (intrageneric) and comparison of *Dinebra* with *Diplachne*, *Disakisperma* and *Leptochloa* in combination (intergeneric).

of *Dinebra* (intrageneric) and comparison (intergeneric) of *Dinebra* with *Diplachne*, *Disakisperma* and *Leptochloa* is shown in Figure 3. Only ITS shows a significant threshold of *Dinebra* versus *Diplachne*, *Disakisperma* and *Leptochloa*. The lowest values belong to the *rbcL* marker, which shows very little change between intergeneric and intrageneric distances. *MatK* and *rpl32-trnL* markers show some intergeneric variation between *Dinebra* and *Diplachne* (higher threshold), although it is not significant ($p < 0.05$).

The threshold variation between intra- and intergeneric sequence divergence (uncorrelated *p*-value) of the four barcode markers among 16 species of *Dinebra* (intrageneric) and comparison of *Dinebra* with *Diplachne*, *Disakisperma* and *Leptochloa* in combination (intergeneric) is seen in Figure 4. In contrast to the plastid markers, there are significant threshold levels of 0.0682 for ITS and 0.0732 for ITS + *rpl32-trnL* between intra- and

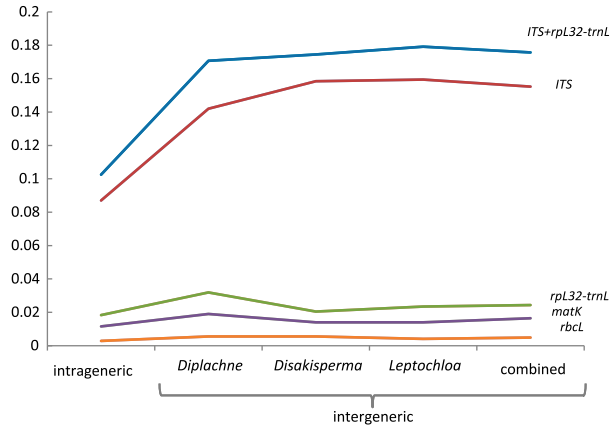

Figure 3. Sequence divergence (uncorrected *p*-value) of the four barcode markers among 16 species of *Dinebra* (intrageneric) and comparison (intergeneric) of *Dinebra* with *Diplachne*, *Disakisperma*, *Leptochloa*, and the three genera in combination.

intergeneric sequence divergences within the 16 species of *Dinebra* and between *Dinebra* and *Diplachne*, *Disakisperma* and *Leptochloa* (in combination). The intrageneric and intergeneric distances rise gradually among plastid markers from *rbcL* (the lowest) to *matK* and *rpl32-trnL* (Figure 4). Inclusion of the ITS data improves these values drastically. ITS also provides significant threshold between intrageneric and intergeneric variations, i.e. 11.3 times that of the best plastid marker, *rpl32-trnL*.

A comparison of five Bayesian phylogenetic trees derived from 50 grass taxa using the evolutionary conservative coding locus *rbcL* (top left), the fast evolving coding region *matK* (top middle), intergenic spacer *rpl32-trnL* (top right), nuclear DNA ITS (bottom left), and combined ITS + *rpl32-trnL* (bottom right), is shown in Figure 5. The *rbcL* barcode does a poor job of resolving relationships within *Leptochloa* s.l. and shows a polytomy of many genera not known to be closely related.

The *matK* barcode provides support for tribe Cynodonteae (PP = 0.95) and subtribe Eleusininae (PP = 1.00), and within the Eleusininae it supports the monophyly of *Diplachne* (PP = 1.00) and *Disakisperma* (PP = 1.00), excluding the unsuccessful sequencing of *Disakisperma obtusiflorum* (Hochst.) P.M. Peterson & N. Snow shown as a dashed box. Additionally, three separate but strongly supported clades of *Dinebra* (PP = 0.98–1.00) were resolved in polytomy. *Dinebra chinensis*, the Federally listed noxious weed, and other representatives of *Leptochloa* s.l. were unsuccessfully sequenced and they are shown as dashed boxes (Figure 5).

The *rpl32-trnL* barcode displays a high level of discriminatory power, depicting many congruent phylogenetic alignments that correspond to our previous analyses (Peterson et al. 2012). Strong support is given for Cynodonteae (PP = 1.00), the Eleusininae (PP = 1.00), *Diplachne* (PP = 1.00), *Disakisperma* (PP = 1.00),

