Marine Ecosystems

Development and validation of a quantitative PCR assay for the invasive mussel, *Mytella strigata* (Hanley, 1843)

Wiltshire, K.H., Giblot-Ducray, D., Deveney, M.R.

SARDI Publication No. F2021/000305-1
SARDI Research Report Series No. 1107

SARDI Aquatics Sciences
PO Box 120 Henley Beach SA 5022

September 2021

Government of South Australia
Department of Primary Industries and Regions

Australian Government
Department of Agriculture, Water and the Environment

SARDI
SOUTH AUSTRALIAN RESEARCH AND DEVELOPMENT INSTITUTE

Development and validation of a quantitative PCR assay for the invasive mussel, *Mytella strigata* (Hanley, 1843)

Wiltshire K.H., Giblot-Ducray, D., Deveney, M.R.

SARDI Publication No. F2021/000305-1
SARDI Research Report Series No. 1107

September 2021

This publication may be cited as:
Wiltshire, K.H., Giblot-Ducray, D. and Deveney, M.R. (2021). Development and validation of a quantitative PCR assay for the invasive mussel, *Mytella strigata* (Hanley, 1843). South Australian Research and Development Institute (Aquatic Sciences), Adelaide. SARDI Publication No. F2021/000305-1. SARDI Research Report Series No. 1107. 53pp.

DISCLAIMER

The authors warrant that they have taken all reasonable care in producing this report. The report has been through the SARDI internal review process, and has been formally approved for release by the Research Director, Aquatic Sciences. Although all reasonable efforts have been made to ensure quality, SARDI does not warrant that the information in this report is free from errors or omissions. SARDI and its employees do not warrant or make any representation regarding the use, or results of the use, of the information contained herein as regards to its correctness, accuracy, reliability and currency or otherwise. SARDI and its employees expressly disclaim all liability or responsibility to any person using the information or advice. Use of the information and data contained in this report is at the user's sole risk. If users rely on the information they are responsible for ensuring by independent verification its accuracy, currency or completeness. The SARDI Report Series is an Administrative Report Series which has not been reviewed outside the department and is not considered peer-reviewed literature. Material presented in these Administrative Reports may later be published in formal peer-reviewed scientific literature.

Author(s):	Wiltshire, K.H., Giblot-Ducray, D., Deveney, M.R.
Reviewer(s):	Tanner, J. and Bastos de Oliveira, H.
Approved by:	Goldsworthy, S. Science Leader – Marine Ecosystems
Signed:	
Date:	20 September 2021
Distribution:	DAWE, SARDI Aquatic Sciences, Parliamentary Library, State Library and National Library
Circulation:	OFFICIAL

ALL ENQUIRIES

South Australian Research and Development Institute - SARDI Aquatic Sciences
2 Hamra Avenue West Beach SA 5024
PO Box 120 Henley Beach SA 5022
P: (08) 8207 5400 **F:** (08) 8207 5415 **E:** pirsa.sardiaquatics@sa.gov.au
W: http://www.pir.sa.gov.au/research

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This project was funded by the Australian Government Department of Agriculture, Water and the Environment (DAWE) and we are grateful to Cian Foster-Thorpe and Brett Herbert for assistance and support. Samples of *Mytella strigata* from Thailand were provided by Andy Shinn (Inve Thailand) and from Uruguay by Noelia Kandratavicius (Faculty of Sciences, Univsersidad de la Republica de Uruguay). At SARDI Aquatic Sciences, Mandee Theil, Leonardo Mantilla and Ian Moody assisted with plankton and scrape sample collection and processing and preparation of mussel tissue for extraction, and A/Prof. Tim Ward and Dr. Mike Steer provided guidance and support. At SARDI Sustainable Systems, the Molecular Diagnostics laboratory staff, including Russell Burns, Danuta Pounsett, Tracy How, Ina Dumitrescu and Nigel Percy, undertook the extractions of environmental samples, inhibition testing and subsampling of extracted DNA for further testing, with guidance and support from Drs Alan McKay and Kathy Ophel-Keller. Some mussel tissue extraction, sequencing, assay design and validation testing was carried out by Concierge Genetics. Important input was received from the ACRF Cancer Genomics Facility. Valuable comments on the manuscript were provided by SARDI staff Jason Tanner and Hugo Bastos de Oliviera.

We are very grateful to everyone who assisted with this project.

Cover photos:

Top left: *Mytella strigata* shell interior after tissue sampling (SARDI)

Lower left: *Mytella strigata* specimen from Thailand (SARDI)

Right: *Mytella strigata* fouling on prawn farm infrastructure, Phaiwat, Kanchanadit District, Surat Thani, Thailand. Courtesy of Dr Andy Shinn.

EXECUTIVE SUMMARY

The charru mussel (*Mytella strigata*) is a fouling species that has formed invasive populations in the south-eastern United States, Philippines, Taiwan, Singapore, India and throughout the Gulf of Thailand. Due to the species' rapid spread throughout Asia and the potential of introduction to Australia, *M. strigata* was added to the Australian Priority Marine Pest List in 2020. Surveillance for exotic pests such as *M. strigata* is important to facilitate early detection and effective management of incursions. This project designed and validated a qPCR assay for *M. strigata* that can be applied for molecular surveillance of this species. The molecular surveillance system developed by SARDI that uses plankton tows tested by qPCR assays provides greater survey confidence for lower cost than traditional survey methods (dive visual, traps, trawl sampling) for detection of target pests. The availability of a validated assay for *M. strigata* allows inclusion of this pest in the suite of species surveyed by this method with known survey confidence (= likelihood of detection in at least one sample).

Nine putative assays were developed, but the assays designed to target gene regions other than CO1 were found by initial validation assessment to lack specificity. Of the five putative qPCR assays targeting CO1 gene regions, the two best performing, which each had assay efficiency > 90% and analytical limit of detection < 2 fg DNA μL^{-1}, were selected for further validation. DNA extracted from plankton samples collected around Australia in 2015 - 2020 was tested to assess field specificity, with no detections by the two selected assays occurring in these samples.

To determine diagnostic performance, plankton and scrape samples to which *M. strigata* tissue was added prior to extraction were tested by the two selected assays. Results were analysed using latent class models to provide estimates of diagnostic sensitivity (DSe), i.e., the likelihood of a detection in a sample containing target DNA, and diagnostic specificity (DSp), i.e., the likelihood of non-detection in a sample without target DNA, for both candidate assays. The best performing assay, MstrigCO1-5, showed DSe of 72.6% in plankton and 71.6% in scrape samples, with the other validated assay, MstrigCO1-3, having a similar DSe (71.9%) in scrapes, but slightly lower DSe (65.5%) in plankton. Both assays had high DSp (99.7%) regardless of sample type. Performance of the MstrigCO1-5 assay in plankton is within the range demonstrated by other assays currently used by SARDI for marine pest surveillance, and therefore a target survey confidence for detection of *M. strigata* would be achieved with the same sampling effort as for other surveyed species. The ability of the assays to reliably detect *M. strigata* DNA provides confidence that the lack of detections in the 2015 – 2020 plankton samples is due to absence of

the target. We additionally assessed a high throughput sequencing approach used for CO1 barcoding and established that is it suitable for confirmatory testing.

PCR inhibition, as measured by a scale factor, occurred in some samples of each type, and negatively impacted DSe of both assays, with a higher scale factor resulting in increased cycle threshold (C_T) value, i.e., more PCR cycles were required for amplification of target DNA, which may impair detection. The effect was, however, relatively minor, with detections occurring reliably in both sample types even at scale factor > 10.

The best performing assay has been validated, permitting implementation of an assay for *M. strigata* in SARDI's testing system for use in molecular surveillance. Assay performance will continue to be assessed during implementation, and, if required, the PCR conditions adjusted, or the assay further refined to improve performance. The assay is suitable for application to DNA extracted from plankton, as used in the SARDI molecular surveillance system, or to scrapes and, by extension, settlement plates, which typically have similar composition to scrapes. The utility of scrapes and settlement plates for surveillance will however also depend on the effectiveness of these sample types at capturing targets, which has not yet been assessed.

Keywords: Marine pests, *Mytella strigata,* qPCR, environmental DNA, assay validation.

1. INTRODUCTION

1.1. Background

The risk of introduction of non-native aquatic species by shipping, via propagules in ballast water or hull-fouling, is increasing with accelerating global trade (Minchin *et al.* 2009; Hewitt and Campbell 2010; Banks *et al.* 2015; Sardain *et al.* 2019). Although not all introduced species become pests, those that do can have wide ranging impacts on ecosystems, marine industries, infrastructure and amenity (Hayes *et al.* 2005a; Schaffelke and Hewitt 2007; Molnar *et al.* 2008; Katsanevakis *et al.* 2014). Aquatic species are rarely able to be eradicated once established, and costs for management or reparation are ongoing (Arthur *et al.* 2015b; Beric and MacIsaac 2015). Australia's National Strategic Plan for Marine Pest Biosecurity, MarinePestPlan 2018-2023, (DAWE 2018b) therefore aims to minimise the risk of pest introduction, prevent further spread of pests already present in Australia, and strengthen surveillance to support these aims and to assist with early detection of new pest incursions.

Species that are amenable to shipping-mediated transport, which have the potential to establish in Australia, and which have demonstrated invasive characteristics and impacts when introduced elsewhere, pose a high risk (Hayes and Sliwa 2003; Hayes *et al.* 2005a). Identified high risk species are included in the Australian Priority Marine Pest List (APMPL), with these species prioritised for surveillance and for management action if detected (DAWE 2021).

The mussel *Mytella strigata* (formerly *Mytella charunna,* known as the charru mussel) is native to both Atlantic and Pacific coasts of central and southern America (Spinuzzi *et al.* 2013; Lim *et al.* 2018). This mussel has formed invasive populations in the south-eastern United States, the Philippines, Singapore, Taiwan, India and throughout the Gulf of Thailand, and has overgrown or out-competed other species, with demonstrated impacts on aquaculture operations and native species (Spinuzzi *et al.* 2013; Lim *et al.* 2018; Jayachandran 2019; Sanpanich and Wells 2019; Fuertes *et al.* 2021; Huang *et al.* 2021). Shipping is the most likely vector for *M. strigata* introductions (Spinuzzi *et al.* 2013; Lim *et al.* 2018). The presence of charru mussels in source ports for shipping to Australia, combined with demonstrated impacts, mean this species is considered a risk to Australia and is included on the APMPL.

The importance of surveillance for introduced marine species is recognised in Australia's marine pest plan (DAWE 2018b). Ports areas at high risk of shipping-mediated introductions and provide typically suitable conditions for establishment of marine pests, and may also act as nodes for

further spread of pest species (Glasby *et al.* 2007; Ojaveer *et al.* 2014; Lehtiniemi *et al.* 2015; Couton *et al.* 2019). Surveillance of ports is therefore important for the early detection and prevention of spread of marine pest species (Bott *et al.* 2010b; Lehtiniemi *et al.* 2015), but the typically high cost of traditional surveillance methods, e.g., using diver visual surveys, trawls, trapping, is an impediment to performing regular monitoring (Arthur *et al.* 2015a).

Molecular techniques for marine pest surveillance are based on detection of pest DNA in environmental samples, and can offer cost and time savings over traditional survey techniques (Bott *et al.* 2010b; DAWE 2018b; Wiltshire *et al.* 2019a). The South Australian Research and Development Institute (SARDI) has developed quantitative polymerase chain reaction (qPCR) assays for DNA-based detection of ten marine pest species (Ophel-Keller *et al.* 2007; Bott *et al.* 2010a; Bott and Giblot-Ducray 2011a, b, 2012; Bott *et al.* 2012), and has developed and refined plankton sampling and preservation methods and quality controls for molecular marine pest surveillance (Giblot-Ducray and Bott 2013; Deveney *et al.* 2017). This surveillance method was field validated by comparison with parallel traditional surveys and is fit-for-purpose for marine pest surveillance (Wiltshire *et al.* 2019a). The molecular method provided higher survey confidence than traditional techniques, while requiring less than half the field costs and person-hours for collection and post-collection processing.

PCR-based approaches, such as those applied in the SARDI surveillance system, perform best of the currently available molecular technologies for detection of specific target species (Darling and Blum 2007; Zaiko *et al.* 2018; Wiltshire *et al.* 2019b). In addition to the assays developed by SARDI, qPCR assays have been developed for several other priority marine pests (e.g. Simpson *et al.* 2018), expanding the range of species that can be targeted with molecular surveillance. No assay, however, had been developed previously for *M. strigata*. We therefore aimed to develop and validate a qPCR assay for this invasive mussel.

Validation of molecular assays is essential before results can be applied in management frameworks because validation data is needed to design surveillance that provides adequate confidence of detection, and for appropriate interpretation of results (Darling and Mahon 2011; Darling *et al.* 2017; Trebitz *et al.* 2017; DAWE 2018a). The World Organisation for Animal Health (Office International des Epizooties, OIE) provides guidelines for validation of molecular tests for animal diseases (OIE 2019); these form the basis of guidelines for validation of tests for marine pest detection in Australia (DAWE 2018a). A validation scale for environmental DNA assays recently proposed by Thalinger *et al.* (2021) also largely aligns with the recommendations of the

OIE and Australian guidelines. Key steps in assay design and validation include (DAWE 2018a; OIE 2019; Thalinger *et al.* 2021):

- *In silico* assessment using DNA from the target species and close relatives to identify sequences suitable for primer and probe development. Suitable sequences need to be identified that do not vary within the target species and that provide adequate mis-matches with sequences from non-target species to achieve specificity.
- Laboratory validation to determine analytical performance, i.e., assay efficiency and limit of detection. Assays with low efficiency (< 80%) or that are unable to detect low levels of target DNA will typically not perform well in the field and so should be redesigned where possible. Ideally, efficiency should be > 90% and the analytical limit of detection in the order of 10 fg DNA μL^{-1}. Laboratory validation typically also includes testing the assay against genomic DNA from relevant non-target species to check for cross-reactions.
- Field validation to assess ability of the assay to detect the target in environmental samples and to check for cross-reactions with non-target DNA.
- Statistical assessment of assay diagnostic performance to quantify (DSe), i.e., the likelihood of a detection in a sample containing target DNA, and diagnostic specificity (DSp), i.e., the likelihood of non-detection in a sample without target DNA. Estimates of DSe and DSp are required in order to design surveillance with known survey confidence, i.e., likelihood of detection in at least one sample, and to appropriately interpret surveillance results.

Field validation involves testing DNA from environmental samples (e.g. plankton), ideally from locations both with and without target pest occurrence (Thalinger *et al.* 2021). Although assays are designed to be specific based on *in silico* analyses and may be tested against DNA of other species during laboratory validation, sequences or DNA are not available for all relevant species. To check for potential cross-reactions with non-target DNA that would result in false positives it is therefore important to trial assays using samples from geographical regions where they will be applied, but from sites where the target is absent (Goldberg *et al.* 2016; Thalinger *et al.* 2021). Assays should also demonstrate detection in samples from sites where the target is known to be present (Thalinger *et al.* 2021), noting that, where the target is rare because of low abundance or temporal factors, e.g., due to reproductive seasonality, detection will not necessarily occur in all samples (Díaz-Ferguson and Moyer 2014; Trebitz *et al.* 2017).

Diagnostic performance in field samples can be assessed by applying an assay to samples with known pest DNA presence or absence, e.g., samples to which target DNA or tissue has been added, and calculating DSe as the proportion of samples containing the pest with a detection, and DSp as the proportion or samples without the pest having no detection. This method can be used where there is only one test available for a species, but, where multiple tests are available, latent class modelling (LCM) approaches are preferred, because these can better quantify the level of uncertainty around DSe and DSp estimates and formally compare the performance of the tests used. LCM can be applied using constructed samples, or to sample sets of unknown target DNA status that comprise samples both with and without target DNA (Branscum *et al.* 2005; Chambert *et al.* 2015; DAWE 2018a). Results from at least two tests are required to apply LCM.

Field validation is complicated for assays for species exotic to Australia, such as *M. strigata*, because environmental samples naturally containing target DNA would need to be sourced from overseas. Testing environmental samples from Australia for these species can provide evidence of field specificity, with a lack of detections demonstrating that the assay is not cross-reacting with and detecting DNA from a non-target species. Testing Australian samples, however, does not provide information on the ability of the assay to detect the target. For validation of tests for *M. strigata*, therefore, samples were constructed by adding known quantities of *M. strigata* tissue or DNA to plankton and scrape samples to assess the ability of candidate assays to detect the target in Australian environmental samples. Constructed samples with added *M. strigata* tissue were also used for quantification of assay diagnostic performance. Environmental samples from both tropical and temperate Australia were tested for the purpose of establishing field specificity across a wide geographic area.

Even where field specificity is checked and DSp of an assay is quantified, there is a chance of cross-reaction with non-target DNA whenever an assay is applied to a new geographic region. In these cases, it is useful to be able to confirm detections by sequencing. High throughput sequencing (HTS) approaches can be applied for confirmatory testing of qPCR results (Díaz-Ferguson and Moyer 2014) and provide information on any non-target sequences detected that can be useful for assay redesign where this is warranted. We therefore also aimed to assess a confirmatory HTS testing approach for *M. strigata*. The HTS confirmatory testing approach targets specific sequences and therefore differs from HTS testing using degenerate primers, which aims to detect a broad range of taxa and is typically applied for characterisation of biodiversity. Multi-species detection approaches via HTS may have useful applications in biosecurity, particularly in investigating impacts of bio-invasions on community structure (Zaiko *et al.* 2018), but these

methods do not have the same level of validation, standardisation or diagnostic performance as qPCR for detection of target species (Darling *et al.* 2017; Zaiko *et al.* 2018; Wiltshire *et al.* 2019b), and we did not assess the application of HTS for this purpose.

