DNA barcoding for identification of exotic grass species present in eastern Australia

Aisuo Wang[1,2], David Gopurenko[1,2], Hanwen Wu[1,], Rex Stanton[1,2] and Brendan J. Lepschi[3]
[1] Department of Primary Industries, Wagga Wagga Agricultural Institute, Private Mail Bag, Wagga Wagga, NSW 2650, Australia
[2] Graham Centre for Agricultural Innovation, Locked Bag 588, Wagga Wagga, NSW 2678, Australia
[3] Australian National Herbarium, Centre for Plant Biodiversity Research, GPO Box 1600, Canberra, ACT 2601, Australia
(aisuo.wang@dpi.nsw.gov.au)

Summary Grass identification traditionally relies on morphological examination of floral material, and often requires specialist knowledge of morphological characteristics diagnostic for the various genera and species. DNA barcoding may provide a rapid genetic screening tool to identify selected noxious grass weeds at the vegetative growth stage. We tested the utility of 18 chloroplast and nuclear genes as potential DNA barcodes for species identifications of introduced grasses present in Eastern Australia. Grasses examined (N = 417) included *Nassella neesiana* (Trin. & Rupr.) Barkworth (Chilean needle grass), *Nassella trichotoma* (Nees) Hack. ex Arechav. (serrated tussock), *Eragrostis curvula* (Schrad.) Nees (African love grass) and 26 other weed species collected from New South Wales (NSW), Australian Capital Territory (ACT) and other parts of Australia. Our preliminary results revealed three chloroplast genes, *matK*, *ndhK* and *petL*, which exhibit potentials as DNA barcodes for distinguishing and identifying weeds species of interest.

Keywords DNA barcoding, weeds, *Nassella neesiana*, *Nassella trichotoma*, *Eragrostis curvula*.

INTRODUCTION

Invasive grass weeds can cause severe losses to farmers by reducing productivity of grazing land and livestock, and increasing weed control costs. Accurate identification of invasive species, at all stages of growth, allows for efficient control and management of potential infestations.

Grass identification traditionally relies on examination of floral material, and often requires specialist knowledge of the morphological characteristics exhibited in the grasses. The misidentification of weeds might lead to either control measures needlessly being imposed on desirable native species, or lack of prompt control of introduced species.

DNA barcoding (Hebert *et al.* 2003), an expedient sequence based means of identifying species across all stages of development and even from trace amounts of tissue (Pradosh and Sankar 2013), will be tested here to develop libraries of diagnostic sequences for identifying the key noxious grasses and other morphologically similar grasses present in eastern Australia. Once developed, these DNA barcodes could be used to rapidly screen query specimens collected in the field for identification purposes. This will enable targeting control activities to noxious weed populations in a timely manner, thereby reducing unnecessary spraying and removal of visually similar, more desirable grasses.

In this paper some preliminary results are reported.

MATERIALS AND METHODS

A total of 417 samples of 29 species were included in this study. Material includes live plants sampled from the field in NSW and the ACT (N = 333) as well as herbarium specimens from across mainland Australia, sampled from the collections of the Australian National Herbarium in Canberra (N = 84). Particular emphasis was given to three major weed species, *Nassella neesiana* (N = 85), *Nassella trichotoma* (N = 62) and *Eragrostis curvula* (N = 73). Each specimen was allocated a unique specimen ID for DNA analyses.

DNA was extracted from leaf tissue of each specimen (<1 mg) using a Corbett Research 1820 X-tractor Gene robotic system. Established DNA extraction protocols were followed (Gopurenko *et al.* 2013).

In order to identify the most suitable markers for DNA barcoding of these weed species, we screened 18 chloroplast and nuclear genes targets (*atpF*, *cemA*, *G3pdh*, *infA*, *ITS*, *matK*, *ndhK*, *petA*, *petL*, *psbK*, *rbcL*, *rpl16*, *rps14*, *tRNA-Leu*, *tRNA-Ser*, *tRNA-Thr*, *trnH* and *YCF6*) to determine levels of PCR fidelity across genera and species (Table 1) and identify intra/interspecific sequence differences among taxa. All PCR procedures followed Gopurenko *et al.* (2013) with the exception of primers used in PCR.

