

RESEARCH ARTICLE

# Evaluation of six candidate DNA barcode loci for identification of five important invasive grasses in eastern Australia

Aisuo Wang<sup>1,2\*</sup>, David Gopurenko<sup>1,2</sup>, Hanwen Wu<sup>1,2</sup>, Brendan Lepschi<sup>3</sup>

**1** Wagga Wagga Agricultural Institute, NSW Department of Primary Industries, Wagga Wagga, Australia, **2** Graham Centre for Agricultural Innovation (An alliance between NSW Department of Primary Industries and Charles Sturt University), Wagga Wagga, New South Wales, Australia, **3** Australian National Herbarium, Centre for Australian National Biodiversity Research, Canberra, Australian Capital Territory, Australia

\* [aisuo.wang@dpi.nsw.gov.au](mailto:aisuo.wang@dpi.nsw.gov.au)



## Abstract

Invasive grass weeds reduce farm productivity, threaten biodiversity, and increase weed control costs. Identification of invasive grasses from native grasses has generally relied on the morphological examination of grass floral material. DNA barcoding may provide an alternative means to identify co-occurring native and invasive grasses, particularly during early growth stages when floral characters are unavailable for analysis. However, there are no universal loci available for grass barcoding. We herein evaluated the utility of six candidate loci (*atpF* intron, *matK*, *ndhK-ndhC*, *psbE—petL*, ETS and ITS) for barcode identification of several economically important invasive grass species frequently found among native grasses in eastern Australia. We evaluated these loci in 66 specimens representing five invasive grass species (*Chloris gayana*, *Eragrostis curvula*, *Hyparrhenia hirta*, *Nassella neesiana*, *Nassella trichotoma*) and seven native grass species. Our results indicated that, while no single locus can be universally used as a DNA barcode for distinguishing the grass species examined in this study, two plastid loci (*atpF* and *matK*) showed good distinguishing power to separate most of the taxa examined, and could be used as a dual locus to distinguish several of the invasive from the native species. Low PCR success rates were evidenced among two nuclear loci (ETS and ITS), and few species were amplified at these loci, however ETS was able to genetically distinguish the two important invasive *Nassella* species. Multiple loci analyses also suggested that ETS played a crucial role in allowing identification of the two *Nassella* species in the multiple loci combinations.

## OPEN ACCESS

**Citation:** Wang A, Gopurenko D, Wu H, Lepschi B (2017) Evaluation of six candidate DNA barcode loci for identification of five important invasive grasses in eastern Australia. PLoS ONE 12(4): e0175338. <https://doi.org/10.1371/journal.pone.0175338>

**Editor:** Shilin Chen, Chinese Academy of Medical Sciences and Peking Union Medical College, CHINA

**Received:** November 22, 2016

**Accepted:** March 24, 2017

**Published:** April 11, 2017

**Copyright:** © 2017 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All DNA sequences files are available from the GenBank database (accession number(s) KX281066 - KX281111; KX290952 - KX291002; KX360607 - KX360637; KX434529 - KX434573; KY079282 - KY079335).

**Funding:** This work was supported by NSW Weeds Action Programs (WAP).

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

A variety of invasive grasses in Eastern Australia, including *Nassella trichotoma* (Nees) Hack. ex Arechav. (Serrated tussock), *Nassella neesiana* (Trin. & Rupr.) Barkworth (Chilean Needle Grass), *Hyparrhenia hirta* (L.) Stapf (Coolatai grass), *Eragrostis curvula* (Schrad.) Nees (African lovegrass), *Chloris gayana* Kunth (Rhodes grass) and others have caused serious damage to the

environment and agricultural industry [1]. Both *N. trichotoma* and *N. neesiana* are native to South American countries (Argentina, Uruguay, Peru etc.), and have become naturalised weeds in Australia (particularly in eastern states such as New South Wales, Australian Capital Territory, Victoria and Tasmania)[2, 3]. They also occur as weeds in New Zealand, South Africa, USA and several European countries [4]. As these two weeds are capable of displacing palatable grasses from pastures, decreasing productivity of grazing livestock, and significantly degrading biodiversity in native grasslands, they have been listed as WONS (Weeds of National Significance in Australia)[4] requiring nationally coordinated control efforts. *E. curvula* is native to South Africa. As an invasive weed in Australia, New Zealand, and some parts of USA, it competes with native pasture species, reduces livestock productivity and poses fire hazard to environment [5]. Being a native in tropical and temperate Africa and the Mediterranean region, *H. hirta* has become naturalized in Australia, Mexico, the Caribbean and parts of South America [6]. It poses a major threat to natural biodiversity in stock routes, nature reserves and National Parks due to its tolerance to drought, fire and herbicide. Originated from Africa, *C. gayana* is a widely naturalised weed infesting pastures and native vegetation zones in Australia, New Zealand, USA, and several Pacific islands (Weeds of Australia, <https://keyserver.lucidcentral.org/weeds/>).

Identification of invasive and native grasses is crucial in weed management as misidentification of weeds may delay the control of invasive weeds while causing unnecessary eradication of native grasses with similar morphology[1]. Unfortunately, the five invasive grasses can be easily misidentified to a variety of native grasses in Australia. For examples, *N. trichotoma* and *N. neesiana* are both very similar in appearance to *Austrostipa* spp and in some cases similar to, *Poa* and *Eragrostis* species; *E. curvula* is alike to other tussock-like grasses; *H. hirta* is comparable to kangaroo grass (*Themeda australis*) and *C. gayana* is similar to feathertop Rhodes grass (*Chloris virgata*) (NSW WeedWise, <http://weeds.dpi.nsw.gov.au/>).

Weed abatement officers in eastern Australia and elsewhere generally identify grasses based mainly on visual examination of diagnostic floral characters present in mature specimens. As a result, grasses at earlier stages of development may not be readily identifiable, and in those instances, invasive grass infestations can go undetected during weed abatement surveys. DNA barcoding [7] provides a potential alternative method for distinguishing invasive grasses from native grasses without the dependency of suitable growth-stage samples. As a genetic based method, DNA barcoding can accurately and rapidly identify samples to species, even from trace amounts or degraded sample tissue [8, 9]. The success of DNA barcoding relies on the amplification of specific barcoding locus or loci in the genomes of the target species. The mitochondrial locus, *COI* (Cytochrome *c* Oxidase subunit I), has been successfully applied to provide species-level resolution in many animal species [10, 11]. In land plants, however, *COI* has proved to be less effective due to conservative rates of nucleotide substitution in land plant mitochondrial DNA [12].

The lack of universal barcoding locus in plants impelled the search for alternative DNA barcoding regions outside the mitochondrial genome [13–15]. As a result, the Consortium for the Barcode of Life (CBOL) recommended use of plastid coding genes maturase K (*matK*) and ribulose-1,5-bisphosphate carboxylase oxygenase (*rbcL*) as core plant barcodes[14] and the nuclear internal transcribed spacers 1–2 (ITS)[14] as supplementary loci. Nevertheless, finding more robust loci to increase the resolution in distinguishing particular groups of plant species (particularly higher plants) is still an ongoing process [16].

Syme, *et al.* (2013) [4] assessed the accuracy of standard barcoding loci (*rbcL*, *matK* and ITS) in a study aiming to test the efficiency of three sequence matching algorithms (BLAST, Neighbour Joining and Bayesian Likelihood) for genetic identification of stipoid grasses. They reported the best DNA barcode accuracy using ITS, partial accuracy using *matK*, and least

accuracy using *rbcL*. Wang, *et al.* (2014) [1] screened 18 loci for the possibility of using DNA barcoding technology to identify invasive weeds and native grasses (up to 29 grass species) collected from eastern Australia. Based on PCR reliability and polymorphism levels, they advocated the use of *matK* and two other cpDNA loci [*ndhK-ndhC* intergenic spacer (referred as *ndhK*) and *psbE-petL* intergenic spacer (referred to here as *psbE*)] as preferred grass DNA barcode targets over *rbcL* and ITS.

