Marine Ecosystems

Assessment of analytical and field performance of qPCR assays for priority marine pests

Wiltshire, K.H., Giblot-Ducray, D., Hill, K., Herdina, Willan, R.C., Ellard, K. and Deveney, M.R.

SARDI Publication No. F2023/000164-1
SARDI Research Report Series No. 1189

SARDI Aquatics Sciences
PO Box 120 Henley Beach SA 5022

Aug 2023

Report to the Department of Agriculture, Fisheries and Forestry

Government of South Australia
Department of Primary Industries and Regions

Australian Government
Department of Agriculture, Fisheries and Forestry

SARDI
SOUTH AUSTRALIAN RESEARCH AND DEVELOPMENT INSTITUTE

Assessment of analytical and field performance of qPCR assays for priority marine pests

Report to the Department of Agriculture, Fisheries and Forestry

Wiltshire, K.H., Giblot-Ducray, D., Hill, K., Herdina, Willan, R.C., Ellard, K. and Deveney, M.R.

SARDI Publication No. F2023/000164-1
SARDI Research Report Series No. 1189

Aug 2023

The South Australian Research and Development Institute respects Aboriginal people as the state's first people and nations. We recognise Aboriginal people as traditional owners and occupants of South Australian land and waters. We pay our respects to Aboriginal cultures and to Elders past, present and emerging.

This publication may be cited as:
Wiltshire, K.H., Giblot-Ducray, D., Hill, K., Herdina, Willan, R.C., Ellard, K. and Deveney, M.R. (2023). Assessment of analytical and field performance of qPCR assays for priority marine pests. Report to the Department of Agriculture, Water and the Environment. South Australian Research and Development Institute (Aquatic Sciences), Adelaide. SARDI Publication No. F2023/000164-1. SARDI Research Report Series No.1189. 111pp.

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., Hill, K., Herdina, Willan, R.C., Ellard, K. and Deveney, M.R.
Reviewer(s):	Sarah Catalano (SARDI) Kevin Mark (SARDI)
Approved by:	Simon Goldsworthy Science Leader – Marine Ecosystems
Signed:	
Date:	4th August 2023
Distribution:	DAFF, 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, Fisheries and Forestry (DAFF) and we are grateful to Nick Fahy and Jason Bayly-Stark for assistance and support. This project would not have been possible without the support of the institutions and individuals who provided specimens or material for testing. We gratefully acknowledge:

- Kevin Spiby from Two Oceans Aquarium, South Africa, for specimens of *Perna perna*
- Justin McDonald, Matthew Hewitt and Claire Wellington from Department of Primary Industries and Regional Development, Western Australia, for specimens of *Arcuatula senhousia* and tissue samples of *Didemnum vexillum, Eriocheir sinensis* and *Rhithropanopeus harrisii*.
- Richard Stafford-Bell from The Department of Jobs, Precincts and Regions, Victoria, and Craig Sherman from Deakin University for specimens, tissue and extracted DNA of *Hemigrapsus sanguineus*.
- Gavin Dally at The Museum and Art Gallery of the Northern Territory for specimens of *Mytilopsis sallei* and *Perna viridis*, and of bivalves for specificity testing.
- Andy Shinn of Benchmark Inve, Thailand, for additional specimens and tissue samples of *Mytilopsis sallei* and *Perna viridis*.
- Cathryn Abbott and Liane Stenhouse of The Department of Fisheries and Oceans, Canada, for tissue samples of *Didemnum vexillum*.
- Woodbridge School Marine Discovery Centre (Tasmanian Department of Education) and The Department of Natural Resources and Environment, Tasmania, for specimens of *Maoricolpus roseus* and *Mya japonica.*
- Russel Pfau of Tarleton State University, Texas, and Elizabeta Briski and Gregor Steffen from Helmholtz Centre for Ocean Research Kiel for specimens of *Rhithropanopeus harrisii*.
- Sally South and Shirley Sorokin of the South Australian Museum for tissue samples of *Charybdis japonica* and of native crab species for specificity testing.
- Alex Chalupa of Biosecurity South Australia for additional material of *Charybdis japonica*.

Plankton sampling in Melbourne was carried out by staff from The Department of Jobs, Precincts and Regions, Victoria. Sharon Drabsch from SARDI Aquatic Sciences assisted with fieldwork in Adelaide, specimen identification, and laboratory sample processing and preparation. Additional field support was provided by Jason Nichols and Jonathon Western. At SARDI Aquatic Sciences, Leonardo Mantilla also assisted with specimen identification and with arranging specimen transport, and we thank Prof Simon Goldsworthy and Dr Mike Steer for their support. At SARDI Crop Sciences, 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.

This report was reviewed by Sarah Catalano and Kevin Mark and cleared for release by Dr Mike Steer, Research Director SARDI Aquatic & Livestock Sciences. We thank Dr Thomas Prowse from the University of Adelaide for his review of statistical methodology used in this report.

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

EXECUTIVE SUMMARY

Molecular techniques for marine pest surveillance offer cost and time savings over traditional techniques, but for molecular surveillance to be routinely applied and used in management frameworks, it is important to understand the performance of the survey system. SARDI has developed a molecular surveillance system using plankton tows tested with species specific quantitative polymerase chain reaction (qPCR) assays for detection of priority marine pests. This molecular surveillance system has been shown to be fit-for-purpose, but it is important to assess the accuracy of each qPCR assay applied, because performance characteristics can vary between these. It is important to understand diagnostic sensitivity (DSe: likelihood of detection where pest DNA in present) and diagnostic specificity (DSp: likelihood of non-detection where pest DNA is absent) of each assay used so that surveys can be adequately designed and appropriately interpreted. Validation of assay performance comprises several stages, including assessment of analytical characteristics in the laboratory, which is typically done during assay development, followed by quantification of diagnostic performance in field samples.

Nineteen qPCR assays for marine pest species have been implemented in the SARDI testing system to date. Field diagnostic performance had been determined for the qPCR assays for six pests: *Asterias amurensis, Carcinus maenas, Magallana gigas, Mytella strigata, Sabella spallanzanii* and *Undaria pinnatifida*, with DSe being 73 to 91%, and DSp > 99% for each assay. The qPCR assays for the 13 other marine pest species had all undergone at least initial analytical assessment but required further validation. This project carried out additional analytical assessment of these 13 assays where required and determined diagnostic performance (DSe and DSp) of each assay in plankton samples. For *Hemigrapsus sanguineus*, an invasive crab recently established in Melbourne, this project also tested plankton samples from areas where this species is recorded to provide additional data on assay field performance.

As a result of this study, data on diagnostic performance is now available for all assays that have been implemented in the SARDI testing system, allowing surveillance to be designed and appropriately interpreted for 17 pest species. Of the 13 assays assessed in this project, those for eleven marine pests: *H. sanguineus, Eriocheir sinensis, Rhithropanopeus harrisii, Charybdis japonica*, *Didemnum vexillum*, *Perna perna, P. viridis, P. canaliculus, Mytilopsis sallei, Mya japonica* and *Maoricolpus roseus* are suitable for application to routine surveillance. DSe of these assays in plankton samples ranged from > 79 to 99% and DSp estimates for these assays were 93.4 – 100.0%. The lowest estimate was for the *P. perna* assay, which also detects

P. canaliculus DNA. Field testing showed that DSp of this assay is likely to be close to 100% in samples from Australia where *P. canaliculus* is absent, but the assay should not be applied without also testing for *P. canaliculus* to avoid ambiguity.

Hemigrapsus sanguineus DNA was successfully detected in plankton samples collected in Melbourne, including in archived DNA from samples collected in November 2017 and March 2018. Spring – summer appears to be more suitable for detection of this species than winter, although further investigation is needed to confirm seasonal patterns. Plankton samples from areas with known occurrence of the pests that are exotic to Australia, from either their native or current invaded ranges, should be obtained and tested if possible. This would provide confirmation of the efficacy of plankton samples for environmental detection of these pests and, if samples can be obtained from multiple seasons, also provide preliminary data on seasonality of detections in plankton.

An approach for confirming detections using high throughput sequencing (HTS) was trialed for the newly developed assays for *Mya japonica, Maoricolpus roseus* and *H. sanguineus*, and for assays for the invasive bivalves *Arcuatula senhousia* and *Varicorbula gibba* that were suspected to detect non-target DNA in plankton samples from some locations, creating uncertainty about the occurrence of these two species. HTS testing of samples from locations with known occurrence of *H. sanguineus, Mya japonica* and *Maoricolpus roseus* confirmed that assay detections were of the target species in each case, further supporting the specificity of each assay.

The assays for *A. senhousia* and *V. gibba* had DSe of 93.2 and 90.0% respectively but DSp was lower than other assays (90.5% and 70.5%). HTS data were inconclusive in confirming target species presence, but data suggest that the 28S gene region targeted by these two assays does not provide sufficient resolution for separating the target species from related taxa. Several sequences were identified that do not match those of the target species but would likely be amplified by these assays. The identity of most potential cross-reacting taxa could not be determined. In addition to unidentified taxa, cross-reaction is likely for the *V. gibba* assay with DNA of bivalves in the Arcidae (including *Anadara trapezia*), and the *A. senhousia* assay with other Mytilidae. Given that detections by these assays may represent cross-reactions and that HTS approaches were unsuitable for confirmatory testing, these two assays will be re-designed.

Keywords: Marine pests, molecular surveillance, plankton sampling, qPCR, assay validation.

GLOSSARY

Table 1. Definition of abbreviations and technical terms used in the text.

Term	Definition
28S	A gene barcode region that makes up part of the large subunit in ribosomal DNA, and can be used for species identification in some taxa.
3' end / 5' end	Nucleic acid strands are inherently unidirectional, with the terms 3' end / 5' end used to specify direction. For the determination of *in silico* specificity of primer sequences, the location of mismatches relative to the 3' or 5' end is important, as well as the total number of mismatches. Mismatches at the 3' end are much more likely to prevent binding than mismatches towards the 5' end.
Amplicon	A DNA fragment produced by amplification, e.g., by PCR.
AP	Apparent prevalence. The proportion of samples returning a detection or, equivalently, the likelihood of detection in a sample.
Barcode	A gene region used for molecular taxonomic identification.
Base pair (bp)	Two complementary DNA nucleotide bases that pair together in a double-stranded nucleic acid molecule. Used as a measure of DNA sequence length.
Beta	A probability distribution that can be used to model probabilities / proportions.
Binary / Binomial / Bernoulli	Data with two possible outcomes / categories, e.g., yes / no, detected / non-detected, died / survived. Binary/Binomial data refers to the outcome of a single trial or test, e.g., 0,1, while Bernoulli data refers to binomial data from multiple tests, e.g., z detections from N samples.
Bioinfomatic pipeline	A series of algorithms used for the processing of HTS data.
BLAST	Basic local alignment search tool. A program for comparing DNA sequences
cloglog	Complementary log-log. A link used in modelling binary data: $\eta = \log(-\log(1 - \pi)$ where η is the linear predictor for probability π.
COI	Cytochrome c oxidase subunit 1, a gene barcode region found in the mitochondrial genome that is widely used for species identification.
confounding sequence / cross-reaction	A confounding sequence of DNA is a sequence from an organism that is not the target of a qPCR assay, but which has sufficient matches with the assay primers and probe that amplification, and hence, non-target detection is likely. A cross-reaction is the reaction of the assay with a confounding sequence that results in non-target detection.
C_T	Cycle threshold. In qPCR, the PCR cycle at which fluorescence exceeds a threshold and a detection is recorded. C_T is inversely related to the quantity of target DNA present (i.e., greater DNA content results in faster amplification and lower C_T).
Cyt-B	Cytochrome B, a gene barcode region found in the mitochondrial genome that can be used for species identification in some taxa.
Delaunay triangulation	The most efficient set of non-overlapping triangles that cover a specified area. Constrained refined Delaunay triangulation enforces alignment of edges and location of vertices to achieve a maximum edge length.
DNA	Deoxyribonucleic acid. The molecule carrying genetic information for an organism.

Term	Definition
DNA yield	In qPCR, the sample DNA content as calculated from the C_T value using a standard curve. Depending on the type of standard used, DNA yield may be expressed in pg g^{-1} or kDNA copies g^{-1}.
dsDNA	Double-stranded DNA.
DSe	Diagnostic sensitivity = likelihood of detection by a test when target is present.
DSp	Diagnostic specificity = likelihood of non-detection by a test when target is absent.
Efficiency	In PCR, the proportion of target DNA molecules successfully amplified at each cycle. For an assay with perfect (100%) efficiency, the amount of target DNA will double at each PCR cycle.
Exotic	Invasive species with potential to be introduced but not known to be present in a region, or, if present, not established and subject to eradication.
FN	False negative, failure to detect a target that is present.
FP	False positive, apparent detection of a target not actually present.
Gamma	A probability distribution that can be used to model continuous strictly positive data.
Gaussian	The Normal probability distribution.
gDNA	genomic DNA extracted from organism tissues.
Genbank	Repository of publicly available DNA sequences. .https://www.ncbi.nlm.nih.gov/genbank/
GLM / GLMM	Generalised linear model / Generalised linear mixed model. A modelling technique where a link function is applied, and data can be modelled using a range of error distributions .
GMRF	Gaussian Markov Random Field. The spatial field (random effects) estimated by R-INLA using the SPDE approach. The GMRF has zero mean and Matérn correlation structure between points.
HDI	Highest density interval, the smallest interval containing a given probability mass (e.g., 95%) of a distribution. See Kruschke (2014) for details.
HTS	High throughput sequencing. A method for obtaining sequences from many (potentially millions of) DNA fragments simultaneously HTS can be used to simultaneously sequence all amplicons present in a sample.
IGS	Intergenic spacer, a barcode region in the large subunit in ribosomal DNA that can be used for species identification in some taxa.
Inhibition	In PCR, the presence of compounds that slow or prevent amplification and so may result in non-detection or increased C_T in qPCR.
INLA / R-INLA	Integrated Nested Laplacian Approximation – a method for fitting Bayesian models using numerical approximation / the R software package for INLA model fitting.
JAGS	Software used for model fitting using MCMC.
Library	In HTS, a set of amplicons to be sequenced.
LCM	Latent class model, a standard modelling approach for statistical estimation of DSe and DSp.
LM / LMM	Linear model / Linear mixed model. A modelling method assuming a linear relationship and normally distributed errors.