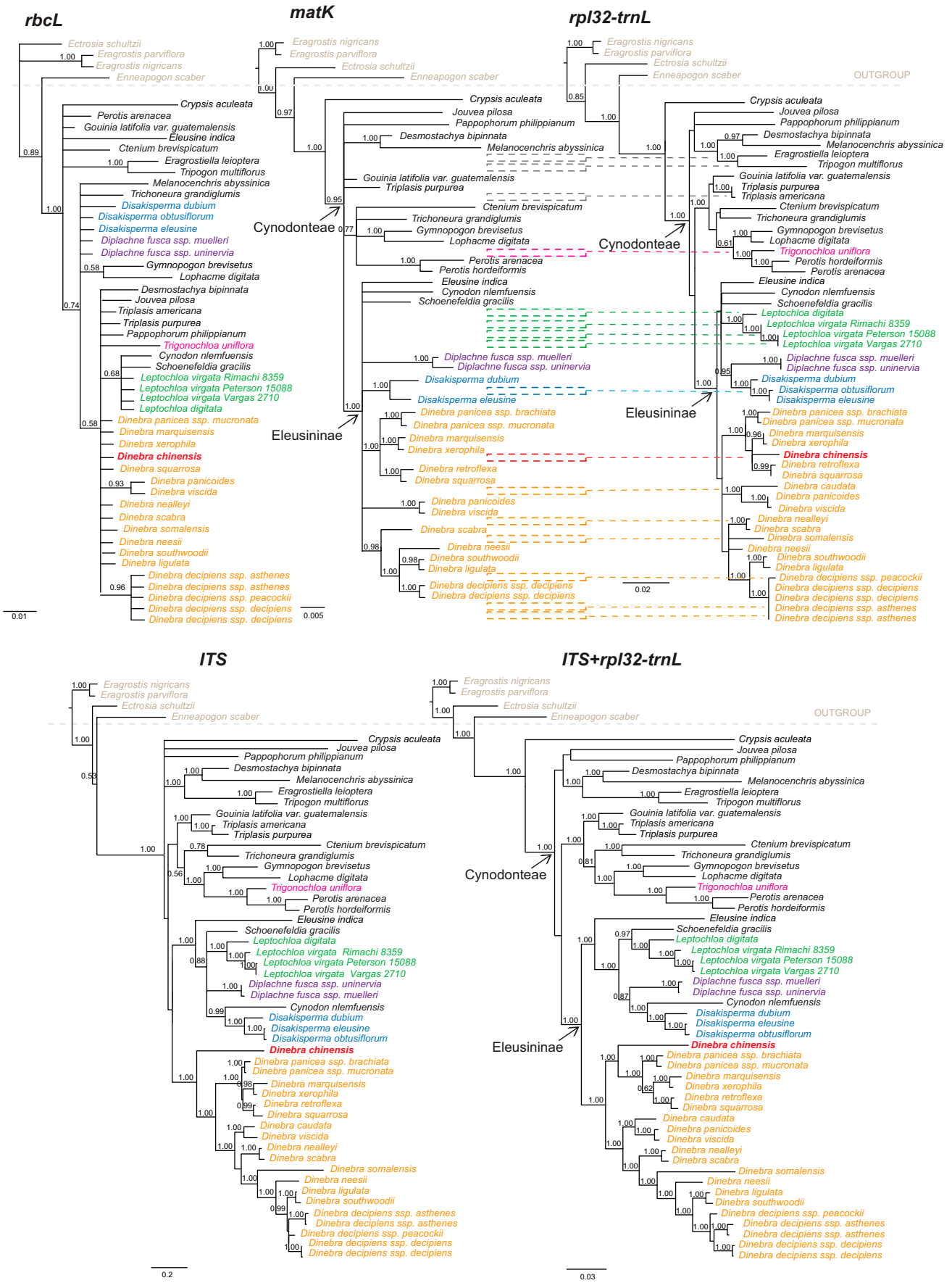

Figure 5. Bayesian majority-rule consensus tree of *rbcL* (upper left), *matK* (upper middle), *rpl32-trnL* (upper right) plastid regions, nuclear DNA internal transcribed spacer (ITS) and combined ITS + *rpl32-trnL* data set. Numbers above branches are Bayesian posterior probability values. Empty boxes on *matK* tree indicate unsuccessful sequencing of the species as shown on the *rpl32-trnL* tree. *Dinebra chinensis* is marked in red and discussed in the text.

Leptochloa (PP = 1.00), *Trigonochloa* (PP = 1.00) as sister to *Perotis* Aiton, and four separate clades (PP = 1.00) within an unsupported polytomy of *Dinebra* (Figure 5). *Dinebra chinensis* is firmly embedded in one of the strongly supported clades within *Dinebra*.

The ITS barcode provides impressive resolution at terminal nodes, especially in *Dinebra* where monophyly is strongly supported (Figure 5, PP = 1.00). *Leptochloa*, *Diplachne* and *Disakisperma* are also supported as monophyletic with high support (PP = 1.00). However, the ITS tree lacks backbone support and does not support a monophyletic Eleusininae or Cynodonteae.

The combined ITS + *rpl32-trnL* barcodes provide good resolution of all groups previously discussed (Figure 5). It provides strong backbone support as well as support for all genera derived from *Leptochloa* s.l. Within *Dinebra* there is maximum support for all terminal nodes (PP = 1.00). Of the inner nodes within *Dinebra* only relationships between *Dinebra marquisensis* (F. Br.) P.M. Peterson & N. Snow and *Dinebra xerophila* (P.M. Peterson & Judz.) P.M. Peterson & N. Snow; and *Dinebra retroflexa* (Vahl) Panz. and *Dinebra squarrosa* (Pilg.) P.M. Peterson & N. Snow remained poorly supported (Figure 5, PP = 0.62).