Here, we evaluated the chloroplast and nuclear loci (ITS, *matK*, *ndhK* and *psbE*), which were recommended by previous studies, on five major invasive grasses (*N. trichotoma*, *N. neesiana*, *E. curvula*, *H. hirtai*, *C. gayana*) and several co-occurring native grasses. We also tested two new loci, *atpF* intron (referred as *atpF*) and external transcribed spacer (ETS), which were reported to be effective in the genetic diversity study of *Lolium perenne* and other related grass species [17], and the phylogenetic studies of *Poa* [18]. We hope the availability of robust DNA barcoding loci will help weeds abatement officers to identify the five invasive weeds at early growth stages, which is important to Australian biosecurity (including border entry points for quarantine purposes) and weed control agencies as early detection in bio-surveillance for weed control will ensure an early intervention by control agencies.

## Materials and methods

### Sample collection

A total of 66 specimens representing five invasive grass species in Eastern Australia (*Chloris gayana*, *Eragrostis curvula*, *Hyparrhenia hirta*, *Nassella neesiana*, *Nassella trichotoma*) and seven native grass species (*Austrostipa densiflora*, *Anthosachne scabra*, *Microlaena stipoides*, *Poa sieberiana*, *Rytidosperma caespitosum*, *Rytidosperma pallidum* and *Themeda triandra*) were collected from the Australian Capital Territory, New South Wales, Queensland and Victoria in eastern Australia under the permit issued by Department of Primary Industry (Permit No. INT14/8307) (see [S1 Table](#) for collection information and GenBank accession numbers). The seven native grasses were selected for examination here, as each is morphologically similar to one or more of the invasive grasses and potentially affected by their presence in areas of Eastern Australia where they overlap. These samples included field sampled specimens (N = 62) and vouchered herbarium specimens from the Australian National Herbarium (N = 2) and the National Herbarium of Victoria (N = 2). Leaf samples (appropriately 0.3 cm<sup>2</sup> in size) were preserved in > 70% EtOH and stored at the Wagga Wagga Agricultural Institute (NSW Department of Primary Industries) and allocated unique specimen identifiers (e.g. ww00001) for sample tracking. All specimen records and associated gene sequences have been submitted into the Barcode of Life Data systems (BOLD) [19].

### DNA extraction, target loci PCR, and sequencing

In preparation for DNA extractions, leaf tissue (< 1 mg) of each specimen was incubated (55°C) overnight in 280 µl of DXT tissue digest reagent (QIAGEN, Doncaster, Australia) with 1% added proteinase K (Sigma—Aldrich). Genomic DNA was isolated from specimen digestions using a Corbett Research 1820 X-tractor Gene robot and associated DX buffers (Qiagen, Doncaster, Australia).

Loci specific forward and reverse primers used in Polymerase Chain Reaction (PCR) amplification are listed in ([Table 1](#)) and were modified by addition of 17 bp forward (GTAAAACGACGGCCAGT) and reverse (CAGGAAACAGCTATGAC) vector M-13 5' tails to simplify subsequent sequencing.

PCR (15 µl aliquots) for each locus (except ITS) contained Invitrogen™ reagents (1 × PCR buffer 2.9 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 0.4 U of Platinum Taq polymerase), 0.1 µM each

**Table 1. Loci targeted for PCR using sourced primers.** Forward (-F) and reverse (-R) primer directions indicated by suffix. Original primer sequences modified by addition of 17 bp vector M-13 5' tails (tail sequences not shown here, refer to [Materials and methods](#)).

| Locus              | Primer name  | Primer sequence 5'-3'       | Source |
|--------------------|--------------|-----------------------------|--------|
| <i>atpF</i> intron | TeaCpSSR27FP | AATGCCGAATCGACGACCTA        | [17]   |
|                    | TeaCpSSR27RP | CAATGGTCCCTCTACGCAAT        |        |
| <i>matK</i>        | 390-F        | CGATCTATTCATTCAATATTC       | [20]   |
|                    | 1326-R       | TCTAGCACACGAAAGTCGAAGT      |        |
|                    | 3F-Kimf      | CGTACAGTACTTTTGTGTTTACGAG   |        |
|                    | 1R-Kimr      | ACCCAGTCCATCTGGAAATCTTGGTTC |        |
| <i>ndhK-ndhC</i>   | TeaCpSSR29FP | GGTACCAATCCATAACGATC        | [17]   |
|                    | TeaCpSSR29RP | GCGCTAGTTTTTGTGTTTTT        |        |
| <i>psbE-petL</i>   | TeaCpSSR31FP | GGTCGTGGAATGCTTTTCTT        |        |
|                    | TeaCpSSR31RP | TCCACGAATCTCAATGACCA        |        |
| ETS                | Rets4-F      | TTGGCTACGCGAGCGCATGAG       | [21]   |
|                    | 18S-R        | AGACAAGCATATGACTACTGGCAGG   | [22]   |
| ITS                | 26SE         | TAGAATTCCCGGTTCGCTCGCCGTTAC | [23]   |
|                    | S3           | AACCTGCGGAAGGATCATTG        | [24]   |
|                    | ITS 5a-F     | TATCATTTAGAGGAAGGAG         | [25]   |
|                    | ITS 4-R      | GCATATCAATAAGCGGAGGA        |        |

<https://doi.org/10.1371/journal.pone.0175338.t001>

primer, and 2 µl of specimen DNA. PCR thermal cycling consisted of 2 min at 95°C, 40 x [94°C (30s), 50°C (30s), 72°C (1 min)] and 72°C (5 min). All PCRs were conducted using an Eppendorf Mastercycler EP Gradient S Thermal Cycler.

ITS PCR modifications [18] included the addition of 0.75 µl of DMSO (100%) (Sigma—Aldrich) and 1 µl of specimen DNA in 15 µl reactions, with thermal cycling set as 95°C (5 min), 35 x [94°C (30 s), 55°C (30 s), 72°C (1 min)], and 72°C (5 min).

PCR products stained with SYBR<sup>®</sup> safe DNA Gel Stain (Invitrogen) were examined under UV light in a Bio—Rad Universal Hood II following electrophoresis through a 1.5% agarose gel, and in the presence of 1 kb size markers and negative controls. Successful PCR products were sequenced at the Brisbane, Queensland node of the Australian Genome Research Facility (AGRF).