1.2. Objectives

This project aimed to:

- Develop candidate qPCR assays for *M. strigata.*
- Validate qPCR assays for *M. strigata*, including:
    - Assess laboratory limit of detection and assay efficiency;
    - Test for cross-reactions with non-target DNA;
    - For selected assays:
        - Determine field specificity when applied to samples from around Australia;
        - Assess efficiency and limit of detection in environmental samples;
        - Quantify diagnostic performance in environmental samples.
- Assess a HTS confirmatory testing approach

2. METHODS

2.1. *Mytella strigata* specimens

A total of 86 *M. strigata* specimens were obtained, comprising 76 from the invaded range (Thailand) and ten from the native range (Uruguay) as shown in Table 1. In all cases, mussels were preserved immediately using 90% analytical grade ethanol (AR EtOH). Preserved whole mussels were shipped to the SARDI Aquatic Sciences South Australian Aquatic Sciences Centre (SAASC) in a small volume of 70% AR EtOH.

Table 1. Details of *Mytella strigata* specimens obtained for this project

Location	Latitude, Longitude	Site	Date	Number
River Mae Klong, Samut Songkhram Province, Thailand	13.384186, 99.984440	*Perna viridis* farm	April 2020	41
Bang Laem, Petchaburi District, East peninsular coast, Thailand	13.265267, 99.933958	Blood cockle farm	October 2020	10
Ta Thong, Kanchanadit District, Surat Thani, Thailand	9.234719, 99.505727	*Perna viridis* farm	October 2020	5
Phaiwat, Kanchanadit District, Surat Thani, Thailand	9.191933, 99.488176	Prawn farm	October 2020	20
La Paloma, Rocha Department, Uruguay	-34.659401, -54.143106	Adjacent Puerto La Paloma	April 2021	10

At SAASC, mussel samples were assigned a unique identifier. Tissue was removed from the shell of each specimen and a subsample of 25 mg was taken from the adductor muscle and foot for DNA extraction. Subsamples for extraction and remaining tissue were each stored separately in labelled tubes of 70% AR EtOH until further processing.

Shells from 20 mussels (samples 1-20) from River Mae Klong were sent to Dr Richard Willan at the Museum and Art Gallery of the Northern Territory (MAGNT) for taxonomic identification. Identity of the specimens was confirmed as *M. strigata*, and specimens were registered in the MAGNT collection with accession number P.62013.

2.2. Assay development and laboratory validation

2.2.1. DNA extraction and sequencing

DNA was extracted from 36 *M. strigata* tissue subsamples using the Macherey-Nagel NucleoMag Tissue kit for DNA purification from cells and tissue, following manufacturer's instructions. Genomic DNA (gDNA) extracts were amplified using the HCO2198-LCO1490 primers, and Sanger sequenced using an Applied Biosystems capillary 3730 DNA Analyzer. The electropherograms were analysed using the Sanger Sequencing and Fragment Analysis Software (Applied Biosystems).

2.2.2. qPCR assay design

Candidate qPCR assays were designed to target three diagnostic gene regions: 28S structural ribosomal RNA (28S, 2 assays), histone H3 encoding gene (H3, 2 assays) and cytochrome oxidase 1 (CO1, 5 assays). The CO1 gene region is commonly used for DNA barcoding and species identification (Deagle *et al.* 2014), and several existing SARDI qPCR assays target this region (e.g. Bott *et al.* 2010a; Bott and Giblot-Ducray 2012; Bott *et al.* 2012). For a few taxa, however, CO1 does not show sufficient variation to delineate species (Krück *et al.* 2013), and, in some cases, diagnostic regions that discriminate species may not have cytosine and guanine content within the range (~40 – 60%) that is recommended for efficient assay performance (Langlois *et al.* 2020). Additional gene regions were therefore also investigated. 28S and H3 gene regions have been utilised for species discrimination and development of molecular phylogenies in several taxonomic groups, including mussels (Colgan *et al.* 2000; Armbruster *et al.* 2005; Crous *et al.* 2006; Wakimura *et al.* 2016; Kartavtsev *et al.* 2018; Ruvindy *et al.* 2018; Guo and Pooler 2021). These two regions were therefore chosen as potentially suitable target regions in addition to CO1.

Sequences obtained for each gene region from *M. strigata* samples, and sequences for *M. strigata* and other bivalves were downloaded from GenBank, grouped by family, and aligned in BioEdit (Hall 1999) using Clustal W alignment to delineate heterologous species. Sequences uploaded to GenBank are not extensively reviewed and may sometimes be recorded against an incorrect species. Sequences that were assigned to *M. strigata* but which were clearly substantially different from the majority of *M. strigata* sequences were considered to be erroneously assigned and were not included in the analysis. Specific primers and TaqMan® Minor Groove Binder (MGB) probe sequences for *M. strigata* with suitable thermodynamic properties

were generated using Primer3 Plus software (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and tested *in silico* for specificity using PrimerBlast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

2.2.3. Laboratory validation

Laboratory validation for each putative qPCR assay included preparation of a standard curve to determine the analytical limit of detection and assay efficiency, and initial assessment of assay specificity.

For preparation of the standard curve, *M. strigata* gDNA remaining from sequencing was quantified using an Invitrogen Qubit Flex fluorometer and Quant-iT™ PicoGreen™ reagent (ThermoFisher). gDNA was diluted or concentrated to 200 pg μl^{-1}. Ten-fold serial dilutions of the 200 pg μl^{-1} concentrate from 10^{-1} to 10^{-6} were prepared in ultrapure water. Ten replicates of each dilution were tested (see section 2.2.4) using each qPCR assay.

Efficiency of a qPCR assay is defined as the proportion of target DNA that is successfully amplified at each cycle (Rebrikov and Trofimov 2006; Kralik and Ricchi 2017). Assay efficiency was determined from the slope of a linear regression of cycle threshold (C_T) value against log concentration (Burns and Valdivia 2008; Caraguel *et al.* 2011). C_T is the number of PCR cycles before fluorescence crosses a threshold and a detection is recorded; a lower C_T therefore indicates greater target DNA content of the sample. The lowest DNA concentration resulting in detection in at least five of ten replicates was considered the analytical limit of detection (LoD) for each assay (OIE 2019).

Relatively few Australian marine species have been sequenced, and it was therefore important to check for potential cross-reactions with non-target DNA that may be present in Australian samples but which did not have sequences available to inform assay design. Field specificity, i.e., a lack of detection where target DNA is absent, can be assessed by testing samples from locations where the target species does not occur. Field specificity of each assay was assessed initially by testing (see Section 2.2.4) DNA from a set of 24 plankton samples collected in Adelaide in 2015 (Table 2). *Mytella strigata* has not been recorded in, and is regarded as absent from, Australia. It is, furthermore, a predominantly tropical species and is unlikely to establish in Adelaide. Any detections in these samples were therefore regarded as being due to cross-reactivity with identical sequences in non-target DNA and indicative of a lack of assay specificity. Assays returning detections in these samples were therefore not considered for further development and

optimisation. From the remaining assays, the two best performing assays (highest efficiency and lowest LoD) were selected for further validation.

Analytical specificity of the two selected assays was assessed using gDNA extracted from the bivalves: *Mytilus galloprovincialis*, *Crassostrea gigas*, *Saccostrea glomerata*, *Ostrea angasi*, *Perna canaliculus*, *Perna viridis*, *Trichomya hirsutus* and seven non-bivalve taxa (see Table 6).

2.2.4. qPCR methods

qPCR was carried out in 384 well plates using a BioRad CFX384 Touch™ real time PCR system for analysis using an activation temperature of 95°C and an annealing temperature of 60°C; cycling parameters were 15 min at 95°C followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Assay concentrations were: 12.5 µl QIAGEN® QuantitTect® Probe master mix, 0.5 µM forward and reverse primer, 0.1 µM TaqMan® Minor Groove Binder (MGB) probe, 2 µl template, and 7.75 µl of nuclease-free water to yield a 25 µl final reaction volume.

2.2.5. High throughput sequencing

We used a modification of the two-step PCR approach described by Wilkinson *et al.* (2017) to amplify and sequence DNA barcodes. This standard approach for HTS detection of marine organisms was modified by selection of primers optimised for mollusca. Specifically, the Leray *et al.* (2013) primers were used to amplify a 313 bp CO1 barcode abutting the 5' end of the CO1. The first PCRs were performed in 20 µL volumes containing 1x MyFi mix (Bioline, Aust.), with both buffer and polymerase, 0.2 nM of each forward and reverse primer and 5 ng of template DNA. PCR products were purified using the Agencourt AMPure XP purification kit (Beckman Coulter, USA) at a ratio of 0.6x beads to PCR product. The second PCR used Nextera 96 index adapter sequences (Illumina, Australia) to add identifying sequences to the amplification products from the first PCR. This was achieved by combining 1x MyFi Buffer (Bioline, Australia), 1.6U MyFi Polymerase (Bioline, Australia), 0.4 nM of paired Nextera 96 Index Sequences and 4 µL of purified PCR product in a 12.5 µL reaction volume. Amplicons were purified using AMPure XP beads (Beckman Coulter, USA) at a ratio of 0.6x beads to PCR product. PCR products were quantified by qPCR with reference to known PhiX standards (Illumina, Australia) using the SYBR FAST qPCR Kit (Kapa Biosystems, South Africa) on a RotorGene RG-6000 machine (Corbett, Australia). Concentration of each of these amplified products was standardized. 5 pM of each sample was 300 bp paired-end sequenced on a MiSeq V3 sequencer at AGRF using a 600-cycle Version 3 kit (Illumina, Australia).

2.2.6. Bioinformatics

HTS sequences were processed using a bioinformatics pipeline coded using MOTHUR software (Schloss *et al.* 2009). Paired-end reads were merged into single sequences, primer sequences were used to identify the barcodes and then trimmed, sequences with ambiguities (i.e., uncertainties at one or more bases) or lengths exceeding the expected 313 bp barcode length by more than 100 bp and singletons were discarded. To facilitate the alignment and/or classification of amplicons, we used previously published metazoan reference alignments for CO1 (147,308 sequences; Machida *et al.* 2017) after adding the *M. strigata* sequence obtained by this study, and constructed a custom alignment for the 5' *M. strigata* CO1 region from NCBI public databases. Raw sequences were classified after alignment to the custom *M. strigata* barcode alignment, and a detection was recorded for a sequence if it most closely matched the reference sequence for *M. strigata*, using MOTHUR's 'classify.seqs' command with default settings. This method considers all taxonomies represented in the template, calculates the probability a sequence from a given taxonomy would contain different 8-mers (i.e., runs of 8 base pairs), calculates the probability a query sequence would be in a given taxonomy based on the 8-mers it contains, and assigns the query sequence to the taxonomy with the highest probability. This strict classification rule greatly diminishes the likelihood of a misidentification based on sequencing error. Using this method, we generated results that recorded the number of detections for *M. strigata* from samples in the plankton and scrape spiking experiments that were tested by qPCR. Positive samples were therefore samples that provided 2 or more reads that matched 99.5% or greater with reference *M. strigata* sequences.

2.3. Environmental samples for validation testing

Two sample types were used for validation testing: plankton samples, as used for the SARDI-developed molecular method (see section 2.3.1), and scrape samples (see section 2.3.2), which are a common method for sampling fouling organisms such as mussels (Hewitt and Martin 2001; Hoedt *et al.* 2001; Campbell *et al.* 2007). Scrape samples were also used as a proxy for settlement plates, which typically contain a similar faunal composition to scrapes (Marraffini *et al.* 2017), and are also a common tool for monitoring fouling pest species (Tait and Inglis 2016; Marraffini *et al.* 2017; Koziol *et al.* 2019; McDonald *et al.* 2020).

Field validation of *M. strigata* assays comprised three stages. DNA extracted from 363 unspiked plankton samples collected at ports around Australia between 2015 and 2020 was tested to assess field specificity, i.e., to check for any cross-reactivity (false positives) with the candidate

M. strigata assays across a large geographical region. Secondly, two sets each of 45 plankton and scrape samples were prepared, with one set of each sample type being spiked with a dilution series of *M. strigata* tissue prior to extraction and the second set being spiked with a dilution series of *M. strigata* DNA post-extraction. These samples were used to assess assay efficiency and limit of detection in environmental samples. Thirdly, a set of 180 samples of each sample type was prepared by spiking with *M. strigata* tissue prior to extraction. These samples were used for the statistical analysis of assay diagnostic performance (see section 2.4.2).

2.3.1. Plankton samples

Plankton samples used for validation testing comprised a total of 630 samples, including 592 collected between 2015 and 2020 for molecular surveillance projects as shown in Table 2, and 38 additional samples collected in Port Adelaide in June 2020 for this project. A map of locations for the previously collected samples is shown in Figure 1, and further detail on collection sites for these samples is available in the reports referenced in Table 2. Collection sites for the samples collected for this project are shown in Figure 2.

All plankton samples were collected based on the methods developed by Giblot-Ducray and Bott (2013) and refined by Deveney *et al.* (2017). A conical mesh plankton net with mouth diameter 0.5 m, length 1.5 m and 50 µm mesh (Sea-Gear 90-50x3-50 or Aquatic Research Instruments AQ-150-50-50) fitted with a flowmeter (Sea-Gear MF315) was towed behind a vessel at a speed of ~1 – 1.5 m s^{-1} and depth of 0.5 – 1 m for a target distance of 100 m. After collection, plankton samples were concentrated down to a volume of ~40 mL by filtering through the mesh windows of the plankton net cod end and transferred to 120 mL tubes containing 80 mL sulfate-based preservation buffer (similar to Stanford University 2015). Samples were kept cool in an insulated container with gel ice packs or refrigerator after collection and for delivery to SAASC where they were stored in a cool room at ≤ 4 °C until processing. Samples collected outside of Adelaide were also kept cool during overnight transport to SAASC using gel ice packs and an insulated container.

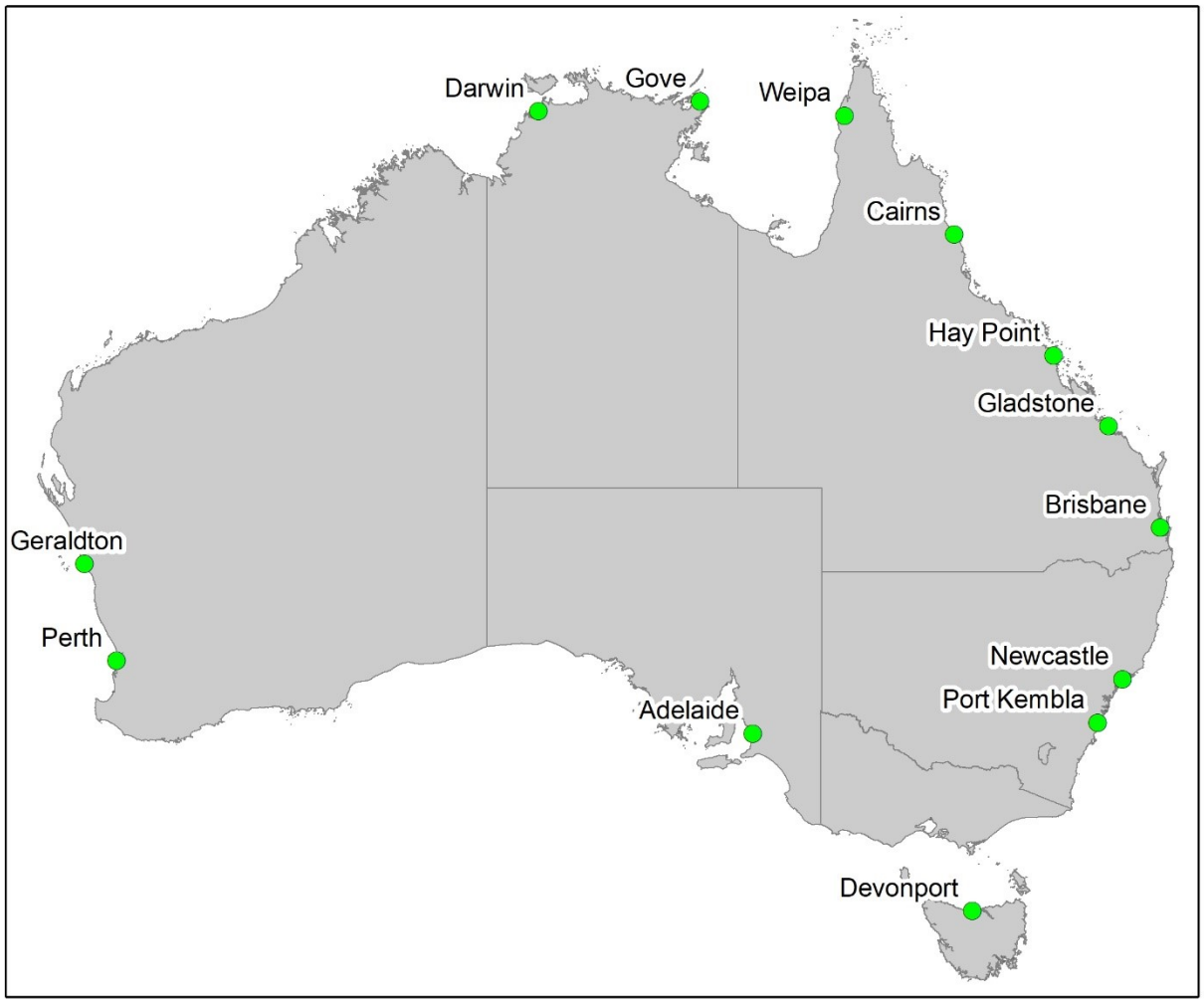

Figure 1. Map of source locations of plankton samples used in this project.