The position of *Dinebra chinensis* differs in the ITS, *rpl32-trnL*, and ITS + *rpl32-trnL* trees. In the ITS tree *D. chinensis* is sister to the all remaining species of *Dinebra*, whereas in the *rpl32-trnL* tree it forms an unresolved trichotomy with *D. marquisensis–D. xerophila* and *D. retroflexa–D. squarrosa* pairs (Figure 5). In the combined ITS + *rpl32-trnL* tree *D. chinensis* is sister (PP = 1.00) to a larger clade that contains *Dinebra panicea* (Retz.) P. M. Peterson & N. Snow, which is sister to *D. marquisensis–D. xerophila* and *D. retroflexa–D. squarrosa*.

Overall, the number of successfully recovered nodes with strong support (PP = 0.95–1.00) for different barcode markers is: *rbcL*, 3; *matK*, 19; *rp32-trnL*, 27; ITS, 34; and combined ITS + *rpl32-trnL*, 43.

Discussion

The present state of barcode "machinery" does not meet the basic criteria for developing barcode libraries: (1) universality (ease of amplification and sequencing); (2) sequence quality, and (3) discriminatory power (identification at the species level and/or correct phylogenetic placement of taxa). In grasses, the standard *rbcL* DNA barcode primers sometimes do not yield target amplicons. In some grasses the PCR product is contaminated with a mitochondrial copy (paralogous pseudogene) in the *rbcL* gene (Cummings et al. 2003), which renders many sites ambiguous with double peaks in the chromatograms. We have developed a revised protocol and a new set of primers that amplify only the chloroplast copy of the gene. Even with a set of orthologous sequences of *rbcL*, we were not able to provide suitable resolution for the model data set in our study of chloridoid grasses with 23 species of *Leptochloa* s.l. Therefore, the utility of *rbcL* in grasses remains questionable.

The fast-evolving coding region of *matK* has been widely applied in phylogenetic studies of Poaceae and Chloridoideae, in particular (Hilu et al. 1999; Hilu and Lawrence 2001; GPWGII 2012). In our study, *matK* was more variable than *rbcL*, and was useful in our tree-based analyses. The *matK* marker correctly recovered clades corresponding to major phylogenetic groups and provided back-bone support for the tree. The primary problem with using *matK* as a barcode marker is the low success of sequencing (Hollingsworth et al. 2011). Our data set suffered significantly from the low rate of amplification in the *matK* barcoding region, especially in *Dinebra* and *Leptochloa* s.s.

The fast evolving region of *trnH-psbA* has been suggested as a supplement for the existing two-barcode set of *rbcL* and *matK* (Kress et al. 2009; Drumwright et al. 2011; Cai et al. 2012). However, in grasses the intergenic spacer is not present in its entire length due to a deletion and contains a complete duplicated *rps19* gene derived from the inverted repeat B at the end of the large single copy unit of the plastid genome (Wang et al. 2008; Romaschenko et al. 2012). The region is characterized by a significantly lower rate of nucleotide substitution, which makes it ineffective in resolving low level relationships in the Poaceae. Instead, we use the supplemental barcode, *rpl32-trnL*. The *rpl32-trnL* region shows the highest level of sequence divergence among our plastid barcode sequences and has been suggested as a barcode because it contains a large amount of polymorphic sites in non-coding regions (Shaw et al. 2007). The *rpl32-trnL* region is very easy to amplify and was successfully used in phylogenetic reconstructions in Chloridoideae and Pooideae (Romaschenko et al. 2008, 2010, 2011, 2012, 2014; Peterson et al. 2010a, 2010b; 2011, 2012, 2014). The *rpl32-trnL* barcode marker shows a high level of sequence divergence among taxonomically closely related species and rendered a unique sequence ensuring the correct identification of *Dinebra chinensis*, a Federally listed noxious weed. Within *Dinebra*, the discriminatory power of *rpl32-trnL* proved to be more than eight times greater than that of the *matK* region. With the *rpl32-trnL* barcode we were able to correctly reconstruct the phylogenetic relationships within the *Leptochloa* s.l. complex and differentiate among the five recently segregated genera (Peterson et al. 2012).