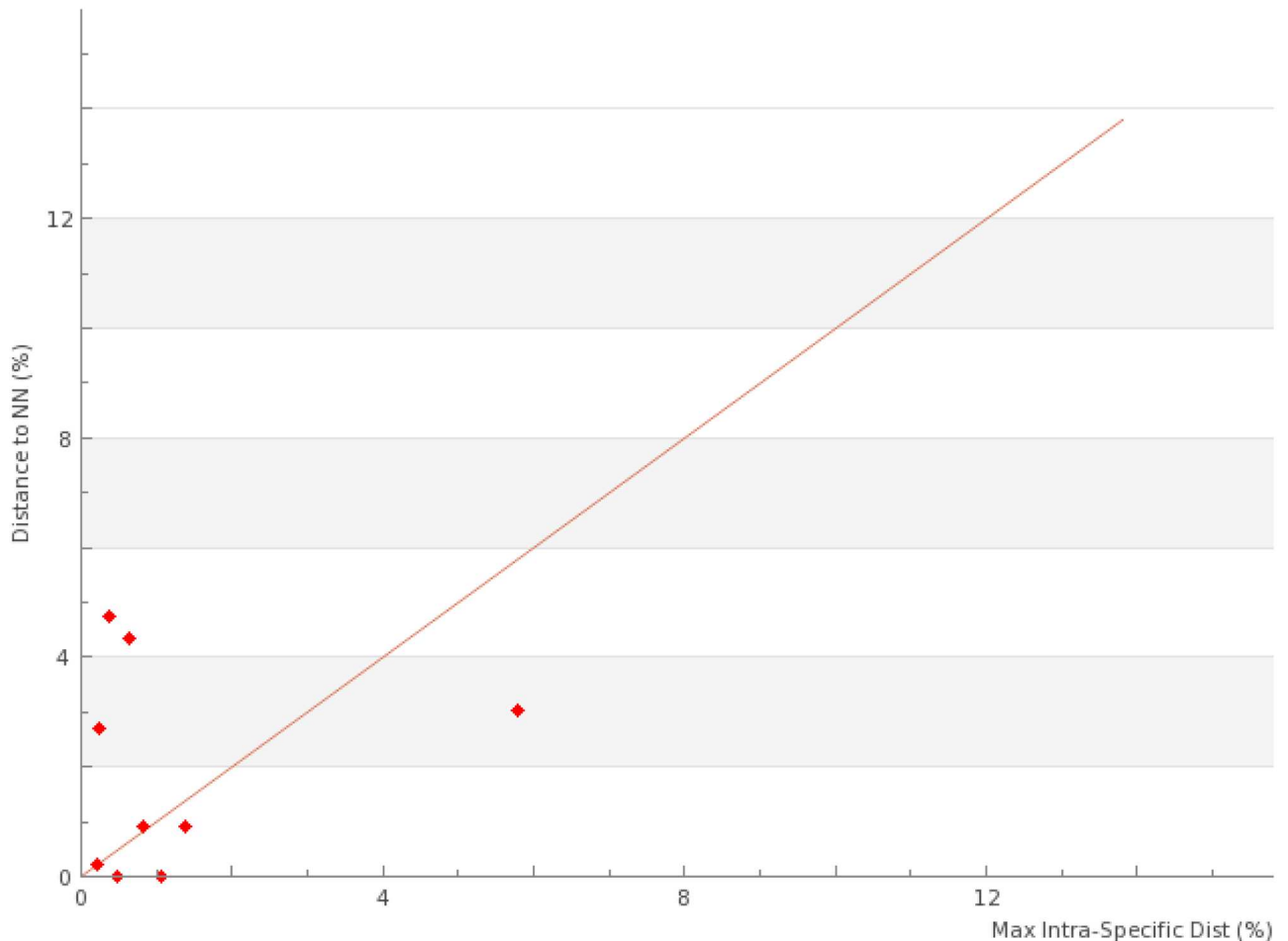
## Data analysis

Forward and reverse sequence chromatograms at each locus were assembled and checked for signal quality using SeqMan (DNA STAR package, DNASTar Inc., Madison, WI, USA). At each locus, specimen consensus sequences were exported into BioEdit [26] for alignment using ClustalW [27] with default parameters. The aligned sequences were also manually edited in BioEdit to remove primer reads.

DNA barcode gap analysis was determined at each locus by plotting maximum intraspecific distance ( $D_{intra}$ ) against minimum nearest neighbour distance ( $D_{NN}$ ). Intra and inter-specific pairwise genetic distances used in barcode gap analyses were generated at BOLD and adjusted by the Kimura 2-parameter (K2P) model of nucleotide evolution.

At each locus, the nearest neighbour minimum interspecific  $p$ -distances ( $D_{NN}$ ) was plotted against maximum intraspecific  $p$ -distances ( $D_{intra}$ ) to determine presence or absence of a DNA barcode gap among species (Fig 1).

Species monophyly was tested at each locus using both Neighbor-joining (NJ) and Maximum Likelihood (ML) tree construction methods. K2P pairwise distances used in NJ tree constructions were computed in MEGA 6.0 [28]. ML trees were constructed using PhyML 3.1



**Fig 1. Maximum percent intraspecific distance ( $D_{intra}$ ), and percent distance to nearest genetic neighbour species for six loci across five invasive and seven native grass species. A: *atpF*; B: *psbE*; C: *matK*; D: ETS; E: *ndhK*; F: ITS.**

<https://doi.org/10.1371/journal.pone.0175338.g001>

[29], incorporating a GTR nucleotide substitution model (plus Gamma distribution). Bootstrap replication (N = 1000) was used to assess confidence of NJ clusters and ML clades.

The same phylogenetic analyses were also performed on multi locus combinations (di, tri, tetra and penta) of the six loci to determine if any locus combination outperformed single loci for resolving species monophyly.

## Results

### Sequence characteristics of the six loci

While only one primer set was tested for each of four loci (ETS, *atpF*, *ndhK* and *psbE*), two primer sets were applied to amplify the ITS and *matK* sequences across the 12 species (Table 1). PCR success rate varied across different loci and grass species (Table 2). *atpF* was the only locus that successfully amplified all tested grass species. High rates of PCR success were also apparent at *matK* and *ndhK*. ITS had the least PCR success, with only three of the twelve grass species (*C. gayana*, *E. curvula* and *N. trichotoma*) being successfully amplified at this locus (Table 2). The aligned sequence matrix of ETS was 491 nucleotides (nt) in length with 314 parsimony informative sites and 325 variable sites (Table 3), which represents the highest

**Table 2. PCR success rate of each locus against each species tested (Invasive grasses are indicated with \*).**

| Species (specimen No.)              | <i>atpF</i> | ETS  | ITS  | <i>matK</i> | <i>ndhK</i> | <i>psbE</i> |
|-------------------------------------|-------------|------|------|-------------|-------------|-------------|
| <i>Austrostipa densiflora</i> (2)   | 100%        | 100% | 0%   | 100%        | 100%        | 0%          |
| * <i>Chloris gayana</i> (3)         | 100%        | 67%  | 100% | 0%          | 33%         | 0%          |
| <i>Anthosachne scaber</i> (4)       | 75%         | 0%   | 0%   | 100%        | 25%         | 75%         |
| * <i>Eragrostis curvula</i> (18)    | 39%         | 0%   | 100% | 78%         | 89%         | 33%         |
| * <i>Hyparrhenia hirta</i> (6)      | 100%        | 33%  | 0%   | 33%         | 83%         | 83%         |
| <i>Microlaena stipoides</i> (3)     | 100%        | 0%   | 0%   | 100%        | 67%         | 0%          |
| * <i>Nassella neesiana</i> (12)     | 100%        | 92%  | 0%   | 83%         | 92%         | 92%         |
| * <i>Nassella trichotoma</i> (8)    | 100%        | 100% | 13%  | 75%         | 88%         | 75%         |
| <i>Poa sieberiana</i> (2)           | 100%        | 0%   | 0%   | 100%        | 0%          | 0%          |
| <i>Rytidosperma caespitosum</i> (2) | 100%        | 0%   | 0%   | 100%        | 0%          | 0%          |
| <i>Rytidosperma pallidum</i> (4)    | 100%        | 0%   | 0%   | 100%        | 0%          | 0%          |
| <i>Themeda triandra</i> (2)         | 100%        | 0%   | 0%   | 100%        | 0%          | 0%          |

<https://doi.org/10.1371/journal.pone.0175338.t002>

percentage value of the parsimony informative or variable nucleotide sites against the total length (66% and 68% respectively). The shortest alignment (356 nt) was at *ndhK*, which also had the least amount of parsimony informative and variable sites against the total length (11% and 15%) respectively. DNA barcode gap minimum interspecific p—distances [represented as nearest neighbor distance ( $D_{NN}$ )] and maximum intraspecific p—distances ( $D_{intra}$ ) among the tested species across the six loci were presented in Table 4 and in Fig 1.