Term	Definition
lnSF	Natural logarithm of scale factor. Used as a covariate in modelling. See also scale factor and inhibition.
LoD	Limit of detection. The lowest concentration of target DNA able to be reliably detected (typically defined as 95% likelihood of detection under laboratory conditions) by an assay.
Matérn correlation	A distance-based correlation function forming the basis of the correlation matrix of spatial random effects in R-INLA.
MCMC	Markov chain Monte Carlo, a method for obtaining model estimates.
Mixed model	A model including both fixed and random effects.
Nucleotides	The basic structural units of nucleic acid.
OTU	Operational taxonomic unit. A grouping of highly similar sequences from HTS.
PC prior	Penalised complexity prior. A type of prior implemented in R-INLA that can be used for strictly positive parameters including variance. See Simpson *et al.* (2017).
PCR / qPCR	Polymerase chain reaction / quantitative PCR. A method for amplifying and detecting target DNA. PCR uses a pair of primers (forward and reverse) for amplification. Each PCR cycle comprises a denaturing stage where double-stranded DNA is separated by heating, and an annealing stage, when temperature is reduced and primers attach to the target DNA and initiate amplification for that cycle. Amplicons produced by PCR can be used for sequencing (e.g. HTS). qPCR uses a primer pair for amplification but also measures target DNA amplification at each PCR cycle using a fluorescent probe. See also Primers, Probe and C_T.
Prevalence	(True) prevalence is the proportion of samples containing a target (equivalently, likelihood of a given sample containing the target). See also AP.
Primers	Short single-stranded DNA sequences that bind to target DNA regions and initiate amplification in PCR. Generic (or universal) primers are designed to provide amplification of a target gene region for a wide range of taxa (e.g., for community-level sequencing), while primers for species-specific assays (including qPCR) are designed to bind only to target species DNA. Semi-generic or targeted primers are designed to amplify a selected range of taxa or sequences.
Priors	In Bayesian analysis, prior information needs to be specified for each parameter to be estimated. Vague or uninformative priors have little to no effect on model estimates, while informative priors may influence estimates. The influence of a prior depends on the relative strength of the prior and on the amount of data used, with the use of more data reducing influence of priors.
Probe	In qPCR, a short sequence attached to a fluorophore than produces fluorescence when it binds to target DNA.
R	Open-source statistical software.
Range	In R-INLA spatial models, range is the distance at which the Matérn correlation between two points decays to 0.1, i.e., beyond which spatial locations are not considered correlated.
Reads	The number of sequences (total or for a specific OTU) returned by HTS.
Scale factor	A measure of inhibition used in PCR. The scale factor is calculated from the DNA yield of an internal control as the ratio of DNA yield in reference samples (known to have no inhibition) to that in test samples.

Term	Definition
Sensitivity	Ability of an assay to detect a target. Apparent sensitivity is the proportion of detections occurring in samples considered to contain target DNA. Analytical or laboratory sensitivity refers to assay LoD and efficiency. See also DSe.
Sequencing	Determining the order of nucleotides in a DNA fragment.
SPDE	Stochastic Partial Differential Equations. An approach used in R-INLA for the estimation of spatial random effects.
Specificity	Ability of an assay to accurately distinguish target from non-target DNA. Apparent specificity is the proportion of non-detections occurring in samples considered free of target DNA. Analytical or laboratory specificity refers to non-detection when testing non-target gDNA, while *in silico* specificity refers to computer-based comparison of assay primer and probe sequences with non-target sequences. See also DSp.
Standard curve	In qPCR, used to relate C_T value to DNA yield in a sample, based on the C_T values returned by a dilution series of a DNA standard for the assay.
Survey confidence	The likelihood of detecting a target in at least one sample in a survey.
ZAG	Zero added Gamma. A modelling approach to simultaneously estimate coefficients for a binary and continuous, positive (Gamma distributed) component, e.g., likelihood of detection and DNA yield in samples with detection by qPCR.

1. INTRODUCTION

1.1. Background

1.1.1. Surveillance for marine pests

Shipping is a major vector for aquatic species introductions via propagules in ballast water or hull-fouling (Hewitt *et al.* 2007; Molnar *et al.* 2008; Minchin *et al.* 2009; Hewitt and Campbell 2010), and while not all introduced species become pests, those that do have wide ranging impacts on ecosystems, marine industries, infrastructure, and amenity (Hayes *et al.* 2005; Schaffelke and Hewitt 2007; Molnar *et al.* 2008; Katsanevakis *et al.* 2014). Ports are at high risk of shipping-mediated introductions because they provide suitable conditions for establishment of marine pests and ports may also act as nodes for further spread of established pests (Hayes *et al.* 2005a; Schaffelke and Hewitt 2007; Molnar *et al.* 2008; Katsanevakis *et al.* 2014). Understanding the occurrence of pests at ports is consequently important for their management and prevention of spread, and for enabling early detection of new incursions (Bott *et al.* 2010b; Lehtiniemi *et al.* 2015). Surveillance for invasive marine species at ports is therefore a key component in managing marine pest risks.

Molecular, i.e., DNA-based, methods are of interest for surveillance because they can provide results rapidly and are considerably cheaper than traditional surveillance methods that are based on collection and morphological identification of specimens (Bott *et al.* 2010b; Darling and Mahon 2011; Darling *et al.* 2017; Trebitz *et al.* 2017; DAWE 2018). Molecular surveillance is therefore being considered for surveillance to inform the risk of domestic ballast water transport for key marine pests (Wiltshire 2021, 2023). Where surveillance is used to inform management, it is important that surveillance can be designed to provide adequate confidence of detection, and that results are appropriately interpreted (Darling and Mahon 2011; Darling *et al.* 2017; Trebitz *et al.* 2017; DAWE 2018). To facilitate survey design and the interpretation of results, including the 'sight un-seen' detections that are provided by molecular surveys, an understanding of the performance of the survey system, specifically, the likelihood of survey errors, is required (Darling and Mahon 2011; Goldberg *et al.* 2016).

1.1.2. Understanding survey performance

There are two types of survey error: false negatives (failure to detect a species where present), and false positives (apparent detection of a species that is not present) (Figure 1). Knowledge of

the false negative rate is required to enable the design of surveillance to achieve a target survey confidence, i.e., likelihood of detection in at least one sample of a survey (Hayes *et al.* 2005b; Kean *et al.* 2015; Wiltshire 2021). Understanding survey false negative and false positive rates enables appropriate interpretation of surveillance results by allowing estimation of how likely a target may be to occur despite non-detection, and of the relative likelihood of detections being true or false positives (Royle and Link 2006; Low-Choy 2013; Stanaway 2015; Wiltshire 2023).

For a species that is present to be detected in a survey, it firstly needs to be collected by the sampling method used. The likelihood of a survey collecting the target in at least one sample depends on the number of samples collected and the likelihood of a given sample capturing the target. The per-sample capture likelihood will depend on the density or concentration of the target relative to the sample area or volume, and the effectiveness of the method at capturing the target (Hayes *et al.* 2005b; Royle and Dorazio 2009; Kean *et al.* 2015). False negatives may therefore result from insufficient or ineffective sampling in either molecular or traditional surveillance. In molecular surveillance, false negatives may also occur if captured DNA is lost due to degradation or extraction failure (Díaz-Ferguson and Moyer 2014; Goldberg *et al.* 2016).

The South Australia Research and Development Institute (SARDI) has developed a molecular surveillance system using plankton tows tested with quantitative polymerase chain reaction (qPCR) assays (Giblot-Ducray and Bott 2013; Deveney *et al.* 2017). Wiltshire *et al.* (2019a) assessed the performance of the molecular surveillance system relative to appropriate traditional methods for six priority pests (diver visual surveys, traps and trawls) and demonstrated that the molecular method was able to detect those pests present in surveyed ports more effectively than traditional methods. Their analysis of results demonstrated that the higher performance of the molecular method was due predominantly to the much greater relative abundance of the planktonic stages targeted by this method in comparison to the adult pest stages targeted by traditional methods. The greater relative abundance of planktonic targets leads to a greater per-sample capture likelihood by the molecular method, meaning that target species can be detected with lower sampling effort than by traditional surveillance (Wiltshire *et al.* 2019a). The planktonic abundance of many species varies seasonally, however, meaning that selection of an appropriate sampling time is important to minimise the risk of sampling false negatives when using the molecular method (de Souza *et al.* 2016; Buxton *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire 2021).

Following successful capture in a sample, targets need to be accurately identified for detection to occur. In traditional surveillance where physical specimens are collected, morphological identification can provide definitive identification, subject to the availability of suitable taxonomic expertise. False negatives may occur if target species are present but not identified, and false positives may occur if non-target species are mistakenly identified as targets. These possibilities are minimised by the use of relevant expertise, but access to taxonomists can be a limiting factor for effective traditional surveillance (Bott *et al.* 2010b; Bott 2015; Zaiko *et al.* 2018). In molecular surveillance, succesful and accurate detection of collected DNA depends on the diagnostic performance of the assays used (Darling and Mahon 2011; Díaz-Ferguson and Moyer 2014; Goldberg *et al.* 2016; Trebitz *et al.* 2017). The likelihood of an assay detecting target DNA when present is known as diagnostic sensitivity (DSe), and assay DSe therefore contributes to the likelihood of survey false negatives in molecular surveillance.

In PCR-based technologies such as qPCR, assay DSe may be reduced where compounds are present that inhibit the PCR reaction (= PCR inhibition) (Bessetti 2007; Kralik and Ricchi 2017). Inhibition may occur in environmental samples (Goldberg *et al.* 2016), and understanding the effect of inhibition on assay DSe, which may vary between assays (Lance and Guan 2020), is therefore also important in assessing the likelihood of survey false negatives.

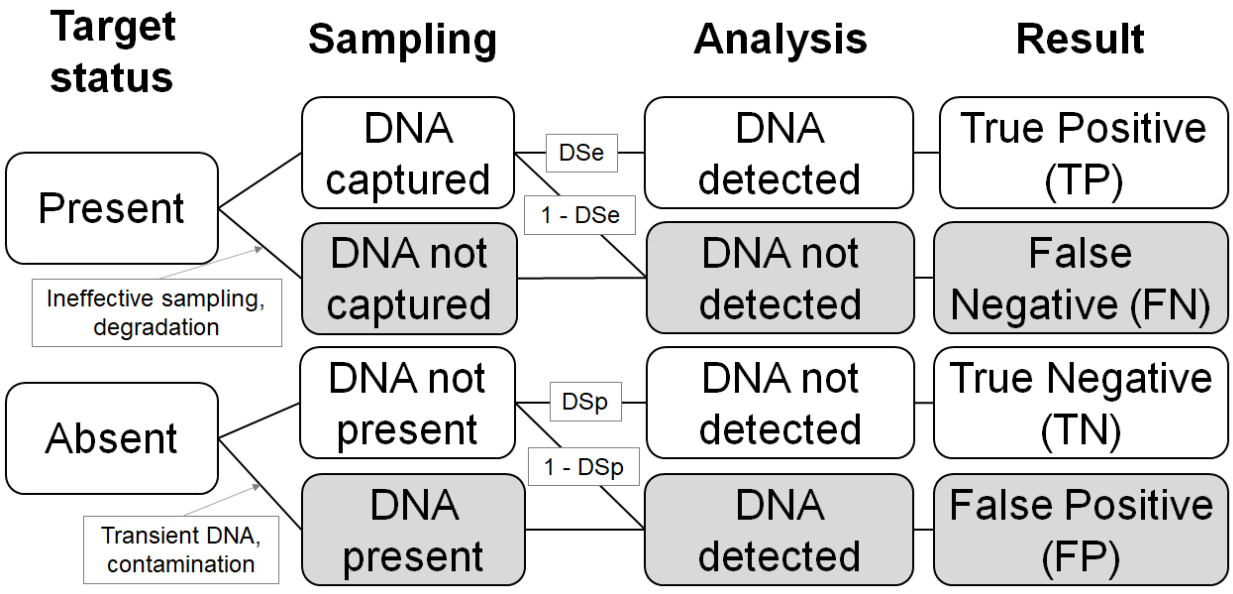

Figure 1. Process for accurate detection of a target by molecular surveillance showing factors at sampling and analysis stages that may contribute to survey errors (false negatives, false positives). Steps leading to survey error are shown in grey boxes. DSe = Diagnostic sensitivity, DSp = Diagnostic specificity.

Where a target is absent, false positives in molecular surveillance could also occur at both sampling and analysis stages. At the sampling stage, target DNA may occur without species presence due to cross-contamination during sample collection or processing, or because, potentially non-viable, target DNA is present from a transient source such as ballast water release, wastewater, or excretion from predators (Díaz-Ferguson and Moyer 2014; Goldberg *et al.* 2016; Hinlo *et al.* 2017; Stoeckle *et al.* 2017; Baillie *et al.* 2019). At the analysis stage, false positives may occur if the assay cross-reacts with non-target DNA in a sample, resulting in a detection despite the absence of target DNA. The likelihood of an assay returning a non-detection where target DNA is absent is known as DSp, and assay DSp therefore contributes to the likelihood of survey false positives in molecular surveillance.

1.1.3. Assay performance assessment for molecular surveys

Wiltshire *et al.* (2019a) demonstrated the overall performance of the molecular surveillance system, showing that the methods for sampling, DNA preservation and extraction were all effective and fit-for-purpose for marine pest surveillance. It is important, however, that performance of the individual assays applied is assessed, to account for any differences in assay DSe and DSp across target species of a survey.

Assay performance is assessed in stages, with initial laboratory assessment ensuring assays have suitable analytical characteristics before verification of assay performance using relevant field samples (Goldberg *et al.* 2016; Darling *et al.* 2017; DAWE 2018; Thalinger *et al.* 2021). Analytical assessment is typically carried out during assay development and includes: design to ensure *in silico* specificity (i.e., selection of a suitably diagnostic gene region to detect the target while not detecting other organisms); testing genomic DNA (gDNA) of the target and non-target species to verify that the assay can detect the target and does not cross-react with DNA of other species such as co-occurring relatives; and development of a standard curve to determine assay efficiency and limit of detection (DAWE 2018; Thalinger *et al.* 2021). Assays that show suitable analytical characteristics (i.e., a limit of detection that would enable the assay to detect target DNA in typical environmental samples and a lack of cross-reactivity with non-target DNA) can progress to implementation, but still require validation of their field performance before application to surveillance to support management (Darling *et al.* 2017; DAWE 2018; Thalinger *et al.* 2021), while assays with inadequate analytical performance require re-design and reassessment before application.

Field validation involves the application of assays to typical environmental samples, i.e., plankton samples in the case of the SARDI plankton tow based molecular method. Initial field assessment includes verification that the assay can detect target DNA in typical environmental samples and does not cross-react with DNA from non-target organisms (Thalinger *et al.* 2021), noting that initial assessment of specificity covers only a limited number of species in comparison to the myriad found in environmental samples, many of which have not been sequenced and may not be described. Operational validation then requires characterisation of DSe and DSp, so that knowledge of these parameters can be used to effectively design and interpret molecular surveys based on the target species, and therefore assays applied, in a survey (DAWE 2018; Wiltshire *et al.* 2019b; Wiltshire 2021, 2023).

Assay DSp should be as near to 100% as possible, because false positive detections complicate the interpretation of surveillance results (Darling *et al.* 2020; Sepulveda *et al.* 2020; Wiltshire 2023). Assay DSe determines the number of samples required to achieve a target survey confidence, with more samples required when using an assay with lower DSe (Goldberg *et al.* 2016; Wiltshire 2021). A survey design tool, which calculates the number of samples required for the plankton sampling method, while accounting for assay DSe, was developed by Wiltshire (2021) and is being applied to design molecular surveys for Australian ports (e.g., Wiltshire *et al.* 2022). Data on assay DSe is therefore required to accurately apply this tool and determine a suitable sampling effort, dependent on the target species and therefore assays applied to a survey. It should be noted that, while collection of additional samples can compensate for lower DSe, if assay DSe is too low, the number of samples required may become infeasible to collect, or surveillance may become cost ineffective. A minimum DSe of ~70% is therefore desirable for assays that are applied to routine surveillance.

1.2. Validation status of assays for marine pests

Nineteen qPCR assays for priority marine pests have been implemented in the SARDI molecular testing system to date (Table 2). All these assays were designed for *in silico* specificity, tested for specificity against gDNA extracts, and assessed for analytical performance as part of their development (see references in Table 2). Validation to an operational level, was, however, complete for only six of the 19 assays as of June 2022 (Table 2). The validation steps completed by previous projects are summarised in this section, with steps requiring completion to operationally validate each assay (as shown in Table 2) detailed in section 1.3, and the specific objectives of this project summarised in section 1.4.