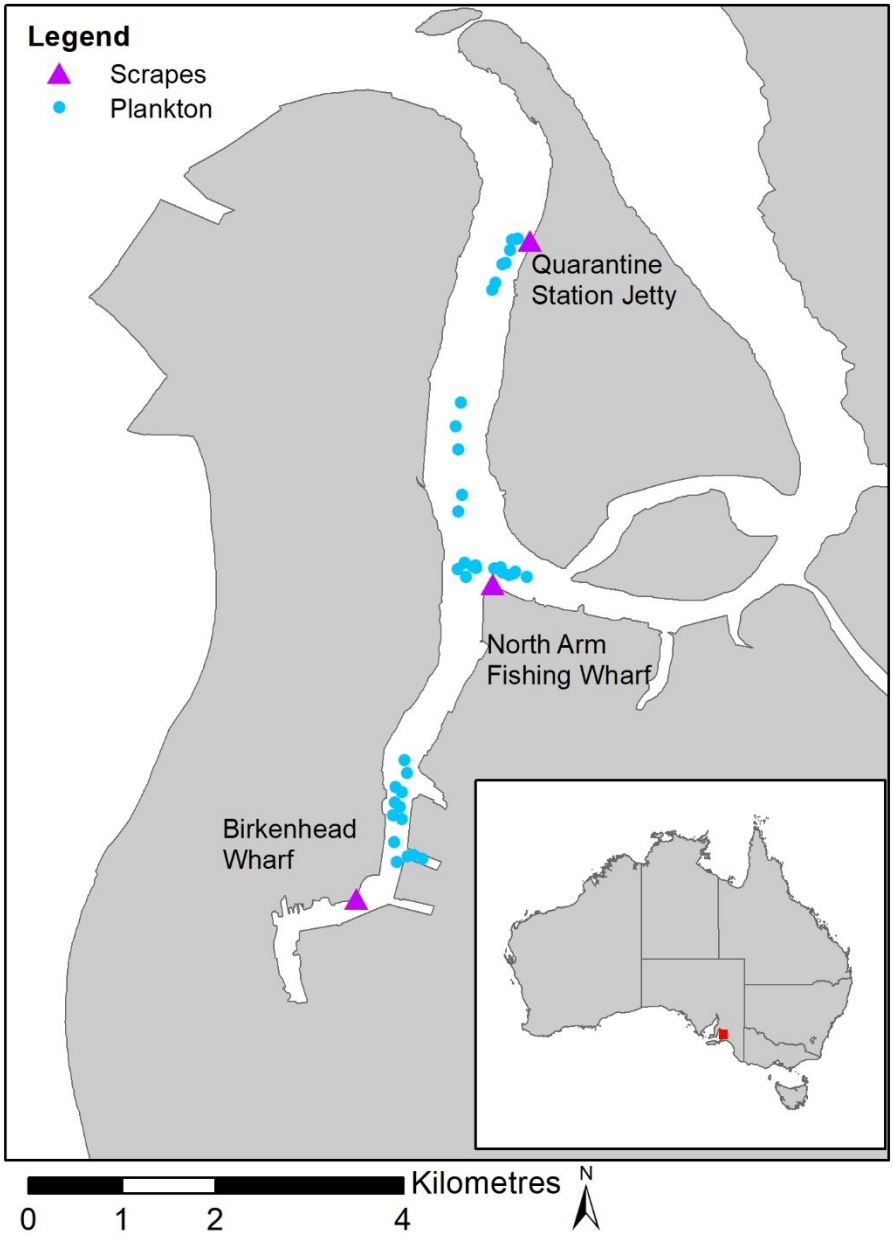

Figure 2. Map of collection sites for plankton and scrape samples collected for this project.

Plankton samples were filtered in the laboratory at SAASC using a manifold and sterile single-use filter cups with 0.45 µm filters (Thermo Scientific™ Nalgene™). Filter papers were transferred to 50 mL centrifuge tubes, frozen at −20 °C and freeze-dried until completely dehydrated prior to tissue spiking of relevant samples and DNA extraction.

Table 2. Summary of plankton samples used for field validation of selected *Mytella strigata* qPCR assays and reference for reports detailing sample collection. (*24 samples from this set were used for initial specificity testing)

Validation stage		
Location	Samples	Reference(s)
Field specificity		
Cairns	30	Deveney *et al.* (2017)
Darwin	20	Deveney *et al.* (2017)
Geraldton	15	Wiltshire *et al.* (2020b)
Gladstone	48	Wiltshire *et al.* (2019a)
Gove	14	Wiltshire *et al.* (2019c)
Hay Point	48	Wiltshire *et al.* (2019c)
Newcastle	19	Wiltshire *et al.* (2019c)
Perth (Garden Island, Kwinana)	88	Deveney *et al.* (2017); Wiltshire *et al.* (2020b)
Adelaide*	38	Deveney *et al.* (2017; 2020); Wiltshire *et al.* (2020a)
Port Kembla	25	Wiltshire *et al.* (2019c)
Weipa	15	Wiltshire *et al.* (2019c)
Field efficiency and Limit of detection		
DNA spiking		
Port Adelaide	45	Wiltshire *et al.* (2020a)
Tissue spiking		
Gove	3	Wiltshire *et al.* (2019c)
Hay Point	7	Wiltshire *et al.* (2019c)
Adelaide	15	this study
Brisbane	3	Wiltshire *et al.* (2019a)
Gladstone	17	Wiltshire *et al.* (2019a)
Assay diagnostic performance		
Brisbane	18	Wiltshire *et al.* (2019a)
Devonport	4	Wiltshire *et al.* (2019c)
Gladstone	65	Wiltshire *et al.* (2019a)
Gove	7	Wiltshire *et al.* (2019c)
Hay Point	39	Wiltshire *et al.* (2019c)
Adelaide	33	Deveney *et al.* (2020); this study
Weipa	4	Wiltshire *et al.* (2019c)

The samples used for field specificity testing and for post-extraction DNA spiking were extracted and tested for inhibition, and for target pests, at the time of collection for the relevant studies shown in Table 2. Remaining extracted DNA from these samples was stored at −20 °C until use in this project. Plankton collected for this study and surplus plankton samples collected during surveillance (Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019c; Deveney *et al.* 2020) were used for

tissue spiking to determine field assay efficiency and limit of detection, and for validation of diagnostic performance. The previously collected samples were filtered, frozen and freeze-dried after collection, but not extracted prior to their use in this project.

2.3.2. Scrape samples

Scrape samples were collected from three wharves in the Port Adelaide region: Quarantine Station Jetty, North Arm Fishing Wharf and Birkenhead Wharf, as shown in Figure 2. Each scrape sample comprised all epifauna from a 10 cm x 10 cm area. Scrape samples were kept cool in the field using an insulated container with ice. On return to the laboratory, scrape samples were weighed to the nearest 0.001 g, placed in pre-weighed foil trays, and frozen at −20 °C. Scrape samples were freeze-dried and dry weights recorded prior to tissue spiking of relevant samples and DNA extraction.

2.3.3. Spiking with *Mytella strigata* tissue and DNA

Remaining tissue from 27 specimens of *M. strigata* that had been subsampled for sequencing (see section 2.1) was patted dry on absorbent paper towel, weighed and freeze-dried. Freeze-dried tissue was homogenised into acid-washed sand by milling to make four stocks of spiked sand, each with a different concentration (Stocks C – F; Table 3). A further two concentrations were prepared as 1:10 and 1:100 dilutions of one of the stock sands (Stocks A – B; Table 3). Aliquots of stock sands were added to freeze-dried plankton and scrape samples to provide a total of ten different doses (Table 3), with eight of the ten tissue dose levels used per sample type. Plankton samples were spiked with doses of level 1 – 8, and scrape samples with doses of level 3 – 10 Table 3). The lowest dose (level 1) was based on the tissue mass of a single D-stage bivalve larva (Waldbusser *et al.* 2013), with successive doses each having an approximately five-fold increase, and dose 3 having the approximate tissue mass of a newly settled mussel spat (Nevejan *et al.* 2007). For determination of assay field efficiency and limit of detection, doses were applied to five samples of each type, with the remaining five samples of each type having no *M. strigata* tissue added. For the assessment of assay diagnostic performance, 30 plankton samples were spiked with each of the doses: 1, 2, 3, 5 and 7; 30 scrape samples were spiked with each of the doses: 3, 4, 5, 7 and 9 (Table 3); and 30 samples of each type had no tissue added.

gDNA extracted from 20 *M. strigata* specimens (see section 2.3.1) was used to spike extracted DNA from plankton and scrape samples. DNA doses were a series of five-fold dilutions,

comprising eight levels between 0.0016 and 125 ng DNA for plankton, and between 0.04 and 3125 ng DNA for scrape samples. Stock DNA dilutions were prepared in ultrapure water by adding the required aliquot of 200 pg μL^{-1} *M. strigata* DNA (see section 2.2.3) in a total volume of 100 µL and vortexing to mix. A 5 µL aliquot of each stock dilution was added to the extracted DNA subsamples from five samples of each type, with the DNA from the remaining five samples of each type having no *M. strigata* DNA added.

Table 3. Doses used for pre-extraction tissue spiking of plankton and scrape samples. Tissue and dose masses are shown as wet weights of *Mytella strigata* tissue. The ten dosing levels were prepared using aliquots of six stock sands (A-F) of differing tissue concentration. See text for further details.

Tissue spiking level (sand stock)	spike type and mass	Total mass in sand	Aliquot (g) added to sample	Dose (mg)	Used for
1 (A)	0.1 g of stock C (0.22 $mg.g^{-1}$)	10	0.01	4.8E-04	Plankton only
2 (A)	0.1 g of stock C (0.22 $mg.g^{-1}$)	10	0.05	2.4E-03	Plankton only
3 (B)	1.0 g of stock C (0.22 $mg.g^{-1}$)	10	0.02	9.3E-03	Plankton and scrapes
4 (B)	1.0 g of stock C (0.22 $mg.g^{-1}$)	10	0.1	0.047	Plankton and scrapes
5 (C)	0.22 g tissue	100	0.1	0.22	Plankton and scrapes
6 (C)	0.22 g tissue	100	0.5	1.10	Plankton and scrapes
7 (D)	0.47 g tissue	100	1	4.74	Plankton and scrapes
8 (E)	1.8 g tissue	100	1.5	26	Plankton and scrapes
9 (E)	1.8 g tissue	100	7	121	Scrapes only
10 (F)	3.7 g tissue	50	8	600	Scrapes only

2.3.4. Extraction and inhibition testing of environmental samples

DNA was extracted from plankton and scrape samples and tested for inhibition by the SARDI Molecular Diagnostics laboratory. Scrape samples were homogenized by milling with the addition of 200 g acid washed, oven-dried paving sand. DNA was extracted from a 10 g subsample. Plankton samples were extracted without additional pre-processing, and the entire sample was extracted.

DNA was extracted using the method developed by SARDI Molecular Diagnostics, with 20 mL of DNA extraction buffer containing an internal control (exogenous organism added to each sample at a standardised amount) added to each sample before physical disruption (Ophel-Keller *et al.* 2008). The efficiency and consistency of SARDI's method to extract DNA from environmental samples have been confirmed in comparison to commercial methods (Haling *et al.* 2011). Final volume of the DNA was 160 µL in elution buffer. Each DNA extract was then tested in singleplex quantitative polymerase chain reaction (qPCR) performed on QuantStudio7 real-time PCR system (Applied Biosystems, Foster City, CA, USA) for the internal control organism used to assess PCR inhibition. For each PCR analysis batch, reference samples that are known to not cause inhibition were also extracted after addition of the inhibition control organism and tested by qPCR. A scale factor was calculated for each sample by comparing the yield of inhibition control DNA detected in the reference samples to that in the sample. A sample with lower yield of the internal DNA control than reference samples, indicative of inhibition, therefore has scale factor > 1, while a sample without inhibition has scale factor = 1.

2.3.5. Environmental validation testing

The two selected qPCR assays were applied to testing the three sets of environmental samples. PCRs were run using the same procedure as for laboratory validation testing (section 2.2.4), with each of the two selected assays applied in triplicate to all samples. A positive detection was considered to occur for a sample when one or more replicates returned a detection. The field limit of detection was determined from the spiked samples as the lowest DNA concentration or tissue dose returning detections in ≥ 50% of replicates.

2.4. Analysis of assay performance

2.4.1. Response of C_T value to dose and inhibition

The C_T value of a qPCR detection decreases approximately linearly with increasing log DNA concentration, although this relationship can be impacted by inhibition in environmental samples (Kralik and Ricchi 2017). Regression analysis can be used to explore these relationships but can be complicated by C_T values being censored, since samples where no target is detected do not return a C_T value (Burns and Valdivia 2008; Caraguel *et al.* 2011). Initial data exploration showed that, contrary to expectation, there was lower correlation of C_T values with DNA concentration for samples with DNA added than with tissue dose for samples with tissue added (0.55 *c.f.* 0.88 in plankton and 0.75 *c.f.* 0.94 in scrapes). Better correlation in samples with DNA added was

expected because these would not be subject to variability due to varying DNA content per mg tissue or due to extraction efficiency, which could occur is samples with added tissue. Regression analysis was therefore performed only for the tissue spiked samples. To determine the effect of tissue dose and inhibition on C_T of each assay in field samples, we used Bayesian censored linear regression, a method that can account for missing C_T values, following Kruschke (2014) with the upper cut off for C_T set as 40 (see Appendix section 7.1 for details).

The natural logarithm of tissue dose was used as the predictor in the regression, with assay and the natural logarithm of scale factor (lnSF) as covariates, and sample barcode as a random effect to account for replicated testing. Field efficiency, which demonstrates how effectively DNA from added tissue is amplified per PCR cycle in environmental samples, was determined for each assay from the coefficient (β) for tissue dose for that assay as: $\exp(-1/\beta) - 1$. The field efficiency shows the relative performance of the assays in environmental samples but is not directly comparable to the amplification efficiency (section 2.2.3) of the assays. A higher field efficiency would be expected to result in a lower C_T value for a given tissue dose, and therefore, a better likelihood of detection for doses near the limit of detection. The coefficient for scale factor was used to assess the effect of inhibition, with the inhibition effect considered significant if the 95% highest density interval (HDI) for this coefficient did not include zero. For scale factor, a positive coefficient would indicate higher C_T for the same tissue dose, indicative of impaired PCR efficiency and potentially detection. Vague normal priors (mean = 0, precision = 0.0001) were used for covariates, and uniform priors bounded at 0.01 and 10 were used for variance components of both fixed and random effects. The natural logarithm of scale factor was used because scale factor has a multiplicative effect; samples with no inhibition have scale factor = 1, hence lnSF = 0 for these samples.

2.4.2. Diagnostic performance

Latent class modelling (LCM) was applied to assess DSe and DSp of the two selected qPCR assays used to test the validation sample sets. LCM was run separately for each sample type using the results of the three replicate tests for each assay.

LCM was carried out in a Bayesian framework with models that allow for covariance between tests, using code adapted from Wang *et al.* (2019) (see Appendix section 7.2 for details). Uniform priors were used for covariance, with bounds set based on minimum and maximum possible values, which depend on the test sensitivities and specificities (Gardner *et al.* 2000; Wang *et al.* 2019). Prevalence, which is the likelihood of a sample containing DNA, was set to 0.83 (=

proportion of samples with *M. strigata* tissue added). The effect of inhibition was assessed by including lnSF as a covariate as per the regression models. An uninformative Beta(1,1) prior was used for the intercept of DSe, i.e. DSe in the absence of inhibition, for each assay. We used relatively informative Beta(361,1) priors for DSp of each assay, based on results of field specificity testing. The addition of *M. strigata* tissue was included as a reference test having perfect DSe and DSp and no effect of inhibition.

2.4.3. Model fitting

Model fitting for both regression and LCM was performed using Markov chain Monte Carlo (MCMC) simulation in JAGS v. 4.3.0 (Plummer 2017) with three chains for 10,000 iterations, thinned at a rate of 10, following 2,000 iterations for adaptation and 10,000 iterations for burn-in. The JAGS code for each analysis is presented in the Appendix. JAGS was run using the 'R2jags' package (Su and Yajima 2015) in R. Convergence was assessed using the Gelman-Rubin convergence statistic and confirmed by visual inspection of trace, density and autocorrelation plots generated using the 'MCMCplots' package (McKay Curtis 2015). Differences between parameters were assessed based on the mean and 95% HDI of the difference between MCMC values, and Bayesian *p*-values were calculated as the proportion of MCMC simulations where one parameter was greater than the other (Caraguel *et al.* 2011). HDIs were calculated using the 'HDInterval' package (Meredith and Kruschke 2018).

2.4.4. Calculation of combined test sensitivity and specificity

The SARDI testing system for marine pests typically applies qPCR assays for each target pest in singlicate, but, where replicate testing with a single assay is applied, or where multiple assays for a target are used, the effective DSe and DSp of the test combination depends on the case definition used and on the covariances between tests (Gardner *et al.* 2000). Where two tests or replicates are used, DSe is maximised by considering a sample positive where either test or replicate returns a detection, known as the 'or' case definition. Using this case definition, the overall DSe is given by:

$$DSe = 1 - ((1 - DSe_1) \times (1 - DSe_2) + covp_{1,2})$$

and overall DSp by:

$$DSp = DSp_1 \times DSp_2 + covn_{1,2}$$

Where DSe1 and DSe2 are the individual test sensitivities, DSp1 and DSp2 the individual specificities, $covp_{1,2}$ is the covariance between the two tests in positive samples, and $covn_{1,2}$ the covariance in negative samples. This case definition results in a decrease in DSp.

The alternative is to only a consider a sample positive where both tests or replicates return a detection, known as the 'and' definition. This case definition maximises DSp while decreasing DSe. Using this case definition, the overall DSe is given by:

$$DSe = DSe_1 \times DSe_2 + covp_{1,2}$$

and overall DSp by:

$$DSp = 1 - ((1 - DSp_1) \times (1 - DSp_2) + covn_{1,2})$$

with parameters defined as above.