The nuclear ribosomal ITS has been reported as an obvious choice for a supplementary barcode (Thomas 2009; Chen et al. 2010; Hollingsworth et al. 2011). However, there are several problems associated with the use of ITS since paralogous copies are common in polyploids, genetic mutation saturation occurs rapidly (making alignments difficult beyond closely related genera), and often the PCR product contains fungal contamination (Jani et al. 2010). All of these problems have prevented this marker from being widely used in barcode studies. We feel it is extremely valuable at discriminating

among taxonomically closely related species and are employing it as a supplementary barcode for the grasses. The risk of fungal contamination is reduced by using ITS5a, a forward primer designed specifically for plants. Even when facing alignment problems, the ITS barcode provides the highest level of sequence divergence, i.e. discrimination, among the 50 taxa in our study, and is several times higher than *rpl32-trnL*, our best plastid marker. In our analysis the discriminatory power of ITS barcode is 96%, and this is three times greater than *rpl32-trnL* and 33 times greater than *matK*. Within *Dinebra* and its closely related sisters, *Diplachne*, *Disakisperma* and *Leptochloa*, the ITS barcode showed a significant threshold in intrageneric and intergeneric divergence not seen in the other plastid barcodes. The ITS region is indispensable in our barcode studies where precise identification is necessary and on a large scale this outweighs problems associated with paralogues. Used in tandem with *rpl32-trnL,* these two barcodes are able to provide a fairly robust phylogeny that can be used for biogeographic and evolutionary studies, in addition to taxon identification. The tree-based method for identification is supplementary to BLAST identification and can help researchers make correct determinations if a taxon or group of species is poorly represented in GenBank. The use of two supplementary barcode markers (*rpl32-trnL* and ITS) in the grasses will help to provide insights into species-level taxonomy and is essential in identifying unknown species, especially in situations where morphology-based identification is challenging. Use of these barcodes represents a powerful tool in cases where correct identification is essential, e.g. for the recognition of noxious weeds or when distinguishing invasive species from the native flora.

Acknowledgements

This study received funding from the United States Department of Agriculture and all laboratory work was done at the Smithsonian's Laboratory of Analytical Biology. We thank Jeff Saarela, Riccardo M. Baldini, and one anonymous reviewer for their constructive comments on an earlier draft.

References

Altschul SF, Madden TL, Schäffer AA, Zhang A, Zhang Z, Miller W, Lipman DJ. 1997. Gapped Blast and psi-Blast: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.

Barkworth ME, Capels KM, Long S, Anderton LK, Piep MB. 2007. Flora of North America, North of Mexico, vol. 24. Magnoliophyta: Commelinidae (in part): Poaceae, Part 1. Oxford: Oxford University Press.

Barkworth ME, Capels KM, Long S, Piep MB. 2003. Flora of North America, North of Mexico, vol. 25. Magnoliophyta: commelinidae (in part): Poaceae, Part 2. Oxford: Oxford University Press.

Barrett RDH, Hebert PDN. 2005. Identifying spiders through DNA barcodes. Can J Zool. 83:481–491.

Cai ZM, Zhang YX, Zhang LN, Gao LM, Li DZ. 2012. Testing four candidate barcoding markers in temperate woody bamboos (Poaceae: Bambusoideae). J Syst Evol. 50:527–539.

CBOL (Consortium for the Barcode of Life, Plant Working Group). 2009. A DNA barcode for land plants. Proc Natl Acad Sci USA. 106:12794–12797. doi:10.1073/pnas.0905845106.

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, et al. 2010. Validation of the ITS2 region as a novel DNA barcode for indentifying medicinal plant species. PLOS ONE. 5:e8613. doi:10.1371/journal.pone.0008613.

Collins RA, Boykin LM, Cruickshank RH, Armstrong KF. 2012. Barcoding's next top model: an evaluation of nucleotide substitution models for specimen identification. Meth Ecol Evol. 3:457–465.

Crawford K, Land J, Rudgers J. 2010. Fungal endophytes of native grasses decrease insect herbivore preference and performance. Oecologia. 164:431–444.

Cuénoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. Amer J Bot. 89:132–144.

Cummings MP, Nugent JM, Olmstead RG, Palmer JD. 2003. Phylogenetic analysis reveals five independent transfers of the chloroplast gene rbcL to the mitochondrial genome in angiosperms. Curr Gen. 43:131–138.

Drumwright AM, Allen BW, Huff KA, Ritchey PA, Cahoon AB. 2011. Survey and DNA barcoding of poaceae in flat rock cedar glades and barrens state natural area, murfreesboro, tennessee. Castanea. 76:300–310.

Espejo Serna A, López-Ferrari AR, Valdes Reyna J. 2000. Poaceae. In: Espejo Serna A, López-Ferrari AR, editors. Las monocotiledóneas Mexicanas: una synopsis floristica,Partes IX–XI, Pandanaceae a Zosteraceae [The Mexican monocotyledons: a floristic synopsis, parts IX–XI, Pandanaceae to Zosteraceae]. Mexico, D.F: Consejo Nacional de la Flora de Mexico and Comisión Nacioanl para el Conocimieto y Uso de la Biodiversidad; p. 8–236.

Federal Noxious Weeds List. 2012. United States Department of Agriculture, Natural Resources Conservation Service, Plants Database. [Cited 2014 16 May]. Available from: http://plants.usda.gov/java/noxious?rptType=Federal

GPWG (Grass Phylogeny Working Group). 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). Ann Missouri Bot Gard. 88:373–457.