Table 2. Assays for priority marine pests available in the SARDI testing system showing validation status as of June 2022. ‡Species considered in risk tables for domestic ballast water management (Zhao *et al.* 2012). †Species on the Australian priority marine pest list. *Species on the National Priority List of Exotic Environmental Pests, Weeds and Diseases. Numbers show reference for completed validation steps, '~' shows partially completed steps (see text for details) and 'x' shows steps requiring validation that are addressed in the current project.

Species	Common name	Assay name	Gene target region	Design & laboratory assessment	Field specificity	Detection in plankton	DSe/DSp
Arcuatula (=*Musculista*) *senhousia*‡	Asian bag mussel	Asen	28S	1	2,3~	2,3	x
Perna canaliculus†*	NZ green mussel	Pcan	IGS	1	2,3	x	x
Perna perna†*	Brown mussel	Pper	COI	4	x	x	x
Perna viridis†*	Asian green mussel	Pvir	COI	4	x	x	x
Mytella strigata (=*M. charruanna*)†	Charru mussel	Mstr	COI	5	5	5	5
Mytilopsis sallei†*	Black-striped false mussel	Msal	COI	6	2,3	x	x
Varicorbula (= *Corbula*) *gibba*‡	European basket shell	Vgib	28S	1, 7	2,3~	2,3	x
Mya japonica	Japanese soft-shell clam	Mjap	COI	8	x	8	x
Magallana (=*Crassostrea*) *gigas*‡	Pacific oyster	Mgig	COI	9	2,3	2,3	10
Maoricolpus roseus	NZ screw shell	Mros	COI	8	x	8	x
Asterias amurensis†‡	Northern Pacific seastar	Aamu	COI	11, 12	2,3	2,3	10
Carcinus maenas†‡	European shore crab	Cmae	COI	12, 13	2,3	2,3	10
Hemigrapsus sanguineus	Asian shore crab	Hsan	COI	14, 15~	x	x	x
*Charybdis japonica**	Asian paddle crab	Cjap	COI	4	x	x	x
Rhithropanopeus harrisii†*	Harris crab	Rhar	COI	4	x	x	x
Eriocheir sinensis†*	Mitten crab	Esin	Cyt-B	16~	x	x	x
Sabella spallanzanii‡	European fanworm	Sspa	COI	7	2,3	2,3	10
*Didemnum vexillum**	Carpet sea squirt	Dvex	COI	4	x	x	x
Undaria pinnatifida†‡	Japanese seaweed	Upin	COI	12, 13	2,3	2,3	10

1. Bott and Giblot-Ducray (2011b), **2**. Deveney *et al.* (2017), **3.** Wiltshire *et al.* (2019a), **4.** Simpson *et al.* (2018), **5.** Wiltshire *et al.* (2021b), **6.** Bott *et al.* (2012), **7.** Ophel-Keller *et al.* (2007), **8.** Giblot-Ducray *et al.* (2022), **9.** Bott and Giblot-Ducray (2012), **10.** Wiltshire *et al.* (2019b), **11.** Bax *et al.* (2006), 1**2**. Bott *et al.* (2010a), **13.** Bott and Giblot-Ducray (2011a), **14.** Knudsen and Møller (2020), **15.** Wiltshire *et al.* (2021a), **16.** Andersen *et al.* (2018).

Field performance of nine assays that were developed prior to 2015 was assessed by Deveney *et al.* (2017), who applied these assays to plankton samples collected in Darwin, Cairns, Sydney, Melbourne, Hobart, Adelaide and Perth. Results showed that seven of the assays (those for pests of domestic ballast water concern: Aamu, Cmae, Sspa, Asen, Mgig, Vgib and Upin) were able to detect the target pests in locations where they are known to occur, while the other two target pests (*Perna canaliculus* and *Mytilopsis sallei*) were not detected. The latter two species are exotic to Australia, and therefore not expected to be detected in these samples. The majority of the assays showed high field specificity, with no detections recorded in areas without known occurrence, except for the Vgib assay, which returned multiple detections from Darwin, and the Asen assay, which returned a single detection each in Cairns and Darwin (Deveney *et al.* 2017). The results for Vgib strongly suggest a cross-reaction with non-target DNA, because environmental conditions in Darwin are unsuitable for this species (Summerson *et al.* 2016), but the Asen results were inconclusive, given the lack of other surveillance for this species in these ports, and singular detections, which could result from transient DNA sources rather than viable species presence.

Further validation of assay field performance was carried out by Wiltshire *et al.* (2019a), using parallel molecular and traditional surveillance in Gladstone, Brisbane, Melbourne and Hobart. These surveys applied assays for six of the ballast water priority species: Aamu, Cmae, Sspa, Asen, Mgig and Upin, and also included the Pcan and Msal assays. The pests present in Australia were reliably detected by the molecular method (Wiltshire *et al.* 2019a). The exotic *Perna canaliculus* and *Mytilopsis sallei* were not detected by either molecular or traditional sampling, providing further confidence that the Pcan and Msal assays do not cross-react with non-target DNA in Australian samples, and that a lack of molecular detection is due to species absence rather than assay failure. The Asen assay, however, returned multiple detections in Gladstone, where this species is not recorded and was not found in traditional surveillance, suggesting that this assay may cross-react with non-target DNA from a native tropical relative in this location (Wiltshire *et al.* 2019a).

The surveys carried out by Deveney *et al.* (2017) and Wiltshire *et al.* (2019a) demonstrated the suitability of the molecular system using plankton tows and qPCR assays for marine pest surveillance and showed that five of the assays performed well (Aamu, Cmae, Sspa, Mgig and Upin), while identifying probable specificity of the Pcan and Msal assays and specificity problems with the Vgib and Asen assays. Data from these surveys, however, do not facilitate calculation of assay DSe and DSp. DSe and DSp were characterised for the five qPCR assays that performed well in the initial field validation using high throughput sequencing (HTS) as a comparative testing method (Wiltshire *et al.* 2019b). These assays all showed high (> 99%) DSp, while DSe was 73 to 91% (Wiltshire *et al.* 2019b).

The project that developed the Mstr assay included validation of assay field performance in plankton samples (Wiltshire *et al.* 2021b). Following assay design and analytical assessment, two candidate assays for *Mytella strigata* were applied to DNA from 360 plankton samples from around Australia and to 180 plankton samples to which target tissue had been added, confirming field specificity, and allowing calculation of DSe and DSp (Wiltshire *et al.* 2021b). The assay that performed best (MstrigCO1-5 in that publication, Mstr herein) was selected for implementation based on those results. Six of the 19 available assays are therefore operationally validated including assessment of their diagnostic performance, but the remaining 13 assays are at various stages of validation (Table 2).

Simpson *et al.* (2018) partially validated field performance of their Dvex, Pvir, Cjap and Rhar assays by testing water, sediment, and settlement plate samples to which target pest DNA or tissue had been added, but a limited number of samples were used in this testing, and no testing was done for the Pper assay due to a lack of available target species material. Plankton samples were not included and the ability of these assays to detect targets in this sample type is therefore untested. The Pper, Pvir, Cjap and Rhar assays were implemented by SARDI in 2020 and applied, along with the earlier developed assays, to plankton samples from four Western Australian ports (Wiltshire *et al.* 2020b). Results from this survey provide initial field specificity data for validation of relevant assays, however, further data were needed for the newly implemented assays because the survey samples were from a limited geographic area. The available data also do not demonstrate detection in plankton, because these species are not established in Australia and detection is not expected.

Giblot-Ducray *et al.* (2022) applied the Mros and Mjap assays to plankton samples from Tasmania, where these species are established, demonstrating that these assays could detect the targets in environmental samples, but these assays needed to be applied more widely to assess field specificity, and the available data do not permit estimation of DSe or DSp.

1.3. Validation steps required for currently implemented assays

1.3.1. gDNA specificity

Assessment of gDNA specificity was incomplete for two assays (Hsan and Esin, Table 2). The Hsan assay for was adapted from an assay developed by Knudsen and Møller (2020), with the adapted design ensuring that the new assay was at least as specific as the original (Wiltshire *et al.* 2021a). Original specificity testing results for this assay are therefore still applicable, but the species used were primarily European, with few of relevance to Australia. The European designed Esin assay (Andersen *et al.* 2018) was similarly tested primarily against European

species. Related crab species from Australia are well represented in Genbank, and designs of each assay provide specificity against these, but a few native related species have not been sequenced. We therefore carried out additional laboratory specificity testing for these two assays using gDNA from Australian relevant species, including native relatives and other invasive crabs that may occur in surveyed Australian locations, prioritising the inclusion of species with no sequences recorded in Genbank. Where specimens of these crabs and other target pests were available from different locations, we also tested gDNA of these targets with relevant assays to verify the ability of each assay to detect the species from different source locations.

1.3.2. Field specificity

Field specificity needed to be established for the nine assays implemented since 2019 (Pper, Pvir, Cjap, Rhar, Esin, Dvex, Mjap, Hsan and Mros*;* Table 2). Three of these nine target species (*Perna perna, Rhithropanopeus harrisii* and *Eriocheir sinensis*) are not recorded in Australia. Isolated Australian reports exist for a further three species (*P. viridis, Charybdis japonica, and Didemnum vexillum*), but none have established to date. Specifically, a few *C. japonica* have been recorded in Adelaide (Wiltshire *et al.* 2020a) and Perth (Hourston *et al.* 2015), but these appear to be isolated occurrences (Hewitt *et al.* 2018; Wiltshire *et al.* 2020a). *Perna viridis* has been recorded in Cairns (Stafford *et al.* 2007) and Western Australia (McDonald 2012), but despite evidence of spawning in these locations, *P. viridis* has not established in Australia (Wells 2017). Colonies of *D. vexillum* have been recently detected in limited locations within Western Australia and New South Wales, with responses to these detections currently underway (Commonwealth of Australia 2023). These three species, along with *P. perna, R. harrisii* and *E. sinensis* are therefore not expected to be detected in Australian plankton samples. The remaining three species are established in parts of Australia: *Mya japonica* has formed a population on Tasmania's east coast (Dann *et al.* 2020), *H. sanguineus* is established in Port Phillip Bay, Victoria, (DAWE 2021) and *Maoricolpus roseus* is established in Tasmania and the east coast of mainland Australia from eastern Victoria (Wilsons Promontory) to Sydney (Bax *et al.* 2003). We therefore applied these nine assays to DNA from plankton samples collected around Australia, including areas without known occurrence of the established pests, and areas unlikely to be suitable for establishment of the exotic pests.

Problems with specificity identified in field sampling for the Asen and Vgib assays prompted consideration of re-design for these assays (Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019b). Re-design of these assays is complicated, however, because there is a lack of sequence information available for native related species, and hence a risk that any new assay will remain non-specific. We therefore carried out further laboratory specificity testing for these two assays using gDNA from a range of bivalve species, including several native tropical relatives that were

not available for testing during assay development and for which sequences were not available. This approach aimed to identify species that cross-react with these assays, which could then be sequenced where necessary to inform re-design. An approach to assess field detections by these assays using HTS was also trialed. The HTS approach aimed to identify DNA sequences in the samples that would be detected by each assay to either confirm that detections are of the target or identify confounding sequences, i.e., sequences that are likely to cross-react with each assay. The HTS approach was also used to further assess specificity of the newly developed assays for *Hemigrapsus sanguineus, Maoricolpus roseus* and *Mya japonica* when applied to plankton.

1.3.3. Diagnostic performance and field detection

The ability to detect the target in plankton samples needed to be established for nine assays (Table 2), eight of which are for species exotic to Australia (*Mytilopsis sallei, P. canaliculus, P. perna, P. viridis, C. japonica, R. harrisii, E. sinensis and D. vexillum*). For *H. sanguineus*, which is established in Port Philip Bay, Victoria, we tested plankton samples collected from the port of Melbourne, but samples with naturally occurring DNA of other target species would need to be sourced from overseas, which was impractical for the current project. The assessment of assay diagnostic performance, however, also demonstrates the ability of each assay to detect the target in plankton, and so this aspect was not separately investigated for the exotic species. Samples from areas with occurrence of these pests should, however, be tested in future if the opportunity arises.

Diagnostic performance (DSe and DSp) in field samples can be assessed by applying assays to samples with known pest DNA presence or absence, e.g. samples to which target DNA or tissue has been added, or by using latent class modelling (LCM) to analyse the results of multiple tests applied to sample sets of unknown target DNA status but which comprise samples with and without target DNA (Branscum *et al.* 2005; Chambert *et al.* 2015; DAWE 2018). DSe and DSp needed to be assessed for 13 assays (Table 2), the majority of which are for species exotic to Australia, and alternative tests are also not available for most of these species. We therefore used a set of plankton samples to which pest tissue was added to determine DSe and DSp for each assay.

1.4. Objectives

This project aimed to complete validation for 13 partially validated qPCR assays. Specific aims were to:

- Assess gDNA specificity of assays for *Hemigrapsus sanguineus* and *Eriocheir sinensis* using gDNA extracts of native related species.

- Verify the ability of the *H. sanguineus* assay to detect this species in plankton from areas where this species occurs.
- Assess field specificity of assays for *Perna perna, P. viridis, Mya japonica, Maoricolpus roseus, Charybdis japonica, H. sanguineus, Rhithropanopeus harrisii, Eriocheir sinensis* and *Didemnum vexillum* by applying these assays to DNA from plankton samples collected around Australia.
- Characterise diagnostic performance (DSe and DSp) of the assays for *A. senhousia, V. gibba, Mytilopsis sallei, P. canaliculus, P. perna, P. viridis, Mya japonica, Maoricolpus roseus, C. japonica, H. sanguineus, R. harrisii, E. sinensis* and *D. vexillum* using plankton samples with added pest tissue.
- Investigate specificity of assays for *Arcuatula senhousia* and *Varicorbula gibba* using gDNA from tropical native and other bivalves not previously tested, and HTS testing of plankton.
- Further investigate field specificity of the newly developed assays for *H. sanguineus, Maoricolpus roseus* and *Mya japonica* using HTS applied to plankton samples.

2. METHODS

2.1. Laboratory gDNA testing

Laboratory specificity testing was carried out for *Hemigrapsus sanguineus, Eriocheir sinensis, Arcuatula senhousia* and *Varicorbula gibba.* This testing involved applying the Hsan, Esin, Asen and Vgib assays to genomic DNA (gDNA) extracts of relevant related species to check for cross-reactivity. Both *H. sanguineus* and *E. sinensis* are Varunidae, so testing for these assays focused on Varunidae and related families (superfamily Grapsoidea), and other common native crabs and invasive crabs that may occur in surveyed areas. The Asen and Vgib assays were tested using gDNA from the same set of bivalves used to test the Mjap assay during its development (Giblot-Ducray *et al.* 2022).

gDNA extracts from specimens of the target species: *H. sanguineus, E. sinensis, Rhithropanopeus harrisii, Charybdis japonica, A. senhousia, V. gibba, Perna canaliculus, P. perna, P. viridis, Mytilopsis sallei,* and *Didemnum vexillum* were also tested to assess the ability of the assay for each species to detect target DNA using specimens from a range of locations.

