Northern Australia Environmental Resources Hub

National Environmental Science Programme

Development of an eDNA assay for fanwort (*Cabomba caroliniana*)

Report

by Richard C. Edmunds and Damien Burrows

TropWATER
Centre for Tropical Water and Aquatic Ecosystem Research

JAMES COOK UNIVERSITY
AUSTRALIA

This report should be cited as:

Edmunds, R.C. and Burrows, D. 2019. Development of an eDNA assay for fanwort (*Cabomba caroliniana*). Report 19/09, Centre for Tropical Water and Aquatic Ecosystem Research (TropWATER), James Cook University, Townsville.

Cover photographs

Front cover: *Cabomba caroliniana* (photo: Northern Territory Government).

Back cover: *Cabomba caroliniana* infestation (photo: Leslie J. Mehrhoff).

This report is available for download from the Northern Australia Environmental Resources (NAER) Hub website at nespnorthern.edu.au

The Hub is supported through funding from the Australian Government's National Environmental Science Program (NESP). The NESP NAER Hub is hosted by Charles Darwin University.

ISBN 978-1-925800-27-2

June, 2019

Printed by Uniprint

Contents

Acronyms....iv

Abbreviations.... v

Acknowledgements....vi

Executive summary.... 1

1. Introduction.... 2

2. Methodology.... 3

2.1 Primer design.... 3

2.2 *In silico* validation.... 4

2.3 *In vitro* validation.... 5

2.4 *In situ* validation.... 8

3. Results.... 10

3.1 Primer design.... 10

3.2 *In silico* validation.... 10

3.3 *In vitro* validation.... 11

3.4 *In situ* validation.... 15

4. Discussion.... 16

References.... 17

List of tables

Table 1. Non-target Australian aquatic plant species for which chloroplast maturase K (*matK*) nucleotide sequences were obtained from GenBank (NCBI) and used to guide C.caroliniana_matK assay development. 3

Table 2. Species against which C.caroliniana_matK was tested *in silico* using targeted PrimerBLAST search of Australian aquatic plant sequences available within NCBI "nr" database. 4

Table 3. Primer information for *C. caroliniana* eDNA assay (C.caroliniana_matK). Asterisk (*) and highhat (^) indicate melting temperature as determined by Geneious (ver. R11) and PrimerBLAST (Ye, et al., 2012), respectively. 5

Table 4. Non-target species against which C.caroliniana_matK was tested empirically. Nanograms of gDNA template loaded into duplicate wells of *in vitro* test provided within brackets (s: gDNA also extracted from seedpod). Non-native Australian species indicated by asterisks (*). All tissue samples collected from Darwin River, NT except for *Azolla pinnata* (^), which was collected from Howard River, NT. 8

Table 5. Species with ≤ 5 mismatches to C.caroliniana_matK based on targeted PrimerBLAST search of Australian aquatic plant sequences in NCBI "nr" databse. 10

Table 6. Species with ≤ 1 mismatch to C.caroliniana_matK based on non-targeted PrimerBLAST search of entire NCBI "nr" databse. 10

List of figures

Figure 1. Artificial double stranded DNA (aDNA) replica of the *C. caroliniana matK* gene region wherein C.caroliniana_matK targets. Total aDNA fragment length is 385 bp. Bolded blue nucleotides indicate location of forward (C.caroliniana_matK-F) and reverse (C.caroliniana_matK-R) primers (Table 3). Underlined region indicates 3' extensions beyond reverse primer binding site for C.caroliniana_matK-R, which was included to promote efficient primer binding to and amplification of aDNA fragment. 7

Figure 2. Amplification curves (A), linear regression of gDNA $\log_{10}$ standards (B), and amplicon dissociation temperature curves (T_m; C) generated by qPCR during C.caroliniana_matK *in vitro* Test 1. 12

Figure 3. Amplification curves (A), linear regression of aDNA $\log_{10}$ standards (B), and amplicon dissociation temperature curves (T_m; C) generated by qPCR during C.caroliniana_matK *in vitro* Test 2. 13

Figure 4. Amplification curves (A) and amplicon dissociation temperature curves (T_m; B) from C.caroliniana_matK *in vitro* Test 3 (non-target gDNA amplification). Vertical black line in panel B indicates combined average T_m of gDNA and aDNA standards (75.85 ± 0.67˚C). All produced amplicons exhibited ΔT_m within 99.7% CI of *C. caroliniana* gDNA and aDNA standards; however, the observed putative positive detections for non-target species (*Hygrophila angustifolia*, *Blyxa aubertii*, and *Pandanus aquaticus*) were

confirmed to be false positives by bidirectional Sanger sequencing (i.e., *C. caroliniana* amplification due to cross-contamination during plate loading). 14

Figure 5. *In situ* validation of C.caroliniana_matK using PPLPP method with OneStep PCR Inhibitor Removal Kit (Zymo Research, USA) purified eDNA collected from Ross River at Douglas Bridge. Amplification curve at common fluorescence threshold 0.2 (A) and dissociation curve (T_m) for all n = 48 amplicons (B). Note that, despite late amplification (C_t = 45.42, panel A), this amplicon exhibited ΔT_m within 99.7% CI of gDNA and aDNA standards (Figure 2 and Figure 3), respectively. All Sanger sequenced representative amplicons (n = 12) were positive for *C. caroliniana.* 15

Acronyms

BLAST..... Basic Local Alignment Search Tool

NESP..... National Environmental Science Program

Abbreviations

aDNA..... Artificial deoxyribonucleic acid

bp ..... Base pair

C_f ..... Final concentration

C_t ..... Threshold cycle

CI ..... Confidence interval

eDNA...... Environmental deoxyribonucleic acid

EDTA...... Ethylenediaminetetraacetic acid

gDNA...... Genomic deoxyribonucleic acid

IDT..... Integrated DNA Technologies

LOD ..... Limit of detection

MEEL..... Molecular Ecology and Evolution Laboratory

NaCl..... Sodium chloride

NCBI..... National Center for Biotechnology Information

qPCR..... Quantitative polyermerase chain reaction

SDS ..... Sodium dodecyl sulfate

TE. ..... Trisaminomethane and ethylenediaminetetraacetic acid

T_m..... Melting temperature

Acknowledgements

Development of *Cabomba caroliniana* eDNA assay (C.caroliniana_matK) was conducted as part of the Australian Government's National Environmental Science Program (NESP), Northern Australia Environmental Resources Hub Project 4.3: "The Northern Australia eDNA Program – Revolutionising Aquatic Monitoring and Field Surveys in Tropical Waters".

We thank Christopher Collins (Weed Management Branch, Department of Environment and Natural Resources, Northern Territory Government) for tissue samples of aquatic plants endemic to Darwin River, NT used for *in vitro* assay validation.

Executive summary

A new environmental DNA (eDNA) assay was developed for the detection of invasive aquatic plant *Cabomba caroliniana* (fanwort) in environmental water samples using Sybr-based quantitative real-time polymerase chain reaction (qPCR). This eDNA assay targets *C. caroliniana* chloroplast gene maturase K (*matK*). Here we present the *in silico*, *in vitro*, and *in situ* validations undertaken during development of *C. caroliniana* eDNA assay.

In silico, *in vitro* and, *in situ* validations confirmed that this eDNA assay is specific to (i.e., detects only) *C. caroliniana* despite co-occurrence with native species. Limit of detection (LOD) was determined to be 9 ± 3 copies per qPCR assay under optimal qPCR conditions (65˚C annealing temperature, 900 nM each primer). Moreover, bidirectional Sanger sequencing confirmed all representative putative positive detections from *in vitro* and *in situ* validations (gDNA standards and Ross River water samples) to be positive for target species detection while all putative negative detections from *in vitro* validation (non-target species amplifications) were confirmed negative (i.e., false positives), respectively. Collectively, these validations demonstrate the readiness of this assay for screening environmental water samples for determination of *C. caroliniana* eDNA presence.