GPWGII (Grass Phylogeny working Group II). 2011. New grass phylogeny resolves deep evolutionary relationships and discovers C_4 origins. New Phytologist. 193:304–312.

Hawlitschek O, Nagy ZT, Berger J, Glaw F. 2013. Reliable DNA barcoding performance proved for species and island populations of comoran squamate reptiles. PLOS ONE. 8: e73368. doi:10.1371/journal.pone.0073368.

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. Proc R Soc London B. 270:313–321. doi:10.1098/rsbl.2009.0848.

Hilu KW, Lawrence AA. 2001. A Phylogeny of Chloridoideae (Poaceae) based on matK sequences. Syst Bot. 26:386–405.

Hilu KW, Lawrence AA, Hongping L. 1999. Phylogeny of Poaceae inferred from matK sequences. Ann Missouri Bot Gard. 86:835–851.

Hiratsuka J, Hiroaki S, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, et al. 1989. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA gene accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet. 217:185–194.

Hollingsworth PM, Graham SW, Little DP. 2011. Choosing and using a plant DNA barcode. PLOS ONE. 6:e19254. doi:10.1371/journal.pone.0019254.

Huelsenbeck JP, Ronquist FR. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 17:754–755.

Jani AJ, Faeth SH, Gardner D. 2010. Asexual endophytes and associated alkaloids alter arthropod community structure and increase herbivore abundances on a native grass. Ecology Letters. 13:106–117.

Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjur O, Bermingham E. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc Nat Acad Sci. 106:18621–18626.

Liu J, Zhou S, Li C, Dong W, Yang R. 2014. DNA barcoding of Leymus (Poaceae). Quest Journals. 2:1–11.

Meier R, Kwon S, Vaidya G, Ng PKL. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst Biol. 55:715–728.

Munch K, Boomsma W, Huelsenbeck JP, Willerslev E, Nielsen R. 2008a. Statistical assignment of DNA sequences using Bayesian phylogenetics. Syst Biol. 57:750–757. doi:10.1080/10635150802422316.

Munch K, Boomsma W, Willerslev E, Nielsen R. 2008b. Fast phylogenetic DNA barcoding. Phil Trans R Soc B. 363:3997–4002.

Peterson PM, Romaschenko K, Barker NP, Linder HP. 2011. Centropodieae and *Ellisochloa*, a new tribe and genus in Chloridoideae (Poaceae). Taxon. 60:1113–1122. Available from: http://www.ingentaconnect.com/content/iapt/tax/2011/00000060/00000004/art00014

Peterson PM, Romaschenko K, Johnson G. 2010a. A classification of the Chloridoideae (Poaceae) based on multi-gene phylogenetics trees. Mol Phylogenet Evol. 55:580–598. Available from: http://dx.doi.org/10.1016/j.ympev.2010.01.018

Peterson PM, Romaschenko K, Johnson G. 2010b. A phylogeny and classification of the *Muhlenbergiinae* (Poaceae: Chloridoideae: Cynodonteae) based on plastid and nuclear DNA sequences. Amer J Bot. 97:1532–1554. Available from: http://dx.doi.org/10.3732/ajb.0900359

Peterson PM, Romaschenko K, Herrera Arrieta Y. 2014. A molecular phylogeny and classification of the *Cteniinae, Farragininae, Gouiniinae, Gymnopogoninae, Perotidinae*, and *Trichoneurinae* (Poaceae: Chloridoideae: Cynodonteae). Taxon. 63:275–286. Available from: http://www.ingentaconnect.com/content/iapt/tax/2014/00000063/00000002/art00006

Peterson PM, Romaschenko K, Snow N, Johnson G. 2012. A molecular phylogeny and classification of *Leptochloa* (Poaceae: Chloridoideae: Chlorideae) *sensu lato* and related genera. Ann Bot. London. 109:1317–1329. Available form: http://dx.doi.org/10.1093/aob/mcs077

Romaschenko K, Garcia-Jacas N, Peterson PM, Soreng RJ, Vilatersana R, Susanna A. 2014. Miocene−Pliocene speciation, introgression, and migration of *Patis* and *Ptilagrostis* (Poaceae: Stipeae). Mol Phylogenet Evol. 70:244–259. Available form: http://dx.doi.org/10.1016/j.ympev.2013.09.018

Romaschenko K, Peterson PM, Soreng RJ, Futorna O, Susanna A. 2011. Phylogenetics of *Piptatherum* s.l. (Poaceae: Stipeae): evidence for a new genus. Piptatheropsis and resurrection of Patis. Taxon. 60:1703–1716. Available form: http://www.ingentaconnect.com/content/iapt/tax/2011/00000060/00000006/art00015

Romaschenko K, Peterson PM, Soreng RJ, Garcia-Jacas N, Futorna O, Susanna A. 2008. Molecular phylogenetic analysis of the American Stipeae (Poaceae) resolves *Jarava sensu lato* polyphyletic: evidence for a new genus. Pappostipa. J Bot Res Inst Texas. 2:165–192.