Bidirectional sequencing was conducted at the Australian Genome Research Facility for successful PCR products. All ABI trace sequences were quality

Table 1. Summary of specimen sampling and molecular analyses.

Genus	No. of species	No. of specimens	No. of genes surveyed	Genes surveyed	Partially sequenced genes
Austrostipa	2	12	2	*ndhK*, *petL*	
Chloris	6	50	9	*atpF*, *cemA*, *infA*, *ITS*, *ndhK*, *petL*, *rbcL*, *rps14*, *tRNA-Ser*	*atpF*, *infA*, *rps14*, *ndhK*, *cemA*, *petL*
Digitaria	2	8	2	*ndhK*, *petL*	*ndhK*, *petL*
Enteropogon	1	4	9	*atpF*, *cemA*, *infA*, *ITS*, *ndhK*, *petL*, *rbcL*, *rps14*, *tRNA-Ser*	*atpF*, *infA*, *rps14*, *ndhK*, *cemA*, *petL*
Eragrostis	2	77	9	*atpF*, *cemA*, *infA*, *ITS*, *ndhK*, *petL*, *rbcL*, *rps14*, *tRNA-Ser*	*atpF*, *infA*, *rps14*, *ndhK*, *cemA*, *tRNA-Leu*, *matK*, *tRNA-Thr*, *psbK*, *G3pdh*, *petL*
Hyparrhenia	1	14	2	*ndhK*, *petL*	*ndhK*, *petL*
Nassella	6	165	18	*atpF*, *cemA*, *G3pdh*, *infA*, *ITS*, *matK*, *ndhK*, *petA*, *petL*, *psbK*, *rbcl*, *rpl16*, *rps14*, *tRNA-Leu*, *tRNA-Ser*, *tRNA-Thr*, *trnH*, *YCF6*	*atpF*, *infA*, *ndhK*, *petL* , *tRNA-Leu*, *matK*, *tRNA-Thr*, *psbK*
Poa	6	21	18	*atpF*, *cemA*, *G3pdh*, *infA*, *ITS*, *matK*, *ndhK*, *petA*, *petL*, *psbK*, *rbcl*, *rpl16*, *rps14*, *tRNA-Leu*, *tRNA-Ser*, *tRNA-Thr*, *trnH*, *YCF6*	*atpF*, *infA*, *rps14*, *ndhK*, *cemA*, *petL*, *tRNA-Leu*, *matK*, *tRNA-Thr*, *psbK*
Rytidosperma	3	20	18	*atpF*, *cemA*, *G3pdh*, *infA*, *ITS*, *matK*, *ndhK*, *petL*, *psbK*, *rbcl*, *rpl16*, *rps14*, *tRNA-Leu*, *tRNA-Ser*, *tRNA-Thr*, *trnH*, *YCF6*	*atpF*, *infA*, *rps14*, *matK*, *tRNA-Thr*, *psbK*

checked and assembled using Lasergene SeqMan Pro ver. 8.1.0(3) (DNASTAR Inc., Maddison, WI, USA). Assembled sequences were exported to BioEdit (Hall 1999) and aligned using ClustalW (Larkin *et al.* 2007). Sequence alignments were imported into MEGA 6.0 (Tamura *et al.* 2013) and analysed as genetic distance trees using the neighbour-joining (NJ) method. Pairwise distances among sequences were adjusted as per the Kimura two-parameter model, and statistical support for all nodes in trees were estimated by bootstrap replication (N = 1000 replicates).

RESULTS AND DISCUSSION

Our preliminary results identified three promising chloroplast genes: *matK*, *ndhK* and *petL*, which PCR amplified successfully across species, and provided levels of sequence polymorphisms useful to this study. These gene targets separate taxa to genus level and in most cases to species level (Figure 1, Figure 2, Figure 3). There are, however, clear differences in the ability of each gene to act as reliable stand-alone DNA barcode for species identification. For example, *N. neesiana* and *N. trichotoma* are marginally separated in the NJ tree inferred from *matK* (Figure 1) and well separated in the *petL* tree (Figure 3); however these two species are poorly separated at the *ndhK* tree (Figure 2). In contrast, the *ndhK* tree provided a clear separation of the *Digitaria* species [*Digitaria divaricatissima* (R.Br.) Hughes and *Digitaria brownii* (Roem. & Schult.) Hughes] but not at the *petL* gene (Figure 2, Figure 3).