2.1.1. Specimens used

Specimens or tissue samples of relevant crab species for specificity testing were obtained from the collection held at the South Australian Research and Development Institute (SARDI) and from the Australian Biological Tissue Collection (ABTC) of the South Australian Museum, while relevant bivalve specimens were from the SARDI collection and from the Museum and Art Gallery of the Northern Territory (MAGNT). All specimens were identified to the lowest taxonomic level possible (family, genus, or species) prior to subsampling appropriate tissue for DNA extraction. Multiple samples were included for some taxa comprising distinct individuals, in some cases from different locations (see Results section 3.1). Specimens of the target species were also included in the testing. Tissue of *H. sanguineus* was provided by Deakin University, and of *E. sinensis* and *A. senhousia* by Department of Primary Industries and Regional Development (DPIRD), Western Australia. A sample of gDNA from *H. sanguineus* was also provided by Deakin University. Specimens of *V. gibba* and additional specimens of *A. senhousia* were available in the SARDI collection, and archived gDNA of these species was also available for each from earlier assay development work.

Target species samples that were tested in addition to those tested during the specificity assessments were: *H. sanguineus* specimens from Victoria provided by Agriculture Victoria; tissue of *C. japonica* specimens from Port Adelaide provided by ABTC; *R. harrisii* specimens from

Texas, USA, provided by Tarelton State University, from Luebeck, Germany, provided by Helmholtz Centre for Ocean Research Kiel, and tissue from Estonia provided by DPIRD; *P. perna* specimens from South Africa provided by Two Oceans Aquarium; *P. viridis* specimens from Thailand provided by Benchmark Inve, from Singapore provided by MAGNT, and from vessel fouling in SA provided by Biosecurity SA; *Mytilopsis sallei* specimens from Darwin and Indonesia provided by MAGNT and from Thailand provided by Benchmark Inve; tissue of *D. vexillum* from Canada provided by Fisheries and Oceans Canada, and from WA provided by DPIRD. All specimens were tested using the assay for that species. Specimens used for the assessment of assay diagnostic performance (see section 2.3.2) were selected following this testing.

2.1.2. DNA extraction from tissue samples

DNA extraction and testing were carried out by the SARDI Molecular Diagnostics laboratory. For specimens preserved only in ethanol (i.e., not formalin-fixed), DNA was extracted using the Qiagen Blood and Tissue (B&T) DNA extraction kit. For specimens known to have been fixed in formalin or where the original fixative was unknown, the Qiagen formalin fixed – paraffin embedded (FFPE) extraction kit was used. Several archived extracted gDNA samples from related species held by SARDI were also used in the testing. All gDNA extracts were quantified using Invitrogen™ Quant-iT™ PicoGreen™ dsDNA Assay Kits (ThermoFisher Scientific) and used at 200 pg μL^{-1} with qPCR conditions as shown in section 2.1.3.

2.1.3. qPCR methods

qPCRs were conducted in a 10 µL total volume consisting of 5 µL of 1x QIAGEN® QuantitTect® Probe master mix, 400 nmol L^{-1} forward and reverse primer of the relevant assay, 200 nmol L^{-1} probe and 4 µL template DNA. Each test comprised a single replicate.

qPCRs were carried out in 384 well plates for analysis using a ViiA 7 or QuantStudio7 real-time PCR systems (Applied Biosystems). PCR cycling parameters were 15 min at 95 °C (activation) followed by 45 cycles of 15 s at 95 °C (denaturation) and 1 min at 60 °C (annealing). Each plate included six replicates of the synthetic oligonucleotide standard at 2,000 copies μL^{-1} and duplicate negative control samples.

2.2. Field validation

Field validation included testing plankton samples from Melbourne with the Hsan assay to assess ability of this assay to detect *Hemigrapsus sanguineus* where present, and plankton samples

from around Australia (Figure 2) for nine pests: *D. vexillum, E. sinensis*, *H. sanguineus*, *R. harrisii*, *C. japonica*, *P. perna*, *P. viridis*, *Mya japonica* and *Maoricolpus roseus* to assess field specificity.

2.2.1. *Hemigrapsus sanguineus* testing

Samples for testing using the Hsan assay included DNA from 128 plankton samples previously collected from the port of Melbourne by Wiltshire *et al.* (2019a), comprising 32 samples from each of 4 sample sets collected approximately seasonally between July 2017 and June 2018, and 30 samples collected for the current project by Agriculture Victoria on 22 February 2022 (Table 3). A map of the sample locations is provided in the results (section 3.2).

Table 3. Sample collection dates of plankton sample sets from Melbourne tested for *Hemigrapsus sanguineus.* Collections in 2017-18 were by Wiltshire *et al.* (2019a) and in Feb 2022 by this project.

Sample set	Samples collected	Analysed for current project	Start date	End date
Jul 2017	124	32	29-Jul-2017	01-Aug-2017
Nov 2017	126	32	08-Nov-2017	12-Nov-2017
Mar 2018	126	32	22-Mar-2018	27-Mar-2018
Jun 2018	126	32	25-Jun-2018	28-Jun-2018
Feb 2022	30	30	22-Feb-2022	22-Feb-2022

2.2.2. Field specificity samples

Samples tested for field specificity were collected 2015 – 2022 by the projects as shown in Table 4. Each assay was applied to DNA extracts from between 184 and 230 of these samples. The Rhar, Cjap, Pper and Pvir assays were applied to a total of 184 samples by this project, having been previously applied to samples from WA by Wiltshire *et al.* (2020b). The Mros, Mjp and Hsan assays were applied to 202 samples each, comprising those from locations without each species presence (Table 4). Samples from areas with species presence (Melbourne for *Hemigrapsus* sanguineus, SE Tasmania for *Mya japonica* and *Maoricolpus roseus*) have been tested by the relevant assays by this project (Hsan) and Giblot-Ducray *et al.* (2022) (Mjap, Mros), but those results were not used for field specificity assessment. The Esin and Dvex assays were applied to all 230 DNA extracts. Sampling locations are shown in Figure 2, and maps of sampling points within each location are provided in the report relating to each collection (see Table 4). Each project collected samples around port areas, and specifically, wharves used by ballast water carrying vessels, except in the case of samples collected in SE Tasmania, where sampling targeted areas with known occurrence of *Maoricolpus roseus* (D'Entrecasteaux Channel) and *Mya japonica* (Orford – Triabunna region) (Giblot-Ducray *et al.* 2022). Samples from that project

that were tested for the other seven pests comprised 13 from the Orford – Triabunna region and 15 from D'Entrecasteaux Channel.

To provide additional field specificity data for assays that had been previously applied to molecular surveillance (see section 1.2), results were extracted from the database compiled by Wiltshire (2021). For pests with known occurrence in parts of Australia, we did not consider results from areas with either confirmed pest presence or suspected presence based on being within the known range of an established pest. We also did not consider results from areas with unconfirmed occurrence, i.e., those where molecular detections have occurred without confirmation by other data, but where occurrence is plausible based on the species environmental tolerance and there has been a lack of relevant surveillance for confirmation of occurrence.

All 19 currently implemented assays were applied by Wiltshire *et al.* (2022) to 490 samples, comprising 70 from each of Portland, Victoria, and Adelaide Outer Harbor, Adelaide Inner Harbor, Klein Point, Port Giles, Port Lincoln and Thevenard, SA. These results were not included in the field specificity data set that was used in modelling to estimate DSe and DSp (see section 2.4.3), because Wiltshire *et al.* (2022) applied the DSe and DSp estimates generated to analysis of their results. Use of the Wiltshire *et al.* (2022) qPCR results in this project would therefore have led to inappropriate re-use of these data.

Table 4. Collecting locations, number of samples used, and original project reports for plankton samples used for specificity testing. Samples were tested for 9 species: *Didemnum vexillum* (Dvex), *Eriocheir sinensis* (Esin), *Hemigrapsus sanguineus* (Hsan), *Rhithropanopeus harrisii* (Rhar), *Charybdis japonica* (Cjap), *Perna perna* (Pper), *Perna viridis* (Pvir), *Mya japonica* (Mjap) and *Maoricolpus roseus* (Mros), or for a subset of these species as indicated.

Location	No. samples	Tests applied	Collected by
Darwin, NT	22	9 species	Deveney *et al.* (2017)
Gove, NT	13	9 species	Wiltshire *et al.* (2019c)
Weipa, Qld	11	9 species	Wiltshire *et al.* (2019c)
Cairns, Qld	22	9 species	Deveney *et al.* (2017)
Hay Point, Qld	14	9 species	Wiltshire *et al.* (2019c)
Brisbane, Qld	16	9 species	Deveney *et al.* (2017)
Newcastle, NSW	9	9 species	Wiltshire *et al.* (2019c)
Botany Bay, NSW	15	9 species	Wiltshire *et al.* (2019c)
Melbourne, Vic	30	8 (not Hsan)	this project
SE Tasmania	28	7 (not Mjap or Mros)	Giblot-Ducray *et al.* (2022)
Bunbury, WA	23	5 (Dvex, Esin, Hsan, Mjap, Mros)	Wiltshire *et al.* (2020b)
Fremantle, WA	23	5 (Dvex, Esin, Hsan, Mjap, Mros)	Wiltshire *et al.* (2020b)

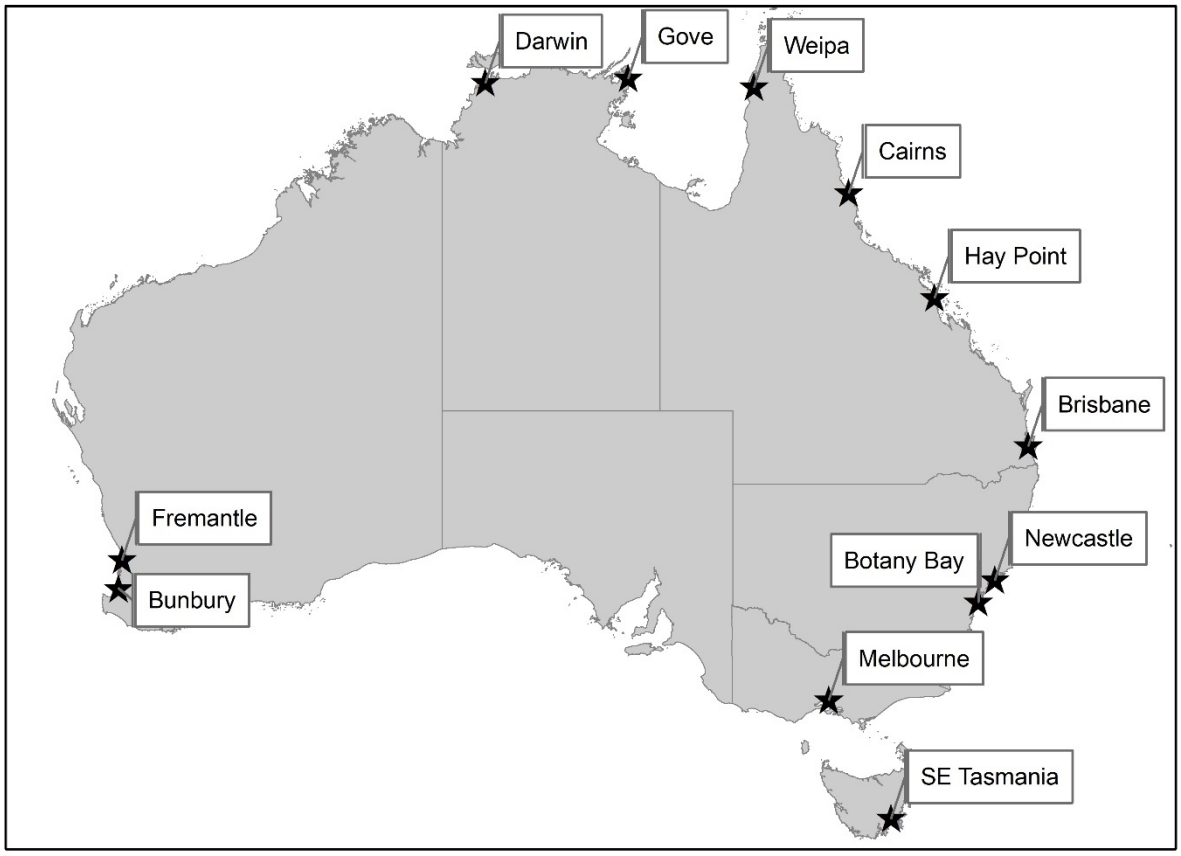

Figure 2. Map showing locations around Australia where plankton samples used in field specificity testing were collected. See reports listed in Table 4 for specific sampling points within each location.

2.2.3. Plankton sampling

Plankton samples were collected by the projects listed in Table 4 based on the methods developed by Giblot-Ducray and Bott (2013) and refined by Deveney *et al.* (2017). Sampling used 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) that 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. Samples collected in Melbourne for the current project used the same tow length but the net specification varied because samples were collected opportunistically during plankton surveys carried out by Agriculture Victoria. The net used in Melbourne was 0.9 m long, with a mouth diameter of 0.3 m and 20 µm mesh. After collection in all cases, 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 Camacho-Sanchez *et al.* 2013). Samples were kept cool in an insulated container with gel ice packs or refrigerator after collection and for overnight delivery to SARDI Aquatic Sciences where they were stored in a cool room at ≤ 4 °C until processing.

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

2.2.4. DNA extraction and qPCR testing of plankton samples

DNA was extracted from plankton samples 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). Final volume of the DNA was 160 µL in elution buffer. A relatively large elution volume is used for DNA extracted from plankton because the bulk sample extraction method applied results is a relatively high total DNA yield, with resulting DNA concentration in elution buffer being typically between 1700 and 4400 pg μL^{-1} (SARDI data). The efficiency and consistency of SARDI's method to extract DNA from environmental samples has been confirmed in comparison to commercial methods (Haling *et al.* 2011), and previous assessment of the method applied to plankton samples has demonstrated the efficacy of the method for this sample type (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019b). Each DNA extract was tested in singleplex qPCR with conditions as shown in section 2.1.3 using assays for the relevant target species (i.e., Hsan for Melbourne samples and as per Table 4 for field specificity testing) and the internal control organism used to assess PCR inhibition. Testing comprised a single replicate of each assay per sample.

Reference samples that are known to not cause inhibition were also extracted after addition of the inhibition control organism and tested by qPCR. A scaling factor was calculated for each sample as the ratio of inhibition control DNA yield in the reference samples to that in the plankton sample. A scale factor of 1 therefore indicates no inhibition, with increasing scale factor indicating greater PCR inhibition. DNA yield (in pg g^{-1} for Asen, Pcan, Msal and Vgib, and kDNA copies g^{-1} for other assays) was calculated from the C_T value in samples with detection using the standard curve for each assay. The scale factor is considered a multiplier to correct calculated DNA yield for the effect of inhibition, although it may not be a perfect correction factor given differing assay responses to inhibition (Wiltshire *et al.* 2019a; Lance and Guan 2020). Raw DNA yield, without correction by the scale factor, was therefore used in further analyses of the data (see section 2.4)

2.3. Constructed samples for assay diagnostic performance

Constructed samples were used to assess the diagnostic performance in plankton samples of assays for 13 pests: *A. senhousia, V. gibba, Mytilopsis sallei, P. canaliculus, P. perna, P. viridis,*

Mya japonica, Maoricolpus roseus, C. japonica, H. sanguineus, R. harrisii, E. sinensis and *D. vexillum.*

2.3.1. Plankton samples

Plankton samples used for assessment of assay diagnostic performance included 82 samples from Gladstone collected by Wiltshire *et al.* (2019a) that were not processed as part of that project, and 158 samples collected for the current project, comprising 133 from the Adelaide metropolitan coast (North Haven to Grange), seven from Thevenard, eight from Klein Point and five each from Port Giles and Port Adelaide (Outer Harbor). These samples were collected, filtered and freeze-dried as per section 2.2.3.