Romaschenko K, Peterson PM, Soreng RJ, Garcia-Jacas N, Futorna O, Susanna A. 2010. Phylogenetics of *Stipeae* (Poaceae, Pooideae) based on plastid and nuclear DNA sequences. In: Seberg O, Petersen G, Barfod AS, Davis JI, editors. Diversity, Phylogeny, and Evolution in the Monocotyledons. Denmark: Aarhus University Press; p. 511–537.

Romaschenko K, Peterson PM, Soreng RJ, Garcia-Jacas N, Futorna O, Susanna A. 2012. Systematics and evolution of the needle grasses (Poaceae: Pooideae: Stipeae) based on analysis of multiple chloroplast loci, ITS, and lemma micromorphology. Taxon. 61:18–44. Available form: http://digital.csic.es/handle/10261/72333

Saarela JM, Sokoloff PC, Gillespie LJ, Consaul LL, Bull RD. 2013. DNA barcoding the Canadian arctic flora: core plastid barcodes (rbcL + matK) for 490 vascular plant species. PLOS ONE. 8:e77982. doi:10.1371/journal.pone.0077982.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Amer J Bot. 94:275–288.

Snow N, Peterson PM. 2012. Systematics of Trigonochloa (Poaceae, Chloridoideae, Chlorideae). PhytoKeys. 13:25–38.

Snow N, Peterson PM, Romaschenko K. 2013. Systematics of *Disakisperma* (Poaceae, Chloridoideae, Chlorideae). PhytoKeys. 26:21–70.

Soreng RJ, Davidse G, Peterson PM, Zuloaga FO, Judziewicz EJ, Filgueiras TS, Morrone O, Romaschenko K. 2014. Catalogue of New World grasses. [Cited 2014 May]. Available form: http://tropicos.org/Project/CNWG.

Srivathsan A, Meier R. 2011. On the inappropriate use of Kimura 2 parameter (K2P) divergences in the DNA barcoding literature. Cladistics. 28:190–194.

Stanford AM, Harden R, Parks CR. 2000. Phylogeny and biogeography of *Juglans* (Juglandaceae) based on matK and ITS sequence data. Amer J Bot. 87:872–882.

Swofford DL. 2000. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sunderland, Massachusetts: Sinauer Associates Inc.

Syme AE, Murphy DJ, Holmes GD, Gardner S, Fowler R, Cantrill DJ. 2012. An expanded phylogenetic analysis of *Austrostipa* (Poaceae: Stipeae) to test infrageneric relationships. Austral. Syst. Bot. 25:1–10.

Thomas C. 2009. Plant barcode soon to become reality. Science. 325:526. doi:10.1126/science.325_526.

Wang R, Cheng C, Chang C, Wu C, Su T, Chaw SM. 2008. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genome of monocots. B.M.C. Evol. Biol. 8:36. doi:10.1186/1471-2148-8-36.

Ward J, Gilmore SR, Robertson J, Peakall R. 2009. A grass molecular identification system for forensic botany: a critical evaluation of the strengths and limitations. J Forensic Sci. 54:1254–1260.

White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M, Gelfand D, Sninsky J, White T, editors. PCR protocols: a guide to methods and applications. New York, NY: Academic Press Inc; p. 315–322.

Zhang YX, Xu XX, Ma PF, Zang LN, Li DZ. 2013. Selection of potential plastid DNA barcodes for Bambusoideae (Poaceae). Plant Div Res. 35:743–750.

Appendix 1

List of specimens sampled, voucher (collector, number and where the specimen is housed), country of origin, and GenBank accession number for each DNA barcode sequence. All accessions marked in **bold** are newly submitted to GenBank. *Data obtained from GenBank.