ETS and ITS were the only loci to show evidence of a clear DNA barcode gap, as exemplified by the absence of overlap between  $D_{intra}$  and  $D_{NN}$ . In contrast these two measures overlapped at each of the remaining four cpDNA loci, indicating instances where more variation was present within particular species than between their nearest genetic neighbor species.

Among the six loci, ETS showed good distinguishing power to separate all targeted invasive grasses (except for *E. curvula* where it failed to amplify) as clear barcode gaps separating maximum intraspecific and minimum interspecific distances were identified between the invasive species (Fig 1D). While the remaining loci differed in their abilities to separate different weeds species, they shared the same feature that they failed to distinguish two important invasive grasses *N. neesiana* and *N. trichotoma* (Fig 1A, 1B, 1C, 1E and 1F).

### Monophyly tests of species based on phylogenetic trees

For *psbE*, the monophyly of two invasive grasses, *H. hirta* and *E. curvula*, were well supported (95% and 99% bootstrap support) by the NJ analysis. Although the genus *Nassella* was supported as monophyletic, species within the genus (*N. neesiana* and *N. trichotoma*) were paraphyletic. All four invasive grasses which amplified at this gene were clearly distinguished from a single native species (*A. scabra*) (Fig 2B). Similar results were obtained from the ML tree of *psbE* (Fig 3B).

**Table 3. Information used to evaluate the utility of the six loci.**

| Items                            | ETS     | ITS     | <i>atpF</i> | <i>matK</i> | <i>ndhK</i> | <i>psbE</i> |
|----------------------------------|---------|---------|-------------|-------------|-------------|-------------|
| No. of primers screened          | 1       | 2       | 1           | 2           | 1           | 1           |
| Aligned length (bp)              | 491     | 687     | 457         | 838         | 356         | 687         |
| Informative sites/variable sites | 314/325 | 140/243 | 95/114      | 168/195     | 38/54       | 396/401     |
| No. of Indels                    | 70      | 48      | 82          | 6           | 10          | 74          |

<https://doi.org/10.1371/journal.pone.0175338.t003>

**Table 4. Percent distance of maximum percent intraspecific distance ( $D_{intra}$ ) and nearest genetic neighbour species ( $D_{NN}$ ) across five invasive and seven native grass species (N/A: not available for single specimen).**

| Locus                   | Species                         | $D_{intra}$           | Nearest species                 | $D_{NN}$ |
|-------------------------|---------------------------------|-----------------------|---------------------------------|----------|
| <b>matK</b>             | <i>Anthosachne scabra</i>       | 0.9                   | <i>Rytidosperma pallidum</i>    | 0.38     |
|                         | <i>Austrostipa densiflora</i>   | 0.12                  | <i>Nassella trichotoma</i>      | 1.43     |
|                         | <i>Eragrostis curvula</i>       | 1.28                  | <i>Hyparrhenia hirta</i>        | 4.49     |
|                         | <i>Hyparrhenia hirta</i>        | 0.56                  | <i>Eragrostis curvula</i>       | 4.49     |
|                         | <i>Microlaena stipoides</i>     | 1.28                  | <i>Austrostipa densiflora</i>   | 6.21     |
|                         | <i>Nassella neesiana</i>        | 0.66                  | <i>Nassella trichotoma</i>      | 0.11     |
|                         | <i>Nassella trichotoma</i>      | 0.97                  | <i>Nassella neesiana</i>        | 0.11     |
|                         | <i>Poa sieberiana</i>           | 0.45                  | <i>Anthosachne scabra</i>       | 4.84     |
|                         | <i>Rytidosperma caespitosum</i> | 9.71                  | <i>Rytidosperma pallidum</i>    | 0        |
|                         | <i>Rytidosperma pallidum</i>    | 9.45                  | <i>Rytidosperma caespitosum</i> | 0        |
|                         | <i>Themeda triandra</i>         | 6.47                  | <i>Rytidosperma pallidum</i>    | 0        |
| <b>ETS</b>              | <i>Austrostipa densiflora</i>   | 0.77                  | <i>Nassella neesiana</i>        | 7.09     |
|                         | <i>Chloris gayana</i>           | 0                     | <i>Hyparrhenia hirta</i>        | 27.31    |
|                         | <i>Hyparrhenia hirta</i>        | 0.7                   | <i>Chloris gayana</i>           | 27.31    |
|                         | <i>Nassella neesiana</i>        | 0.21                  | <i>Nassella trichotoma</i>      | 2.92     |
|                         | <i>Nassella trichotoma</i>      | 0.21                  | <i>Nassella neesiana</i>        | 2.92     |
| <b>ITS</b>              | <i>Chloris gayana</i>           | 0.87                  | <i>Eragrostis curvula</i>       | 16.74    |
|                         | <i>Eragrostis curvula</i>       | 5.21                  | <i>Chloris gayana</i>           | 16.74    |
|                         | <i>Nassella trichotoma</i>      | N/A                   | <i>Eragrostis curvula</i>       | 20.7     |
| <b>atpF</b>             | <i>Anthosachne scabra</i>       | 0.25                  | <i>Austrostipa densiflora</i>   | 2.7      |
|                         | <i>Austrostipa densiflora</i>   | 0.2                   | <i>Nassella trichotoma</i>      | 0.22     |
|                         | <i>Chloris gayana</i>           | 0                     | <i>Eragrostis curvula</i>       | 3.03     |
|                         | <i>Eragrostis curvula</i>       | 5.8                   | <i>Chloris gayana</i>           | 3.03     |
|                         | <i>Hyparrhenia hirta</i>        | 0.38                  | <i>Eragrostis curvula</i>       | 4.73     |
|                         | <i>Microlaena stipoides</i>     | 0.63                  | <i>Nassella neesiana</i>        | 4.34     |
|                         | <i>Nassella neesiana</i>        | 1.05                  | <i>Nassella trichotoma</i>      | 0        |
|                         | <i>Nassella trichotoma</i>      | 0.48                  | <i>Nassella neesiana</i>        | 0        |
|                         | <i>Poa sieberiana</i>           | 0                     | <i>Themeda triandra</i>         | 2.45     |
|                         | <i>Rytidosperma caespitosum</i> | 0.83                  | <i>Rytidosperma pallidum</i>    | 0.9      |
|                         | <i>Rytidosperma pallidum</i>    | 1.38                  | <i>Rytidosperma caespitosum</i> | 0.9      |
| <i>Themeda triandra</i> | 0                               | <i>Poa sieberiana</i> | 2.45                            |          |
| <b>ndhK</b>             | <i>Anthosachne scabra</i>       | N/A                   | <i>Austrostipa densiflora</i>   | 2.13     |
|                         | <i>Austrostipa densiflora</i>   | 0                     | <i>Nassella neesiana</i>        | 0.26     |
|                         | <i>Chloris gayana</i>           | N/A                   | <i>Eragrostis curvula</i>       | 2.53     |
|                         | <i>Eragrostis curvula</i>       | 2.79                  | <i>Chloris gayana</i>           | 2.53     |
|                         | <i>Hyparrhenia hirta</i>        | 0.75                  | <i>Eragrostis curvula</i>       | 3.2      |
|                         | <i>Microlaena stipoides</i>     | 0.82                  | <i>Eragrostis curvula</i>       | 3.37     |
|                         | <i>Nassella neesiana</i>        | 2.83                  | <i>Nassella trichotoma</i>      | 0        |
|                         | <i>Nassella trichotoma</i>      | 4.85                  | <i>Nassella neesiana</i>        | 0        |
| <b>psbE</b>             | <i>Anthosachne scabra</i>       | 0.47                  | <i>Nassella neesiana</i>        | 42.41    |
|                         | <i>Eragrostis curvula</i>       | 0.63                  | <i>Hyparrhenia hirta</i>        | 5.16     |
|                         | <i>Hyparrhenia hirta</i>        | 1.16                  | <i>Eragrostis curvula</i>       | 5.16     |
|                         | <i>Nassella neesiana</i>        | 2.07                  | <i>Nassella trichotoma</i>      | 0        |
|                         | <i>Nassella trichotoma</i>      | 0.29                  | <i>Nassella neesiana</i>        | 0        |