Effective DSe and DSe were calculated using these equations for each case definition and sample type, assuming two tests were applied in the combinations: each selected assay in duplicate, and one replicate of each assay.

3. RESULTS

3.1. Candidate assays

Table 4 shows the forward and reverse primers and probes for each putative assay, and the nucleotide content, percentage cytosine and guanine (GC %), and melting temperature (Tm) of the primers and probes, which are important for determining the reaction conditions of qPCR experiments.

3.2. Initial assay assessment

All putative assays showed efficiency > 80%, with four having efficiency > 90% (Table 5). The limit of detection (i.e., where > 95% amplification success is achieved) was less than 100 fg of the target molecule per reaction, with the best performing assays having a limit of detection < 2 fg μL^{-1} (Table 5).

Initial specificity testing showed that assays targeting 28S (assays Mstrig28S-1 & Mstrig28S-2) or H3 (assays MstrigH3-1 & MstrigH3-2) regions returned detections from the Port Adelaide plankton samples. The Port Adelaide samples were used for initial testing due to having exceptionally low likelihood of *M. strigata* presence. These detections are therefore due to the assays cross-reacting with non-target DNA, and these assays were not considered further given this lack of field specificity. None of the five putative assays targeting CO1 (assays MstrigCO1-1 to -5) returned detections in the Port Adelaide samples. Of these assays, two (MstrigCO1-3 and MstrigCO1-5-) had efficiency > 90% and limit of detection < 2 fg μL^{-1} (Table 5). These two assays were selected for further validation.

The two selected assays were shown to be specific against gDNA extracts from eight bivalve and seven other marine taxa (Table 6).

3.3. Sample mass and inhibition in environmental samples

PCR inhibition, as measured by scale factor, was apparent in some environmental samples of each type. The plankton samples used for field specificity testing were tested for inhibition when initially analysed. Previous studies did not find any clear patterns in inhibition, but samples from some locations, particularly tropical ports, had inhibition in a relatively high proportion of samples (Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019c, 2020b). Because we wanted to include samples from tropical locations in the field specificity testing, 105 of the 363 plankton samples used for testing in this project had some inhibition (scale factor >2), with 38 samples having scale factor >

5. The plankton samples used for DNA spiking had minor inhibition (scale factor > 2) in 12 samples, with a maximum scale factor of 3.2.

Table 4. Design of candidate qPCR assays for detection of *Mytella strigata* showing assay name (with format "Mstrig" plus gene region targeted and sequential number for each assay for that gene region), sequences for Forward (Fwd) and Reverse (Rev) primers and probes (Prb), sequence length, melting temperature (Tm), and percentage cytosine and guanine (GC %). *Assays selected for validation.

Assay	Type	Sequence	Length	Tm (°C)	GC %
Mstrig28S-1	Fwd	GGTGTTTAGGACGTCAACTG	20	60.4	50.0
	Prb	AAAGCGGGTGTCAGGCCTTTACA	23	68.4	52.2
	Rev	TGGGCTGCATTCTCAAAC	18	60.3	50.0
Mstrig28S-2	Fwd	GGAGTCGGGTTGTTTGAGAA	20	62.0	50.0
	Probe	CAGCCCAAAGAGGGTGGTAGACTC	24	67.5	58.3
	Rev	GGACTCGTGCCGGTATTTAG	20	61.9	55.0
MstrigH3-1	Fwd	AGTCAAGAAACCACACAGATAC	22	60.3	40.9
	Prb	CTCTGAGAGCAACAGTTCCTGGCC	24	68.1	58.3
	Rev	TCTCTGACTAATCTCTGGAAGG	22	60.4	45.5
MstrigH3-2	Fwd	CTCTCAGAGAAATCAGAAGATACC	24	60.0	41.7
	Prb	CCAGGACTTCAAGACTGATCTTAGATTCCA	30	67.0	43.3
	Rev	AGAGCCATGACAGCAGA	17	60.1	52.9
MstrigCO1-1	Fwd	GTCAGGAACTGGGTGAACAA	20	62.0	50.0
	Prb	TCATACTGGTCCTTCGGTAGACATTCT	27	66.4	44.4
	Rev	CCTTAGACCAATCAAGTGGAGAG	23	62.0	47.8
MstrigCO1-2	Fwd	GCTCCTAACGCCTTATACTTATT	23	60.2	39.1
	Prb	AATTGTTCACCCAGTTCCTGACCC	24	66.6	50.0
	Rev	CTACCGAAGGACCAGTATGA	20	60.2	50.0
MstrigCO1-3*	Fwd	CGGTGCATTTGGGAATTGAC	20	62.0	50.0
	Prb	TCTTCCTTTATGTATCGGTGGTTGTGA	27	65.8	40.7
	Rev	AGCATTAGGAGCCAACCAAA	20	62.2	45.0
MstrigCO1-4	Fwd	TTGGTTGGCTCCTAATGC	18	59.8	50.0
	Prb	ATTGTTCATCCTGTCCCTGCTCCC	24	67.9	54.2
	Rev	ATCTACTGCAGGACCAGTAT	20	59.9	45.0
MstrigCO1-5*	Fwd	GTAGTGGTAACTACTCATGCTTTA	24	60.1	37.5
	Prb	CGGTGCATTTGGGAATTGACTTCTTCC	27	67.7	48.1
	Rev	AATCACAACCACCGATACATAA	22	60.1	36.3

The mean dry weight (± s.e.) of scrapes was 46.5 ± 0.8 g. Scrapes were randomly assigned to treatments, and mass was therefore similar across treatments (Table 7). Of the total 270 scrape

samples, 35 had some inhibition (scale factor >2), with five of these samples having scale factor > 5. The distribution of samples with inhibition across treatments is shown in Table 7.

Table 5. Analytical performance: efficiency and limit of detection (LoD) of putative *M. strigata* assays. *assays selected for further validation

Assay	R^2	Efficiency (%)	LoD (fg μl^{-1})
Mstrig28S-1	0.891	88	<2
Mstrig28S-2	0.824	83	58
MstrigH3-1	0.903	94	<2
MstrigH3-2	0.802	89	<10
MstrigCO1-1	0.911	91	<10
MstrigCO1-2	0.846	84	72
MstrigCO1-3*	0.900	91	<2
MstrigCO1-4	0.884	87	<10
MstrigCO1-5*	0.901	93	<2

Table 6. Results of specificity testing against genomic DNA from non-target species for two candidate *Mytella strigata* assays. ND = not detected. NTC = no template (negative) control

Phylum	Species	DNA (pg μl^{-1})	MstrigCO1-3 C_T	MstrigCO1-5 C_T
Mollusca	*Mytella strigata*	200	20.1	19.3
	Mytilus galloprovincialis	200	ND	ND
	Crassostrea gigas	200	ND	ND
	Saccostrea glomerata	200	ND	ND
	Ostrea angasi	200	ND	ND
	Perna canaliculus	200	ND	ND
	Perna viridis	200	ND	ND
	Trichomya hirsuta	200	ND	ND
Echinodermata	*Asterias amurensis*	200	ND	ND
Chordata	*Ciona intestinalis*	200	ND	ND
	Ascidiella sp.	200	ND	ND
Arthropoda	*Portunus pelagicus*	200	ND	ND
	Nectocarcinus tuberculosus	200	ND	ND
	Ovalipes australiensis	200	ND	ND
Annelida	*Sabella spallanzanii*	200	ND	ND
NTC	nil	0	ND	ND

Table 7. Mean mass (dry weight) ± s.e. of scrape samples and occurrence of inhibition across scrape treatments. See text for a summary of inhibition in plankton samples.

Validation stage	Samples	Scale factor >2 (>5)	Spike level	sample DW (g)
Field efficiency and Limit of detection	5	0 (0)	0	55.0 ± 10.5
	5	2 (0)	3	50.7 ± 24.6
Tissue spiking	5	0 (0)	4	50.8 ± 25.0
	5	0 (0)	5	53.0 ± 8.5
	5	1 (0)	6	42.9 ± 4.7
	5	1 (0)	7	51.9 ± 5.6
	5	1 (1)	8	41.5 ± 15.3
	5	1 (0)	9	47.5 ± 20.1
	5	1 (0)	10	42.1 ± 11.7
DNA spiking	45	8 (0)	-	45.2 ± 11.7
Assay diagnostic performance	30	0 (0)	0	45.7 ± 11.0
	30	6 (2)	3	46.6 ± 11.5
	30	4 (0)	4	50.5 ± 10.0
	30	3 (0)	5	49.5 ± 13.9
	30	5 (2)	7	41.6 ± 12.4
	30	2 (0)	9	44.2 ± 12.4

Of the 45 plankton samples used for field efficiency testing with tissue spiking, eight had some inhibition, with five of these having scale factor > 5. A maximum of two of the five samples for each spiking level had inhibition. Of the 180 plankton samples used for assay diagnostic performance validation, 13 had some inhibition, with seven having scale factor > 5. A maximum of three of the 30 samples per spiking level had inhibition.

3.4. Assay performance in plankton and scrape samples

In plankton spiked with *M. strigata* DNA, both MstrigCO1-3 and MstrigCO1-5 detected 0.04 ng target DNA in > 50% of samples, with MstrigCO1-5 returning detections in a greater number of replicates. In scrape samples spiked with *M. strigata* DNA, both assays detected 1 ng target DNA in > 50% of samples with the same number of replicate detections per assay (Table 8).

In plankton spiked with *M. strigata* tissue, both MstrigCO1-3 and MstrigCO1-5 detected 0.0024 mg tissue, with MstrigCO1-5 having detections in a greater proportion of replicates. In scrape samples spiked with *M. strigata* tissue, both assays detected 0.22 mg tissue in > 50% of samples with the same number of replicate detections per assay (Table 8).

Table 8. Results of testing constructed plankton and scrape samples (environmental samples spiked with *M. strigata* DNA or tissue) by two assays in triplicate. The lowest dose providing detection in >50% of samples (= field limit of detection) is shown in bold. Detections are shown per sample (= detection by one or more of the replicates) and in total (total replicates with a detection).

Set	Sample Type	Spike Qty	Spike type	Samples	Samples with detection (total replicate detections)	
					MstrigCO1-3	MstrigCO1-5
DNA spiking	Plankton	0	ng DNA	5	0 (0)	0 (0)
		0.0016	ng DNA	5	0 (0)	0 (0)
		0.008	ng DNA	5	0 (0)	0 (0)
		0.04	**ng DNA**	**5**	**3 (3)**	**5 (11)**
		0.2	ng DNA	5	5 (14)	5 (14)
		1	ng DNA	5	5 (14)	5 (15)
		5	ng DNA	5	5 (15)	5 (15)
		25	ng DNA	5	5 (15)	5 (15)
		125	ng DNA	5	5 (15)	5 (14)
	Scrape	0	ng DNA	5	0 (0)	0 (0)
		0.04	ng DNA	5	0 (0)	0 (0)
		0.2	ng DNA	5	0 (0)	0 (0)
		1	**ng DNA**	**5**	**5 (12)**	**5 (12)**
		5	ng DNA	5	5 (15)	5 (14)
		25	ng DNA	5	5 (15)	5 (15)
		125	ng DNA	5	5 (15)	5 (15)
		625	ng DNA	5	5 (14)	5 (14)
		3125	ng DNA	5	5 (15)	5 (15)
Tissue spiking	Plankton	0	mg Tissue	5	0 (0)	0 (0)
		0.00048	mg Tissue	5	0 (0)	0 (0)
		0.0024	**mg Tissue**	**5**	**5 (6)**	**5 (11)**
		0.0093	mg Tissue	5	5 (15)	5 (15)
		0.05	mg Tissue	5	5 (13)	5 (14)
		0.22	mg Tissue	5	5 (15)	5 (15)
		1.1	mg Tissue	5	5 (15)	5 (15)
		4.74	mg Tissue	5	5 (14)	5 (15)
		25.9	mg Tissue	5	5 (15)	5 (15)
	Scrape	0	mg Tissue	5	0 (0)	0 (0)
		0.0093	mg Tissue	5	0 (0)	0 (0)
		0.05	mg Tissue	5	0 (0)	0 (0)
		0.22	**mg Tissue**	**5**	**5 (12)**	**5 (12)**
		1.1	mg Tissue	5	5 (14)	5 (14)
		4.74	mg Tissue	5	5 (15)	5 (15)
		25.9	mg Tissue	5	5 (13)	5 (14)
		121	mg Tissue	5	5 (15)	5 (15)
		600	mg Tissue	5	5 (15)	5 (15)

The regression of C_T against log tissue dose in plankton samples showed that slope of the response, and therefore field efficiency, was similar between the two assays (Table 9, Figure 3), with a Bayesian *p*-value of 0.17 for the comparison: field efficiency MstrigCO1-3 > MstrigCO1-5. The positive coefficient for scale factor (Table 9) shows that inhibition led to an increase in C_T for the equivalent tissue dose, with both assays being similarly affected (Table 9, Figure 4); the

Bayesian *p*-value for the comparison: scale factor coefficient MstrigCO1-3 > MstrigCO1-5 was 0.83. A Bayesian *p*-value of 0.09 for the comparison: field efficiency MstrigCO1-3 > MstrigCO1-5, showed that field efficiency was also similar between the assays in scrape samples, but with some evidence of better performance by MstrigCO1-5 (Table 9, Figure 5). The coefficient for scale factor in scrapes was also positive (Table 9) demonstrating increasing C_T with inhibition (Figure 6); the increase was similar for both assays, with a Bayesian *p*-value of 0.91 for the comparison: scale factor coefficient MstrigCO1-3 > MstrigCO1-5.

Table 9. Modelled assay performance (field efficiency) and coefficient (coef) for the scale factor effect for environmental samples dosed with *M. strigata* tissue and tested by two qPCR assays.

Sample type	Assay: / Parameter	MstrigCO1-3	MstrigCO1-5	Bayesian *p*-value MstrigCO1-3 > MstrigCO1-5
Plankton	Field efficiency	70.8 (58.6 – 82.8)	75.9 (62.7 – 89.6)	0.17
	Scale Factor coef	0.957 (0.456 – 1.51)	0.764 (0.203 – 1.28)	0.83
Scrapes	Field efficiency	44.7 (38.5 – 51.0)	47.6 (40.8 – 54.7)	0.09
	Scale Factor coef	1.95 (1.12 – 2.81)	1.59 (0.783 – 2.44)	0.91

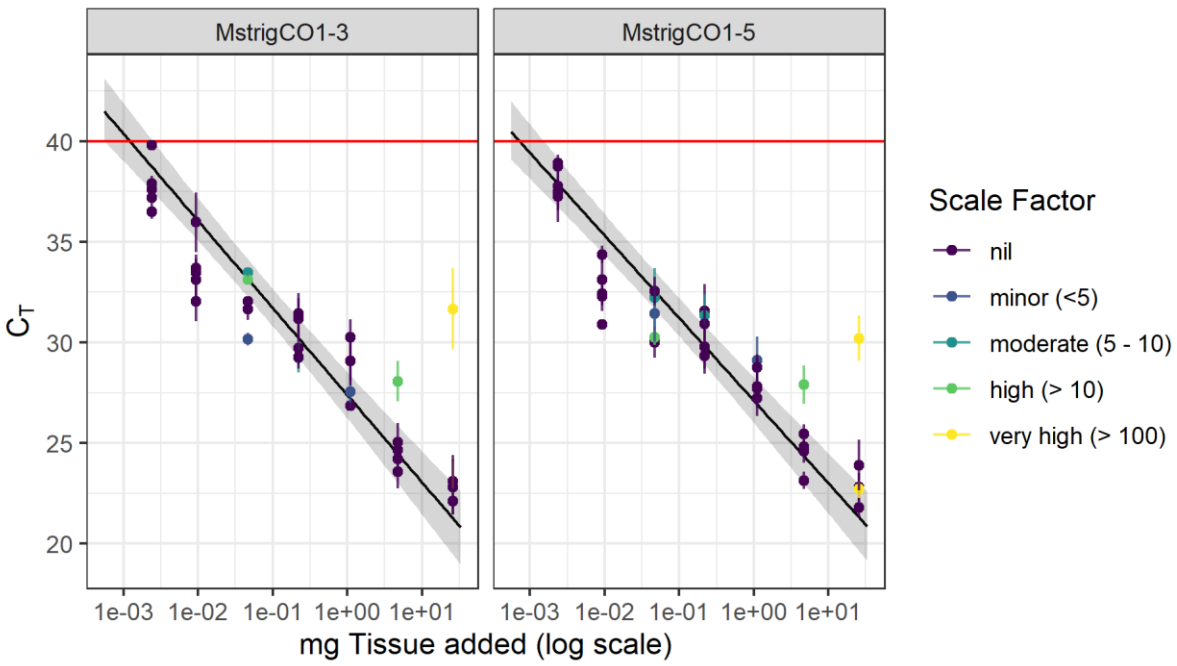

Figure 3. Regression of C_T value against log tissue dose for two qPCR assays applied to plankton samples. Line shows fitted model with shaded area showing 95%HDI. Points show the mean C_T of each sample for each assay, coloured by Scale Factor, bars show standard error (n = 3). The red line shows C_T = 40, the number of PCR cycles used for analysis.