In light of invasion fronts consisting of only a few individuals, waterways previously invaded and waterways suspected of potential invasion should be monitored regularly for *C. caroliniana* eDNA so as to monitor changes in abundance over time (i.e., eradication or invasion front monitoring). If detected, this sensitive eDNA assay can be subsequently utilized in conjunction with spatial water sampling to guide *C. caroliniana* management or eradication efforts.

1. Introduction

Cabomba caroliniana (Angiosperms, Nymphaeales, Cabombaceae, *Cabomba* aka Carolina fanwort, Carolina water shield, green cabomba, fanwort, fish grass, and Washington grass) is a fully submerged aquatic perennial herbaceous plant that is native to North America (east and west coasts) and South America (southern Brazil, Paraguay, Uruguay, and north-eastern Argentina) (*Cabomba caroliniana* Factsheet). In late summer, *C. caroliniana* stems become brittle and break apart, which facilitates distribution and invasion into new water bodies. More specifically, the main methods of invasive spreading are vegetative (e.g., seed broadcasting) and fragmentation (e.g., ≥ 1 cm floating stem fragments or leaves) as leaf nodes can take root and grow into new plants (Sanders, 1979; Schooler and Julien, 2006; Bickel, 2012); however, to date viable seed production has only been reported in Australia in the top End of the Northern Territory (Schooler, et al., 2009). Moreover, *C. caroliniana* is fast growing with 50 mm growth attainable in a single day (Lake Macdonald, Queensland Australia; Global Species Invasion Database). Establishment of *C. caroliniana* populations in stagnant to slow-flowing waterbodies (e.g., streams, smaller rivers, lakes, ponds, sloughs, and ditches) commences with roots growing into muddy substrate followed by stem growth to waterbody surface (*Cabomba caroliniana* Weed Management Guide). Of note is that *C. caroliniana* can survive up to eight weeks free-floating as long as it doesn't completely dry out (*Cabomba caroliniana* Weed Management Guide).

Cabomba caroliniana is grown commercially in Florida, USA for domestic distribution only whereas *C. caroliniana* is grown commercially in Asia and exported worldwide. As such, some global invasion introductions are suspected to have been mediated by aquarists. In Australia *C. caroliniana* is considered a "Weed of National Significance" given its invasion potential (i.e., ability to spread rapidly) and impacts to both native environments (e.g., native flora and fauna such as platypus and water rats) and commercial entities (e.g., water retention and purification). Moreover, *C. caroliniana* can reduce capacity and quality of water facilities if invaded (*Cabomba caroliniana* Factsheet; *Cabomba caroliniana* Weed Management Guide).

In order to accurately monitor Australian waterways for the invasion of this aggressive aquatic plant a sensitive method for *C. caroliniana* detection is needed. Environmental DNA (eDNA), or DNA shed by all living organisms into their local environment (Goldberg, et al., 2016), allows for *C. caroliniana* detection in any water source known or suspected to have been inhabited by *C. caroliniana* (intact or fragmented pieces). Here we describe the development of a sensitive Sybr-based quantitative real-time PCR (qPCR) assay that can be used to detect *C. caroliniana* eDNA present in field-collected water samples (unfiltered or filtered).

2. Methodology

2.1 Primer design

The presence of multiple mitochondria or chloroplast within each cell makes mitochondrial or chloroplast DNA more abundant and thus more detectable than nuclear DNA within environmental water samples (Goldberg, et al., 2016; Braukmann, et al., 2017; Kuzmina, et al., 2017). Moreover, mitochondrial *16S* and chloroplast *matK* genes are commonly targeted by barcoding studies and thus nucleotide sequence information from a broad range of species is available within the National Center for Biotechnology Information public database (GenBank; www.ncbi.nlm.nih.gov). Geneious analysis software (version R11; Kearse et al., 2012) was used to obtain all available *16S* or *matK* nucleotide sequences from Australian fish (*n* = 86), frog (*n* = 172), or aquatic plant (*n* = 71) species (Table 1) and align using ClustalW (Thompson, et al., 2003). Regions of conservation within tilapia, *R. marina*, and *C. caroliniana* sequences (*n* = 11, 27, and 13), respectively, were identified, annotated, and assessed by eye for regions wherein target species exhibited ≥ 1 base pair mismatch with non-target species (Table 1) and human.

GenBank (NCBI) was mined for available maturase K (*matK*) nucleotide sequences from Australian aquatic plant species (*n* = 39; Table 1). This chloroplast gene is commonly targeted by DNA barcoding studies and, as a result, has nucleotide sequences available for a wide range of plants. Moreover, *matK* has been targeted previously for the detection of eDNA from a variety of invasive aquatic plant (Scriver, et al., 2015). Targeting *matK*, or other chloroplast gene(s), also precludes Human DNA amplification.

All available *matK* sequences were downloaded from NCBI (GenBank accessions: AF543729, AH007211, DQ185527, LC219190, LC219198, LC219331, NC031505) into Geneious (ver. R11) and aligned using ClustalW algorithm (Kearse, et al., 2012). Regions of conservation were identified, annotated, and visually assessed for regions wherein *C. caroliniana* exhibited ≥ 3 base pair mismatches with other Australian aquatic plant *matK* sequences. Primers were assessed for quality and probability of accuracy and efficiency based on the following criteria: 1) melting temperature (T_m): 55–65˚C with < 4˚C difference between primer pair, 2) G/C content: 40–80%, 3) length: 16–25 base pairs (bp), 4) amplicon size: 80–350 bp, 5) self-dimer T_m: < 30˚C, 6) hairpin T_m: < 30˚C, 7) overall self-complementarity: PrimerBLAST score < 6, and 8) 3' self-complementarity: PrimerBLAST score < 6.

*Table 1. Non-target Australian aquatic plant species for which chloroplast maturase K (*matK*) nucleotide sequences were obtained from GenBank (NCBI) and used to guide C.caroliniana_matK assay development.*

Non-target Australian aquatic plant species	
Ceratophyllum demersum	*Najas tenuifolia*
Cyperus difformis	*Nelumbo nucifera*
Cyperus involucratus	*Nelumbo nucifera subsp. nucifera*
Cyperus polystachyos	*Nymphaea caerulea*
Cyperus polystachyos var. hahnianus	*Nymphaea nouchali*
Eichhornia crassipes	*Nymphaea pubescens*
Eleocharis dulcis	*Nymphoides indica*
Elodea canadensis	*Ottelia alismoides*

Hydrilla verticillata	*Persicaria barbata*
Hymenachne amplexicaulis	*Persicaria decipiens*
Ipomoea aquatica	*Persicaria lapathifolia*
Lemna aequinoctialis	*Persicaria orientalis*
Lemna trisulca	*Pistia stratiotes*
Ludwigia adscendens	*Potamogeton crispus*
Ludwigia octovalvis	*Potamogeton octandrus*
Ludwigia peploides	*Sagittaria platyphylla*
Marsilea mutica	*Salvinia molesta*
Myriophyllum aquaticum	*Typha domingensis*
Myriophyllum simulans	*Typha orientalis*
Myriophyllum verrucosum	*Vallisneria asiatica*

2.2 *In silico* validation

Cabomba caroliniana has been the subject of a previous eDNA-based study wherein one primer pair was designed; however, the designed *C. caroliniana* assay was determined to cross-amplify non-target species of Canadian aquatic plants during *in vitro* testing (see Section 2.3) and therefore were not empirically tested *in situ* (see Section 2.4) for eDNA detectability (Scriver, et al., 2015). As such, a new *C. caroliniana matK* assay was designed (hereafter referred to as C.caroliniana_matK) to specifically exclude detection of eDNA from aquatic plant species endemic to Australia (Table 3).