<https://doi.org/10.1371/journal.pone.0175338.t004>

The higher PCR success rates in *atpF* (Table 2) made it possible to examine the distinguishing power of this locus across a wider range of grass species. Both NJ and ML trees of *atpF* supported the monophyly of three invasive species (*C. gayana*, *E. curvula*, *H. hirta*) and four native species (*M. stipoides*, *A. scabra*, *T. triandra*, *P. Sieberiana*), which means the majority of the examined taxa was distinguished by this locus. Although *atpF* failed to separate two invasive *Nassella* species (*N. neesiana* and *N. trichotoma*), it confirmed the monophyly of the *Nassella* genus and separated this genus from the native stipa species *A. densiflora* (Figs 2A and 3A). Similar results were obtained from the *ndhK* data (Figs 2E and 3E) but species coverage dropped slightly in the NJ and ML trees of *ndhK* (*C. gayana* and *A. scabra* were each represented by a single specimen, and *T. triandra* and *P. sieberiana* were not represented at all).

Similar to *atpF*, *matK* distinguished the invasive grasses of *H. hirta* and *E. curvula* from the tested native grasses (*A. scabra*, *P. sieberiana*, *T. triandra*, *M. stipoides*, *R. pallidum* and *R. caespitosum*). The monophyly of these native grasses were also confirmed except for the *T. triandra*, *R. pallidum* and *R. caespitosum* (Figs 2C and 3C). In addition, both NJ and ML trees of *matK* provided weak support to the monophyly of *N. trichotoma* and *N. neesiana*, and strong support for the separation of native stipa species (*A. densiflora*) from the two invasive *Nassella* species.

Two nuclear loci, ITS and ETS, were tested for their potential in distinguishing invasive grasses from native grasses in present study. While the PCR success rate for these two loci were relatively lower than that of the plasmid loci, the available sequence data from these two loci provided good distinguishing power in identifying different grass species. The NJ and ML trees of ITS clearly supported the monophyly of *E. curvula* and *C. gayana* whilst its counterparts of ETS provided strong support for the monophyly of *N. neesiana*, *N. trichotoma*, *C. gayana* and *H. hirta* (100% in both NJ and ML) (Figs 2D, 2F, 3D and 3F). The NJ and ML trees of ETS also separated the native stipa species (*A. densiflora*) from the invasive *Nassella* species with strong bootstrap support (100%).

## Multiple loci analysis

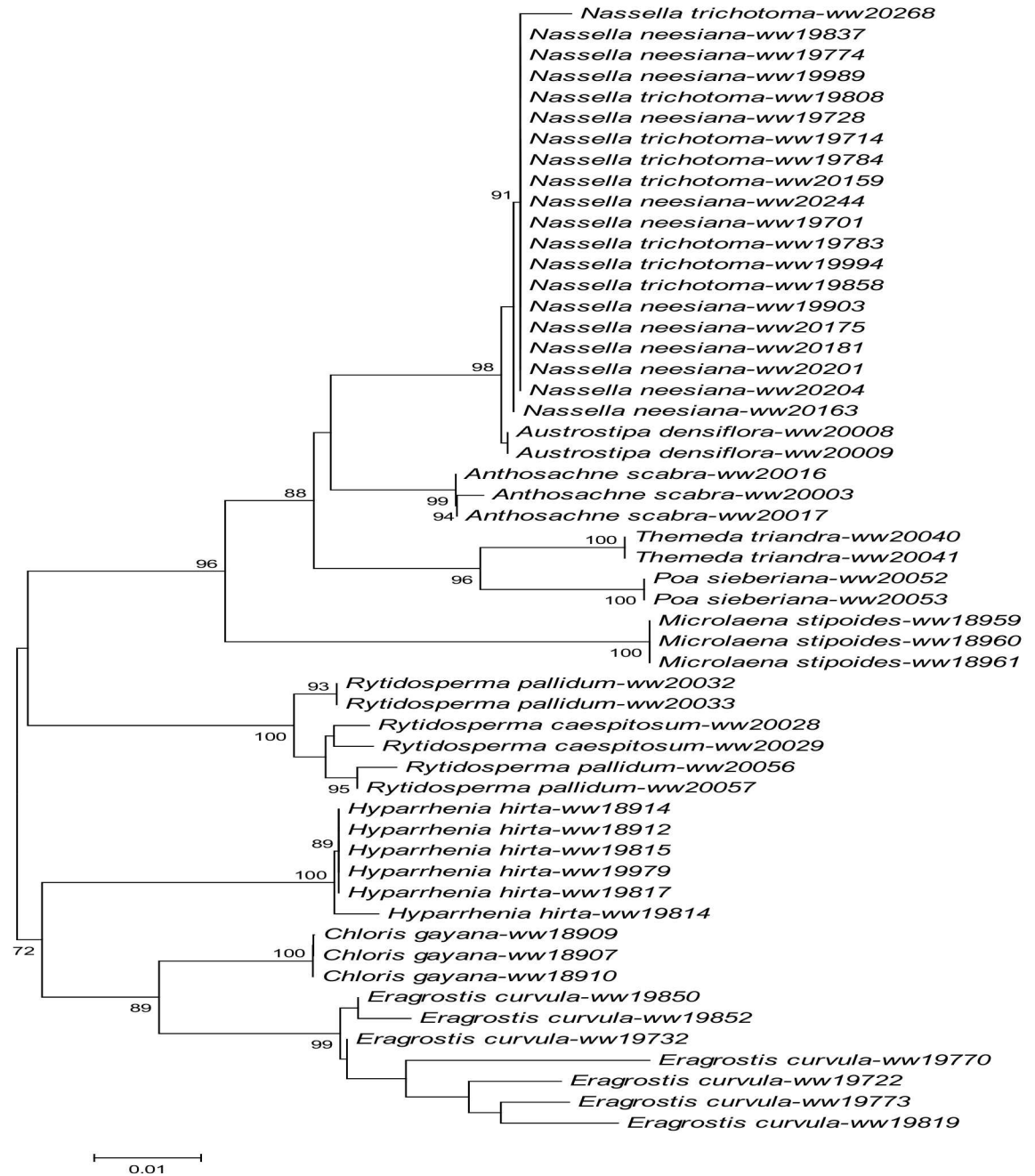
Four ETS related two loci combinations (ETS—*matK*, ETS—*ndhK*, ETS—*psbE* and ETS—*atpF*) strongly confirmed the monophyly of *N. neesiana* and *N. trichotoma*, which is an improvement relative to the results of single locus (*matK*, *ndhK*, *psbE* and *atpF*). On the contrary, *matK* related two loci combinations (except for *matK*—*psbE*) remain weak in distinguishing *N. neesiana* from *N. trichotoma*, although some combinations (*matK*—*atpF*, *matK*—*ndhK*, *matK*—*psbE*) were effective in confirming the monophyly of *E. curvula*, *H. hirta* and several native grass species. Monophyly of *E. curvula* was confirmed by all ITS related two loci combinations except for ETS—ITS (Table 5) (Trees not show).