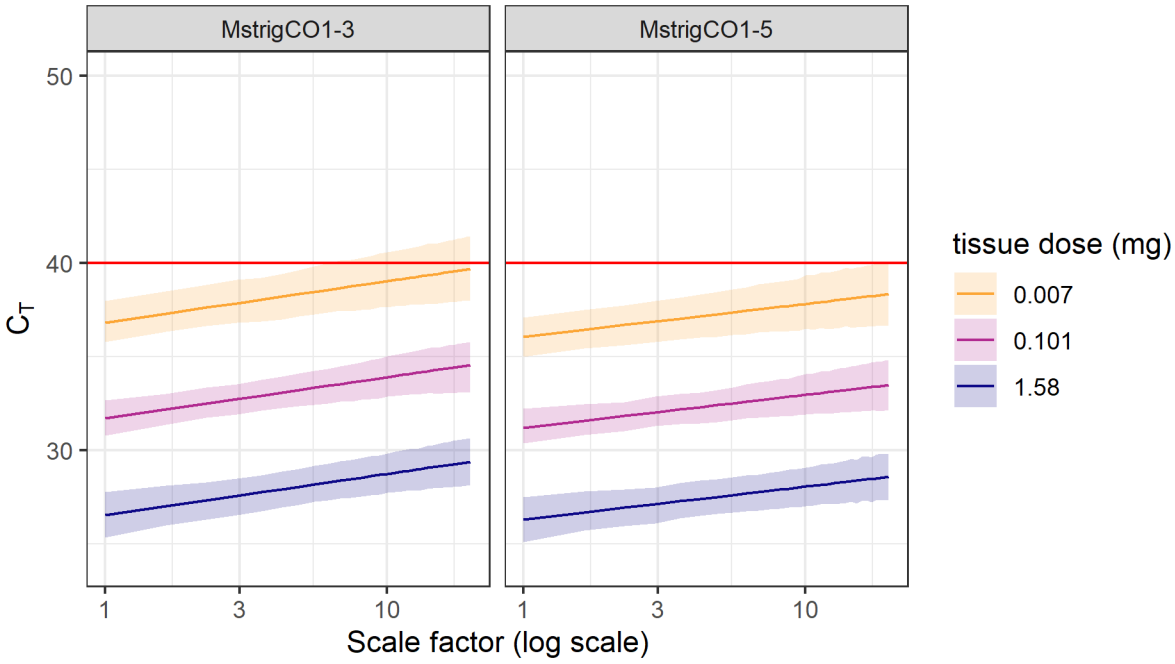

Figure 4. Modelled scale factor response for two qPCR assays applied to plankton samples with added *M. strigata* tissue. Model fit is shown for three doses corresponding to the 25th, 50th and 75th percentiles of the actual doses applied. The red line shows C_T = 40, the number of PCR cycles used for analysis.

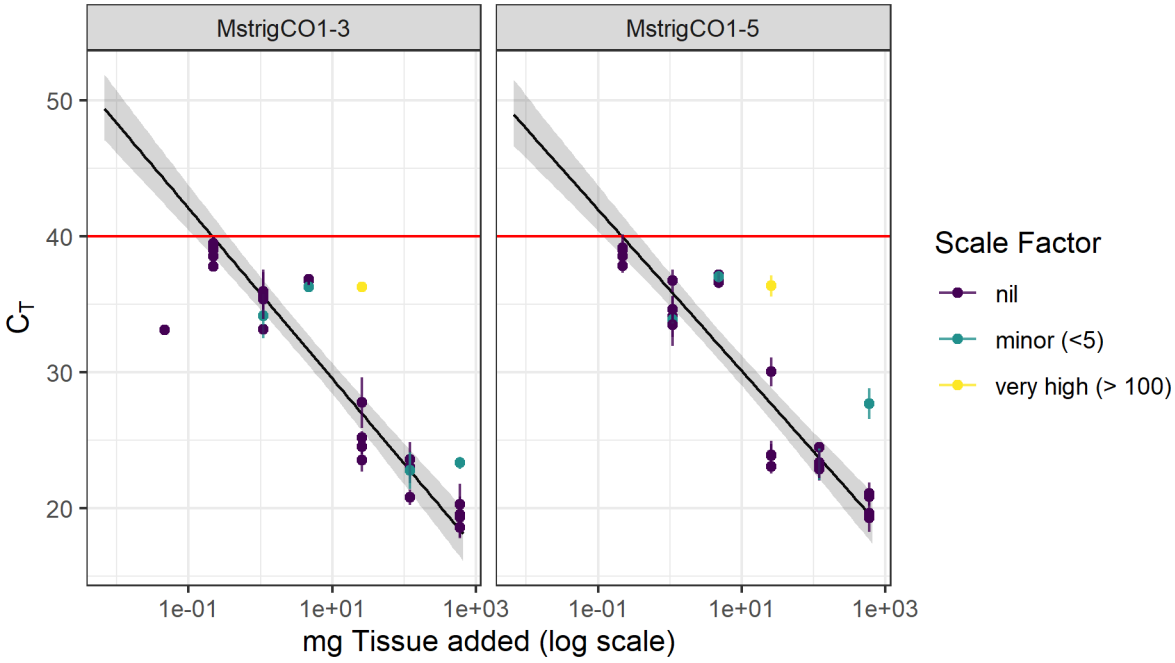

Figure 5. Regression of C_T value against log tissue dose for two qPCR assays applied to scrape samples. Line shows fitted model with shaded area showing 95%HDI. Points show results for each assay coloured by log Scale Factor, bars show standard error (n = 3). The red line shows C_T = 40, the number of PCR cycles used for analysis.

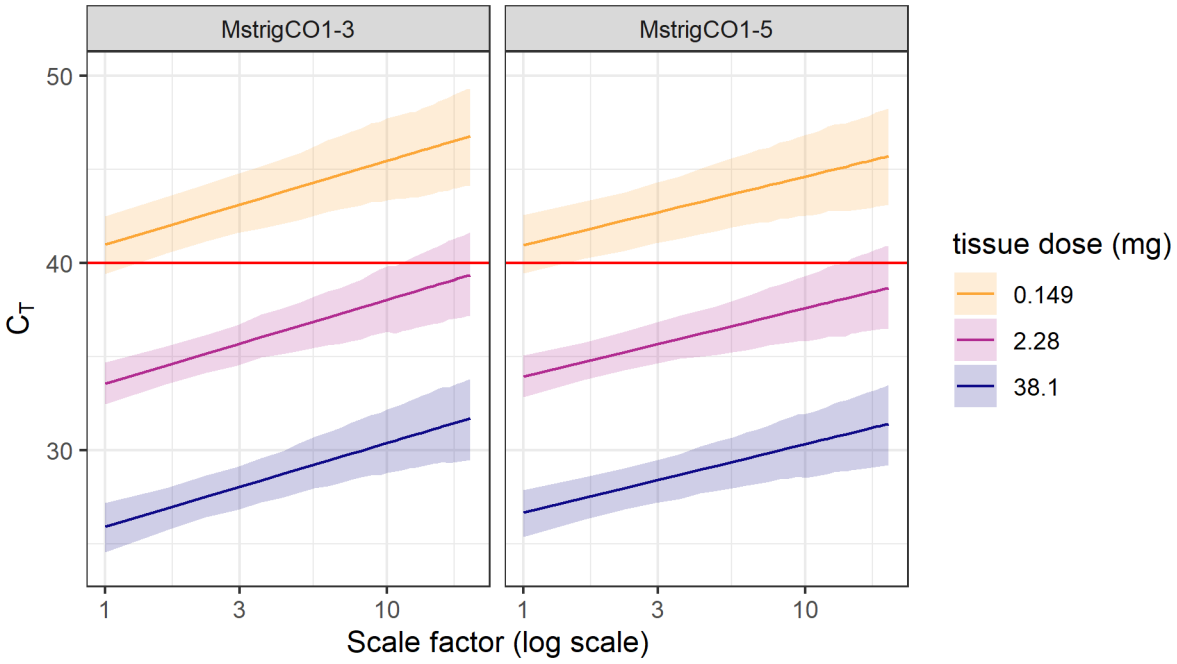

Figure 6. Modelled scale factor response for two qPCR assays applied to scrape samples with added *M. strigata* tissue. Model fit is shown for three tissue doses corresponding to the 25th, 50th and 75th percentiles of the actual doses applied. The red line shows C_T = 40, the number of PCR cycles used for analysis.

3.5. High throughput sequencing

HTS was applied to ten spiked samples selected for being positive by qPCR with C_T > 35. More than two *M. strigata* CO1 barcodes were consistently identified in samples with more than 250,000 reads, providing HTS detections in those samples (Table 10).

3.6. Diagnostic performance of qPCR assays

There were no detections in the 363 unspiked plankton samples from around Australia, or in the unspiked (control) replicates of the constructed sample sets for either plankton or scrapes by the two selected assays (Table 11). There were few detections in samples of either type spiked with the lowest dose level for each, but detections occurred in most samples spiked at higher doses. MstrigCO1-5 returned detections in all replicates for plankton spiked with ≥ 0.0024 mg and both assays had detections in all replicates for scrape samples spiked with ≥ 0.05 mg *M. strigata* tissue (Table 11). MstrigCO1-3 returned detections in most plankton samples spiked at ≥ 0.0024 mg with a greater proportion of replicate samples having detections as the dose increased (Table 11).

Table 10. High throughput sequencing results of ten high C_T positive samples showing C_T values for qPCR detections, total number of HTS reads, and detection (D) or non-detection (ND) by HTS.

Sample	Scale factor	MstrigCO1-3			MstrigCO1-5			HTS reads	D/ND
		rep 1	rep 2	rep 3	rep 1	rep 2	rep 3		
1	1.29	ND	ND	ND	ND	38.5	ND	300,000	D
2	1.27	38.1	38.9	35.2	35.4	39	34.9	250,000	D
3	138.91	37.4	38.8	38.3	37.6	37.8	39.6	180,000	ND
4	1.68	39.8	ND	38.7	ND	39.8	ND	300,000	D
5	1.61	36.7	38.9	37.4	39.8	38.1	ND	275,000	D
6	1.41	39.1	ND	ND	38.4	39.1	39.8	190,000	ND
7	1.53	ND	ND	ND	38.4	ND	ND	475,000	D
8	1.47	ND	36.9	ND	38.2	39.1	37.9	320,000	D
9	1.62	38.5	35.9	ND	37.7	38.1	36.6	150,000	ND
10	1.51	ND	ND	37.6	39.2	36.3	ND	480,000	D

The LCM for plankton samples provided estimates (mean and 95% credible intervals) of DSe of 65.5% (60.6 – 70.2) for MstrigCO1-3 and 72.6% (67.7 – 77.2) for MstrigCO1-5 in the absence of inhibition (Table 12). The Bayesian *p*-value demonstrated that DSe was statistically higher for MstrigCO1-5 (0.022 for comparison DSe MstrigCO1-3 > MstrigCO1-5). Both assays had DSp of 99.7% (99.2 – 1). There was covariance in positive samples between replicates of both assays: 21.5% (19.0 – 23.7) for MstrigCO1-3, and 19.7% (17.5 – 21.8) for MstrigCO1-5 (Table 12). There was covariance between the two assays of 9.51% (7.99 – 11.0). Inhibition, as measured by scale factor, decreased DSe, with the effect being similar between assays (Table 12, Figure 7). Although there was a negative effect of inhibition, MstrigCO1-5 maintained DSe > 70% and MstrigCO1-3 maintained DSe > 60% to scale factors > 10 (Figure 7), and detections were returned by both assays in spiked samples having very high (> 1,000) scale factor.

In scrape samples, the LCM estimated DSe of 71.9% (69.8 – 73.7) for MstrigCO1-3 and 71.6% (69.6 – 73.7) for MstrigCO1-5 in the absence of inhibition, and DSp of 99.7% (99.2 – 1) for both assays (). DSe was not statistically different between the assays in scrape samples (Bayesian *p*-value 0.620). There was covariance in positive samples of 19.5% (18.1 – 20.9) between MstrigCO1-3 replicates, 19.1% (16.4 – 21.0) between MstrigCO1-5 replicates, and 19.7% (18.4 – 20.8) between assays 7 and 9 (Table 12). Inhibition led to decreasing DSe for both assays (Table 12), although both maintained DSe > 65% to scale factors > 10 (Figure 8). Detections by both assays occurred in the sample having the highest inhibition (scale factor > 600) of the set.

In both sample types, covariance between all test combinations in negative samples was negligible (≤ 0.14%) due to high DSp (Table 12).

Table 11. Results of testing unspiked plankton samples for field specificity and constructed plankton and scrape samples (environmental samples spiked with *M. strigata* DNA or tissue) by two assays in triplicate for determination of diagnostic performance. Detections are shown per sample (= detection by one or more of the replicates) and in total (total replicates with a detection).

Set	Sample Type	Spike Qty	Spike type	Samples	Samples with detection (total replicate detections)	
					MstrigCO1-3	MstrigCO1-5
Field specificity	Plankton	0		363	0 (0)	0 (0)
Diagnostic performance	Plankton	0	mg Tissue	30	0 (0)	0 (0)
		4.77E-04	mg Tissue	30	3 (3)	6 (8)
		2.38E-03	mg Tissue	30	28 (63)	30 (90)
		9.30E-03	mg Tissue	30	30 (87)	30 (90)
		0.22	mg Tissue	30	30 (89)	30 (90)
		4.74	mg Tissue	30	30 (90)	30 (90)
	Scrape	0	mg Tissue	30	0 (0)	0 (0)
		9.30E-03	mg Tissue	30	5 (8)	3 (4)
		0.05	mg Tissue	30	30 (90)	30 (90)
		0.22	mg Tissue	30	30 (90)	30 (90)
		4.74	mg Tissue	30	30 (90)	30 (90)
		121	mg Tissue	30	30 (90)	30 (90)

Calculations of effective diagnostic performance for applying tests in combination to plankton showed that applying the assays in combination with the ‘or’ case definition increased DSe to 80.9% (95% HDI: 77.7 – 84.0; Figure 9), but, due to correlation between replicates, applying either assay in duplicate to plankton (Figure 9), or applying any assay combination to scrapes (Figure 10), provided only a minor increase in DSe over application of either assay in singlicate. The ‘or’ case definition decreases effective DSp, but, because DSp of each assay was high, the effect of applying this case definition on DSp was negligible for any test combination (Figures 9,10). The high DSp of the assays also meant that the ‘and’ case definition provided a negligible increase in DSp for either sample type. The ‘and’ case definition decreased effective DSe slightly for each test combination in scrapes (Figure 10), and for either assay in duplicate in plankton (Figure 9), with a greater decrease in DSe for the application of both assays in combination to plankton (Figure 9).

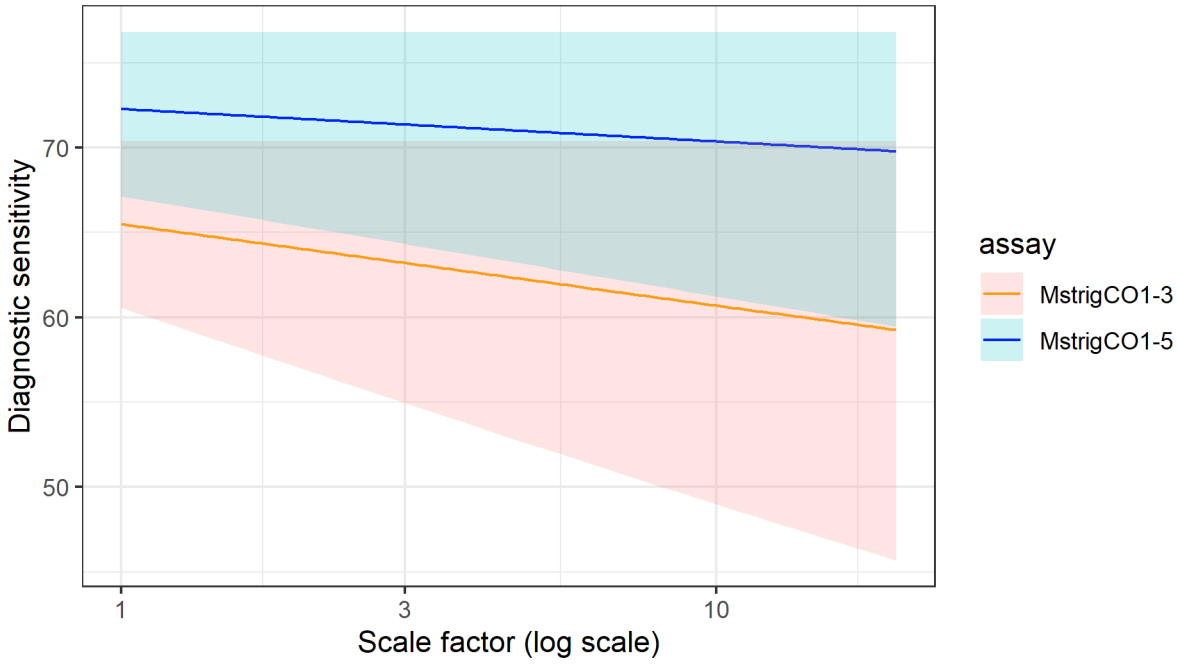

Figure 7. Modelled scale factor effect on Diagnostic sensitivity of two qPCR assays applied to plankton samples with added *M. strigata* tissue. Line shows mean LCM prediction and shaded area shows 95% HDI.