These results suggest that it is unlikely DNA barcoding for reliable identification of these taxa is feasible based on a single ubiquitous gene target; rather, specific genes (and/or combinations of these genes) are likely to be needed to genetically distinguish between closely related species within particular genera.

Although we have identified several targets from the chloroplast genome, which will be useful as DNA barcodes at many of the species surveyed, we have not identified similar informative nuclear genes despite our screening over several of the commonly targeted nuclear DNA barcode gene regions. It would be highly advantageous to identify useful nuclear DNA barcode

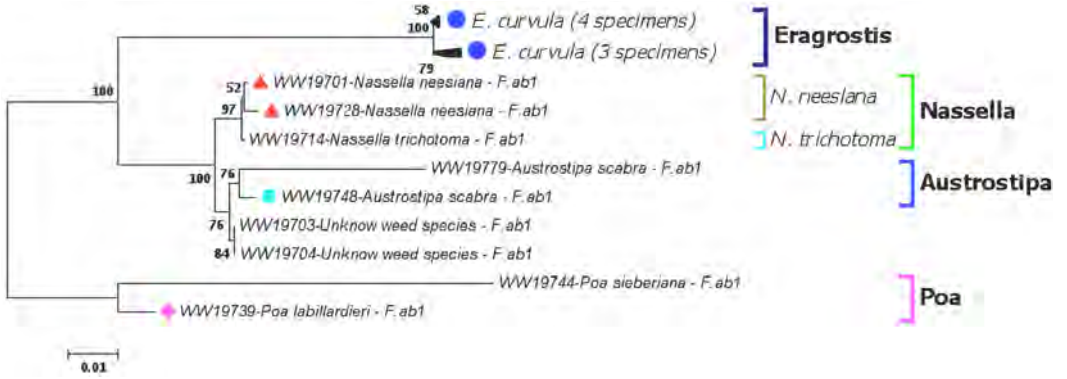

Figure 1. NJ trees constructed on the sequences of *matK* across samples of *Nassella*, *Eragrostis*, *Austrostipa* and *Poa* species.

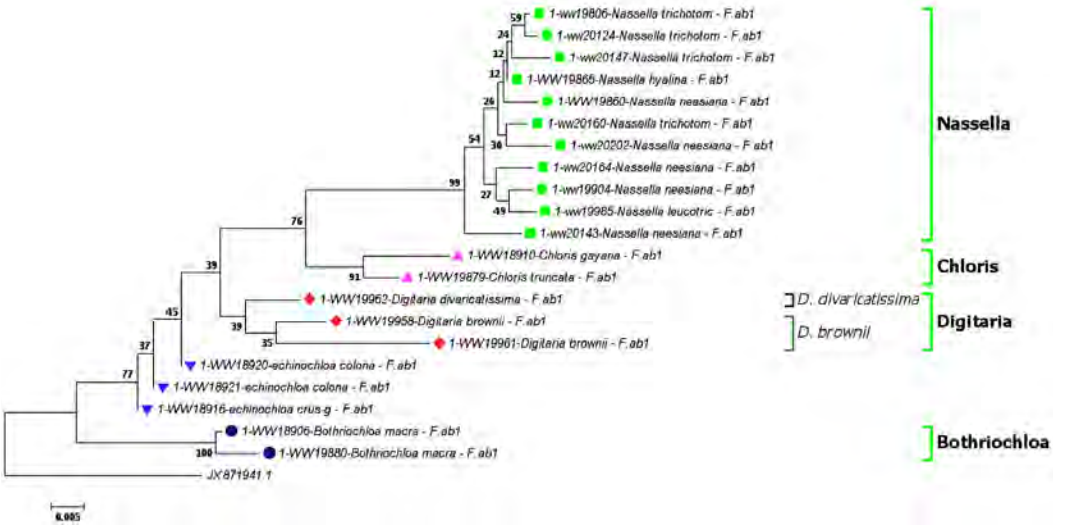

Figure 2. NJ trees constructed on the sequences of *ndhK* across specimens of *Nassella*, *Chloris*, *Digitaria* and *Bothriochloa* species.

regions, as these will provide species delimitations independent of the chloroplast genome and more importantly provide a comparative means to identify instances of gene flow and or hybridization between taxonomically defined morpho-species. To this end, more nuclear gene targets need to be screened for their utility as DNA barcode markers.