	Taxon	Voucher	Country	rpl32-trnL	rbcL	matK	ITS
1	*Crypsis aculeata* (L.) Aiton	Soreng 7940, Johnson, Johnson, Dzyubenko, Dzyubenko & Schilnikov (US)	Russia	JQ345316	**KJ768921**	**KJ768888**	JQ345163
2	*Ctenium brevispicatum* J.G.Sm.	Filgueiras s.n.& Oliveira (MO)	Brazil	JQ345317	**KJ768922**	**KJ768889**	JQ345164
3	*Cynodon nlemfuensis* Vanderyst	Lobo 2099, Cascante, Quesada & Prieto (CR)	Costa Rica	**KJ768976**	**KJ768923**	**KJ768890**	**KJ768881**
4	*Desmostachya bipinnata* (L.) Stapf	Dwyer 13,120 (MO)	Saudi Arabia	JQ345319	**KJ768924**	**KJ768891**	JQ345165
5	*Dinebra caudata* (K. Schum.) P.M. Peterson & N. Snow	Kuchar 24,265 (MO)	Tanzania	JQ345324	**KJ768925**	failed	JQ345170
6	*Dinebra chinensis* (L.) P.M. Peterson & N. Snow	Snow 6909 (MO)	Botswana	JQ345325	**KJ768926**	failed	JQ345171
7	*Dinebra decipiens* subsp. *asthenes* (Roem. & Schult.) P.M. Peterson & N. Snow	Snow 7327 (MO)	Australia	JQ345326	**KJ768927**	failed	JQ345173
8	*Dinebra decipiens* subsp. *asthenes* (Roem. & Schult.) P.M. Peterson & N. Snow	Snow 7355 (MO)	Australia	JQ345328	**KJ768929**	failed	JQ345175
9	*Dinebra decipiens* subsp. *decipiens* (R. Br.) P.M. Peterson & N. Snow	Snow 7328 & Simon (MO)	Australia	JQ345327	**KJ768928**	**KJ768892**	JQ345174
10	*Dinebra decipiens* subsp. *decipiens* (R. Br.) P.M. Peterson & N. Snow	Waterhouse 5948 (MO)	Australia	JQ345330	**KJ768931**	**KJ768893**	JQ345177
11	*Dinebra decipiens* subsp. *peacockii* (Maiden & Betche) P.M. Peterson & N. Snow	Snow 7361 & Simon (MO)	Australia	JQ345329	**KJ768930**	failed	JQ345176
12	*Dinebra ligulata* (Lazarides) P M. Peterson & N. Snow	Snow 7402 (MO)	Australia	JQ345336	**KJ768932**	**KJ768894**	JQ345183
13	*Dinebra marquisensis* (F. Br.) P.M. Peterson & N. Snow	Wood 10,145 (MO)	French Polynesia	JQ345337	**KJ768933**	**KJ768895**	JQ345184
14	*Dinebra nealleyi* (Vasey) P.M. Peterson & N. Snow	Snow 5806 (MO)	USA	JQ345338	**KJ768934**	failed	JQ345185
15	*Dinebra neesii* (Thwaites) P.M. Peterson & N. Snow	Snow 7380 & Simon (MO)	Australia	JQ345339	**KJ768935**	**KJ768896**	JQ345186
16	*Dinebra panicea* subsp. *brachiata* (Steud.) P.M. Peterson & N. Snow	Peterson 22,185 & Saarela (US)	Mexico	GU359810	**KJ768936**	**KJ768897**	GU359146
17	*Dinebra panicea* subsp. *mucronata* (Michx.) P.M. Peterson & N. Snow	Peterson 9546, Annable & Herrera (US)	Mexico	JQ345342	**KJ768937**	**KJ768898**	JQ345188
18	*Dinebra panicoides* (J. Presl) P.M. Peterson & N. Snow	Snow 5810a (MO)	USA	JQ345343	**KJ768938**	**KJ768899**	failed
19	*Dinebra retroflexa* (Vahl) Panz.	Ndegwa 610 (US)	Kenya	GU359778	**KJ768939**	JF729105	GU359332
20	*Dinebra scabra* (Nees) P.M. Peterson & N. Snow	Snow 5788 (MO)	USA	JQ345344	**KJ768940**	**KJ768900**	JQ345189
21	*Dinebra somalensis* (Stapf) P.M. Peterson & N. Snow	Faden 74–991 & Faden (MO)	Kenya	JQ345321	**KJ768941**	failed	JQ345166
22	*Dinebra southwoodii* (N. Snow & B·K. Simon) P.M. Peterson & N. Snow	Snow 7362 (MO)	Australia	JQ345345	**KJ768942**	**KJ768901**	JQ345190

(*Continued*)

(*Continued*).