2.3.2. Pest tissue samples

Specimens or tissue of each target pest were obtained from the sources shown in Table 5. We selected specimens from those available based on the earlier gDNA testing (see section 2.1.3), using additional material from the same individual or individuals as tested where practical, or from other specimens from the same collection where necessary. The prior testing confirmed that specimens used were correctly identified and suitably preserved. Tissue subsamples were taken from each specimen in a PCR-clean laboratory, with all surfaces cleaned using Lookout® DNA Erase between species to prevent cross contamination. Adductor muscle was used from bivalves and leg tissue from crabs where possible, but because some specimens were very small, other tissues were included in subsamples in some cases (Table 5). Five subsamples of varying mass were taken from each species for addition to plankton samples. Wet weights were recorded after gently patting subsamples dry of ethanol with lint-free Kimwipes®.

Tissue subsamples were freeze-dried then homogenised into acid-washed sand by milling to make six stocks of spiked sand, each with a different combination of pest tissue (Table 6). Each of the six sand stocks was added to 40 plankton samples, comprising 10 samples dosed with each of four spiked sand masses: 0.1, 0.5, 1 and 2 g. Each target pest was absent from one of the sand stocks and hence there were 40 samples without added tissue of each pest. Treatments were randomised across the available plankton samples. Following addition of spiked sand, plankton samples were extracted and tested using the method outlined in section 2.2.4 with the assays for the 13 species being assessed (Table 5).

Table 5. Description of pest tissue samples used for determination of assay diagnostic performance.

Species	Source	Provided by	Tissue used
Arcuatula senhousia	Perth, WA	Department of Primary Industries and Development, WA	All tissue from two specimens
Varicorbula gibba	Melbourne, Vic	SARDI collection (from Wiltshire *et al.* 2019a)	Multiple whole specimens
Mytilopsis sallei	Darwin, NT	Museum and Art Gallery of the NT	All tissue from multiple specimens
Perna canaliculus	New Zealand	Purchased commercially	Adductor muscle
Perna perna	South Africa	Two Oceans Aquarium, Cape Town	All tissue from one specimen
Perna viridis	Thailand	Benchmark Inve, Nonthaburi	Adductor and foot from two specimens
Mya japonica	Orford, Tas	Natural Resources and Environment, Tas	Adductor muscle from two specimens
Maoricolpus roseus	D'Entrecasteaux Channel, Tas	Woodbridge school Marine Discovery Centre, Tas	Foot tissue from multiple specimens
Charybdis japonica	Port Adelaide, SA	SARDI collection, provided by Biosecurity SA	Leg tissue from one specimen
Eriocheir sinensis	Portugal	Department of Primary Industries and Regional Development, WA	Leg tissue from one specimen
Hemigrapsus sanguineus	Mount Martha, Vic	Agriculture Victoria, Vic	Claw and leg tissue from one specimen
Rhithropanopeus harrisii	USA	Tarelton State University, Texas	Claw, leg and body tissue from multiple specimens
Didemnum vexillum	Canada	Department of Fisheries and Oceans, Nanaimo, BC	Thoracic tissue from one specimen

Table 6. Tissue wet weight (mg) of each species per g sand for each spiked sand treatment (A – F) and total combined pest tissue wet weight per g sand in each treatment.

Species	A	B	C	D	E	F
Arcuatula senhousia	0.21	0.85	1.81	7.06	-	0.06
Varicorbula gibba	2.57	10.06	-	0.09	0.31	1.00
Mytilopsis sallei	5.05	1.74	0.82	0.19	0.08	-
Perna canaliculus	2.56	9.96	-	0.07	0.25	1.04
Perna perna	-	0.04	0.21	0.99	2.28	7.89
Perna viridis	0.07	-	9.75	2.16	1.07	0.23
Mya japonica	1.01	0.25	0.08	-	9.69	1.95
Maoricolpus roseus	0.06	-	8.42	2.24	0.98	0.33
Charybdis japonica	-	0.33	1.03	2.31	4.14	10.38
Eriocheir sinensis	1.99	0.95	0.19	-	10.10	3.94
Hemigrapsus sanguineus	10.74	5.23	2.66	1.34	0.48	-
Rhithropanopeus harrisii	0.74	1.30	3.61	9.97	-	0.21
Didemnum vexillum	-	0.19	0.82	1.56	3.29	8.28
Total pest tissue (mg g^{-1})	**25.02**	**30.92**	**29.39**	**27.98**	**32.68**	**35.30**

2.4. Statistical analyses and mapping

2.4.1. Patterns of *Hemigrapsus sanguineus* detection and DNA yield

Patterns in detection likelihood and target DNA yield (kDNA copies per sample) for *Hemigrapsus sanguineus* across sample times were explored using spatial zero-added Gamma (ZAG) models. ZAG models, also known as zero-altered Gamma or Gamma hurdle models, consist of a binary component, modelled using the Bernoulli distribution, and a continuous component modelled using the Gamma distribution (Quiroz *et al.* 2015; Zuur *et al.* 2017; Zuur and Ieno 2018). ZAG models were applied because these allow simultaneous investigation of covariate effects on likelihood of detection (binary component), and DNA yield in samples with detection (continuous component), with the Gamma distribution applied because this is a flexible distribution suitable for continuous, strictly positive, data (Zuur *et al.* 2017; Zuur and Ieno 2018). Spatial modelling was applied because plankton samples collected within the same area are likely to be more similar than those further apart, violating the independence assumption of non-spatial models (Quiroz *et al.* 2015; Zuur *et al.* 2017; Bakka *et al.* 2018; Zuur and Ieno 2018). The general modelling approach followed that of other studies involving non-gaussian spatially correlated data (e.g., Arab 2015; Paradinas *et al.* 2015; Quiroz *et al.* 2015; Cavieres and Nicolis 2018; Vilela *et al.* 2021; Izquierdo *et al.* 2022).

A complementary log-log (cloglog) link was used for the binary component with the response being detection/non-detection, and the default log link was used for the Gamma distribution of the continuous component, with the response variable being DNA yield calculated from the assay standard curve for each sample. The data used in model fitting were the results from testing 158 plankton samples from Melbourne with the Hsan assay. These samples were from five sample sets (see section 2.2.1): Jul 2017, Nov 2017, Mar 2018, Jun 2018 (Wiltshire *et al.* 2019a) and Feb 2022 (this project). Sample set was applied as a fixed factor and spatially correlated random effects were included in both model components. Sample set was applied as a factor because DNA concentration and abundance, and hence detectability, varies seasonally for many species (de Souza *et al.* 2016; Buxton *et al.* 2017). Timing of sampling was not consistent between the previously collected (2017-18) samples and those collected for the current project, given that the latter was opportunistic and comprised a single sampling event. Data were not pooled by year or project because this would have resulted in confounding effects of year or project and sample timing. PCR inhibition was not considered as a factor because all except two tested samples had scale factor = 1 (i.e., no inhibition) (see Results section 3.2), meaning the data did not permit estimation of inhibition effects. The cloglog link was applied because resulting coefficient estimates can be interpreted as the relative log-concentration of planktonic DNA-containing

particles in each sample set (see Appendix section 6.1.1 for details). The general form of the model used was therefore:

$$Yi \sim ZAG(\mu_i, \pi_i)$$

$$log(\mu_i) = X_i\beta_1 + u_i$$

$$u_i \sim GMRF(0, \sigma^2_u)$$

$$cloglog(\pi_i) = X_i\beta_2 + v_i$$

$$v_i \sim GMRF(0, \sigma^2_v)$$

Where Y_i is the set of results (detect/non-detect and DNA yield) for observation i, μ_i is the mean of the Gamma component, π_i is the mean of the Bernoulli component, X_i is the matrix of covariates (fixed effects), β_1 and β_2 are the sets of coefficients (estimated by the model) for the fixed effects in the Gamma and Bernoulli components respectively, and u_i and v_i are spatially correlated random effects (spatial field) for the respective components. ZAG models allow for the use of different covariates in each component, but in our case the same set of covariates (i.e., sample set) was used for each. The spatial random effects for each component follow a Gaussian (normal) distribution with zero mean and a Matérn correlation structure, estimated by a Gaussian Markov Random field (GMRF) with variance of σ^2_u and σ^2_v for the continuous and binary model components respectively. The Matérn correlation between two spatially referenced observations depends on the Euclidean distance between the observations and a range parameter estimated by the model, where the range parameter is defined as the distance at which correlation decays below 0.1 (Quiroz *et al.* 2015; Zuur *et al.* 2017), i.e., the distance beyond which two points are not considered correlated.

ZAG models were fitted using a Bayesian hierarchical modelling approach with integrated nested Laplace approximations (Rue *et al.* 2009) for model inference and the stochastic partial differential equation (SPDE) approach (Rue and Held 2005; Lindgren *et al.* 2011) to estimate the GMRF for spatial random effects. Models were run with the *R-INLA* package (Martins *et al.* 2013; Lindgren and Rue 2015; Rue *et al.* 2017) using *R* statistical software (R Core Team 2023) and following Zuur and Ieno (2018). To account for coastal features and prevent smoothing of spatial effects over land, barrier models were applied (Bakka *et al.* 2018; 2019).

Spatial effects were estimated on a constrained refined Delaunay triangulation spatial mesh (Figure 3) built with the *inla.mesh.2d* function and a maximum distance between nodes of 200 m. The mesh was constructed for the entire survey area, with a 2 km buffer around the extent of sample points applied to avoid boundary effects (Lindgren *et al.* 2011). A projector matrix was

generated using the *inla.spde.make.A* function to link results and covariates for each sample to the mesh nodes based on sample co-ordinates (Latitude, Longitude). Following model fitting, which generates posterior predictions for each model component on nodes of the mesh, results were linearly interpolated across the study area using the *inla.mesh.projector* function. The interpolated fields for each model component were converted to rasters for use in mapping using the *Raster* package (Hijmans 2021). Maps showing sample results and spatial fields were generated using *ArcGIS* 10.8 (Esri Inc).

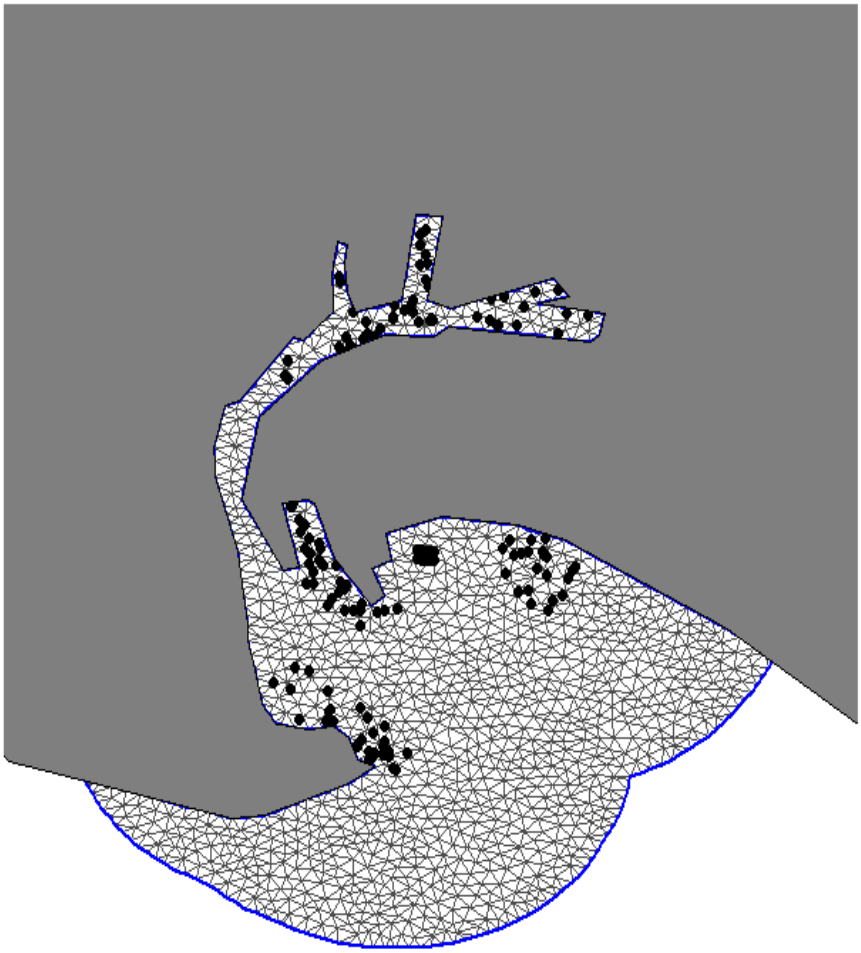

Figure 3. Constrained refined Delaunay triangulation used for estimation of spatially correlated spatial effects in the ZAG modelling for *Hemigrapsus sanguineus*, with sample points and land area, which was used as the barrier layer for the spatial field, overlaid.

Differences between sample sets were assessed using the *inla.make.lincombs* function, which generates linear combinations of coefficients, following Gómez-Rubio (2020). Specifically, linear combinations were generated to calculate the pairwise differences between the estimated coefficients for each sample set in each model component. The difference between coefficients for detection likelihood can be interpreted as a difference in log abundance of DNA-containing particles in plankton, while the difference between coefficients for DNA yield can be interpreted

as the difference in log(DNA yield) between sample sets. See Appendix section 6.1.1 for details. Sample set coefficients were considered significantly different from one another where 95% highest density intervals (HDIs) of the difference between posterior estimates of those factors did not contain zero.

For the range and standard deviation of the spatial random effects, penalised complexity (PC) priors (Simpson *et al.* 2017; Fuglstad *et al.* 2019) were applied. PC priors are based on a type-2 Gumbel distribution but parameterised in *R-INLA* based on user-specified values *U* and α to provide a distribution for the parameter, *z*, such that the probability $z > U = \alpha$ (Simpson *et al.* 2017). The use of PC priors is recommended because these allow the objective setting of uninformative, weakly- or strongly- informative priors while allowing for efficient computation (Zuur *et al.* 2017; Bakka *et al.* 2018; Zuur and Ieno 2018; Fuglstad *et al.* 2019). The PC priors used provided a probability of 0.05 for range < 1 km and for standard deviation of the spatial random effects > 1. Gaussian priors with mean zero and precision of 0.025 were used for fixed effects in the binary component. This precision was chosen because it provides 95% confidence that changes to the linear predictor are between −12 and 12, with resulting probability being between ~$5E^{-6}$ (effectively zero) and ~1, preventing numerical overflow. Default Gaussian priors (mean = 0, precision 0.001) were used for fixed effects in the continuous component of the model.

2.4.2. Effect of tissue dose on detection likelihood and C_T in constructed samples

The effects of tissue dose on the likelihood of detection and resulting C_T value by each assay in the 240 constructed samples (see section 2.3) were examined using binomial generalised linear mixed modelling (GLMM) and linear mixed modelling (LMM) respectively. Models were run using the *R-INLA* package (Martins *et al.* 2013; Lindgren and Rue 2015; Rue *et al.* 2017). A cloglog link was used for the binomial GLMM with the response being detection/non-detection. This link was chosen because it is suitable for modelling a detection process that depends on abundance or concentration (Royle and Nichols 2003; Royle and Dorazio 2009). Target species, the natural logarithm of added dose (*lndose*, where dose = mg tissue in the sample of the target species), and natural logarithm of scale factor (*lnSF*) were included as fixed factors in both models. The interaction terms *Species* x *lndose* and *Species* x *lnSF* were included to determine the response of each assay to tissue dose and scale factor. Sand stock (A-E) was included as an independent identically distributed (iid) random effect in both models. The models used were therefore:

$$Detection_{ij} \sim Bernoulli(\pi_{ij})$$

$$cloglog(\pi_{ij}) = X_{ij}\beta_1 + a_i$$

$$a_i \sim N(0, \sigma^2_{stock1})$$

and

$$C_{Tij} = X_{ij}\beta_2 + b_i + \varepsilon_{ij}$$

$$b_i \sim N(0, \sigma^2_{stock2})$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

Where π_{ij} is the detection likelihood for sample *j* made with sand stock *i*, X_{ij} is the matrix of covariates (fixed effects), β_1 and β_2 are the sets of coefficients (estimated by the model) for the fixed effects in the binomial GLMM and LMM respectively, ε_{ij} is the residual error for sample *ij* in the LMM, a_i and b_i are the iid random effects for sand stock i in the respective models, σ^2_{stock} is the variance of the sand stock random effect in each model, and σ^2 is the residual variance in the LMM.