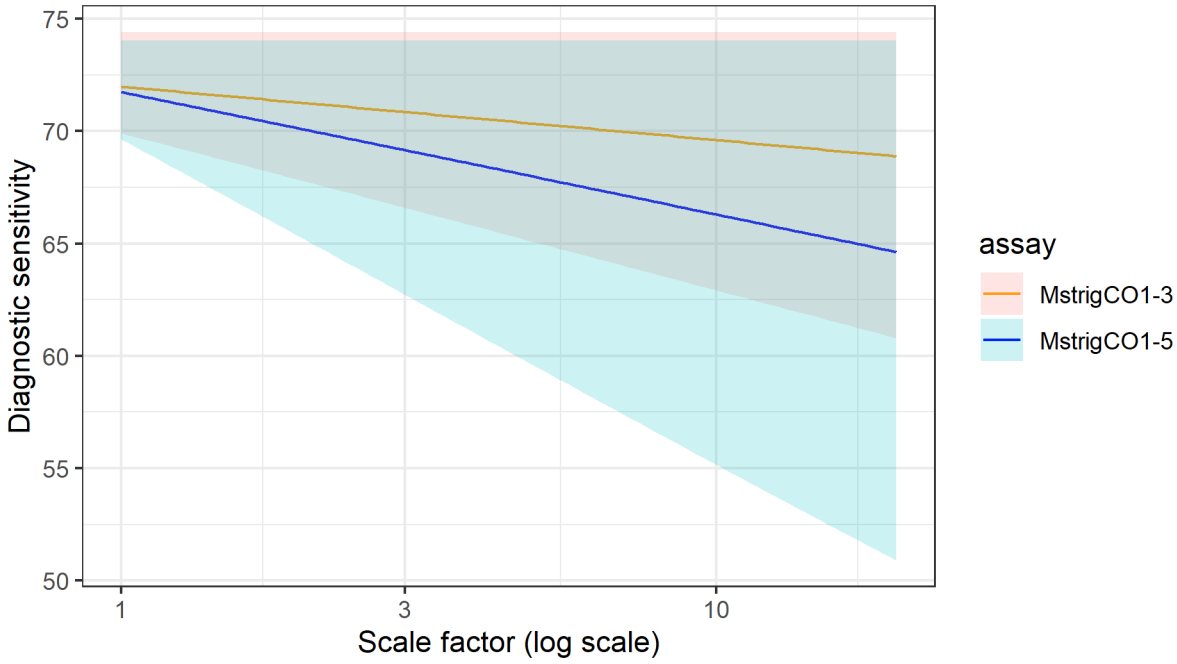

Figure 8. Modelled scale factor effect on Diagnostic sensitivity of two qPCR assays applied to scrape samples with added *M. strigata* tissue. Line shows mean LCM prediction and shaded area shows 95% HDI.

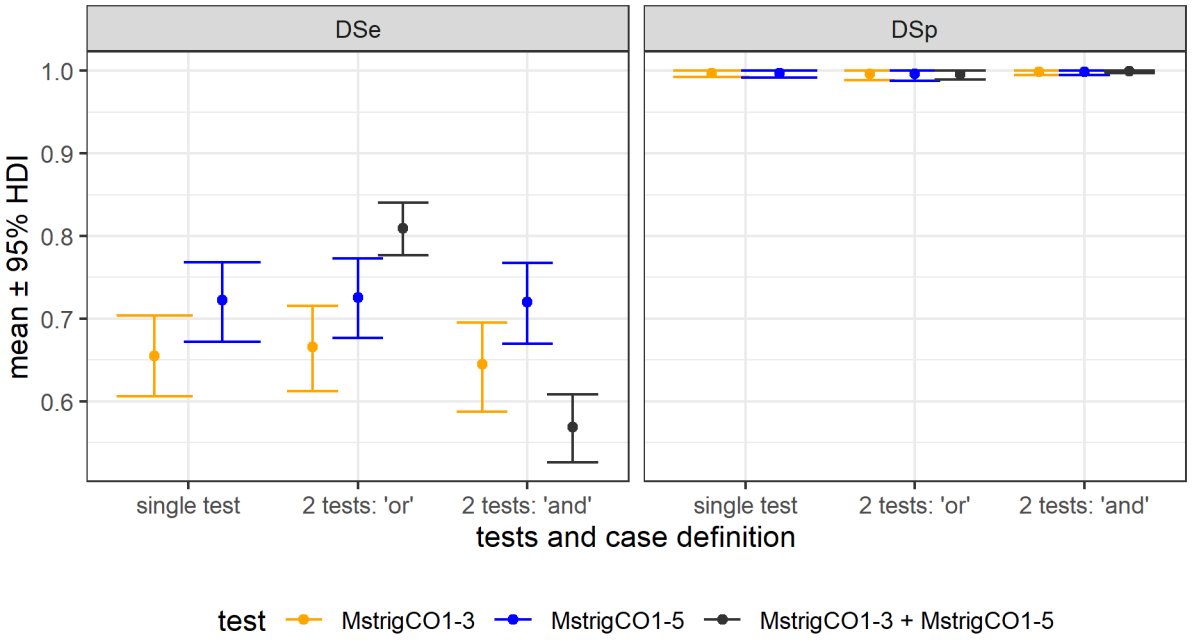

Figure 9. Effective DSe and DSp of applying each assay in singlicate (single test), duplicate (two tests by the same assay) or combination (two tests comprising one by each assay) to plankton with two case definitions.

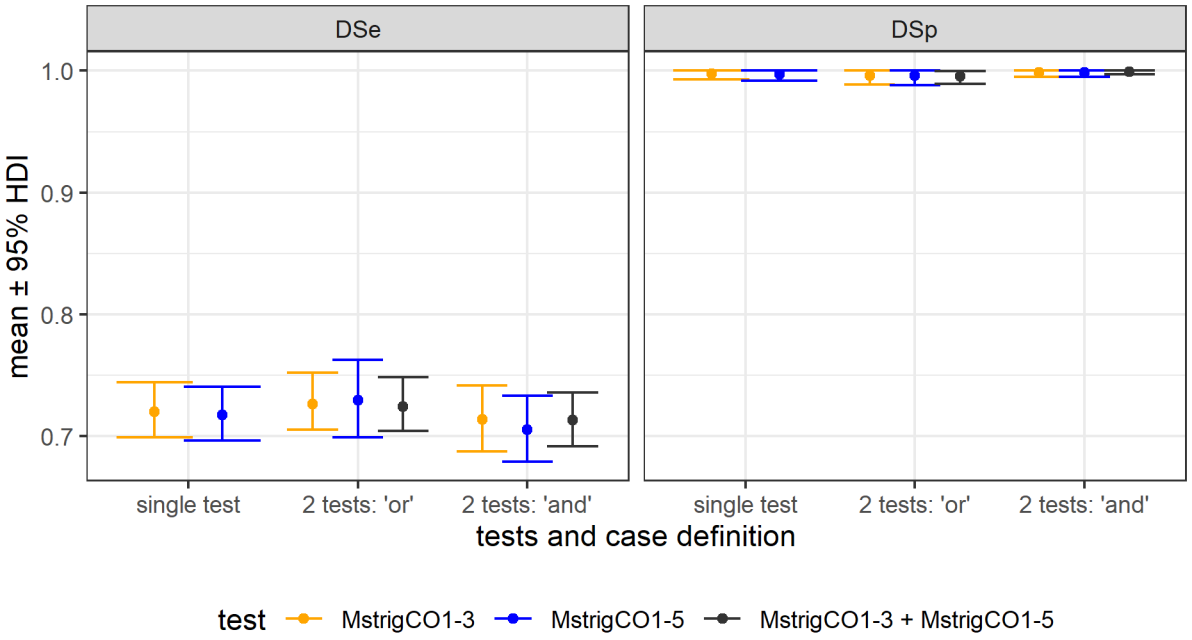

Figure 10. Effective DSe and DSp of applying each assay in singlicate (single test), duplicate (two tests by the same assay) or combination (two tests comprising one by each assay) to scrapes with two case definitions.

Table 12. Modelled diagnostic assay performance (DSe and DSp), covariance in positive (covp) and negative (covn) samples and coefficient (coef) for the scale factor effect on DSe for environmental samples dosed with *M. strigata* tissue and tested by two qPCR assays. NA = Not applicable.

Sample Type	**Mean estimate (95% HDI)**		**Bayesian *p*-value MstrigCO1-3 > MstrigCO1-5**
Parameter	**MstrigCO1-3**	**MstrigCO1-5**	
Plankton			
DSe	65.5 (60.6 – 70.2)	72.6 (67.7 – 77.2)	0.022
DSp	99.7 (99.2 – 100.0)	99.7 (99.2 – 100.0)	0.488
covp (between replicates)	21.5 (19.0 – 23.7)	19.7 (17.5 – 21.8)	NA
covn (between replicates)	0.14 (−0.001 – 0.49)	0.15 (−0.0005 – 0.54)	NA
covp (between assays)	9.51 (7.99 – 11.0)		NA
covn (between assays)	0.08 (−0.002 – 0.29)		NA
Scale Factor coef	−0.057 (−0.142 – 0.000)	-0.022 (−0.064 – 0.000)	0.229
Scrapes			
DSe	71.9 (69.8 – 73.9)	71.6 (69.6 – 73.7)	0.620
DSp	99.7 (99.2 – 100.0)	99.7 (99.2 – 100.0)	0.494
covp (between replicates)	19.5 (18.1 – 20.9)	19.1 (16.4 – 21.0)	NA
covn (between replicates)	0.14 (−0.001 – 0.48)	0.14 (−0.002 – 0.52)	NA
covp (between assays)	19.7 (18.4 – 20.8)		NA
covn (between assays)	0.08 (−0.001 – 0.30)		NA
Scale Factor coef	−0.057 (-0.142 – 0.000)	−0.022 (−0.064 – 0.000)	0.707

4. DISCUSSION

We have developed and validated two qPCR assays for detection of *M. strigata,* with both showing suitable diagnostic performance for application to environmental surveillance using either plankton or scrape samples. The validated assays show amplification efficiency greater than 90% and laboratory limit of detection < 2 fg DNA μl^{-1}. In environmental samples, these assays reliably detected 0.04 ng of target DNA per sample in plankton and 1 ng target DNA per sample in scrapes. Diagnostic performance was similar between the assays in scrape samples with both having DSe > 71%, but MstrigCO1-5 had higher DSe than MstrigCO1-3 in plankton samples (72.6% *c.f.* 65.5%) and also returned detections in a greater proportion of replicates for plankton spiked with low doses of either *M. strigata* tissue or DNA (Tables 8, 11). DSp was high (99.7%) for both assays in both sample types. MstrigCO1-5 is the best candidate assay for implementation given its higher DSe in plankton samples than MstrigCO1-3.

The DSe of the better performing *M. strigata* assay in plankton is comparable to that of previously developed SARDI assays (73 – 91%; Wiltshire *et al.* 2019a), albeit at the lower end. The DSe of an assay is one factor that determines the number of samples required to achieve a given survey confidence for detection of a pest at a target level (e.g., abundance or density). Detecting a species that is absent is impossible, therefore, survey confidence is based on detection of a species at a target threshold level. For the same target abundance or density and sampling method, more samples are required for an assay with lower DSe to achieve the same survey confidence than for an assay with higher DSe.

Wiltshire (2021) developed a tool (https://sardi-mar-biosec.shinyapps.io/surveydesign/) that allows calculation of the number of plankton samples to achieve a user-specified survey confidence for detection of a target concentration of planktonic pests using the SARDI molecular surveillance system. Including the MstrigCO1-5 assay in the suite of tests for plankton samples along with the existing validated SARDI assays (Wiltshire *et al.* 2019a) would not change the number of samples required to achieve a target survey confidence (= likelihood of detection in at least one sample) for detection of a given target planktonic pest concentration (Wiltshire 2021). As an example, 35 samples would be required to detect a concentration of 0.075 planktonic pests m^{-3} with a survey confidence of 80%, or 26 samples to detect 0.1 planktonic pests m^{-3} for the same survey confidence given a DSe of 73% as shown for MstrigCO1-5. These planktonic pest concentrations are indicative of emerging and low abundance established pests respectively, based on an analysis of patterns in detection from previous sampling (Wiltshire 2021). Applying

an assay with DSe of 66%, e.g., MstrigCO1-3, the required sample numbers would be slightly higher: 39 and 29 respectively.

Effective DSe of molecular testing can be increased in some cases by the application of replicate or multiple tests with the 'or' case definition. Covariance between replicate test results for each of the *M. strigata* assays in both sample types, however, means there is little benefit to applying replicated testing of either assay for detection of *M. strigata*, although DSe could be slightly improved in plankton by applying the two assays in combination. The high DSp of the assays means there is little benefit to applying the 'and' case definition for any test combination in either sample type. Applying multiple tests for the same target or replicated testing imposes additional costs and reduces the amount of extracted DNA available for other analyses. The gain in DSe of applying both tests to plankton would mean that a survey confidence of 80% to detect 0.075 planktonic pests m^{-3} could be achieved with 32 rather than 35 samples, but this would require applying two tests for *M. strigata* to each sample, and therefore a total of 64 rather than 35 tests. To detect 0.1 planktonic pests m^{-3} with the same confidence, the required number of samples would reduce from 26 to 24, with 48 tests required instead of 26. The cost of this additional testing is likely to out-weigh the cost of collection of the 2 - 3 additional samples, especially given that the additional samples may need to be collected in any case to achieve the target survey confidence for detection of other species. Although applying combined testing may not be worthwhile in most cases, having two validated assays and an understanding of combined test performance allows for informed consideration of this option.

Both validated assays for *M. strigata* showed similar DSe in scrape samples of 72%, a value only slightly lower than the better performing assay's DSe in plankton. We considered scrape samples as a proxy for settlement plates. Both of these sample types can be used for marine pest surveillance (e.g. Hoedt *et al.* 2001; Tait and Inglis 2016; Marraffini *et al.* 2017; Koziol *et al.* 2019; McDonald *et al.* 2020), and diagnostic performance of the *M. strigata* assays should be similar between scrapes and settlement plates given these typically have similar composition (Marraffini *et al.* 2017). Settlement plates are desirable for monitoring because their deployment is low cost and logistically straightforward (Tait and Inglis 2016; Marraffini *et al.* 2017; McDonald *et al.* 2020). It is, however, difficult to assess the potential performance of either settlement plates or scrapes for detection of *M. strigata* in comparison to plankton because relative performance will depend on the relative likelihood of target occurrence in each sample type, termed 'prevalence' or 'encounter rate', as well as on test diagnostic performance. This relative encounter rate will depend on the relative density of targets (spat for settlement plates or adults for scrapes *c.f.* larvae

or other propagules for plankton) at the threshold trigger level (adult pest abundance or density or planktonic concentration), the area sampled by the scrapes or settlement plates *c.f.* volume sampled by a plankton tow, and effectiveness of each method at collecting targets where they occur.

The SARDI plankton sampling method for molecular surveillance has been field validated (Wiltshire *et al.,* 2019a), and molecular surveillance using this method has been carried out in more than 20 locations since 2015, providing data on the typical planktonic pest concentrations that occur in locations with the target species present (Wiltshire 2021). Data on the method performance and typical planktonic pest concentrations can be used to inform sample number calculations (Wiltshire 2021). Equivalent data is, however, not available for scrape samples or settlement plates. Some studies have been carried out to quantify performance (e.g. Floerl *et al.* 2012) and to optimise the design of settlement arrays (e.g. Tait *et al.* 2018) for marine pest detection, but these assessments were made in limited locations and did not include mussels among the taxa assessed. There is a lack of data from routine application of these sample types in multiple areas of pest occurrence that could be used to assess performance. Where the methods have been applied in locations believed free of target pests (e.g., with the aim of detecting new incursions), detections are not expected, and results do not allow assessment of the method effectiveness. Further investigation is therefore needed to assess the effectiveness of these sample types at collecting target pests in relation to pest population size or density in locations where relevant pests occur before the utility of these sample types for molecular surveillance can be determined.

The DSe of the *M. strigata* assays and their ability to detect ~ 0.2 mg tissue or 1 ng DNA per scrape sample demonstrate that either assay would be suitable to apply to this sample type if spat or adult mussels are effectively captured. If the likelihood of pest occurrence in these sample types is established for a relevant pest population size or density, the required number of samples can be calculated using the general principles of survey design (e.g. Hayes *et al.* 2005b). The relationship between pest population size and occurrence in settlement plates is, however, likely to be complicated (Tait *et al.* 2018). If scrapes or settlement plates have low likelihood of containing targets, then the utility of these methods will be limited despite adequate test diagnostic performance. Species accumulation curves from settlement plate samples suggest that for detection of rare species, such as for emerging invasions, it is likely that very high replication (> 80 plates) would be required (Tait *et al.* 2018).

The limit of detection for both DNA and tissue dose was higher in scrape than plankton samples, but scrape or settlement plate samples would collect spat or adults rather than larvae, and therefore a larger mass and consequently DNA quantity per individual target captured than plankton. A different range of tissue doses was applied to spiking each sample type, with higher doses used for scrapes than plankton because of this difference in life-stage target. Scrape or settlement plate samples result in a larger sample mass than plankton, permitting subsampling prior to extraction, which provides for additional redundancy in the case of extraction failure or inhibition. Target DNA, may, however, comprise only a small fraction of the total DNA extracted, and targets may be difficult to detect in these cases.

The DSe values estimated by this study show the average likelihood of detection across the range of tissue doses used in each sample type. It is important to note, however, that likelihood of detection in a specific sample will depend on the quantity of target DNA present, with detection being more likely with increasing target DNA content. This is also demonstrated by the regression of C_T values against tissue dose, which shows decreasing C_T at higher doses, and by higher dose levels returning detection in more replicates (up to 100%). For samples containing target DNA near the level of detection, conversely, likelihood of detection will be lower than the calculated DSe. The range of tissue doses was chosen to be environmentally relevant based on available data on the tissue mass of bivalve larvae or spat, but ideally samples from a location with *M. strigata* occurrence should be tested to verify the assay's ability to detect target DNA in environmental samples and to determine the DNA yield that occurs in samples of each type from these locations. Because we used a range of tissue doses in the samples tested for DSe estimation, if future investigation shows that some of the dose levels used were either too high or too low, the data can be re-analysed to derive the DSe at a more appropriate range of doses.