Similar results were obtained from other combinations (tri, tetra and penta) of the six loci. Two *Nassella* invasive grasses, *N. neesiana* and *N. trichotoma*, were clearly distinguished by loci combinations consisting the ETS locus (three tri combinations: ETS—*atpF*—*ndhK*, ETS—*matK*—*atpF*, and ETS—*ndhK*—*psbE*; two tetra combinations: ETS—*matK*—*atpF*—*ndhK* and ETS—*atpF*—*ndhK*—*psbE*; and one penta combination: ETS—*matK*—*atpF*—*ndhK*—*psbE*). In addition, multiple loci combinations consisting of ITS failed to separate all grasses species except for *E. curvula* (Table 5) (Trees not show).

## Discussion

In present study, we tested six loci for their utility as DNA barcode targets to distinguish five invasive grass species from seven native grasses which frequently co-occur in eastern Australia. Among these, *matK* is recommended as one of two core loci by CBOL [14] for plant DNA



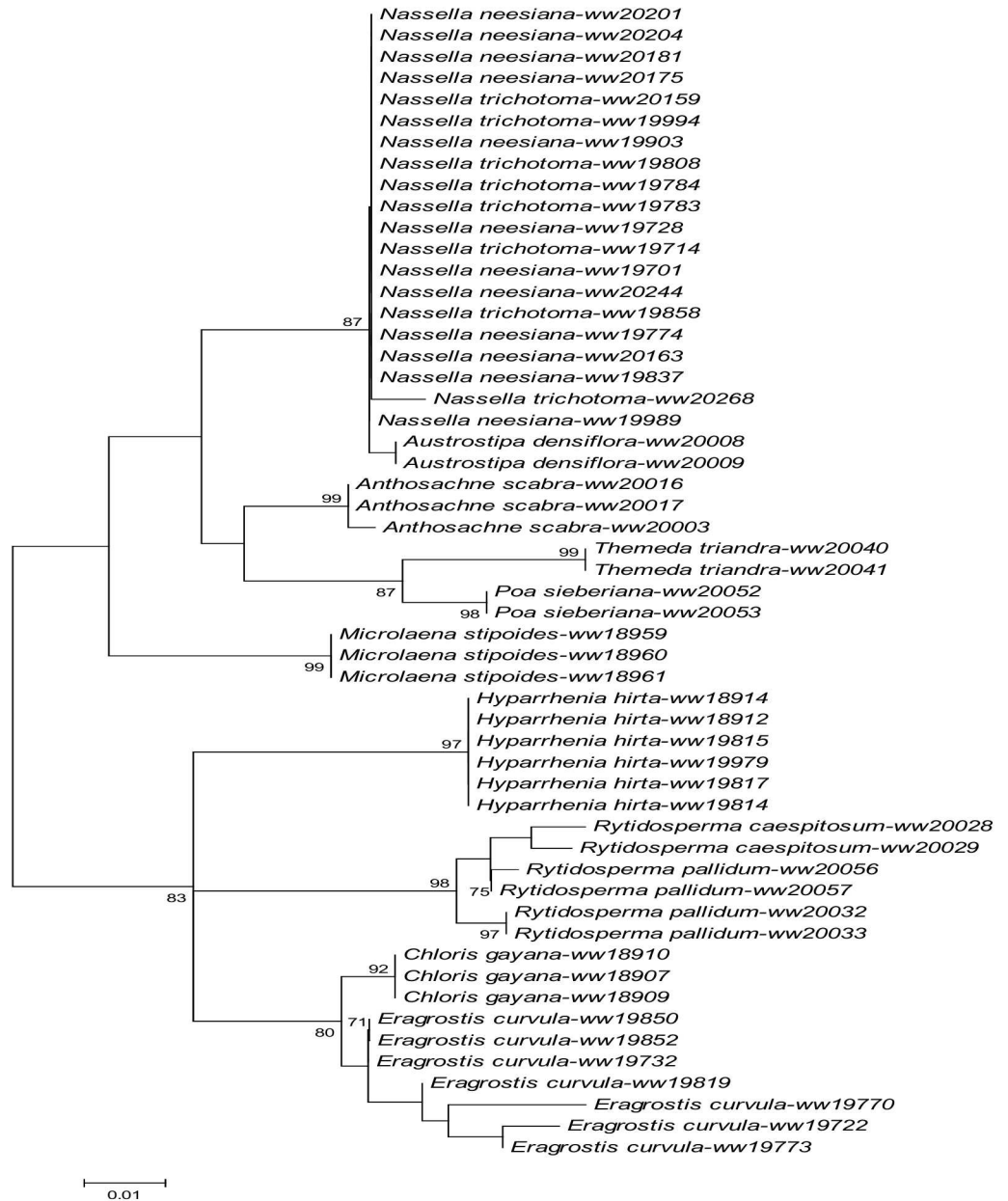


**Fig 2.** NJ trees inferred from six loci. A: *atpF*, B: *psbE*, C: *matK*, D: ETS, E: *ndhK*, F: ITS. Bootstrap support > 70% (N = 1000 replications) for clusters as reported. Scale bars indicate proportion of differences under a K2P model.

<https://doi.org/10.1371/journal.pone.0175338.g002>

barcoding. Our results (DNA barcode gap analysis, NJ and ML phylogenetic analyses) indicated that *matK* was suitable for distinguishing invasive *H. hirta* and *E. curvula* from native grasses (*M. stipoides*, *A. scabra*, *T. triandra*, *P. sieberiana*), but provided no or weak support for distinguishing *N. neesiana* from *N. trichotoma*, which are two important invasive grasses.

Similar results were obtained from *ndhK*, but the length of this locus (356 bp) is shorter than the recommended DNA barcode length [30]. The relatively lower PCR success rate of



**Fig 3. ML trees inferred from six loci. A: *atpF*, B: *psbE*, C: *matK*, D: ETS, E: *ndhK*, F: ITS. Bootstrap supports > 70% (N = 1000 replications) for clades as reported. Scale bars report proportion of differences under a GTR model.**

<https://doi.org/10.1371/journal.pone.0175338.g003>

*ndhK* also limits its use as a general DNA barcode locus across the surveyed species. In contrast highest PCR success was achieved at *atpF*, and most of the examined taxa were distinguished (except for the two *Nassella* species) by this locus, indicating its potentials as a promising DNA barcode locus for the grasses of concern.

The remaining chloroplast locus, *psbE*, had longer sequence length (687 bp) and a high proportion (57.6%) of informative sites, but low PCR success rates across the examined species, which limits its utility as a general DNA barcode for grasses. As evidenced at other plastid loci, *psbE* provided no resolution in distinguishing *N. neesiana* from *N. trichotoma*.