C_T value and *lndose* were used as the response and predictor respectively in the linear model because the C_T value of a qPCR detection decreases approximately linearly with increasing log DNA concentration (Kralik and Ricchi 2017), and the DNA concentration in samples should be proportional to the tissue dose added. Detection likelihood on the cloglog scale should also increase linearly with log concentration (Royle and Nichols 2003; Royle and Dorazio 2009; Wiltshire 2021). Data for unspiked samples, i.e., where dose = 0, for each species were not included in these analyses. The 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. Default priors were used for fixed effects in both models. PC priors for the standard deviation (sd) of random effects were parameterised to provide probability sd > 400 = 0.05.

Predictions of detection likelihood and C_T value were generated for each model using the *inla.make.lincombs* function for the range of tissue doses applied across species and five levels of inhibition: nil (scale factor = 1), minor (2), moderate (5), high (10) and very high (100). Fitted C_T values for samples with no detection were obtained from the linear model. The lowest tissue dose required to give a predicted detection likelihood of at least 95% was determined for each assay from coefficients of the binomial GLMM at each of the five levels of inhibition as:

$$Dose(95\%)_i = exp(cloglog(0.95) - \beta_0)/(\beta_{SF} \times lnSF_i)$$

Where $Dose(95\%)_i$ is the dose required to give 95% detection likelihood at inhibition level $lnSF_i$, β_0 is the intercept and β_{SF} is the coefficient for *lnSF*.

2.4.3. Assessment of diagnostic performance

Diagnostic performance of each assay was assessed using a Bayesian latent class model (LCM) with Markov chain Monte Carlo (MCMC) simulation in JAGS v. 4.3.0 (Plummer 2017) for parameter estimation. LCM is a commonly used and recommended method for the assessment of diagnostic test performance (Branscum *et al.* 2005; Johnson *et al.* 2019; Rahman *et al.* 2019; Wang *et al.* 2020; Cheung *et al.* 2021; Clouthier *et al.* 2021). Code for the model (Appendix section 6.2) was adapted from Wiltshire *et al.* (2021b), whose code was based on a model by Wang *et al.* (2020). LCM estimation of diagnostic performance typically requires results from a minimum of two tests applied to the same set of samples (Branscum *et al.* 2005). Samples were tested with only one qPCR assay per species, but the known pest presence or absence of a species in each sample, based on the addition of spiked sand, was included as a reference test. Informative priors were also applied to relevant assays for some model parameters, based on data from the current and previous projects.

The Beta distribution is used for priors of parameters that lie between 0 and 1, e.g., probabilities such as DSp and DSe, in Bayesian LCM (Branscum *et al.* 2005; Johnson *et al.* 2019; Wang *et al.* 2020). The Beta distribution uses two parameters (*a,b*), with a Beta(a = 1, b = 1) prior providing uniform likelihood and therefore being suitable as an uninformative prior (Kruschke 2014; Johnson *et al.* 2019; Wang *et al.* 2020). Where prior knowledge is available on an aspect of test performance, however, this can be incorporated into an LCM using informative Beta priors (Branscum *et al.* 2005; Johnson *et al.* 2019).

Beta priors for DSp in our model were parameterised following Kruschke (2014). Specifically, given an initial Beta(*a,b*) distribution for a process, and observed data from N trials with z successes, the updated probability distribution for the process can be described with a Beta($z+a$, $N-z+b$) distribution (Kruschke 2014). The data used to parameterise DSp priors were the results of testing plankton samples around Australia, including field specificity testing carried out by this project (see section 2.2.2 and Results section 3.3), and collated results from previous molecular surveillance where relevant assays were applied (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019c, 2020b). We made no assumption about DSp of the assays beyond the information provided by these field specificity data and therefore used an uninformative Beta(1,1) distribution updated by the field specificity data for each assay to provide the Beta prior distributions for DSp used in the model. Beta(*a,b*) prior distributions were therefore parameterised using a = number of samples with no detection + 1, b = number of samples with detection + 1, considering samples from areas without presence of each pest, and noting that “success” for DSp is non-detection given target absence. Samples from areas with potential but unconfirmed pest presence were not included.

DSe for each assay was estimated taking the potential effect of PCR inhibition, as measured by the scale factor, into account. An uninformative Beta(1,1) prior was used for the intercept of DSe, i.e. DSe in the absence of inhibition, for each assay, and the natural logarithm of scale factor (lnSF) was used as a covariate, with a cloglog link, i.e.:

$$cloglog(DSe_{ij}) = \beta_{DSej} + \beta_{SFj} * lnSF_i$$

where DSe_{ij} is effective DSe of assay *j* in sample *i* after accounting for the scale factor (inhibition) effect, β_{DSej} is the DSe of assay *j* in the absence of inhibition (i.e. where $lnSF = 0$) on the cloglog scale, β_{SFj} is the coefficient of the scale factor effect for assay *j*, and $lnSF_i$ is the log(scale factor) for sample *i*. (see code in Appendix section 6.2 for further details). A truncated normal distribution with mean zero, precision of 0.154, and upper limit of zero was used for the prior for the *lnSF* effect (β_{SF}) of each assay. This prior provides 95% confidence that this coefficient is between 0 and −5 on the cloglog scale, which is equivalent to a reduction in DSe of between 0 and 99% per unit increase in *lnSF*.

DSe and DSp of the reference test (presence/absence of added pest tissue) were set to 1, and prevalence (= proportion of samples with pest presence, and, equivalently, likelihood of pest presence in a sample) was set to 200/240 (=0.833), equal to the proportion of samples with added pest tissue for each species. The data used for model fitting were the results (detection/non-detection) of testing the 240 constructed samples with each assay (see section 2.3)

MCMC simulations were obtained from three chains using 10,000 iterations thinned at a rate of 10, following 40,000 iterations for burn-in. JAGS was run using the *R2jags* package in *R*. Convergence was assessed by Gelman-Rubin convergence statistic and confirmed by visual inspection of trace, density and autocorrelation plots generated using the *MCMCplots* package (McKay Curtis 2015). HDIs were calculated using the *HDInterval* package (Meredith and Kruschke 2018) and used to describe parameter estimates. The prior used for the effect of scale factor restricted this parameter to be negative. We therefore considered the scale factor effect to be significant where the upper limit of the HDI for this effect was less than −0.1.

2.5. High throughput sequencing

A targeted high throughput sequencing (HTS) approach was used to provide additional validation of the newly developed Hsan, Mjap and Mros assays, and to investigate detections by the Asen and Vgib assays. The HTS approach has been developed by SARDI Molecular Diagnostic for GRDC project: *UOA1802-019BLX - SARDI Molecular Diagnostic Centre national disease surveillance*. The method was made available for use in this project, but full details are not available for publication at this stage. The aim of the approach is to assess assay specificity by

determining the presence of target sequences in samples with qPCR detection and investigate whether non-target sequences present in environmental samples that could cross-react with the assay due to similarity with the assay primer and probe sequences. HTS targeted the relevant gene region for each assay (COI for Hsan, Mjap and Mros; 28S for Asen and Vgib) and used two semi-generic primer sets per assay to provide results relevant to the forward and reverse primers, referred to in the results as the forward and reverse primer sets respectively. Amplicons produced by each primer set were between 200 and 400 bp in length.

The plankton DNA samples used for HTS testing had been tested using the qPCR assay for the relevant species (Table 7). Samples for *Hemigrapsus sanguineus* were from this project, while samples for *Mya japonica* and *Maoricolpus roseus* were from Giblot-Ducray *et al.* (2022). For these three species, the samples used were predominantly from locations with known species occurrence and comprised a mixture of samples with and without detection by the relevant qPCR assay. Six samples with detections by Mjap or Mros from locations without known species occurrence (Mjap: Port Giles, D'Entrecasteaux Channel; Mros: Hay Point, Gove, Klein Point, Adelaide) were also included in the testing (Table 7). These were samples that returned detections in field specificity testing for this project (see sections 2.2.2 and 3.3), in spiked samples without target tissue added (see section 3.4.1), in the 2021-22 molecular surveillance that applied these assays (Wiltshire *et al.* 2022) or in the testing by Giblot-Ducray *et al.* (2022). No Hsan detections were recorded in samples from locations other than Melbourne and therefore no samples from other locations were included in the *Hemigrapsus sanguineus* HTS run. For *Arcuatula senhousia* and *Varicorbula gibba,* samples with a detection by one or both Asen and Vgib assays were selected for HTS analysis. These samples were collected between 2015 and 2022 by the projects as shown in Table 7 and included samples from some locations with confirmed populations of the target species.

The HTS libraries were prepared using a dual-PCR protocol. The first PCR amplified the relevant target sequence, and the second PCR reaction included the addition of indexing, which enabled pooling of multiple PCR amplicon sets (libraries) into a single run. HTS for *A. senhousia* and *V. gibba* was carried out in two sequencing runs, one including libraries prepared from samples from WA, NT, Qld, NSW, Tas, and samples collected in 2016 from Whyalla and Port Lincoln SA, and the second including samples from Melbourne and 2021-22 samples from SA ports and Portland, Vic. These two runs included testing for both species. The total number of libraries pooled in these two runs was four times the number of plankton samples, with DNA of each sample amplified using two primer sets for each of the two tested species, plus four negative (water) controls. A single run was carried out for each of *H. sanguineus, Mya japonica* and *Maoricolpus roseus*, with these runs including testing of a single target species each. The total number of pooled libraries for these runs was two times the number of plankton samples, with each sample

amplified with two primer sets, plus 1 – 2 negative controls. Sample and library numbers per run are provided in the results (see section 3.5).

Table 7. Collecting locations, number of samples used, and original project reports for plankton samples used for high throughput sequencing (HTS) confirmatory testing. Samples were tested using HTS for: *Hemigrapsus sanguineus* (Hsan), *Mya japonica* (Mjap), *Maoricolpus roseus* (Mros), *Arcuatula senhousia* (Asen) or *Varicorbula gibba* (Vgib). *Indicates confirmed occurrence of the target species in at least some sites at the sampled location. qPCR testing for target species was carried out by the cited project except: † tested by Giblot-Ducray *et al.* (2022), ‡ tested by current project.

Location	No. samples	HTS testing for	Collected by
Brisbane	4	Asen, Vgib	Wiltshire *et al.* (2019a)
Gladstone	15	Asen, Vgib	Wiltshire *et al.* (2019a)
Cairns	1	Asen	Deveney *et al.* (2017)
Hay Point	6	Asen, Vgib	Wiltshire *et al.* (2019c)
Hay Point	1	Mros	Wiltshire *et al.* (2019c) ‡
Weipa	6	Asen, Vgib	Wiltshire *et al.* (2019c)
Gove	6	Asen, Vgib	Wiltshire *et al.* (2019c)
Gove	1	Mros	Wiltshire *et al.* (2019c) ‡
Darwin	4	Asen, Vgib	Deveney *et al.* (2017)
Bunbury	3	Vgib	Wiltshire *et al.* (2020b)
Fremantle	3	Asen*, Vgib	Wiltshire *et al.* (2020b)
Kwinana	12	Asen, Vgib	Wiltshire *et al.* (2020b)
Port Lincoln	3	Vgib	Wiltshire *et al.* (2017)
Port Lincoln	8	Vgib	Wiltshire *et al.* (2022)
Whyalla	3	Vgib	Wiltshire *et al.* (2017)
Thevenard	16	Asen, Vgib	Wiltshire *et al.* (2022)
Klein Point	1	Mros	Wiltshire *et al.* (2022)
Port Giles	1	Mjap	Wiltshire *et al.* (2022)
Pt Adelaide	1	Mros	Wiltshire *et al.* (2022)
Pt Adelaide	1	Vgib	Wiltshire *et al.* (2022)
Devonport	4	Asen*, Vgib*	Wiltshire *et al.* (2019c)
Hobart	9	Vgib*	Wiltshire *et al.* (2019a)
SE Tasmania	35	Mjap*	Giblot-Ducray *et al.* (2022)
SE Tasmania	7	Mjap	Deveney *et al.* (2020)†
SE Tasmania	62	Mros*	Giblot-Ducray *et al.* (2022)
SE Tasmania	31	Mros*	Deveney *et al.* (2020)†
Melbourne	30	Hsan*	this project
Melbourne	16	Hsan*	Wiltshire *et al.* (2019a)‡
Melbourne	2	Asen*, Vgib*	Wiltshire *et al.* (2019a)
Portland	24	Mros	Wiltshire *et al.* (2022)
Portland	2	Vgib	Wiltshire *et al.* (2022)
Port Kembla	3	Vgib	Wiltshire *et al.* (2019c)
Newcastle	3	Vgib	Wiltshire *et al.* (2019c)
Botany Bay	3	Vgib	Wiltshire *et al.* (2019c)

Amplicons were purified after the second PCR reaction using AMPure XP beads (Beckman Coulter, USA). Primer dimerisation in the *Mya japonica* runs required additional amplicon purification. Not all samples however could be successfully cleaned and sequenced, and therefore sequences were not obtained for some samples of interest in the *Mya japonica* run. Final individual libraries were quantified using Invitrogen™ Quant-iT™ PicoGreen™ dsDNA Assay Kits (ThermoFisher Scientific) and pooled to equal concentrations where possible for sequencing. Sequencing of the amplicon libraries was carried out by the Australian Genome Research Facility (AGRF) using the Illumina MiSeq platform.

HTS sequences were processed using a bioinformatics pipeline coded within the *Snakemake* workflow management system (Mölder *et al.* 2021). Quality filtered reads were clustered based on 99% sequence similarity into Operational Taxonomic Units (OTUs). A local version of the Genbank nucleotide database (downloaded on 23rd Jan 2023) was used to perform a local *blastn* search of OTUs and assign taxonomy. Taxonomy was assigned to OTUs with ≥ 97% sequence similarity taken as the cut-off for assignment to species level. Individual sequences were checked for their complementarity to the assay primer and probe sequence, after removing the primer sequences used for library preparation, to determine the potential for the assay to detect non-target sequences. Depending on the number, location, and type of mismatches, sequences were further categorised as potentially confounding or not. Likelihood of non-target amplification decreases with increasing number of mismatches, but sequence mismatches confer greater specificity the nearer they occur to the 3' end of primers (Lefever *et al.* 2013; Wright *et al.* 2014; So *et al.* 2020). The type of nucleotides involved in mismatches also strongly influences the likelihood of mismatches preventing amplification (So *et al.* 2020), as does the length of primer and probe sequences relative to the number of mismatches (Wilcox *et al.* 2013). These factors were therefore all considered in determining the likelihood for non-target amplification of each sequence.