PCR inhibition is a recognised issue for molecular surveillance (Goldberg *et al.* 2016). PCR inhibition occurs when products are co-extracted or inadequately removed from the DNA matrix used for PCR. Inhibitors include products that interact with DNA or interfere with polymerases and their cofactors that drive PCR reactions. These include a wide variety of organic substances including structural proteins, enzymes, alcohols, complex polysaccharides and humates, and inorganic substances such as ionic detergents, calcium, urea, chlorides and detergents (Bessetti 2007). Inhibition, as measured by scale factor, occurred in some samples of both plankton and scrapes, with a small number of samples of each type having high or very high inhibition (scale factor > 10 or > 100). Given the wide range of possible causes, it is typically not possible to determine the specific cause of inhibition in a sample or sample set. DSe of both *M. strigata*

assays decreased with scale factor, and the regression analyses demonstrated that C_T value increased with scale factor for a given tissue dose in each sample type. The effect of scale factor was, however, not large, and moderate to high inhibition would be unlikely to impact detection in most cases. Detections occurred in spiked samples of both types even at scale factors > 100, however, such high inhibition could prevent detection in samples containing only a small quantity of target DNA, because these samples would return a high C_T value even in the absence of inhibition. The impact of inhibition on the *M. strigata* assays was similar to that found for the previously-developed SARDI qPCR assays for several target marine pests, which also showed slightly decreasing DSe with inhibition but with some detections occurring even at scale factors > 100 (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019b). Inhibition should continue to be assessed in environmental sampling, regardless of sample type, to identify cases where detection likelihood may be compromised. Where samples have scale factor > 2 and particularly at scale factor > 10, it is likely that accuracy of DNA quantification is reduced (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019b), and knowledge of inhibition is therefore also important for interpretation of pest DNA yields in samples with detections.

Specificity of the validated assays was assessed using gDNA from a range of molluscs and other relevant marine organisms that are likely to occur in Australian survey areas. The extensive testing against DNA extracted from plankton at a range of Australian sites where *M. strigata* is believed not to occur provides confidence of the specificity of the assay and that the likelihood of detections caused by the assay amplifying DNA of non-target species is negligible. The fact that the assays were reliably able to detect *M. strigata* DNA in spiked samples provides further evidence that the lack of detections in unspiked samples is because the assays are specific rather than being due to low DSe, and supports absence of this species from the tested locations at the time of sampling. The high DSp demonstrated by the assays means that applying replicate or multiple tests with the ‘and’ case definition does not appreciably improve specificity.

The HTS approach can detect and confirm *M. strigata* in samples that provide a C_T of >35. Samples with high C_T were chosen for assessment of the HTS approach because these are typically the most challenging to confirm due to relatively low target DNA yield. The taxonomic resolution of HTS depends on the length and variability of the selected DNA barcode across target and non-target species, and the availability of reference sequences for comparison (Porter and Hajibabaei 2018). Target variability was addressed in this study by careful barcode selection and accurate barcode definition, which is particularly important where HTS is used for diagnostic confirmation. Sequencing depth is always, however, required to achieve sensitivity when longer

sequences are generated. The likelihood of detection by HTS increases with the total number of sequences obtained for a barcode, particularly where pest DNA concentration is low (Wiltshire *et al.* 2019b). At least 250,000 reads per sample were required to detect low concentration *M. strigata* targets; achieving this sometimes required multiple runs per sample. This results in higher per sample analytical cost than if the target reads can be achieved in a single sample run. HTS is cost ineffective as a primary tool for detecting species if qPCR is available for that target (Zaiko *et al.* 2018), but the approach has utility for confirmation. The validated assays show high DSp based on testing of samples from a range of locations, but cross-reactions with taxa that occur in other, as yet untested, areas cannot be ruled out. If detections occur in samples from locations without known *M. strigata* occurrence, the HTS testing approach could be applied to provide additional confidence that the detections are of *M. strigata* DNA or alternatively determine whether a cross-reaction has occurred. The advantage of a HTS confirmatory approach is that, if a cross-reaction occurs, information on the non-target sequence is obtained that can assist with assay refinement.

This study benefited both from archived DNA from other SARDI surveys and the sequence libraries in publicly available databases that have substantially expanded since SARDI began designing marine pest assays in 2004. More sequences for the Mytilidae are available and have better coverage than in the 2000s, but specificity problems were observed in the putative *M. strigata* assays developed to target 28S and H3 gene regions. Species-specific qPCR assays targeting 28S and H3 gene regions have been developed for some taxa (e.g. Crous *et al.* 2006; Ruvindy *et al.* 2018; Guo and Pooler 2021), but these gene regions are not always sufficiently diagnostic for discrimination to species level (Matsuda *et al.* 2014; Kartavtsev *et al.* 2018; Ayyagari and Sreerama 2019). Specifity issues have also been identified in SARDI qPCR assays for *Arcuatula senhousia* and *Varicorbula gibba* that target 28S (Deveney *et al.* 2017; Wiltshire *et al.* 2019a). The SARDI qPCR assay for *Sabella spallanzanii* also targets 28S and has been shown to be specific when applied to Australian samples (Wiltshire *et al.* 2019b), but cross-reacts with DNA of two Sabellidae that are native in New Zealand (Wood *et al.* 2013). The design of 28S and H3 assays for *M strigata* was hampered by the relative scarcity of sequences for these gene targets from Australian marine fauna in publicly available databases, which made the identification of suitable diagnostic sequences difficult. The specificity of any assay should always be assessed by testing both gDNA from species that occur in, and environmental samples from, the geographic region in which it will be applied (Wood *et al.* 2013; Arthur *et al.* 2015a; Thalinger *et al.* 2021), but this is particularly important where reference sequences are lacking. CO1 has extensive coverage

of reference sequence databases, and common and successful DNA sequencing protocols, making this gene target a popular choice for species identification (Deagle *et al.* 2014). Although CO1 may not be diagnostic for a few taxa (Krück *et al.* 2013), the good availability of reference sequences allowed identification of several putative CO1 assays for *M. strigata*, all of which were specific in the initial testing, allowing selection of those showing the best analytical characteristics (efficiency and LoD) for further validation.

The validated assays were designed using TaqMan® MGB chemistry. MGB probes have better sensitivity than other qPCR probes because they are dual-labelled, have more favourable thermodynamic properties and better specificity because they are more sensitive to mismatches (Yao *et al.* 2006). MGB assays also provide good performance in complex environmental samples (Yao *et al.* 2006; Alonso *et al.* 2007; Gasparic *et al.* 2010). All SARDI designed assays, and most implemented in the SARDI testing system, therefore use the MGB format. Amplicon length and the number of mismatches with non-target sequences are also important for assay performance. MstrigCO1-5 produces a short (109 bp) amplicon and its primers and probe have the most mismatches with known sequences of non-target species of the putative assays designed in this study. The target can easily be synthesized to produce laboratory positive control material. The *M. strigata* qPCR assay MstrigCO1-5 therefore satisfies the requirements for a qPCR assay for incorporation into SARDI's detection system. Assay performance will continue to be assessed during implementation, and, if required, the PCR conditions adjusted, or the assay further refined to improve performance. The validation carried out by this study provides confidence in assay performance, and permits calculation of required sample numbers and appropriate interpretation of results where *M. strigata* is included as a target in molecular surveillance using the plankton sampling method. MstrigCO1-5 also shows suitable performance for application to testing scrape samples, but the utility of scrape or settlement plate sampling for *M. strigata* or other pest surveillance will depend on the ability of these sample types to capture targets, which has not yet been assessed.

5. CONCLUSIONS

Two qPCR assays for detection of *M. strigata* were developed and validated in this study. Both validated assays showed adequate specificity and sensitivity for testing both plankton and scrape samples, with MstrigCO1-5 showing greater sensitivity in plankton than MstrigCO1-3, and is therefore preferred for implementation. Performance of MstrigCO1-5 in environmental samples, its laboratory performance (efficiency and analytical limit of detection) and characteristics (primer and probe melt temperatures, MGB probe chemistry) mean this assay is suitable to be implemented at SARDI's Molecular Diagnostics laboratory for use in the panel of marine pest qPCR assays.

Field validation provides confidence that negative results support evidence of absence of the target, but this requires validation of both the sampling method and of the tests applied. The plankton sampling approach combined with qPCR testing has been demonstrated to be fit-for-purpose for marine pest surveillance, and more effective at detecting pests than traditional surveillance using diver visual surveys, dredges or traps (Wiltshire *et al.* 2019a). The *M. strigata* assay MstrigCO1-5 is suitable for application to this sample type and including this assay in the suite of tests would not affect the sample numbers required to provide evidence of absence given its comparable performance to other SARDI assays. Testing plankton samples collected from areas with known *M. strigata* occurrence would however provide additional verification of the assay performance, and permit refinement of the diagnostic performance estimates. MstrigCO1-5 showed similar diagnostic performance in scrape samples to plankton, indicating that the assay is suitable to apply to DNA extracted from scrapes or settlement plates. The utility of these alternative sampling methods for detection will however depend on their effectiveness as capturing targets, which has not yet been assessed. Either sample type may be affected by inhibition. The effect of inhibition on assay performance was minor within the range of inhibition observed, but inhibition in environmental samples should continue to be assessed so that any potential impacts of inhibition on surveillance results can be identified.

Where detections occur in areas without known pest occurrence, the high specificity of the *M. strigata* assay provides confidence that any detection is of this species DNA, with a HTS approach available to provide additional confirmation of positives if required. There are no records of *M. strigata* from Australia, and surveillance using specific, sensitive tools such as the assay developed here can contribute to management steps that will aid in preventing establishment of this important pest in Australia.

6. REFERENCES

Alonso R, Mateo E, Cisterna R (2007) Detection of parC mutations in *Streptococcus pneumoniae* by Real-time PCR and Taqman-MGB probes. *J Microbiol Methods* 69, 214-7.

Armbruster GFJ, BÖHme M, Bernhard D, Schlegel M (2005) The H3/H4 Histone Gene Cluster of Land Snails (Gastropoda: Stylommatophora): Ts/Tv Ratio, Gc3 Drive and Signals in Stylommatophoran Phylogeny. *Journal of Molluscan Studies* 71, 339-348.

Arthur T, Arrowsmith L, Parsons S, Summerson R (2015a) Monitoring for marine pests: A review of the design and use of Australia's national monitoring strategy and identification of possible improvements. Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra.

Arthur T, Summerson R, Mazur K (2015b) A comparison of the costs and effectiveness of prevention, eradication, containment and asset protection of invasive marine species incursions. *ABARES Report, Canberra*.

Ayyagari VS, Sreerama K (2019) Molecular phylogenetic analysis of Pulmonata (Mollusca: Gastropoda) on the basis of Histone-3 gene. *Beni-Suef University Journal of Basic and Applied Sciences* 8.

Banks NC, Paini DR, Bayliss KL, Hodda M (2015) The role of global trade and transport network topology in the human-mediated dispersal of alien species. *Ecol Lett* 18, 188-99.

Beric B, MacIsaac HJ (2015) Determinants of rapid response success for alien invasive species in aquatic ecosystems. *Biological Invasions* 17, 3327-3335.

Bessetti J (2007) An introduction to PCR inhibitors. *Journal of Microbiological Methods* 28, 159-67.

Bott NJ, Giblot-Ducray D (2011a) Molecular tools for detection of marine pests: Development of diagnostic PCR assays for the detection of significant marine pests: *Carcinus maenas, Ciona intestinalis* and *Undaria pinnatifida*. Report prepared for Adelaide & Mt Lofty Ranges Natural Resources Management Board. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2011/000503-1. SARDI Research Report Series No. 587, Adelaide.

Bott NJ, Giblot-Ducray D (2011b) Molecular tools for detection of marine pests: *Musculista senhousia, Corbula gibba* and *Perna canaliculus* quantitative PCR assays. Report prepared for Biosecurity SA. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2010/000991-1. SARDI Research Report Series No. 522, Adelaide.

Bott NJ, Giblot-Ducray D (2012) Development of a qPCR assay for the Pacific Oyster, *Crassostrea gigas*. Report prepared for Biosecurity SA. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2012/000210-1. SARDI Research Report Series No. 641, Adelaide.

Bott NJ, Giblot-Ducray D, Deveney MR (2010a) Molecular tools for detection of marine pests: Development of putative diagnostic PCR assays for the detection of significant marine pests: *Asterias amurensis, Carcinus maenas, Undaria pinnatifida* and *Ciona intestinalis*. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2010/000669-1. SARDI Research Report Series No. 509, Adelaide.

Bott NJ, Giblot-Ducray D, Sierp MT (2012) Development of a qPCR assay for the Black-striped Mussel, *Mytilopsis sallei*. Report prepared for Commonwealth Department of Agriculture, Fisheries and Forestry & Biosecurity SA. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2012/000436-1. SARDI Publication No. 676, Adelaide.

Bott NJ, Ophel-Keller KM, Sierp MT, Herdina, Rowling KP, McKay AC, Loo MGK, Tanner JE, Deveney MR (2010b) Toward routine, DNA-based detection methods for marine pests. *Biotechnology Advances* 28, 706-714.

Branscum AJ, Gardner IA, Johnson WO (2005) Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Prev Vet Med* 68, 145-63.

Burns M, Valdivia H (2008) Modelling the limit of detection in real-time quantitative PCR. *European Food Research and Technology* 226, 1513-1524.

Campbell ML, Gould B, Hewitt CL (2007) Survey evaluations to assess marine bioinvasions. *Marine Pollution Bulletin* 55, 360-378.

Caraguel CGB, Stryhn H, Gagné N, Dohoo IR, Hammell KL (2011) Selection of a Cutoff Value for Real-Time Polymerase Chain Reaction Results to Fit a Diagnostic Purpose: Analytical and Epidemiologic Approaches. *Journal of Veterinary Diagnostic Investigation* 23, 2-15.

Chambert T, Miller DAW, Nichols JD (2015) Modeling false positive detections in species occurrence data under different study designs. *Ecology* 96, 332-339.

Colgan DJ, Ponder WF, Eggler PE (2000) Gastropod evolutionary rates and phylogenetic relationships assessed using partial 28S rDNA and histone H3 sequences. *Zoologica Scripta* 29, 29-63.

Couton M, Comtet T, Le Cam S, Corre E, Viard F (2019) Metabarcoding on planktonic larval stages: an efficient approach for detecting and investigating life cycle dynamics of benthic aliens. *Management of Biological Invasions* 10, 657-689.

Crous PW, Groenewald JZ, Groenewald M, Caldwell P, Braun U, Harrington TC (2006) Species of *Cercospora* associated with grey leaf spot of maize. *Stud Mycol* 55, 189-97.

Darling J, Blum M (2007) DNA-Based Methods for Monitoring Invasive Species: A Review and Prospectus. *Biological Invasions* 9, 751-765.

Darling JA, Galil BS, Carvalho GR, Rius M, Viard F, Piraino S (2017) Recommendations for developing and applying genetic tools to assess and manage biological invasions in marine ecosystems. *Marine Policy* 85, 54-64.

Darling JA, Mahon AR (2011) From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environ Res* 111, 978-88.

DAWE (2018a) Guidelines for development and validation of assays for marine pests. Department of Agriculture and Water Resources, Canberra.

DAWE (2018b) Marine Pest Plan 2018–2023: the National Strategic Plan for Marine Pest Biosecurity. Department of Agriculture and Water Resources, Canberra.

DAWE (2021) The Australian Marine Priority Pests List Accessed: 7 May 2021. URL: https://www.marinepests.gov.au/what-we-do/apmpl

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biol Lett* 10.

Deveney MR, Giblot-Ducray D, Wiltshire KH (2017) Marine pests testing methods: contributing to establishment of an Australian Testing Centre for Marine Pests. SARDI Aquatic Sciences, SARDI Publication No. F2017/000273-1. SARDI Research Report Series No. 963, Adelaide.

Deveney MR, Wiltshire KH, Lieu YN, Buss JJ, Moody NJG, Mohr PA, Ellard KA (2020) Advanced surveillance for OsHV-1 microvariant. In 'Future Oysters CRC-P: Advanced aquatic disease surveillance for known and undefined oyster diseases'. (Eds MR Deveney and KH Wiltshire) pp. 1 - 43. (Fisheries Research and Development Corporation Deakin, ACT)

Díaz-Ferguson EE, Moyer GR (2014) History, applications, methodological issues and perspectives for the use environmental DNA (eDNA) in marine and freshwater environments *Revista de Biología Tropical* 62, 1273-1284.

Floerl O, Inglis G, Peacock L, Plew D (2012) The Efficacy of Settlement Plate Arrays for Marine Surveillance. MPI Technical Paper No: 2012/16. Ministry for Primary Industries, Wellington.

Fuertes VNB, de los Reyes RA, Monteclaro HM (2021) The spread of the non-indigenous mussel species *Mytella strigata* (Hanley, 1843) in the Philippines: Ensuing issues and responses of local communities. *Regional Studies in Marine Science* 41.

Gardner IA, Stryhn H, Lind P, Collins MT (2000) Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Preventive Veterinary Medicine* 45, 107-122.

Gasparic MB, Tengs T, La Paz JL, Holst-Jensen A, Pla M, Esteve T, Zel J, Gruden K (2010) Comparison of nine different real-time PCR chemistries for qualitative and quantitative applications in GMO detection. *Anal Bioanal Chem* 396, 2023-9.

Giblot-Ducray D, Bott NJ (2013) Development of a plankton sampling protocol for molecular testing of marine pests. Report prepared for Biosecurity SA. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2013/000281-1. SARDI Research Report Series No. 734, Adelaide.

Glasby TM, Connell SD, Holloway MG, Hewitt CL (2007) Nonindigenous biota on artificial structures: could habitat creation facilitate biological invasions? *Marine Biology* 151, 887-895.

Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF, McKee A, Oyler-McCance SJ, Cornman RS (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution* 7, 1299-1307.

Guo Y, Pooler M (2021) Real-Time and Conventional PCR Tools for Detection and Discrimination of *Calonectria pseudonaviculata* and *C. henricotiae* Causing Boxwood Blight. *Plant Dis* 105, 164-168.