In silico tests of C.caroliniana_martK, or virtual assessment of binding potential for primer to non-target species (Goldberg, et al., 2016), were conducted using both targeted and non-targeted searches against the NCBI “nr” nucleotide database via PrimerBLAST algorithm (Ye, et al., 2012). Initial targeted *in silico* test specified a list of Australian aquatic plants against which the NCBI “nr” database was queries using PrimerBLAST (Table 2). Species with ≤ 8 base pair mismatches to *C. caroliniana* were documented for the initial targeted *in silico* test (see Section 3.2). The subsequent non-targeted *in silico* test did not specify species against which to assess C.caroliniana_matK but rather assessed these primers against all available sequences within the NCBI “nr” database using. Species with ≤ 4 base pair mismatch to *C. caroliniana* were documented for the subsequent non-targeted *in silico* test (see Section 3.2).

Following satisfactory compliance of F1 and R1 primer sequences and C.caroliniana_matK *in silico* tests, standard desalted oligonucleotides (Table 3) and replica *matK* aDNA fragment (Figure 1) were synthesized by Integrated DNA Technologies (IDT; New South Wales, Australia). Primers were pre-diluted to 100µM in low-EDTA TE buffer by manufacturer while *matK* aDNA fragment was received dry pellet and was resuspended in 50µL 1x TE buffer following manufacturer’s instructions (IDT; New South Wales, Australia).

Table 2. Species against which C.caroliniana_matK was tested in silico *using targeted PrimerBLAST search of Australian aquatic plant sequences available within NCBI "nr" database.*

Australian freshwater aquatic plants	
Azolla spp.	*Gunnera manicata*
Brasenia schreberi	*Hydrilla verticillata*

Cabomba spp.	*Hymenachne amplexicaulis*
Ceratophyllum spp.	*Ipomoea aquatica*
Cyperus spp.	*Lemna spp.*
Eichhornia crassipes	*Ludwigia spp.*
Eleocharis dulcis	*Marsilea mutica*
Elodea canadensis	*Myriophyllum spp.*
Euryale ferox	*Potamogeton crispus*
Najas tenuifolia	*Potamogeton octandrus*
Nelumbo nucifera	*Sagittaria platyphylla*
Nymphaea caerulea	*Salvinia molesta*
Nymphaea spp.	*Sarcandra spp.*
Nymphoides indica	*Typha domingensis*
Nymphoides spp.	*Typha orientalis*
Ottelia alismoides	*Vallisneria asiatica var. biwaensis*
Persicaria spp.	*Vallisneria natans*
Pistia stratiotes	

Table 3. Primer information for C. caroliniana *eDNA assay (C.caroliniana_matK). Asterisk (*) and highhat (^) indicate melting temperature as determined by Geneious (ver. R11) and PrimerBLAST (Ye, et al., 2012), respectively.*

Primer name	Melt temp (˚C)*	Melt temp (˚C)^	GC content (%)	Amplicon (bp)	Oligonucleotide (5'–3')
C.caroliniana_matK-F	60.9	59.3	40	265	GCTCCTTCTTTACATCTATTGCGAT
C.caroliniana_matK-R	62.1	60.0	57.9		GGTGCCACTACAAGACGCT

2.3 *In vitro* validation

Following confirmation of satisfactory *in silico* tests C.caroliniana_matK was empirically tested for specificity to target species (*C. caroliniana*) by attempting to PCR amplify *matK* from genomic DNA (gDNA) extracted from both target (n = 1) and non-target (n = 11 co-occurring aquatic plants) species (i.e., *in vitro* validation; Table 4). Tissue samples for all 12 species of Australian aquatic plants were collected in May 2018 from the Darwin or Howard Rivers in Darwin, Northern Territory by Christopher Collins (Cabomba Eradication Program Manager, Department of Environment and Natural Resources, Darwin, Northern Territory Australia).

Genomic DNA was extracted from each plant tissue sample following standard cetyl trimethylammonium bromide (CTAB) extraction protocol (Gomes, et al., 2017) with a final elution in 500µL MilliQ® water. Following extraction, gDNA for each sample was diluted 1:10 (300µL final volume) for the generation of working stocks for *in vitro* validation. Working stock for each species was split into two neighbouring wells on a standard 96-well plate (150µL per well) and then quantified using the QuantiFluor® fluorometer and associated QuantiFluor® ONE dsDNA System (Promega Co., Australia). Each duplicate gDNA working stock aliquot served as template for *in vitro* Test 3 (see below; Table 4).

In vitro tests included verification of C.caroliniana_matK amplification efficiency and limit of detection (LOD) as nanograms of gDNA per reaction (ng loaded) and copies of artificial *C. caroliniana matK* double stranded DNA (330 bp of aDNA; gBlocks™, IDT Australia; Figure 1) per reaction (copies loaded). For precise quantification of diluted gDNA and stock aDNA (ng/µL ± 99.7% CI) each was measured in duplicate using QuantiFluor® fluorometer with QuantiFluor® ONE dsDNA System (Promega Co., Australia) and then averaged. Duplicate aDNA stock measurements were converted to copies/µL ± 99.7% CI based on the specific nucleotide sequence of synthesized aDNA (Figure 1) and the average weight (ng), upper 99.7% CI weight, and lower 99.7% CI weight (n = 3 calculations) using an online calculator (http://www.endmemo.com/bio/dnacopynum.php).

To test C.caroliniana_matK efficiency and LOD a standard curve was generated by diluting *C. caroliniana* gDNA serially fives time at a factor of 1:10 with MilliQ® water (Standards 1-5) followed four additional serial dilutions at a factor of 1:2 with MilliQ® water (Standards 6-9): Standard 1 = 1.77 ± 0.11 ng/µL, Standard 2 = 1.77 ± 0.11 x 10^{-1} ng/µL, Standard 3 = 1.77 ± 0.11 x 10^{-2} ng/µL, Standard 4 = 1.77 ± 0.11 x 10^{-3} ng/µL, Standard 5 = 1.77 ± 0.11 x 10^{-4} ng/µL, Standard 6 = 8.85 ± 0.11 x 10^{-5} ng/µL, Standard 7 = 4.425 ± 0.11 x 10^{-5} ng/µL, Standard 8 = 2.2125 ± 0.11 x 10^{-5} ng/µL, Standard 9 = 1.1063 ± 0.11 x 10^{-5} ng/µL. Subsequently, C.caroliniana_matK efficiency and LOD were tested using a standard curve generated by diluting *C. caroliniana* stock aDNA (4.644 ± 1.216 ng/µL or 11,751,741,720 ± 3,077,114,111 copies/µL). Standard 1 was generated by diluting *C. caroliniana matK* stock aDNA at a factor of 1:500 with MilliQ® water (Standard 1 = 2.35 ± 0.615 x 10^{7} copies/µL) followed by seven serial dilutions at a factor of 1:10 with MilliQ® water (Standards 2–8 = 2.35 ± 0.615 x 10^{6} – 2.35 ± 0.615 copies/µL) and two final serial dilutions at a factor of 1:2 with MilliQ® water (Standards 9 and 10 = 1.175 ± 0.308 and 0.5875 ± 0.154 copies/µL), respectively. The number of *matK* amplicon copies generated from each technical qPCR replicate run during *in vitro* Tests 1 and 3 (see below) was determined by extrapolation from aDNA standard curve run under the same optimal conditions (see below).