**Table 5. Overview of the multiple loci combinations analyses.**

| Combination of loci  | Shared Specimens Number | Confirmed monophyletic species   | Number of loci |
|--|-------------------------|--|----------------|
| ETS— <i>matK</i>   | 17                      | <i>N. trichotoma</i> , <i>N. neesiana</i> , <i>A. densiflora</i>   | Di             |
| ETS— <i>ndhK</i>   | 21                      | <i>N. trichotoma</i> , <i>N. neesiana</i> , <i>H. hirta</i> , <i>A. densiflora</i>   | Di             |
| ETS— <i>psbE</i>   | 17                      | <i>N. trichotoma</i> , <i>N. neesiana</i> , <i>H. hirta</i>  | Di             |
| ETS— <i>atpF</i>   | 25                      | <i>N. trichotoma</i> , <i>N. neesiana</i> , <i>H. hirta</i> , <i>A. densiflora</i> , <i>C. gayana</i>                      | Di             |
| ITS— <i>matK</i>   | 15                      | <i>E. curvula</i> , <i>N. trichotoma</i>   | Di             |
| ITS— <i>ndhK</i>   | 18                      | <i>E. curvula</i> , <i>N. trichotoma</i> , <i>C. gayana</i>  | Di             |
| ITS— <i>psbE</i>   | 7                       | <i>E. curvula</i> , <i>N. trichotoma</i>   | Di             |
| ITS— <i>atpF</i>   | 11                      | <i>E. curvula</i> , <i>C. gayana</i> , <i>N. trichotoma</i>  | Di             |
| <i>matK</i> — <i>ndhK</i>                                  | 35                      | <i>E. curvula</i> , <i>H. hirta</i> , <i>M. stipoides</i> , <i>A. densiflora</i>   | Di             |
| <i>matK</i> — <i>psbE</i>                                  | 26                      | <i>N. neesiana</i> , <i>N. trichotoma</i> , <i>H. hirta</i> , <i>E. curvula</i> , <i>A. scabra</i>                         | Di             |
| <i>matK</i> — <i>atpF</i>                                  | 34                      | <i>E. curvula</i> , <i>H. hirta</i> , <i>A. densiflora</i> , <i>A. scabra</i> , <i>P. sieberiana</i> , <i>M. stipoides</i> | Di             |
| <i>ndhK</i> — <i>psbE</i>                                  | 27                      | <i>E. curvula</i> , <i>H. hirta</i>  | Di             |
| <i>ndhK</i> — <i>atpF</i>                                  | 35                      | <i>E. curvula</i> , <i>H. hirta</i> , <i>M. stipoides</i> , <i>C. gayana</i> , <i>A. scabra</i>                            | Di             |
| <i>psbE</i> — <i>atpF</i>                                  | 26                      | <i>H. hirta</i> , <i>E. scabe</i> , <i>E. curvula</i>  | Di             |
| <i>atpF</i> — <i>ndhK</i> — <i>psbE</i>                    | 22                      | <i>H. hirta</i>  | Tri            |
| ETS— <i>atpF</i> — <i>ndhK</i>                             | 21                      | <i>N. trichotoma</i> , <i>N. neesiana</i> , <i>A. densiflora</i> , <i>H. hirta</i>   | Tri            |
| ETS— <i>matK</i> — <i>atpF</i>                             | 17                      | <i>N. trichotoma</i> , <i>N. neesiana</i> , <i>A. densiflora</i>   | Tri            |
| ETS— <i>ndhK</i> — <i>psbE</i>                             | 16                      | <i>N. trichotoma</i> , <i>N. neesiana</i>  | Tri            |
| <i>matK</i> — <i>atpF</i> — <i>ndhK</i>                    | 24                      | <i>E. curvula</i> , <i>M. stipoides</i> , <i>A. densiflora</i>   | Tri            |
| <i>matK</i> — <i>ndhK</i> — <i>psbE</i>                    | 22                      | <i>E. curvula</i>  | Tri            |
| <i>ndhK</i> —ITS— <i>matK</i>                              | 14                      | <i>E. curvula</i>  | Tri            |
| <i>psbE</i> — <i>matK</i> — <i>atpF</i>                    | 20                      | <i>A. scabra</i> , <i>H. hirta</i>   | Tri            |
| ETS— <i>matK</i> — <i>atpF</i> — <i>ndhK</i>               | 17                      | <i>N. trichotoma</i> , <i>N. neesiana</i> , <i>A. densiflora</i>   | Tetra          |
| ETS— <i>atpF</i> — <i>ndhK</i> — <i>psbE</i>               | 16                      | <i>N. trichotoma</i> , <i>N. neesiana</i>  | Tetra          |
| ETS— <i>matK</i> — <i>atpF</i> — <i>ndhK</i> — <i>psbE</i> | 14                      | <i>N. trichotoma</i> , <i>N. neesiana</i>  | Penda          |

<https://doi.org/10.1371/journal.pone.0175338.t005>

Among the nuclear loci, ITS has been frequently reported as a potentially useful locus for plant DNA barcoding [31, 32], including its use for identification of stipoid grasses [4]. However, we experienced difficulties in amplifying this locus across multiple grass species. Our previous study [1] reported 75% fungal contamination rates among sequenced PCR products when using non-specific and universal ITS primers (ITS 5aF—ITS4R). Similarly fungal contamination with ITS primers has also been reported by Hollingsworth *et al.* [33]. However, in the present study, we have successfully eliminated the fungal contamination using the new primer set of ITS 26SE—ITS S3 together with modified PCR cycling conditions described previously [18]. Nevertheless, PCR and sequencing success rates using this new primer set remained low (average 20.5%), and the limited number of retrieved sequences failed to distinguish between *N. neesiana* from *N. trichotoma*, despite its success in distinguishing *E. curvula* and *C. gayana*.

The other nuclear locus, ETS, outperformed ITS in many ways, including the relatively higher PCR and sequencing success rate, and the power to distinguish three invasive grass species (*N. neesiana*, *N. trichotoma* and *H. hirta*). This locus could be a promising marker for grass DNA barcoding if more robust primers are designed for this locus to increase its PCR success rate (particularly for *E. curvula*).

Results of our PCR screening and DNA barcode gap analyses indicated no single locus can be universally used as a DNA barcode for distinguishing the grass species examined here. Loci

examined either failed to amplify a portion of species or to resolve genetic limits among the species. Greater species resolution was in some cases obtained when combined loci analyses were employed. For example, *matK* confirmed the monophyly of multiple grass species except for *N. neesiana* and *N. trichotoma*, whilst ETS confirmed the monophyly of *N. neesiana* and *N. trichotoma* but failed to amplify *E. curvula*. When the DNA sequences of these two loci were concatenated and jointly analyzed, species monophyly of *N. trichotoma*, *N. neesiana* and *A. densiflora* were confirmed (Table 5). We noticed that the monophyly of *N. neesiana* and *N. trichotoma* were confirmed by all ETS related loci combinations (except for those combinations consisting both ETS and ITS), but were not confirmed by other loci combinations without the component of ETS. This suggests that ETS plays a crucial role in allowing identification of the two *Nassella* species in the multiple loci combinations, and could be a useful 2<sup>nd</sup> locus in combined analyses to improve accuracy of invasive species identifications.

In summary, the present studies evaluated the distinguishing power of six loci (ETS, ITS, *atpF*, *matK*, *ndhK* and *psbE*) for DNA barcoding of five invasive weeds and seven native grasses, which co-occur in eastern Australia. Among the four plastid loci, *atpF* and *matK* showed higher PCR rates and better distinguishing power than the remaining loci, making them suitable for further consideration as promising DNA barcodes of the targeted grass species. Among the two nuclear loci, ETS showed better potential as a DNA barcode for the separation of two invasive *Nassella* species. We conclude that a dual locus DNA barcode combination of *atpF* and *matK* may be used to genetically distinguish several prominent invasive grass species present in eastern Australia from co-occurring native grasses often mistaken for the invasive types. Furthermore, use of the ETS locus as a DNA barcode for genetic separation and identification of the two *Nassella* spp. may provide some application in future screening of those two WONS species; arguably further optimization of this locus may also allow it to be used for DNA barcode assay and identification of a broader assemblage of native and invasive grass species in Australia.

## Supporting information

**S1 Table. Plant material, collection details and GenBank accession numbers of material used for comparative evaluation of six candidate DNA barcoding regions in five Invasive grasses and seven native grasses.** All specimens are housed at WWAI unless otherwise indicated; N/A—indicates no sequence was obtained.  
(DOCX)

## Acknowledgments

This study was funded by the NSW Weeds Action Program (WAP). Mr Adam Shephard, Mr Peter Deane, Mr Malcolm Ross, Mr Jeremy Crocker and other Council Weeds Officers provided assistances in field sampling. We thank Dr Pina Milne and Mr Wayne Gebert from the National Herbarium of Victoria for providing herbarium samples (MEL 2375629A and MEL 2330503A). Native grasses were provided by Mr Ian Chivers (Native Seeds Nursery in Melbourne) and Dr Peter Orchard (Charles Sturt University).