3. RESULTS

3.1. Genomic DNA testing

Quantification of gDNA extracts confirmed that DNA yield from samples extracted using the FFPE kit was within the range of specimens extracted using the B&T kit, and, in all cases, the target concentration of 200 pg μL^{-1} was achieved and applied for testing.

The Hsan qPCR assay was assessed against 34 gDNA extracts from 25 non-target decapod species, including one extract from *Eriocheir sinensis*, while the Esin assay was assessed using 35 gDNA extracts from 25 non-target decapod species, including two *Hemigrapsus sanguineus* extracts, one being pre-extracted DNA and the other a tissue sample extracted for this project (Table 8). There were no detections by either assay of non-target species DNA. The Hsan qPCR assay detected *Hemigrapsus sanguineus* with a C_T value of 21.9 for the pre-extracted DNA sample, and of 22.2 for the gDNA extracted using the B&T kit. The Esin qPCR detected *Eriocheir sinensis* DNA extracted from tissue using the B&T kit with a C_T value of 19.2 (Table 8). Several of the closely related specimens were preserved with unknown fixative and extracted using the FFPE kit, while target specimens were ethanol preserved and extracted with the B&T kit. This was due to no ethanol-preserved specimens being available for these species, which would have been used in preference. Quantification of gDNA concentrations, however, provides confidence that the non-detections in FFPE-extracted samples were due to assay specificity rather than to a lack of DNA in these samples. Separate specimens of *Brachynotus spinosus*, furthermore, were extracted with each kit with all being un-detected.

The Asen and Vgib qPCR assays were assessed against 52 extracts from 42 non-target species each, including one sample of each target (Table 9). The samples included 15 extracts from tropical species from NT, Qld or northern WA. The Asen and Vgib assays each detected DNA of their respective targets from pre-existing DNA extracts with a C_T of 21.0 in each case (Table 9). There were no non-target detections for the Asen assay, while the Vgib assay returned a high C_T (43.2) detection for the *Anadara trapezia* DNA extract (Table 9). Sequencing of the *A. trapezia* sample showed that it had 100% match to the Vgib assay reverse primer and probe regions with 7 mismatches to the forward primer, but none at the 3' end. These mismatches therefore are likely insufficient to prevent amplification, leading to cross-reactivity, although with a higher C_T value resulting than for target DNA.

The assays for 11 species: *Arcuatula senhousia, Perna canaliculus, P. perna, P. viridis, Mytilopsis sallei, Varicorbula gibba, Charybdis japonica, Hemigrapsus sanguineus, Eriocheir sinensis, Rhithropanopeus harrisii* and *Didemnum vexillum*, were applied to gDNA extracts of between 1 and 6 target species specimens per assay (Table 10). All gDNA extracts were detected by the

relevant assay, with C_T values ranging from 19.2 to 35.8. The FPPE extracted specimen of *C. japonica* was detected with a C_T value of 19.4, providing confidence in the quality of other gDNA samples extracted by this method that were used in specificity testing.

C_T values ≥ 25, above the range typically expected for testing 200 pg μL^{-1} gDNA, were recorded for the Pper assay in 1 of 2 samples, for Pvir and Vgib in 1 of 4 samples each, for Msal in 1 of 6 samples, for Dvex in 3 of 8 samples, and for Asen in all four samples tested (Table 10). The samples with highest C_T in each case were mostly of types more likely to contain mixed DNA or potential inhibitors, e.g., small bivalve specimens where all tissue was used rather than just adductor muscle. The *Didemnum vexillum* extracts with high C_T were all from specimens that were growing intermingled with other fouling organisms, making selection of clean tissue difficult.

3.2. *Hemigrapsus sanguineus* detection in plankton samples

Hemigrapsus sanguineus DNA was detected in 12 of the 30 plankton samples collected in February 2022 for this project, and in 10 of the 128 samples from 2017-18, comprising 6 from 32 samples collected in November 2017 and 4 from 32 samples collected in March 2018 (Figure 4). There were no detections in sets of 32 samples collected in July – August 2017 or June 2018. The detections in 2017-18 occurred in the vicinity of Station Pier, the entrance to Webb Dock, and around the Williamstown wharves. Sampling in February 2022 (this project) did not include the Williamstown area, and detections occurred around Station Pier and in the Yarra River. There was no PCR inhibition in any of the samples collected for this project (scale factor = 1 for all samples). In the previously collected samples, only two samples showed any inhibition: one from Jul 2017 with a scale factor 1.73, and one from Mar 2018 with a scale factor of 11.4. There was, therefore, insufficient variation in scale factor for this to be included as a factor in the analyses.

Table 8. Test results for *Hemigrapsus sanguineus* (Hsan) and *Eriocheir sinensis* (Esin) -specific qPCR assays applied to gDNA extracts from target specimens and non-target decapod species. Testing used DNA at 200 pg μL^{-1}. Source of specimen shows state in Australia or country if not Australia for specimens extracted by this project. DNA kit is not applicable (NA) for existing DNA extracts used in testing, B&T = Blood and Tissue kit, FFPE = formalin fixed-paraffin embedded kit. Test results show ND = not detected, C_T = Cycle threshold value of qPCR where detected, nt = not tested. Multiple results for a taxon represent samples from separate individuals. *Species not in Genbank.

Taxon	Superfamily	Family	Source	DNA Kit	Esin C_T	Hsan C_T
Hemigrapsus sanguineus	Grapsoidea	Varunidae	Existing DNA	NA	ND	21.9
Hemigrapsus sanguineus	Grapsoidea	Varunidae	Vic	B&T	ND	22.2
Eriocheir sinensis	Grapsoidea	Varunidae	Poland	B&T	19.2	ND
Helograpsus haswellianus*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Helograpsus haswellianus*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Helograpsus haswellianus*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Cyclograpsus audonii*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Cyclograpsus audonii*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Cyclograpsus granulosus	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Paragrapsus gaimardii*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Brachynotus spinosus*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Brachynotus spinosus*	Grapsoidea	Varunidae	SA	FFPE	ND	ND
Brachynotus spinosus*	Grapsoidea	Varunidae	SA	B&T	ND	ND
Guinusia chabrus*	Grapsoidea	Plagusiidae	SA	FFPE	ND	ND
Charybdis japonica	Portunoidea	Portunidae	SA	B&T	ND	ND
Charybdis* c.f. *feriata	Portunoidea	Portunidae	SA	B&T	ND	ND
Charybdis helleri	Portunoidea	Portunidae	Existing DNA	NA	ND	ND
***Lissocarcinus* sp.**	Portunoidea	Portunidae	SA	FFPE	ND	ND
Portunus armatus	Portunoidea	Portunidae	Existing DNA	NA	ND	ND
Portunus armatus	Portunoidea	Portunidae	SA	FFPE	ND	ND
***Thalamita* sp.**	Portunoidea	Portunidae	Existing DNA	NA	ND	ND
Carcinus maenas	Portunoidea	Portunidae	Existing DNA	NA	ND	ND

Taxon	Superfamily	Family	Source	DNA Kit	Esin C_T	Hsan C_T
Pilumnus monilifera*	Pilumnoidea	Pilumnidae	SA	FFPE	ND	ND
Macropthalmus* sp.	Ocypooidea	Macrophthalmidae	SA	B&T	ND	ND
Paguristes frontalis*	Paguroidea	Diogenidae	SA	B&T	ND	ND
Cancer irroratus	Cancroidea	Cancridae	Canada	B&T	ND	ND
Homarus amercanus	Nephropoidea	Nephropidae	Canada	B&T	ND	ND
Litocheira bispinosa*	Goneplacoidea	Litocheiridae	Vic	B&T	ND	ND
Ozius truncatus	Eriphioidea	Oziidae	SA	B&T	ND	ND
Halicarcinus* c.f. *quoyi	Hymenosomatoidea	Hymenosomatidae	Tas	B&T	ND	ND
Halicarcinus* c.f. *quoyi	Hymenosomatoidea	Hymenosomatidae	Tas	B&T	ND	ND
Neorhyncoplax* sp.	Hymenosomatoidea	Hymenosomatidae	Vic	B&T	ND	ND
Halicarcinus ovatus	Hymenosomatoidea	Hymenosomatidae	Vic	B&T	ND	ND
Dromiidae	Dromioidea	Dromiidae	SA	FFPE	ND	ND
Dromiidae	Dromioidea	Dromiidae	Vic	B&T	ND	ND

Table 9. Test results for *Arcuatula senhousia* (Asen) and *Varicorbula gibba* (Vgib) -specific qPCR assays applied to gDNA extracts from target specimens and non-target bivalve species. Testing used DNA at 200 pg μL^{-1}. Source of specimen shows state in Australia or country if not Australia for specimens extracted by this project. DNA kit is not applicable (NA) for existing DNA extracts used in testing, B&T = Blood and Tissue kit, FFPE = formalin fixed – paraffin embedded kit. Test results show ND = not detected, C_T = Cycle threshold value of qPCR where detected, nt = not tested. Multiple results for a taxon represent samples from separate individuals. *Species not in Genbank.

Taxon	Superfamily	Family	Source	DNA Kit	Asen C_T	Vgib C_T
Arcuatula senhousia	Mytiloidea	Mytilidae	Existing DNA	NA	21.0	ND
Varicorbula gibba	Myoidea	Corbulidae	Existing DNA	NA	ND	21.0
Mya japonica	Myoidea	Myidae	Tas	B&T	ND	ND
Mya japonica	Myoidea	Myidae	Tas	B&T	ND	ND
Corbula smithiana*	Myoidea	Corbulidae	NT	B&T	ND	ND
Corbula smithiana*	Myoidea	Corbulidae	NT	B&T	ND	ND
***Corbula* sp.**	Myoidea	Corbulidae	Qld	B&T	ND	ND

Taxon	Superfamily	Family	Source	DNA Kit	Asen C_T	Vgib C_T
***Corbula* sp.**	Myoidea	Corbulidae	Qld	B&T	ND	ND
Iridona iridescens*	Tellinoidea	Tellinidae	NT	B&T	ND	ND
Macomona deltoidalis*	Tellinoidea	Tellinidae	SA	FFPE	ND	ND
Donax electilis*	Tellinoidea	Donacidae	SA	B&T	ND	ND
Gafrarium australe*	Veneroidea	Veneridae	NT	B&T	ND	ND
Gafrarium pectinatum	Veneroidea	Veneridae	NT	B&T	ND	ND
Costellipitar inconstans*	Veneroidea	Veneridae	NT	B&T	ND	ND
***Tapes* sp.**	Veneroidea	Veneridae	SA	B&T	ND	ND
Bassina disjecta*	Veneroidea	Veneridae	SA	FFPE	ND	ND
Dreissena polymorpha	Dreissenoidea	Dreissenidae	Existing DNA	NA	ND	ND
Dreissena burgensis	Dreissenoidea	Dreissenidae	Existing DNA	NA	ND	ND
Mytilopsis sallei	Dreissenoidea	Dreissenidae	Existing DNA	NA	ND	ND
Mytilopsis sallei	Dreissenoidea	Dreissenidae	Thailand	B&T	ND	ND
Mytilopsis sallei	Dreissenoidea	Dreissenidae	Thailand	B&T	ND	ND
Pinctada albina	Pteroidea	Pteriidae	SA	B&T	ND	ND
Malleus* meridianus	Pteroidea	Malleidae	Existing DNA	NA	ND	ND
Hiatella australis*	Hiatelloidea	Hiatellidae	SA	B&T	ND	ND
Philobyra robensis*	Limopsoidea	Phylobryidae	Vic	B&T	ND	ND
Anadara trapezia	Arcoidea	Arcidae	SA	B&T	ND	43.2
Glycymeris radians*	Arcoidea	Glycymerididae	SA	B&T	ND	ND
Musculus nanus*	Mytiloidea	Mytilidae	SA	FFPE	ND	ND
***Musculus* sp.**	Mytiloidea	Mytilidae	NT	B&T	ND	ND
***Musculista* sp.**	Mytiloidea	Mytilidae	Qld	B&T	ND	ND
Mytella strigata	Mytiloidea	Mytilidae	Thailand	B&T	ND	ND
Mytella strigata	Mytiloidea	Mytilidae	Thailand	B&T	ND	ND
Mytella strigata	Mytiloidea	Mytilidae	Thailand	B&T	ND	ND

Taxon	Superfamily	Family	Source	DNA Kit	Asen C_T	Vgib C_T
Mytella strigata	Mytiloidea	Mytilidae	Thailand	B&T	ND	ND
Trichomya hirsuta	Mytiloidea	Mytilidae	Existing DNA	NA	ND	ND
Mytilus galloprovincialis	Mytiloidea	Mytilidae	Existing DNA	NA	ND	ND
Brachidontes rostratus	Mytiloidea	Mytilidae	Existing DNA	NA	ND	ND
Brachidontes crebristriatus*	Mytiloidea	Mytilidae	NT	B&T	ND	ND
***Brachidontes* sp.**	Mytiloidea	Mytilidae	Qld	B&T	ND	ND
***Brachidontes* sp.**	Mytiloidea	Mytilidae	Qld	B&T	ND	ND
***Brachidontes* sp.**	Mytiloidea	Mytilidae	Qld	B&T	ND	ND
Perna canaliculus	Mytiloidea	Mytilidae	Existing DNA	NA	ND	ND
Perna viridis	Mytiloidea	Mytilidae	Existing DNA	NA	ND	ND
Septifer bilocularis	Mytiloidea	Mytilidae	WA	B&T	ND	ND
Myochamidae	Myochamoidea	Myochamidae	SA	FFPE	ND	ND
***Laternula* sp.**	NA	Laternulidae	SA	FFPE	ND	ND
***Neotrigonia* sp.**	Trigonioidea	Trigoniidae	SA	FFPE	ND	ND
Monia zelandica*	Anomioidea	Anomiidae	Existing DNA	NA	ND	ND
Mimachlamys asperrima	Pectinoidea	Pectinidae	Existing DNA	NA	ND	ND
Magallana gigas	Ostreoidea	Ostreidae	Existing DNA	NA	ND	ND
Ostrea angasi	Ostreoidea	Ostreidae	Existing DNA	NA	ND	ND
Saccostrea glomerata	Ostreoidea	Ostreidae	Existing DNA	NA	ND	ND
Mactridae	Mactroidea	Mactridae	SA	B&T	ND	ND

Table 10. Test results for gDNA tested with species-specific assays. Testing used DNA at 200 pg μL^{-1}. Specimen details show year and location of specimen collection (limited details were available for some specimens). B&T = Blood and Tissue DNA extraction kit, FFPE = formalin fixed-paraffin embedded kit. C_T = Cycle threshold value of qPCR. Multiple results for a taxon with same collecting year and location represent samples from separate individuals.