Quantitative real-time polymerase chain reaction (qPCR) was used for three separate *in vitro* tests: 1) determination of C.caroliniana_matK efficiency and LOD using 8-step $\log_{10}$ gDNA standard curve, 2) determination of C.caroliniana_matK efficiency and LOD using 8-step $\log_{10}$ aDNA standard curve, and 3) determination of C.caroliniana_matK specificity to target species (i.e., exclusion of co-occurring non-target species; Table 4). *In vitro* Tests 1–3 were run as 10 µL reactions containing the following: 5 µL PowerUP® Sybr Green Master Mix (Thermo Fisher Scientific, Australia), 0.9 µL forward primer (900 nM final), 0.9 µL reverse primer (900 nM final), 3 µL gDNA or aDNA template, and 0.2 µL MilliQ® water. Room temperature master mix (7 µL) was loaded into white 96-well plates (Life Technologies Inc., Australia) with the Multipette® Xstream electronic dispensing pipette (Eppendorf Inc., Australia) fitted with 500µL CombiTip ® (Eppendorf Inc., Australia) in a UV-sterilized PCR cabinet (Esco, Australia) in the dedicated low copy DNA room within the Molecular Ecology and Evolution Laboratory (MEEL) at James Cook University's Australian Tropical Science and Innovation Precinct (ATSIP, Building 145) in Townsville, Queensland Australia.

For *in vitro* Test 1 the 96-well plate containing 7µL master mix was moved (unsealed) to a cleaned bench (wiped thoroughly with 70% ethanol) in the dedicated pre-PCR room within MEEL where 3 µL of each room temperature gDNA standard (see above) was loaded using a manual single channel P10 pipette (Eppendorf, Australia) fitted with Maximum Recovery filter tips (Axygene, Australia).

For *in vitro* Tests 2 and 3 the 96-well plate containing 7µL master mix was moved (unsealed) to a cleaned bench (wiped thoroughly with 70% ethanol) in the dedicated pre-PCR room within MEEL where 3 µL of each *C. caroliniana* gDNA standard (see above) and 3 µL gDNA template of each species (Table 4) was loaded as per *in vitro* Test 1, respectively.

Following DNA loading, plates were sealed with an optical adhesive film (Life Technologies Australia Ltd. Pty.), briefly vortexed (10 sec), pulse spun (10 sec), loaded onto opened tray of QuantStudio3™ Real-Time PCR System (Life Technologies Inc., Australia), and wiped thoroughly with nonabrasive Kimwipe® to ensure a complete removal of any transparency obstructions present on optical seal (e.g., smudges or dust) before closing QuantStudio3™ tray and commencing qPCR run.

All three *in vitro* tests were run under the following qPCR cycling conditions: initial UDG incubation at 50˚C for 2 min then initial denaturation at 95˚C for 2 min followed by 45 cycles of 95˚C for 15 secs and 65˚C for 1 min (ramp rate = 2.7˚C/sec) before terminal dissociation curve generation (65˚C–95˚C, ramp rate 0.15˚C/sec). *In vitro* Tests 1 and 2 (gDNA and aDNA standard curves, respectively) were run in triplicate while *in vitro* Test 3 (target and non-target gDNA amplification) was run in duplicate. QuantStudio™ Design and Analysis Software (version 1.4.2; Life Technologies, Australia) was used to set the threshold fluorescence to 0.2 then analyse and export data (Excel format) following each *in vitro* test.

Single amplification products produced from each technical replicate of C.caroliniana_matK in vitro Tests 1 (gDNA standard) and 2 (aDNA standards) were used to calculate the 99.7% CI for this species-specific eDNA assay. Moreover, given potential for spurious amplification of non-target region within C. caroliniana genome three replicates of gDNA standard 1 (see above) were Sanger sequenced (bidirectionally) to confirm amplification of targeted matK region whereas all aDNA standard amplifications were considered positive detection without Sanger sequencing confirmation given that aDNA template was specifically designed to be an exact replicate of targeted matK region. Non-target species that exhibited amplification during in vitro Test 3 with amplicon T_m inside 99.7% CI of gDNA and aDNA standards were considered to be putative positive detections and Sanger sequenced (bidirectionally) for confirmation.

AAATTGCATTCTCGACGCGATTAGTATCCAGAATTTACGATCTATTCATTCAATATTTCCCTT
TTTAGAGGACAAATTGACACATTTATCTTATGTGTCAGATATACTAATACCCTACCCCATACA
TCTGGAAATCTTGCTTCAAACTCTTCGTACCCGGGTACGAGAT**GCTCCTTCTTTACATCTAT**
TGCGATGCTTTTTATACGAGCATCCAAAGAAAAAAGATTATTCTTGTTCTTGTATAATTCTCA
TGTATATGAATCCGAATCCATATTCGTTTTTATTCGTAAACAATCCTCTCATTTACGGTCAAT
ATCTTTCCTAGCCTTTCTTGAGAGAACACATTTTTATGCAAAAATAA**AGCGTCTTGTAGTGG**
CACCTCGTA

Figure 1. Artificial double stranded DNA (aDNA) replica of the C. caroliniana matK gene region wherein C.caroliniana_matK targets. Total aDNA fragment length is 385 bp. Bolded blue nucleotides indicate location of forward (C.caroliniana_matK-F) and reverse (C.caroliniana_matK-R) primers (Table 3). Underlined region indicates 3' extensions beyond reverse primer binding site for C.caroliniana_matK-R, which was included to promote efficient primer binding to and amplification of aDNA fragment.

Table 4. Non-target species against which C.caroliniana_matK was tested empirically. Nanograms of gDNA template loaded into duplicate wells of in vitro *test provided within brackets (s: gDNA also extracted from seedpod). Non-native Australian species indicated by asterisks (*). All tissue samples collected from Darwin River, NT except for* Azolla pinnata *(^), which was collected from Howard River, NT.*

Freshwater aquatic plants (Darwin River, Northern Territory)	
Vallisneria nana (2.07)	*Hygrophila angustifolia* (2.12)
Nelsonia campestris (0.603)	*Pandanus aquaticus* (2.60; s = 0.227)
Azolla pinnata^ (1.99)	*Staurogyne leptocaulis* (1.99)
*Cabomba caroliniana** (5.31)	*Nymphaea violacea* (3.51)
Blyxa aubertii (3.20)	*Limnophila brownii* (1.83)
Flagellaria indica (7.40)	*Spirodela polyrhiza* (5.70)

2.4 *In situ* validation

Following confirmation via *in vitro* Tests 1–3 that C.caroliniana_matK had acceptable qPCR amplification efficiency and species-specificity (see Sections 2.3 and 3.3), it was validated *in situ* (i.e., tested empirically using eDNA captured in water samples collected from field site wherein *C. caroliniana* is known to occur). More specifically, eDNA for *in situ* validation was captured from Ross River under Douglas Bridge (19° 18' 21.96" S, 146° 45' 38.52" E) wherein *C. caroliniana* is known to occur (Calvert and Liessmann, 2014).

Samples were transported back to MEEL at ambient temperature (≈ 24˚C) and eDNA was extracted using a novel eDNA workflow ("Preserve, Precipitate, Lyse, Precipitate, Purify (PPLPP)"; Edmunds and Burrows, submitted). Briefly, 20 mL samples were precipitated overnight (4˚C) with glycogen (final concentration (C_f) = 4.4 µg/mL), sodium chloride (C_f = 0.44M), and isopropanol (C_f = 40%) then pelleted (3,270 x g for 90 min at 20˚C; Allegra X12R centrifuge with SX4750 swinging-bucket rotor; Beckman Coulter Pty Ltd, Australia), resuspended in lysis buffer (30 mM Tris-HCl pH 8, 30 mM EDTA pH 8, 800 mM guanidium hydrochloride, 0.5% TritonX-100, pH 10; Leaver, et al., 2015), frozen (≤ -20˚C, ≥ 30 min), thawed (≥ 30 min, room temperature), incubated (50˚C, ≥ 3 hours), precipitated overnight (4˚C) with glycogen (C_f = 55.5 µg/mL) and 2 volumes polyethylene glycol (PEG) precipitation buffer (30% PEG in 1.6M NaCl), pelleted (20,000 x g for 30 min at 20˚C; 5430R centrifuge with FA-45-30-11 rotor; Eppendorf Pty Ltd, Australia), washed twice (1 mL 70% ethanol each wash), and purified of inhibitors (OneStep PCR Inhibitor Removal Kit; Zymo Research Inc., USA). Extracted eDNA was eluted in 100 µL water and split equally four-ways when transferred into 96-well plate (Axygene, Australia) so as to allow for rapid loading of eDNA template technical replicates using Xplorer® electronic 12-channel pipette (Eppendorf, Australia).