	Taxon	Voucher	Country	rpl32-trnL	rbcL	matK	ITS
23	*Dinebra squarrosa* (Pilg.) P.M. Peterson & N. Snow	Kayombo 5196 (MO)	Tanzania	JQ345346	**KJ768943**	**KJ768902**	JQ345191
24	*Dinebra viscida* (Scribn.) P.M. Peterson & N. Snow	Peterson 22,184 & Saarela (US)	Mexico	GU359808	**KJ768944**	**KJ768903**	GU359148
25	*Dinebra xerophila* (P.M. Peterson & Judz.) P.M. Peterson & N. Snow	Perlman 18,438 (MO)	French Polynesia	JQ345353	**KJ768945**	**KJ768904**	JQ345196
26	*Diplachne fusca* var. *muelleri* (Benth.) J.M. Black	Badman 1282 (MO)	Australia	JQ345334	**KJ768946**	**KJ768905**	JQ345181
27	*Diplachne fusca* var. *uninervia* (J. Presl) P.M. Peterson & N. Snow	Peterson 20,786, Soreng, Romaschenko & Gonzalez-Elizondo (US)	Peru	JQ345335	**KJ768947**	**KJ768906**	JQ345182
28	*Disakisperma dubium* (Kunth) P.M. Peterson & N. Snow	Peterson 8105 & Annable (US)	Mexico	JQ345332	**KJ768948**	**KJ768907**	JQ345179
29	*Disakisperma eleusine* (Nees) P.M. Peterson & N. Snow	Snow 6982 (MO)	South Africa	JQ345333	**KJ768949**	**KJ768908**	JQ345180
30	*Disakisperma obtusiflorum* (Hochst.) P.M. Peterson & N. Snow	Belsky 527 (MO)	Kenya	JQ345340	**KJ768950**	failed	JQ345187
31	*Ectrosia schultzii* Benth.	Dunlop 8428 (MO)	Australia	**KJ768977**	**KJ768951**	**KJ768909**	**KJ768882**
32	*Eleusine indica* (L.) Gaetrn.	Peterson 21,362, Saarela & Flores Villegas (US)	Mexico	GU359797	**KJ768952**	JF729107	GU359338
33	*Enneapogon scaber* Lehm.	Sachse 008 (MO)	South Africa	JQ345322	**KJ768953**	**KJ768910**	JQ345168
34	*Eragrostiella leioptera* (Stapf) Bor	Chand 7961 (US)	India	GU359827	**KJ768954**	failed	GU359305
35	*Eragrostis nigricans* (Kunth) Steud.	Peterson 21,623, Soreng, LaTorre & Rojas Fox (US)	Peru	GU359790	**KJ768955**	**KJ768911**	GU359299
36	*Eragrostis parviflora* (R. Br.) Trin.	Snow 7357 & Simon (MO)	Australia	**KJ768978**	**KJ768956**	**KJ768912**	**KJ768883**
37	*Gouinia latifolia* var. *guatemalensis* (Hack.) J.J. Ortíz	Martinez S. 29,764, Alvarez & Ramirez (MO)	Mexico	KF827638	**KJ768957**	**KJ768913**	KF827520
38	*Gymnopogon brevisetus* (Hack.) J.P. Sm.	Nee 37,732 (MO)	Bolivia	KF827648	**KJ768958**	**KJ768914**	KF827531
39	*Jouvea pilosa* (J. Presl) Scribn.	Reina G. 2007–1095 (MO)	Mexico	**KJ768979**	**KJ768960**	**KJ768915**	**KJ768884**
40	*Leptochloa digitata* (R.Br.) Domin	Risler 476 (MO)	Australia	JQ345331	**KJ768961**	failed	JQ345178
41	*Leptochloa virgata* (L.) P.Beauv.	Peterson 15,088 & Refulio-Rodriguez (US)	Peru	JQ345349	**KJ768962**	failed	JQ345193
42	*Leptochloa virgata* (L.) P.Beauv.	Rimachi 8359 (US)	Peru	JQ345350	**KJ768963**	failed	JQ345194
43	*Leptochloa virgata* (L.) P.Beauv.	Vargas 2710 (US)	Bolivia	JQ345351	**KJ768964**	failed	JQ345195
44	*Lophacme digitata* Stapf	Smook 1453 (MO)	South Africa	JQ345354	**KJ768965**	**KJ768916**	JQ345197
45	*Melanocenchris abyssinica* (R.Br. ex Fresen.) Hochst.	DeWilde 6912 (MO)	Ethiopia	JQ345355	**KJ768966**	failed	JQ345198
46	*Melanocenchris abyssinica* (R.Br. ex Fresen.) Hochst.	Hilu 5645 (VPI)	India			AF312326*	
47	*Pappophorum philippianum* Parodi	Renvoize 4225, Cope & Beck (MO)	Bolivia	**KJ768980**	**KJ768967**	**KJ768917**	**KJ768885**
48	*Perotis arenacea* (Judz.) P.M. Peterson	Phillipson 4117 & Raharilala (MO)	Madagascar	JQ345358	**KJ768968**	**KJ768918**	JQ345202
49	*Perotis hordeiformis* Nees	Soreng 5717, Peterson & Sun Hang (US)	China	GU359991	**KJ768969**	JF729138	GU359132
50	*Schoenefeldia gracilis* Kunth	Pauwels 7413 (MO)	Benin	**KJ768981**	**KJ768970**	**KJ768919**	**KJ768886**
51	*Trichoneura grandiglumis* var. *grandiglumis* (Nees) Ekman	Smook 4896 (MO)	South Africa	KF827681	**KJ768971**	**KJ768920**	KF827560
52	*Trigonochloa uniflora* (Hochst. ex A. Rich.) P. M. Peterson & N. Snow	Snow 6978, Burgoyne & Gumbi (MO)	South Africa	JQ345348	**KJ768972**	failed	JQ345192
53	*Triplasis americana* P. Beauv.	Kral 12,065 (MO)	USA	KF827690	**KJ768973**	failed	**KJ768887**
54	*Triplasis purpurea* (Walter) Chapm.	Peterson 14,238, Weakley & LeBlond (US)	USA	GU359921	**KJ768974**	JF729155	GU359184
55	*Tripogon multiflorus* Miré & H.Gillet	Spellenberg 7441 (MO)	Yemen	JQ345360	**KJ768975**	failed	JQ345204