Haling RE, Simpson RJ, McKay AC, Hartley D, Lambers H, Ophel-Keller K, Wiebkin S, Herdina, Riley IT, Richardson AE (2011) Direct measurement of roots in soil for single and mixed species using a quantitative DNA-based method. *Plant and Soil* 348, 123-137.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In 'Nucleic acids symposium series' pp. 95-98. ([London]: Information Retrieval Ltd., c1979-c2000.)

Hayes K, Sliwa C, Migus S, McEnnulty F, Dunstan P (2005a) National priority pests: Part II. Ranking of Australian marine pests. CSIRO Marine Research, Australian Government Department of the Environment and Heritage.

Hayes KR, Cannon R, Neil K, Inglis G (2005b) Sensitivity and cost considerations for the detection and eradication of marine pests in ports. *Marine Pollution Bulletin* 50, 823-834.

Hayes KR, Sliwa C (2003) Identifying potential marine pests - a deductive approach applied to Australia. *Marine Pollution Bulletin* 46, 91-98.

Hewitt CL, Campbell M (2010) The relative contribution of vectors to the introduction and translocation of marine invasive species. Department of Agriculture, Fisheries and Forestry, Canberra.

Hewitt CL, Martin RB (2001) Revised protocols for baseline port surveys for introduced marine species: survey design, sampling protocols and specimen handling. CRIMP Technical Report 22. Centre for Research on Introduced Marine Pests, Hobart.

Hoedt F, Choat J, Cruz J, Collins J (2001) Sample collection methods and practical considerations for introduced species' surveys at tropical ports. CRC Reef Research Centre Technical Report No. 35. CRC Reef Research Centre, Townsville.

Huang Y-C, Li Z-K, Chen W-L, Chan C-C, Hsu H-Y, Lin Y-T, Huang Y-S, Han Y-S (2021) First record of the invasive biofouling mussel *Mytella strigata* (Hanley, 1843) (Bivalvia: Mytilidae) from clam ponds in Taiwan. *BioInvasions Records* 10, 304-312.

Jayachandran P (2019) First record of the alien invasive biofouling mussel *Mytella strigata* (Hanley, 1843) (Mollusca: Mytilidae) from Indian waters. *BioInvasions Records* 8, 828-837.

Kartavtsev YP, Sharina SN, Chichvarkhin AY, Chichvarkhina OV, Masalkova NA, Lutaenko KA, Oliveira C (2018) Genetic Divergence of Mussels (Mollusca, Mytilidae) Based on the 28S

rRNA, 18S rRNA, and H3 Nuclear Gene Sequences. *Russian Journal of Genetics* 54, 652-669.

Katsanevakis S, Wallentinus I, Zenetos A, Leppäkoski E, Çinar ME, Oztürk B, Grabowski M, Golani D, Cardoso AC (2014) Impacts of invasive alien marine species on ecosystem services and biodiversity: a pan-European review. *Aquatic Invasions* 9, 391-423.

Koziol A, Stat M, Simpson T, Jarman S, DiBattista JD, Harvey ES, Marnane M, McDonald J, Bunce M (2019) Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources* 19, 366-376.

Kralik P, Ricchi M (2017) A basic guide to real time PCR in microbial diagnostics: definitions, parameters, and everything. *Frontiers in microbiology* 8, 108.

Krück NC, Tibbetts IR, Ward RD, Johnson JW, Loh WKW, Ovenden JR (2013) Multi-gene barcoding to discriminate sibling species within a morphologically difficult fish genus (Sillago). *Fisheries Research* 143, 39-46.

Kruschke J (2014) Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan. (Academic Press)

Langlois VS, Allison MJ, Bergman LC, To TA, Helbing CC (2020) The need for robust qPCR-based eDNA detection assays in environmental monitoring and species inventories. *Environmental DNA* 3, 519-527.

Lehtiniemi M, Ojaveer H, David M, Galil B, Gollasch S, McKenzie C, Minchin D, Occhipinti-Ambrogi A, Olenin S, Pederson J (2015) Dose of truth—Monitoring marine non-indigenous species to serve legislative requirements. *Marine Policy* 54, 26-35.

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* 10, 34.

Lim JY, Tay TS, Lim CS, Lee SSC, Teo SLM, Tan KS (2018) *Mytella strigata* (Bivalvia: Mytilidae): an alien mussel recently introduced to Singapore and spreading rapidly. *Molluscan Research* 38, 170-186.

Machida RJ, Leray M, Ho S-L, Knowlton N (2017) Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data* 4, 170027.

Marraffini M, Ashton G, Brown C, Chang A, Ruiz G (2017) Settlement plates as monitoring devices for non-indigenous species in marine fouling communities. *Management of Biological Invasions* 8, 559-566.

Matsuda T, Morishita M, Hinomoto N, Gotoh T (2014) Phylogenetic analysis of the spider mite sub-family Tetranychinae (Acari: Tetranychidae) based on the mitochondrial COI gene and the 18S and the 5′ end of the 28S rRNA genes indicates that several genera are polyphyletic. *PloS one* 9, e108672.

McDonald JI, Wellington CM, Coupland GT, Pedersen D, Kitchen B, Bridgwood SD, Hewitt M, Duggan R, Abdo DA (2020) A united front against marine invaders: Developing a cost-effective marine biosecurity surveillance partnership between government and industry. *Journal of Applied Ecology* 57, 77-84.

McKay Curtis S (2015) mcmcplots: Create Plots from MCMC Output. R package version 0.4.2. https://CRAN.R-project.org/package=mcmcplots.

Meredith M, Kruschke J (2018) HDInterval: Highest (Posterior) Density Intervals. R package version 0.2.0. https://CRAN.R-project.org/package=HDInterval.

Minchin D, Gollasch S, Cohen AN, Hewitt CL, Olenin S (2009) Characterizing Vectors of Marine Invasion. In 'Biological Invasions in Marine Ecosystems: Ecological, Management, and Geographic Perspectives'. (Eds G Rilov and JA Crooks) pp. 109-116. (Springer Berlin Heidelberg: Berlin, Heidelberg) .

Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the Global Threat of Invasive Species to Marine Biodiversity. *Frontiers in Ecology and the Environment* 6, 485-492.

National System for the Prevention and Management of Marine Pest Incursions (2010) Australian marine pest monitoring manual version 2.0. (Department of Agriculture, Fisheries and Forestry: Canberra) .

OIE (2019) Principles and methods of validation of diagnostic assays for infectious diseases. In 'Manual of Diagnostic Tests for Aquatic Animals'. (World Organisation for Animal Health: Paris) .

Ojaveer H, Galil BS, Minchin D, Olenin S, Amorim A, Canning-Clode J, Chainho P, Copp GH, Gollasch S, Jelmert A, Lehtiniemi M, McKenzie C, Mikuš J, Miossec L, Occhipinti-Ambrogi A, Pećarević M, Pederson J, Quilez-Badia G, Wijsman JWM, Zenetos A (2014) Ten recommendations for advancing the assessment and management of non-indigenous species in marine ecosystems. *Marine Policy* 44, 160-165.

Ophel-Keller K, Loo M, Bott N, McKay A (2007) Development of gene probes for introduced marine pest species (*Musculista senhousia, Corbula gibba, Sabella spallanzanii*). A report for Natural Heritage Trust. South Australian Research and Development Institute, Adelaide.

Ophel-Keller K, McKay A, Hartley D, Herdina, Curran J (2008) Development of a routine DNA-based testing service for soilborne diseases in Australia. *Australasian Plant Pathology* 37, 243-253.

Plummer M (2017) JAGS Version 4.3.0 user manual.

Porter TM, Hajibabaei M (2018) Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. *Molecular Ecology* 27, 313-338.

Rebrikov DV, Trofimov DY (2006) Real-time PCR: A review of approaches to data analysis. *Applied Biochemistry and Microbiology* 42, 455-463.

Ruvindy R, Bolch CJ, MacKenzie L, Smith KF, Murray SA (2018) qPCR Assays for the Detection and Quantification of Multiple Paralytic Shellfish Toxin-Producing Species of *Alexandrium*. *Front Microbiol* 9, 3153.

Sanpanich K, Wells F (2019) Mytella strigata (Hanley, 1843) emerging as an invasive marine threat in Southeast Asia. *BioInvasions Records* 8, 343-356.

Sardain A, Sardain E, Leung B (2019) Global forecasts of shipping traffic and biological invasions to 2050. *Nature Sustainability* 2, 274-282.

Schaffelke B, Hewitt CL (2007) Impacts of introduced seaweeds. *Botanica Marina* 50, 397-417.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* 75, 7537.

Simpson T, Bunce M, Stat M, Huerlimann R, Veileux H, Jerry D, Rojahn J, Gleeson D, Furlan E, Deveney M, McDonald J (2018) Partial validation of molecular diagnostic assays for priority exotic marine pests significant to Australia. Final report to the Department of Agriculture and Water Resources. Department of Agriculture and Water Resources, Canberra.

Spinuzzi S, Schneider KR, Walters LJ, Yuan WS, Hoffman EA (2013) Tracking the distribution of non-native marine invertebrates (Mytella charruana, Perna viridis and Megabalanus coccopoma) along the south-eastern USA. *Marine Biodiversity Records* 6.

Stanford University (2015) Recipe for an RNA-later like buffer solution Accessed: 27 November 2015. URL: http://sfg.stanford.edu/RNAbuffer.pdf

Su Y-S, Yajima M (2015) R2jags: Using R to Run 'JAGS'. R package version 0.5-7. https://CRAN.R-project.org/package=R2jags.

Tait LW, Inglis GJ (2016) Optimising settlement arrays for surveillance of non-indigenous biofouling species. New Zealand Ministry for Primary Industries, Wellington.

Tait L, Inglis G, Seaward K (2018) Enhancing passive sampling tools for detecting marine bioinvasions. *Mar Pollut Bull* 128, 41-50.

Thalinger B, Deiner K, Harper LR, Rees HC, Blackman RC, Sint D, Traugott M, Goldberg CS, Bruce K (2021) A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. *Environmental DNA* 10.1002/edn3.189.

Trebitz AS, Hoffman JC, Darling JA, Pilgrim EM, Kelly JR, Brown EA, Chadderton WL, Egan SP, Grey EK, Hashsham SA, Klymus KE, Mahon AR, Ram JL, Schultz MT, Stepien CA, Schardt JC (2017) Early detection monitoring for aquatic non-indigenous species: Optimizing surveillance, incorporating advanced technologies, and identifying research needs. *Journal of Environmental Management* 202, 299-310.

Wakimura K, Takemon Y, Takayanagi A, Ishiwata S-i, Watanabe K, Tanida K, Shimizu N, Kato M (2016) Characterization of genes for histone H3, 18S rRNA, and cytochrome oxidase subunit I of East Asian mayflies (Ephemeroptera). *DNA Barcodes* 4, 1-25.

Wang C, Lin X, Nelson KP (2019) Bayesian hierarchical latent class models for estimating diagnostic accuracy. *Statistical Methods in Medical Research* 10.1177/0962280219852649, 0962280219852649.

Wilkinson MJ, Szabo C, Ford CS, Yarom Y, Croxford AE, Camp A, Gooding P (2017) Replacing Sanger with Next Generation Sequencing to improve coverage and quality of reference DNA barcodes for plants. *Scientific Reports* 7, 46040.

Wiltshire KH (2021) Designing molecular surveillance for marine pests to support domestic ballast water management. Report to Department of Agriculture, Water and the Environment. South Australian Research and Development Institute (Aquatic Sciences), Adelaide. SARDI Publication No. F2021/000152-1, SARDI Research Report Series No. 1095.

Wiltshire KH, Giblot-Ducray D, Prowse T, Deveney MR (2019a) Validation of molecular surveillance: parallel traditional and molecular surveys for priority marine pests in the ports of Gladstone, Brisbane, Melbourne and Hobart. Report to Department of Agriculture and Water Resources. South Australian Research and Development Institute (Aquatic Sciences), Adelaide, SARDI Publication No. F2019/000040-1. SARDI Research Report Series No. 1019.

Wiltshire KH, Prowse T, Buss JJ, Giblot-Ducray D, Lieu YN, Deveney MR (2019b) Diagnostic performance of molecular tests used to detect priority marine pests in plankton samples. South Australian Research and Development Institute (Aquatic Sciences), Adelaide, SARDI Publication No. F2019/000337-1. SARDI Research Report Series No. 1036.

Wiltshire KH, Theil MJ, Giblot-Ducray D, Deveney MR (2019c) Molecular surveys for priority marine pests in 2018-2019 for the ports of Gove, Weipa, Hay Point, Newcastle, Botany Bay, Port Kembla and Devonport. South Australian Research and Development Institute (Aquatic Sciences), Adelaide, SARDI Publication No. F2019/000406-1. SARDI Research Report Series No. 1044.

Wiltshire KH, Theil MJ, Giblot-Ducray D, Deveney MR (2020a) Molecular surveillance for Asian Paddle Crab in the Port River – Barker Inlet. South Australian Research and Development Institute (Aquatic Sciences), Adelaide, SARDI Publication No. F2020/000289-1. SARDI Research Report Series No. 1068.

Wiltshire KH, THeil MJ, Giblot-Ducray D, Deveney MR (2020b) Molecular surveys for priority marine pests in 2020 for the ports of Bunbury, Kwinana, Fremantle and Geraldton. South Australian Research and Development Institute (Aquatic Sciences), Adelaide.

Wood SA, Smith KF, Banks JC, Tremblay LA, Rhodes L, Mountfort D, Cary SC, Pochon X (2013) Molecular genetic tools for environmental monitoring of New Zealand's aquatic habitats, past, present and the future. *New Zealand Journal of Marine and Freshwater Research* 47, 90-119.

Yao Y, Nellaker C, Karlsson H (2006) Evaluation of minor groove binding probe and Taqman probe PCR assays: Influence of mismatches and template complexity on quantification. *Mol Cell Probes* 20, 311-6.

Zaiko A, Pochon X, Garcia-Vazquez E, Olenin S, Wood SA (2018) Advantages and Limitations of Environmental DNA/RNA Tools for Marine Biosecurity: Management and Surveillance of Non-indigenous Species. *Frontiers in Marine Science* 5.

7. APPENDIX

7.1. JAGS model for censored regression

```
model {
# N = number of results
# y = Ct value (censored)
# X = covariate matrix
  for ( i in 1:N ) {
    isCens[i] ~ dinterval( y[i] , CensLim)
    y[i] ~ dnorm( mu[i] , tau )T(0,)
    mu[i] <- inprod(beta[],X[i,]) + a[re[i]]
  }
  for (i in 1:K) {
   beta[i] ~ dnorm(0,0.0001)
  }
  for (i in 1:Nre){
   a[i] ~ dnorm(0,tau.re)
  }
  tau <- 1/pow(sigma,2)
  sigma ~ dunif(0.001,100)
  tau.re <- 1/pow(sigma.re,2)
  sigma.re ~ dunif(0.001,10)
 }
```

7.2. JAGS latent class model

```
model{
# n = number of samples
# K = number of tests
# pi = prevalence (=proportion of spiked samples)
 for(i in 1:n){
  pi[i] <- 5/6
  for (k in 1:K){
   cloglog(se[i,k]) <- max(min(20, a[assay[k]] + bSF[assay[k]] * lnSF[i]),-20)
```

```
   s1[i,k] <- se[i,k]^x[i,k]*((1-se[i,k])^(1-x[i,k]))
   s2[i,k] <- sp[assay[k]]^(1-x[i,k])*((1-sp[assay[k]])^x[i,k])
  }
  for (j in 1:K){
   for (h in 1:K){
    cop[i,j,h] <- c1[j,h]*(-1)^(x[i,j] + x[i,h])/(s1[i,j]*s1[i,h])
    con[i,j,h] <- c2[j,h]*(-1)^(x[i,j] + x[i,h])/(s2[i,j]*s2[i,h])
   }
  }
  eta[i] <- (prod(s1[i,1:K]) *(1+ sum(cop[i,,])))
  theta[i] <-(prod(s2[i, 1:K]) *(1+sum(con[i,,])))
  prob[i] <- max(min(1-1e-9, (pi[i]*eta[i] + (1-pi[i])*theta[i])),1e-9)
  z[i] ~ dpois( - log(prob[i]))
 }
 # Se and Sp
 # Including inhibition effect (bSF) on Se
# Assay 1 = MstrigCO1-3, Assay 2 = MstrigCO1-5
 for(j in 1:2) {
  a[j] <- max(min(14, cloglog(se.int[j])), -14)
  se.int[j] ~ dbeta(1,1)T(1-sp[j], )
  bSF[j] ~ dnorm(0,0.154)T(,0) #95% prob in range -5 to 0
  sp[j] ~ dbeta(361,1)
  for(k in j:3){
   c1a[j,k] ~ dunif(se.lo[j,k], se.hi[j,k])
   c2a[j,k] ~ dunif((sp[j]-1)*(1-sp[k]), (min(sp[j],sp[k])-sp[j]*sp[k]))
   se.lo[j,k] <- (se.int[j]-1)*(1-se.int[k])
   se.hi[j,k] <- (min(se.int[j],se.int[k])-se.int[j]*se.int[k])
  }
 }
 #Reference test (tissue addition)
 a[3] <- max(min(14, cloglog(se.int[3])), -14)
 se.int[3] ~ dbeta(1000,1)T(1-sp[3], )
 bSF[3] <- 0
 sp[3] ~ dbeta(1000,1)
```

```
# covariance components
# c1 = covariance in true positive samples
# c2 = covariance in true negative samples
 for (l in 1:(K-1)){
  for (h in (l+1):K){
   c1[l,h] <- c1a[assay[l],assay[h]]
   c2[l,h] <- c2a[assay[l],assay[h]]
  }
 }
 for (h in 1:K){
  for (l in h:K){
   c1[l,h] <- 0
   c2[l,h] <- 0
  }
 }
}
```