In situ test of C.caroliniana_matK was run in quadruplicate 10 µL technical qPCR replicates using the same chemistry and cycling conditions as *in vitro* Tests 1–3 (see Section 2.3) except with 3µL of Ross River eDNA as template (see above), 50 PCR cycles, and template addition undertaken on a cleaned bench (10% bleach then 70% ethanol then water) in the dedicated low-copy DNA room within MEEL. To avoid the risk of cross-contamination with high-copy gDNA or aDNA standards, *in situ* test plate was run with only Ross River eDNA template added (i.e., no standards included on same plate). Following eDNA template

loading, *in situ* test plate was sealed, vortexed, spun, run, and analysed as described above for *in vitro* Tests 1–3.

Amplicons produced from C.caroliniana_matK *in situ* test on Ross River eDNA that exhibited T_m within 99.7% CI of both gDNA and aDNA standards combined ($75.85 \pm 0.67^{\circ}C$) were considered putative positive detections and one technical replicate per biological replicate was Sanger sequenced (bidirectionally) for confirmation (Trujillo-Gonzalez, et al., 2019).

3. Results

3.1 Primer design

Based on *matK* sequence alignments of *C. caroliniana* to all available Australian aquatic plant species (Table 1) a 32 bp indel present between *C. caroliniana* and all non-target species was targeted for F1, with ≤ 15 mismatches present within F1 binding region (in sequence on 5' and 3' side of indel) and ≤ 13 mismatches present within R1 binding region (see Section 3.2).

Optimal forward and reverse primer binding regions were identified between base pairs 640–664 and 884–903 of *C. caroliniana matK* nucleotide sequence, respectively, with an amplicon of 265 bp (Table 3). C.caroliniana_matK-F primer exhibited the following characteristics: 1) T_m = 59.3–60.9, 2) GC content = 40%, 3) length = 25 bp, 4) self-dimer T_m = 0˚C, 5) hairpin T_m = 0˚C, 6) self-complementarity score = 4, and 7) 3' self-complementarity score = 4. C.caroliniana_matK-R primer exhibited the following characteristics: 1) T_m = 60.0–62.1, 2) GC content = 60%, 3) length = 19 bp, 4) self-dimer T_m = 0˚C, 5) hairpin T_m = 0˚C, 6) self-complementarity score = 5, and 7) 3' self-complementarity score = 2.

3.2 *In silico* validation

Initial targeted *in silico* PrimerBLAST search of NCBI "nr" database (see Section 2.2) confirmed that C.caroliniana_matK is entirely dissimilar to all Australian aquatic plant species queried against (Table 5). Subsequent non-targeted *in silico* PrimerBLAST test of C.caroliniana_matK also failed to return any similar species from the entire "nr" database (Table 6).

Table 5. Species with ≤ 5 mismatches to C.caroliniana_matK based on targeted PrimerBLAST search of Australian aquatic plant sequences in NCBI "nr" databse.

Forward mismatches	Reverse mismatches	Species
0	0	*Cabomba caroliniana*

Table 6. Species with ≤ 1 mismatch to C.caroliniana_matK based on non-targeted PrimerBLAST search of entire NCBI "nr" databse.

Forward mismatches	Reverse mismatches	Species
0	0	*Cabomba caroliniana*

3.3 *In vitro* validation

C.caroliniana_matK exhibited satisfactory efficiency and LOD based on gDNA and aDNA standard curve (see Section 2.3).

In vitro Test 1 demonstrated that C.caroliniana_matK amplified gDNA Standards 1, 2, 3, 4, and 6 (see Section 2.3) at an efficiency of 97.9% (R_2 = 0.996) with no primer dimerization detected (see Section 2.3; Figure 2). Moreover, this gDNA standard curve demonstrated LOD for C.caroliniana_matK to be Standard 6 (6.875 x 10^{-4} ng/µL or 1–2 copies/µL; Figure 2). Standard 5 (5/6 replicates with T_m inside 99.7% CI), Standard 7 (2/6 replicates with T_m inside 99.7% CI), and Standard 8 (1/6 replicates with T_m inside 99.7% CI) amplified but were not included in the gDNA standard curve because these amplifications did not fall within the expected linear order based on their dilution factor (log_{10} = 3.3 cycle shift, log2 = 1 cycle shift); however, amplified replicates were extrapolated using aDNA standard curve as follows: Standard 5 = 1–7 copies/µL, Standard 7 = 1–4 copies/µL, and Standard 8 = 2 copies/µL. Standard 9, the lowest dilution factor, exhibited amplification in only 1/6 technical replicates; however, this amplicon exhibited T_m outside 99.7% CI so this amplification was considered spurious and not included in gDNA standard curve nor extrapolated using aDNA standard curve. Forward and reverse Sanger sequences for gDNA standard replicates 1, 2, and 3 were aligned and consensus sequences (241, 255, and 240 bp) demonstrated 100%, 99.4%, and 99.2% pairwise identity to *C. caroliniana* chloroplast genome (GenBank accession KY392764), respectively.

In vitro Test 2 demonstrated that C.caroliniana_matK amplified Standards 1–8 (log_{10} portion) at an efficiency of 83.85% (R_2 = 0.997) with no primer dimerization detected (see Section 2.3; Figure 3). This aDNA standard curve demonstrated LOD of C.caroliniana_matK to be Standard 8 (2.35 $\pm$ 0.615 copies/µL; Figure 3). Standard 9 had 5/6 replicates amplify with all amplicons exhibiting T_m inside 99.7% CI of aDNA Standards 1-8 while Standard 10 had 4/6 replicates amplify with 3/6 amplicons exhibiting Tm inside 99.7% CI and 1/6 amplicon exhibiting T_m outside 99.7%, which was considered spurious. None of these amplifications fell within the expected linear order based on their log_2 dilution factor and were therefore not included in the aDNA standard curve but, instead, were extrapolated using aDNA Standards 1-8 (Standard 9 = 1–5 copies/µL and Standard 10 = 0.4–2 copies/µL).

In vitro Test 3 demonstrated that C.caroliniana_matK successfully amplified gDNA template derived from target species (5.31 $\pm$ 0.32 ng gDNA template loaded or 35,261 $\pm$ 866 copies based on aDNA standard curve extrapolation; Figure 3). Sanger sequencing of *Hygrophila angustifolia*, *Blyxa aubertii*, and *Pandanus aquaticus* amplicons (one technical replicate that amplified for each species during *in vitro* Test 3; Figure 4) confirmed all three detections to be false positive detections as sequences matched *C. caroliniana* chloroplast genome (GenBank accession KY392764) with pairwise identity ≥ 96.4%.