## Author Contributions

**Conceptualization:** AW DG HW BL.

**Data curation:** AW.

**Formal analysis:** AW DG.

**Funding acquisition:** AW DG HW.

**Investigation:** AW.

**Methodology:** AW DG HW.

**Project administration:** HW.

**Resources:** AW DG HW.

**Software:** AW.

**Supervision:** AW DG HW BL.

**Validation:** AW DG HW BL.

**Visualization:** AW DG HW BL.

**Writing – original draft:** AW.

**Writing – review & editing:** AW DG HW BL.

## References

1. Wang AS, Gopurenko D, Wu H, Stanton R, Lepschi B. DNA barcoding for identification of exotic grass species present in eastern Australia. In: Proceedings of the 19th Australasian Weeds Conference—Science, Community and Food Security: the Weed Challenge, 1–4 September, Hobart, Tasmania, ed. M. Baker. (Tasmanian Weed Society, Hobart). 2014; pp 444–447.
2. Weller SL, Florentine SK, Sillitoe JF, Grech CJ, McLaren DA. An investigation of the effects of stage of ensilage on *Nassella neesiana* seeds, for reducing seed viability and injury to livestock. *Scientific reports*. 2016; 6:22345. <https://doi.org/10.1038/srep22345> PMID: 26925589
3. Ayres L, Leech F. Serrated tussock—identification and control, Primefact.2006; 44, NSW DPI, Orange
4. Syme AE, Udovicic F, Stajsic V, Murphy DJ. A test of sequence-matching algorithms for a DNA barcode database of invasive grasses. *DNA Barcodes*. 2013; 1: 19–26.
5. Campbell M. Effects of age on the germination of African lovegrass seed, *Grasslands Society Newsletter*, 2002; 17(1)
6. CRC for Australian Weed Management. Coolatai grass (*Hyparrhenia hirta*) Weed Management Guide. 2007;9 p. CSIRO & Bureau of Meteorology (2007) Climate change in Australia Technical Report, - <http://climatechangeinaustralia.com.au/>
7. Hebert PDN, Cywinska A, Ball SL, DeWaard JR. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B*.2003; 270: 313–321. <https://doi.org/10.1098/rspb.2002.2218> PMID: 12614582
8. deWaard JR, Mitchell A, Keena MA, Gopurenko D, Boykin LM, Armstrong KF, et al. Towards a Global Barcode Library for *Lymantria* (Lepidoptera: Lymantriinae) Tussock Moths of Biosecurity Concern. *PLoS One*. 2010; 5.
9. Mahadani P, Ghosh SK. DNA Barcoding: A tool for species identification from herbal juices. In. *DNA Barcodes*. 2013; 35–38 p.
10. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrartes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101: 14812–14817. <https://doi.org/10.1073/pnas.0406166101> PMID: 15465915
11. Smith MA, Woodley NE, Janzen DH, Hallwachs W, Hebert PDN. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103: 3657–3662. <https://doi.org/10.1073/pnas.0511318103> PMID: 16505365
12. Wolfe KH, Li WH, Sharp PM. Rates of Nucleotide Substitution Vary Greatly among Plant Mitochondrial, Chloroplast, and Nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America*. 1987; 84:9054–9058. PMID: 3480529
13. Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, Petersen G, et al. A proposal for a standardised protocol to barcode all land plants. *Taxon*. 2007; 56: 295–299.

14. CBOL Plant Working Group. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106: 12794–12797. <https://doi.org/10.1073/pnas.0905845106> PMID: 19666622
15. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102: 8369–8374. <https://doi.org/10.1073/pnas.0503123102> PMID: 15928076
16. Dong WP, Xu C, Li CH, Sun JH, Zuo YJ, Shi S, et al. *ycf1*, the most promising plastid DNA barcode of land plants. *Scientific Reports*. 2015; 5.
17. Diekmann K, Hodkinson TR, Barth S. New chloroplast microsatellite markers suitable for assessing genetic diversity of *Lolium perenne* and other related grass species. *Annals of Botany*. 2012; 110: 1327–1339. <https://doi.org/10.1093/aob/mcs044> PMID: 22419761
18. Birch JL, Cantrill DJ, Walsh NG, Murphy DJ. Phylogenetic investigation and divergence dating of *Poa* (Poaceae, tribe Poeae) in the Australasian region. *Biological Journal of the Linnean Society*. 2014; 175: 523–552.
19. Ratnasingham S, Hebert PD. Bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes*. 2007; 7: 355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x> PMID: 18784790
20. Ford CS, Ayres KL, Toomey N, Haider N, Stahl JV, Kelly LJ, et al. Selection of candidate coding DNA barcoding regions for use on land plants. *Biological Journal of the Linnean Society*. 2009; 159: 1–11.
21. Gillespie LJ, Soreng RJ, Jacobs SWL. Phylogenetic relationships of Australian *Poa* (Poaceae: Poinae), including molecular evidence for two new genera, *Saxipoa* and *Sylvipoa*. *Australian Systematic Botany*. 2009; 22: 413–436.
22. Starr JR, Harris SA, Simpson DA. Potential of the 5' and 3' ends of the intergenic spacer (IGS) of rDNA in the Cyperaceae: New sequences for lower-level phylogenies in sedges with an example from *Uncinia* pers. *International Journal of Plant Sciences*. 2003; 164: 213–227.
23. Sun Y, Skinner DZ, Liang GH, Hulbert SH. Phylogenetic Analysis of *Sorghum* and Related Taxa Using Internal Transcribed Spacers of Nuclear Ribosomal DNA. *Theoretical and Applied Genetics*. 1994; 89: 26–32. <https://doi.org/10.1007/BF00226978> PMID: 24177765
24. Kass E, Wink M. Molecular phylogeny and phylogeography of *Lupinus* (Leguminosae) inferred from nucleotide sequences of the *rbcl* gene and ITS 1+2 regions of rDNA. *Plant Systematics and Evolution*. 1997; 208: 139–167.
25. Baldwin BG. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the compositae. *Molecular Phylogenetics and Evolution*. 1992; 1: 3–16. PMID: 1342921
26. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*. 1999; 41: 95–98.
27. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007; 23: 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404> PMID: 17846036
28. Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*. 2013; 30: 2725–2729. <https://doi.org/10.1093/molbev/mst197> PMID: 24132122
29. Guindon S, Delsuc F, Dufayard JF, Gascuel O. Estimating maximum likelihood phylogenies with PhyML. *Methods in Molecular Biology*. 2009; 537: 113–137. [https://doi.org/10.1007/978-1-59745-251-9\\_6](https://doi.org/10.1007/978-1-59745-251-9_6) PMID: 19378142
30. Kress WJ, Erickson DL. DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105: 2761–2762. <https://doi.org/10.1073/pnas.0800476105> PMID: 18287050
31. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: the coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One*. 2007; 2: e508. <https://doi.org/10.1371/journal.pone.0000508> PMID: 17551588
32. Zhang JQ, Meng SY, Wen J, Rao GY. DNA barcoding of *rhodiola* (crassulaceae): a case study on a group of recently diversified medicinal plants from the qinghai-tibetan plateau. *PLoS One*. 2015; 10: e0119921. <https://doi.org/10.1371/journal.pone.0119921> PMID: 25774915
33. Hollingsworth PM, Graham SW, Little DP. Choosing and Using a Plant DNA Barcode. *PLoS One*. 2011; 6.