As this study is still at its early stage, the current results are promising. Further study to increase both the taxonomic breadth and the geographic range of species examined, will critically identify levels of intra and inter-specific genetic distances at the assayed genes and determine their reliability as DNA barcodes for accurate species identifications.

ACKNOWLEDGMENTS

This study was funded by the NSW Weeds Action Program. The authors are grateful to Mr Peter Deane (Wagga Wagga Agricultural Institute), Mr Adam Shephard (Wagga Wagga Agricultural Institute) and a number of local council Weeds Officers across New South Wales and Australian Capital Territory for their assistance in field sampling.

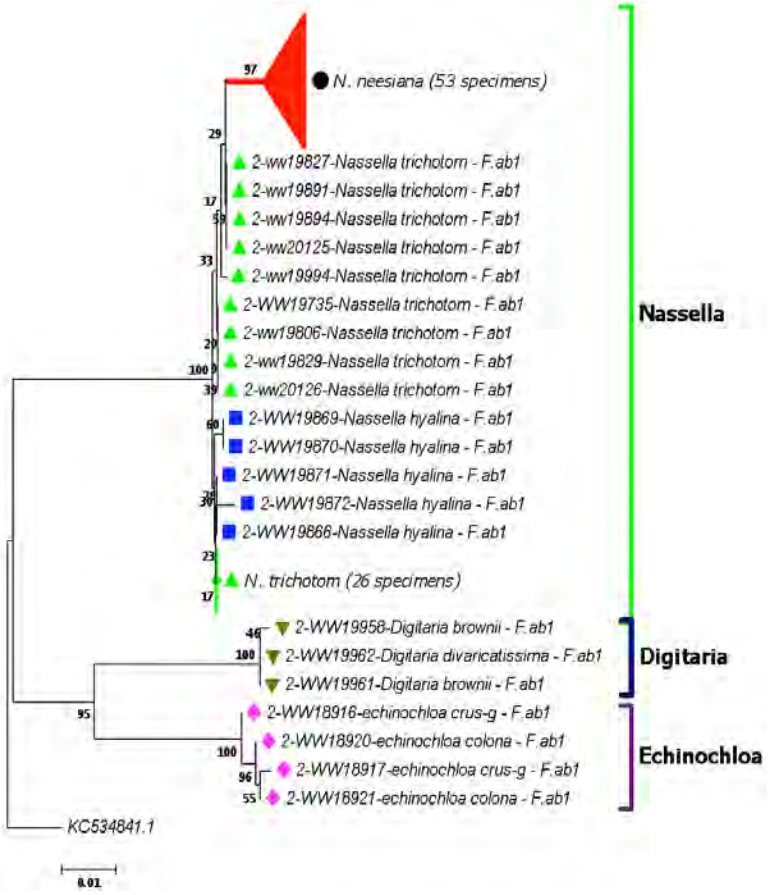

Figure 3. NJ trees constructed on the sequences of *petL* across specimens of *Nassella*, *Digitaria* and *Echinochloa* species.

REFERENCES

Gopurenko, D., Fletcher, M.J., Löcker, H. and Mitchell, A. (2013). Morphological and DNA barcode species identifications of leafhoppers, planthoppers and treehoppers (Hemiptera: Auchenorrhyncha) at Barrow Island. *In* 'Records of the Western Australian Museum, Supplement 83: The Terrestrial Invertebrate Fauna of Barrow Island, Western Australia', eds N.R. Gunawardene, J.D. Majer, C.K. Taylor and M.S. Harvey, pp. 253-85. (Western Australian Museum).

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41, 95-8.

Hebert, P.D., Cywinska, A., Ball, S.L. and DeWaard, J.R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: *Biological Sciences* 270, 313-21.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins D.G. (2007). ClustalW and ClustalX version 2. *Bioinformatics* 23, 2947-8.

Pradosh, M. and Sankar, K.G. (2013). DNA Barcoding: a tool for species identification from herbal juices. *DNA Barcodes* 1, 35-8.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30, 2725-9.