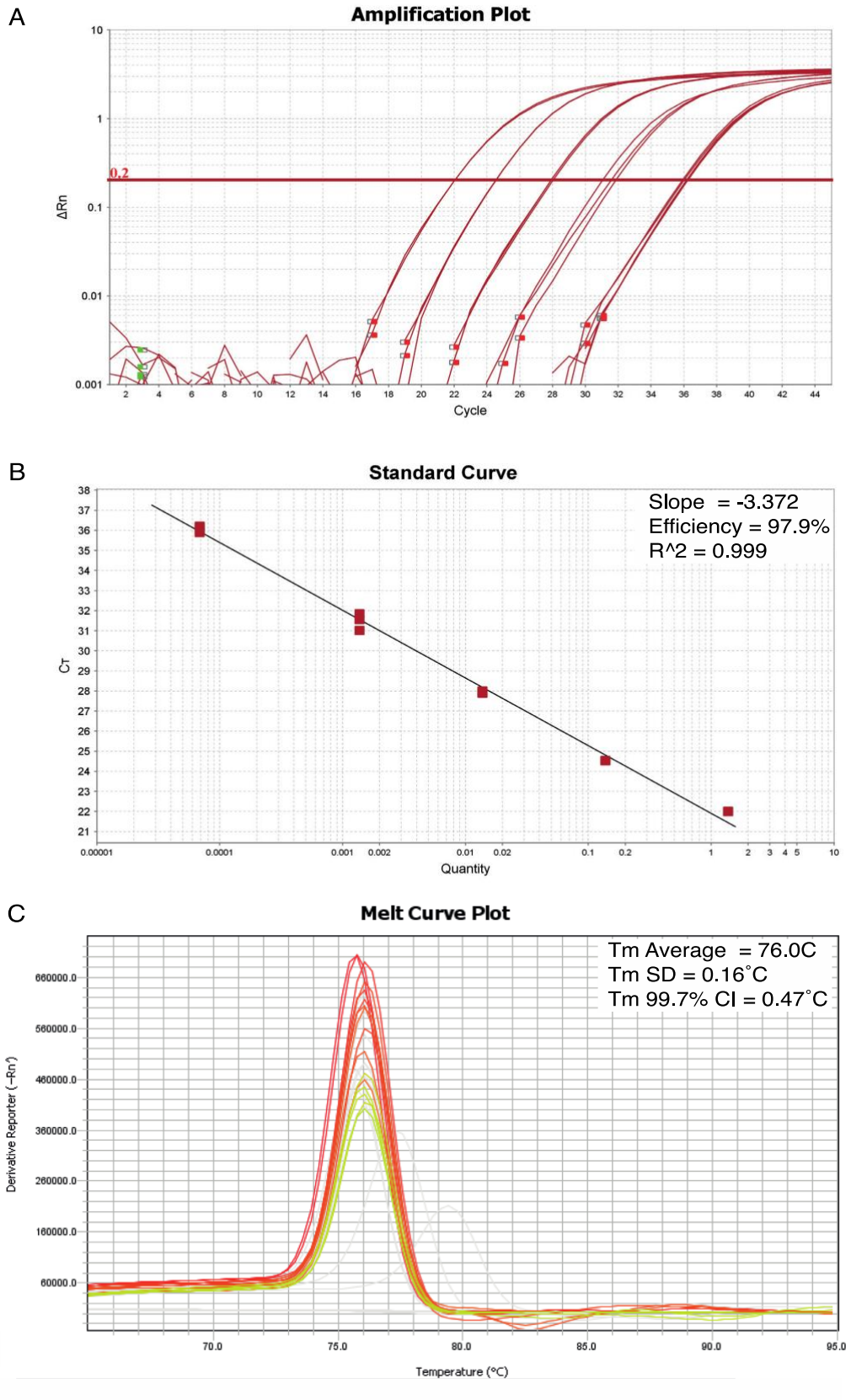

Figure 2. Amplification curves (A), linear regression of gDNA log_{10} standards (B), and amplicon dissociation temperature curves (T_m; C) generated by qPCR during C.caroliniana_matK in vitro *Test 1.*

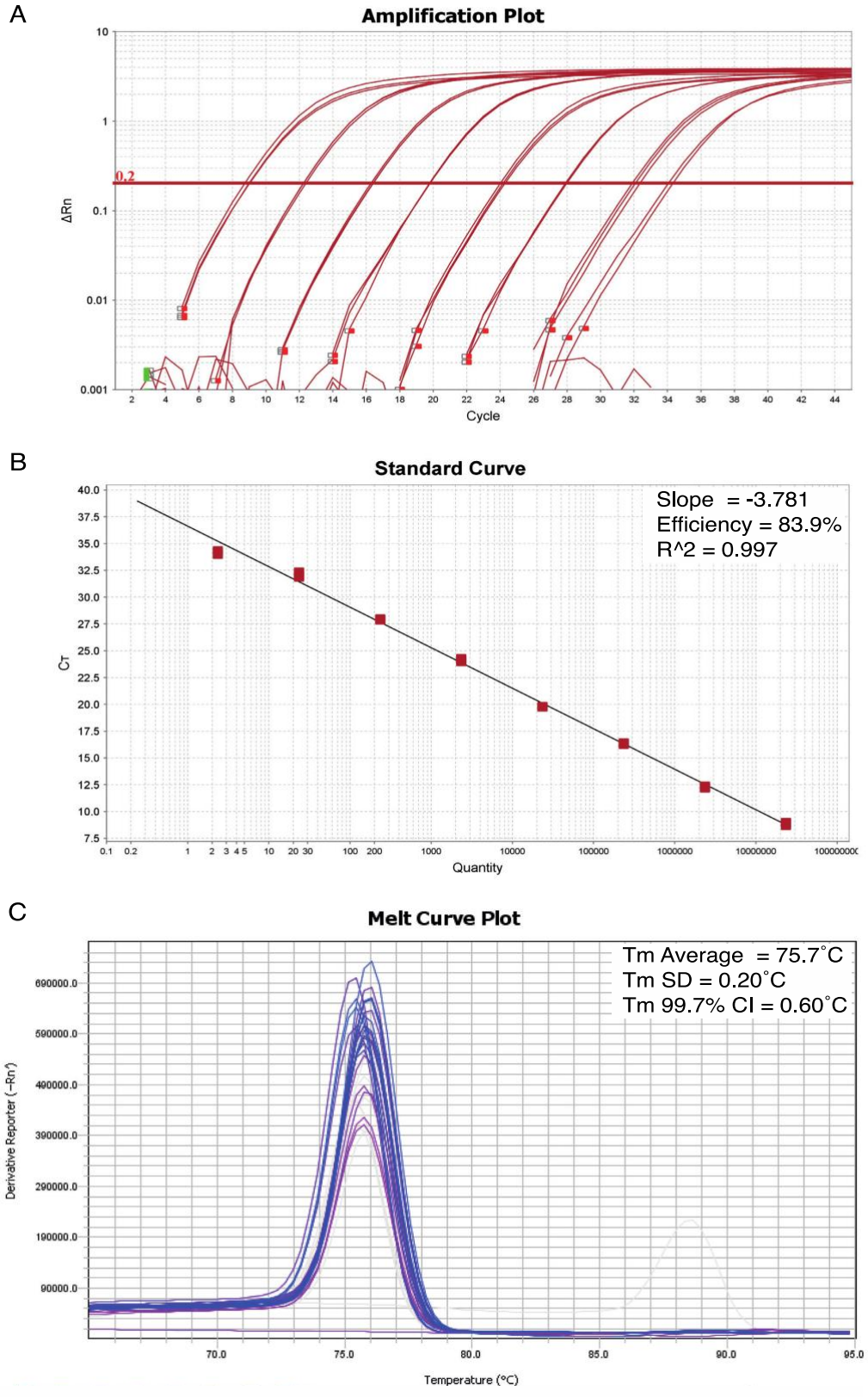

Figure 3. Amplification curves (A), linear regression of aDNA log$_{10}$ standards (B), and amplicon dissociation temperature curves (T_m; C) generated by qPCR during C.caroliniana_matK in vitro *Test 2.*

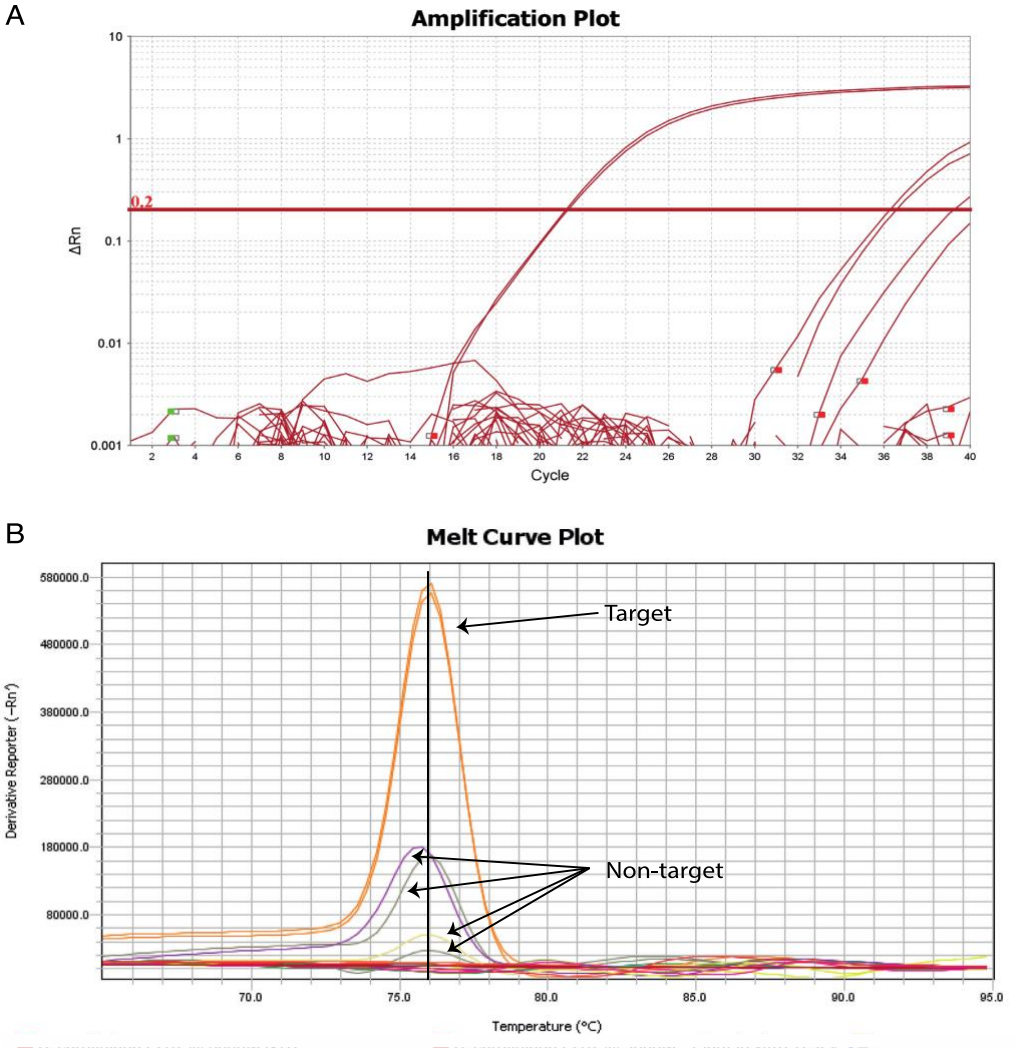

Figure 4. Amplification curves (A) and amplicon dissociation temperature curves (T_m; B) from C.caroliniana_matK in vitro *Test 3 (non-target gDNA amplification). Vertical black line in panel B indicates combined average T_m of gDNA and aDNA standards (75.85 ± 0.67˚C). All produced amplicons exhibited ΔT_m within 99.7% CI of* C. caroliniana *gDNA and aDNA standards; however, the observed putative positive detections for non-target species (*Hygrophila angustifolia, Blyxa aubertii*, and* Pandanus aquaticus*) were confirmed to be false positives by bidirectional Sanger sequencing (i.e.,* C. caroliniana *amplification due to cross-contamination during plate loading).*

3.4 *In situ* validation

C.caroliniana_matK demonstrated 100% putative positive detection across all technical replicates run for each water sample collected from Ross River (Figure 5; see Section 2.4). All amplicons exhibited T_m within 99.7% CI of combined gDNA and aDNA standards (75.85 ± 0.67˚C; Figure 5). Extrapolation of C_t values using aDNA standard curve demonstrated a wide range in *C. caroliniana matK* copies/µL captured from Ross River (C_t 36.124–24.328 correspond to 1–1,788 copies/µL). Consensus sequences generated from aligned forward and reserve Sanger sequences for all twelve validated replicates ranged from 253 to 258 base pairs in length and exhibited ≥ 99% pairwise identity with *C. caroliniana* chloroplast genome (GenBank accession KY392764).

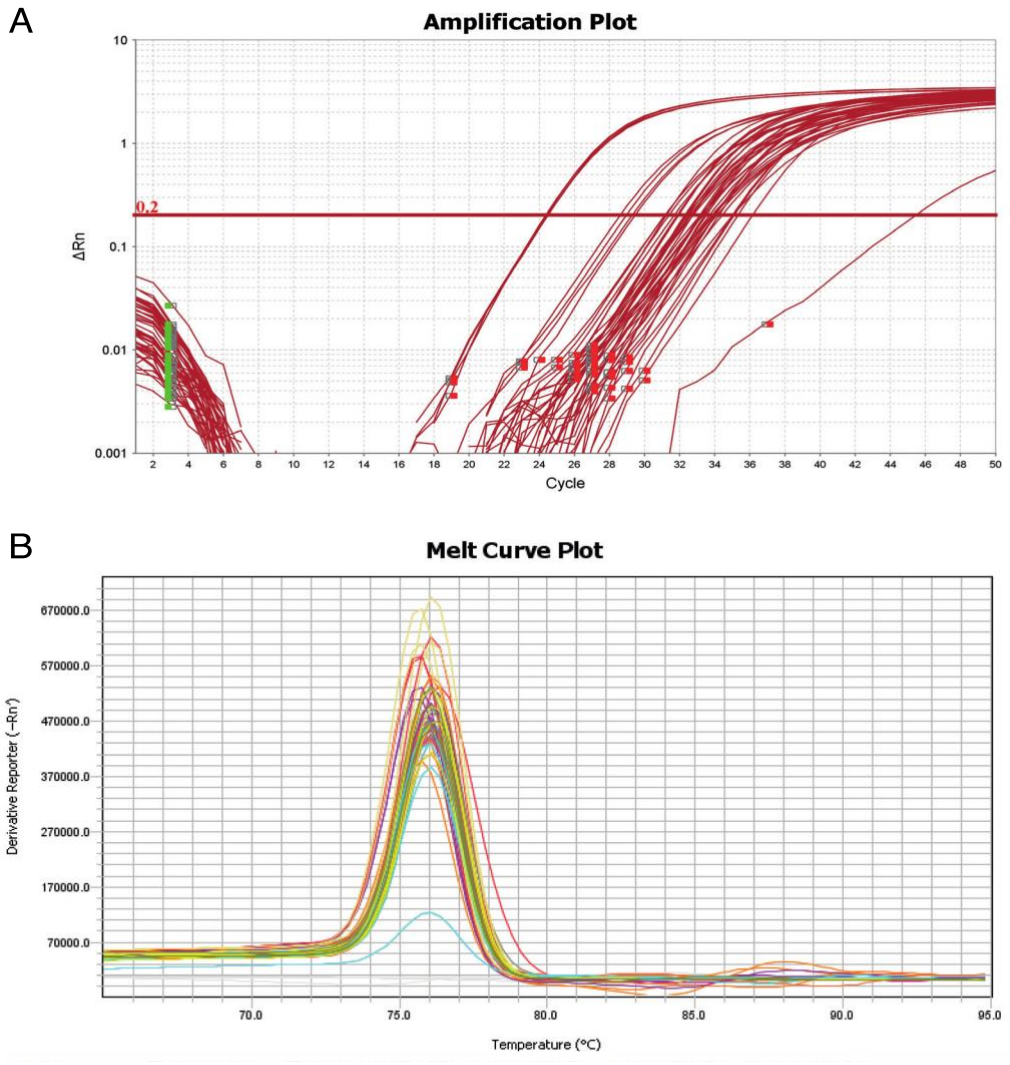

Figure 5. In situ *validation of C.caroliniana_matK using PPLPP method with OneStep PCR Inhibitor Removal Kit (Zymo Research, USA) purified eDNA collected from Ross River at Douglas Bridge. Amplification curve at common fluorescence threshold 0.2 (A) and dissociation curve (T_m) for all n = 48 amplicons (B). Note that, despite late amplification (C_t = 45.42, panel A), this amplicon exhibited ΔT_m within 99.7% CI of gDNA and aDNA standards (Figure 2 and Figure 3), respectively. All Sanger sequenced representative amplicons (n = 12) were positive for* C. caroliniana*.*

4. Discussion

Cabomba caroliniana is a significant invasive pest of waterways in Australia. A new environmental DNA (eDNA) assay was developed for the detection of this species in water samples using Sybr-based quantitative real-time polymerase chain reaction (qPCR). This eDNA assay targets *C. caronlinana* chloroplast gene maturase K (*matK*).

GenBank (NCBI) was mined for available maturase K (*matK*) nucleotide sequences from 39 Australian aquatic plant species. Then, 35 Australian aquatic plants against which the NCBI "nr" database was queried using PrimerBLAST, were tested *in silico*. Following satisfactory *in silico* tests we empirically tested for specificity to *C. caroliniana* by attempting to PCR amplify *matK* from genomic DNA (gDNA) extracted from *C. caroliniana* and 11 non-target aquatic plant species. Finally an *in situ* test was performed using water samples from a site known to contain *C. caroliniana*.

These *in silico*, *in vitro* and, *in situ* validations confirmed that this eDNA assay is specific to *C. caroliniana* despite co-occurrence with native species. Limit of detection (LOD) was determined to be 9 ± 3 copies per qPCR assay under optimal qPCR conditions (65˚C annealing temperature, 900 nM each primer). Moreover, bidirectional Sanger sequencing confirmed all representative putative positive detections from *in vitro* and *in situ* validations to be positive for target species detection while all putative negative detections from *in vitro* validation were confirmed negative, respectively. Collectively, these validations demonstrate the readiness of this assay for screening environmental water samples for determination of *C. caroliniana* eDNA presence.

The *in silico*, *in vitro*, and *in situ* validations undertaken during the development of C.caroliniana_matK described herein demonstrate the readiness of this sensitive eDNA assay for screening environmental water samples. This validated assay can be used to monitor water samples taken from any location wherein *C. caroliniana* is known to be present, suspected to be present (e.g., potential invasion front), or has previously been present (e.g., post-eradication monitoring).

One previous attempt has been made to develop a species-specific eDNA assay for *C. caroliniana* (Scriver, et al., 2015), although *in vitro* validation determined that this assay amplified multiple species and was not utilised for *in situ* detection. However, C.caroliniana_matK can be considered specific to *C. caroliniana* within Australia and, based on non-targeted *in silico* test (see Section 3.2), potentially worldwide. Additional *in vitro* validations are recommended if this assay is to be utilized for screening water samples collected from waterways outside Australia.

References

Bickel, T.O., 2012. Ecology of the submersed aquatic weed Cabomba caroliniana in Australia, in: Eldershaw, V. (Ed.), Eighteenth Australasian Weeds Conference. Weed Society of Victoria, Melbourne, pp. 21–24.

Braukmann, T. W., Kuzmina, M. L., Sills, J., Zakharov, E. V., & Hebert, P. D. (2017). Testing the efficacy of DNA barcodes for identifying the vascular plants of Canada. *PloS one*, *12*(1), e0169515.

Cabomba caroliniana Factsheet, Australian Department of Agriculture and Fisheries: https://www.daf.qld.gov.au/__data/assets/pdf_file/0005/.../IPA-Cabomba-PP30.pdf, accessed 23 May 2018

Cabomba (*Cabomba caroliniana*) – Weed Management Guide; Australian Department of the Environment and Energy: http://www.environment.gov.au/resource/cabomba-cabomba-caroliniana-weed-management-guide, accessed 23 May 2018.

Calvert G. and Liessmann L. (2014) Wetland Plants of the Townsville–Burdekin Flood Plain. Lower Burdekin Landcare Association Inc., Ayr.

Edmunds, R. and Burrows, D. (in review). Got glycogen?: Multi-species validation of PPLPP workflow for eDNA extraction from unfiltered and filtered water samples preserved with Longmire's solution. *eDNA.*

Edmunds, R.C., Villacorta-Rath, C., Collins, C., and Burrows, D. (2019). Monitoring for invasive *Cabomba caroliniana* eDNA in two Darwin River billabongs following eradication attempts. Report 19/25, Centre for Tropical Water and Aquatic Ecosystem Research (TropWATER), James Cook University, Townsville. Available at: nespnorthern.edu.au

Global Species Invasion Database: http://www.iucngisd.org/gisd/, accessed 23 May 2018.

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., ... & Laramie, M. B. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, *7*(11), 1299-130

Gomes, G., Hutson, K., Domingos, J., Chung, C., Hayward, S., Miller, T., & Jerry, D. Use of environmental DNA (eDNA) and water quality data to predict protozoan parasites outbreaks in fish farms. *Aquaculture* **479**, 467–473 (2017)

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S. & Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647-1649.

Kuzmina, M. L., Braukmann, T. W., Fazekas, A. J., Graham, S. W., Dewaard, S. L., Rodrigues, A., ... & Newmaster, S. G. (2017). Using Herbarium-Derived DNAs to Assemble a Large-Scale DNA Barcode Library for the Vascular Plants of Canada. *Applications in plant sciences*, *5*(12), 1700079.

Longmire JL, Maltbie M, Baker RJ (1997) Use of 'lysis buffer' in DNA isolation and its implications for museum collections. *Museum of Texas Tech University*, 163, 1–3.

Renshaw, M. A., Olds, B. P., Jerde, C. L., McVeigh, M. M., & Lodge, D. M. (2015). The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol–chloroform–isoamyl alcohol DNA extraction. *Molecular ecology resources*, *15*(1), 168-176.

Roussel, J. M., Paillisson, J. M., Treguier, A., & Petit, E. (2015). The downside of eDNA as a survey tool in water bodies. *Journal of Applied Ecology*, *52*(4), 823-826

Sanders, D., 1979. The ecology of Cabomba caroliniana, in: Gangstad (Ed.), Weed Control Methods for Public Health Applications. CRC Press Inc, Boca Raton, FL, pp. 133–146

Schooler, S., Julien, M., 2006. Effects of depth and season on the population dynamics of Cabomba caroliniana in south-east Queensland, in: Preston, E., Watts, J.H., Crossman, N.D. (Eds.), Proceedings of the 15th Australian Weed Conference. Weed Management Society of South Australia, Adelaide, Australia, pp. 768–771.

Schooler, S., Cabrera-Walsh, W., Julien, M., 2009. Cabomba caroliniana gray (Cabombaceae), in: Muniappan, R., Reddy, G., Raman, A. (Eds.), Biological Control of Tropical Weeds Using Arthropods. Cambridge University Press, Cambridge, pp. 88–107.

Scriver, M., Marinich, A., Wilson, C., & Freeland, J. (2015). Development of species-specific environmental DNA (eDNA) markers for invasive aquatic plants. *Aquatic botany*, *122*, 27-31.

Williams, K. E., Huyvaert, K. P., & Piaggio, A. J. (2016). No filters, no fridges: a method for preservation of water samples for eDNA analysis. *BMC research notes*, *9*(1), 298.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics*, *13*(1), 134.

Northern Australia Environmental Resources Hub

National Environmental Science Programme

www.nespnorthern.edu.au

This project is supported through funding from the Australian Government's National Environmental Science Program.

TropWATER
Centre for Tropical Water and Aquatic Ecosystem Research

JAMES COOK UNIVERSITY
AUSTRALIA