Taxon	Specimen details	Subsample	DNA Kit	C_T
Arcuatula senhousia	2018, Yarra River, Vic	All tissue	B&T	35.8
Arcuatula senhousia	2018, Webb Dock, Vic	All tissue	B&T	32.6
Arcuatula senhousia	2020, Swan River, WA	Adductor muscle	B&T	28.3
Arcuatula senhousia	2020, Swan River, WA	Adductor muscle	B&T	29.5
Perna canaliculus	2021, NZ (commercially purchased)	Adductor muscle	B&T	21.8
Perna canaliculus	2021, NZ (commercially purchased)	Adductor muscle	B&T	21.7
Perna canaliculus	2021, NZ (commercially purchased)	Adductor muscle	B&T	23.1
Perna perna	2021, Scarborough, South Africa	Adductor muscle	B&T	29.6
Perna perna	2021, Simonstown, South Africa	Adductor muscle	B&T	21.7
Perna viridis	2011, Singapore	Adductor muscle	B&T	21.1
Perna viridis	2009, Port Adelaide (on vessel hull)	Adductor muscle and foot	B&T	25.0
Perna viridis	2021, Sri Racha, Thailand	Adductor muscle	B&T	20.1
Perna viridis	2021, Sri Racha, Thailand	Adductor muscle	B&T	20.0
Mytilopsis sallei	2020, Surat Thani, Thailand	Adductor muscle	B&T	22.3
Mytilopsis sallei	2020, Surat Thani, Thailand	Adductor muscle	B&T	20.7
Mytilopsis sallei	2020, Surat Thani, Thailand	Adductor muscle	B&T	21.5
Mytilopsis sallei	2020, Surat Thani, Thailand	Adductor muscle	B&T	22.8
Mytilopsis sallei	2009, South Sulawesi, Indonesia	Adductor muscle and foot	B&T	20.3
Mytilopsis sallei	1999, Cullen Marina, NT	Adductor muscle and foot	B&T	29.9
Varicorbula gibba	2018, Sullivan Cove, Tas	All tissue	B&T	27.6
Varicorbula gibba	2018, Sullivan Cove, Tas	All tissue	B&T	19.8
Varicorbula gibba	2018, Yarra River, Vic	All tissue	B&T	19.9
Varicorbula gibba	2018, Yarra River, Vic	All tissue	B&T	25.3

Taxon	Specimen details	Subsample	DNA Kit	C_T
Varicorbula gibba	2018, Webb Dock, Vic	All tissue	B&T	24.4
Charybdis japonica	2021, North Arm, SA	Leg tissue	B&T	22.9
Charybdis japonica	2000, Outer Harbor, SA	Leg tissue	FFPE	19.4
Charybdis japonica	2019, Outer Harbor, SA	Leg tissue	B&T	21.5
Charybdis japonica	2020, North Arm, SA	Leg tissue	B&T	21.3
Eriocheir sinensis	2015, Poland	Leg tissue	B&T	19.2
Hemigrapsus sanguineus	2021, Mount Martha, Vic	Leg tissue	B&T	20.0
Hemigrapsus sanguineus	2021, Mount Martha, Vic	Leg tissue	B&T	21.2
Hemigrapsus sanguineus	2021, Mount Martha, Vic	Leg tissue	B&T	22.3
Hemigrapsus sanguineus	2021, Mount Martha, Vic	Eggs from berried female	B&T	21.3
Rhithropanopeus harrisii	2014, Estonia	Leg and body tissue	B&T	20.7
Rhithropanopeus harrisii	2008, Mission River, Texas, USA	Eggs from berried female	B&T	23.3
Rhithropanopeus harrisii	2008, Mission River, Texas, USA	Leg tissue	B&T	20.8
Rhithropanopeus harrisii	2022, Hubbard Creek, Texas, USA	Leg tissue	B&T	20.1
Rhithropanopeus harrisii	2022, Hubbard Creek, Texas, USA	Leg tissue	B&T	21.1
Rhithropanopeus harrisii	2022, Dassower Lake, Luebeck, Germany	Leg tissue	B&T	21.6
Rhithropanopeus harrisii	2022, Dassower Lake, Luebeck, Germany	Leg tissue	B&T	22.0
Rhithropanopeus harrisii	2022, Dassower Lake, Luebeck, Germany	Leg tissue	B&T	21.6
Didemnum vexillum	2022, Garden Island, WA	Thoracic tissue	B&T	27.6
Didemnum vexillum	2022, Garden Island, WA	Thoracic tissue	B&T	25.2
Didemnum vexillum	2022, Garden Island, WA	Thoracic tissue	B&T	27.6
Didemnum vexillum	2010, Desolation Sound, Canada	Thoracic tissue	B&T	23.4
Didemnum vexillum	2010, Desolation Sound, Canada	Thoracic tissue	B&T	24.5
Didemnum vexillum	2010, Desolation Sound, Canada	Thoracic tissue	B&T	22.9
Didemnum vexillum	2010, Desolation Sound, Canada	Thoracic tissue	B&T	21.8
Didemnum vexillum	2010, Lemmens Inlet, Canada	Thoracic tissue	B&T	24.7

The lack of *Hemigrapsus sanguineus* DNA detection in two of the sample sets creates the issue of complete separation (Mansournia *et al.* 2017) in the data , leading to wide HDIs for the ZAG model estimates of detection likelihood for these two data sets (Appendix section 6.1.2, Figure 5). Additionally, no estimate is available for DNA yield for either of these two sample sets given the lack of detections. Pooling data so that each group of samples contained some detections was not appropriate due to the confounding effect of sample timing and year or project that would have resulted. Statistical comparison could therefore only be made between the sample sets with detections, i.e.: Nov 2017, Mar 2018, and Feb 2022. Model results showed that detection likelihood was higher in the Feb 2022 than Mar 2018 sample set, and DNA yield was higher in the Feb 2022 samples than in both the Nov 2017 and Mar 2018 sets (Table 11). The mean estimate of detection likelihood was greater for Feb 2022 than Nov 2017 and Nov 2017 than Mar 2018, while the mean estimate of DNA yield was lower for Nov 2017 than Mar 2018, but the 95% HDIs contained zero in each case, hence the estimates are not statistically different between these sample sets. It should be noted that the Feb 2022 sample set used a different specification plankton net, hence, differences between this and earlier sample sets may be due to different net characteristics, as well as to the different sample timing and potential changes in the *H. sanguineus* population size.

Table 11. Statistical comparison between sample set coefficients for *Hemigrapsus sanguineus* detection likelihood and DNA yield from the zero added gamma spatial model. Sample sets are considered different where the 95% HDI of the difference between coefficients does not contain zero. A positive estimate for the comparison indicates that the parameter is higher for the first sample set in the comparison, with a negative estimate indicating that the parameter is higher for the second sample set. Mean estimated differences are on the link scale: cloglog for detection likelihood, log for DNA yield. The exponent of the mean estimate provides the predicted multiplicative difference in each parameter in the first compared with second sample set (see Appendix section 6.1.1). *Mean estimate statistically different from zero.

Parameter: difference between	Mean estimate (95% HDI)	Multiplicative difference
Detection likelihood: Mar 2018 – Feb 2022	−1.53 (−3.14 – −0.214)*	0.22 (0.04 – 0.81)
Detection likelihood: Nov 2017 – Feb 2022	−1.07 (−2.60 – 0.186)	0.34 (0.07 – 1.20)
Detection likelihood: Nov 2017 – Mar 2018	0.459 (−0.854 – 1.77)	1.58 (0.43 – 5.87)
DNA yield: Mar 2018 – Feb 2022	−2.46 (−4.17 – −0.42)*	0.085 (0.015 – 0.657)
DNA yield: Nov 2017 – Feb 2022	−2.98 (−4.55 – −1.21)*	0.051 (0.011 – 0.298)
DNA yield: Nov 2017 – Mar 2018	−0.527 (−2.47 – 1.40)	0.590 (0.085 – 4.06)

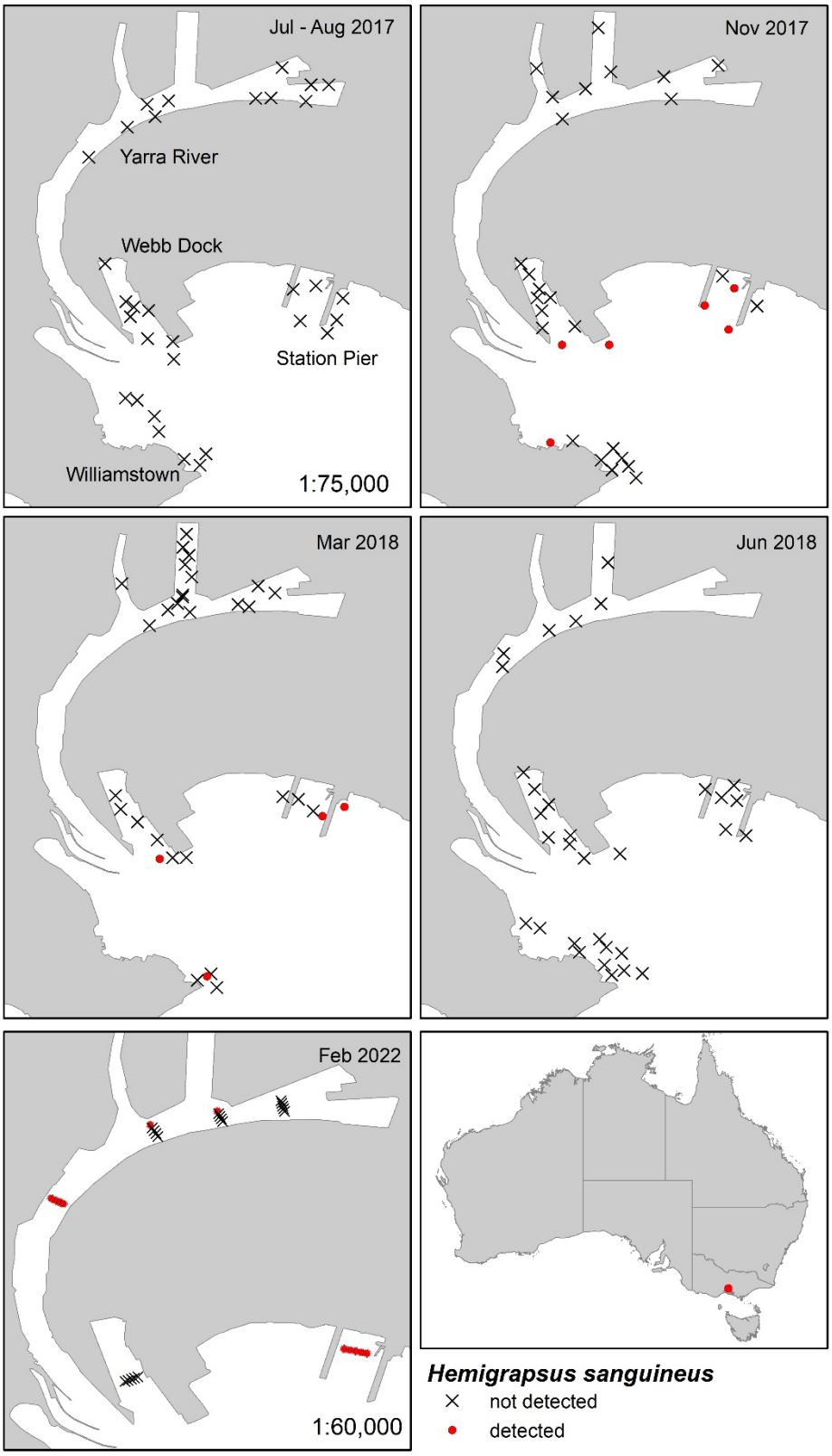

Figure 4. Map of plankton sampling locations and detections of *Hemigrapsus sanguineus* DNA in Melbourne.

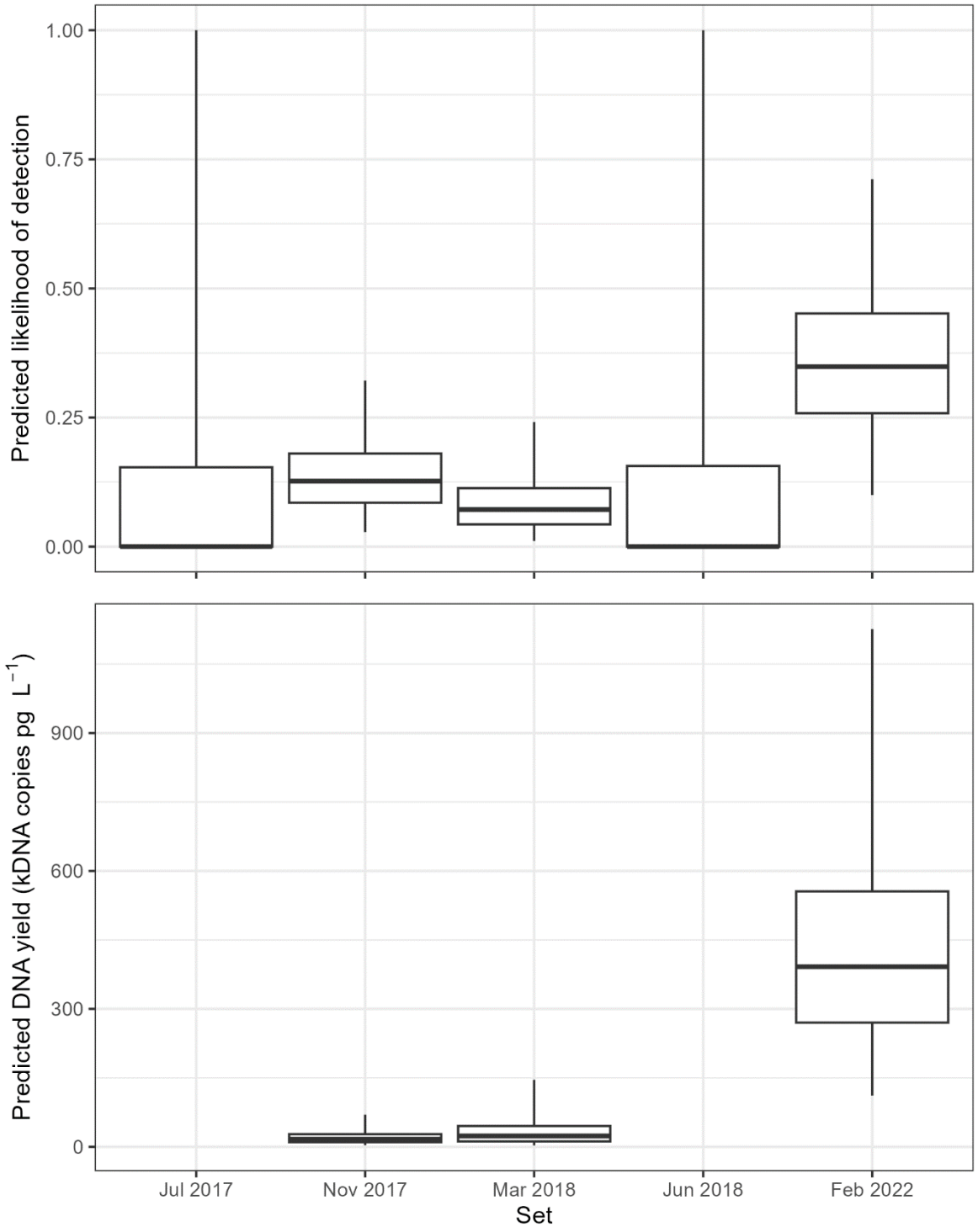

Figure 5. Predicted detection likelihood and DNA yield in samples with detection from ZAG model of results of testing Melbourne plankton samples for *Hemigrapsus sanguineus.*

The estimated range and standard deviation of the spatial fields for detection likelihood and DNA yield estimated by the ZAG model are shown in Table 12. The spatial field showed increased likelihood of detection and higher DNA yield around Station Pier, the entrance to Webb Dock and the downstream portion of the Yarra River. Lower than model average DNA yield and detection likelihood occurred in the upstream area of the Yarra River; DNA yield was also lower in the Williamstown area, and detection likelihood was lower within Webb Dock (Figure 6). Note that the spatial field shows variation from the model average prediction across the model area after accounting for the effect of sample set, with the standard deviation of each field describing the scale of the variation. The detection likelihood spatial random effect (Figure 6) had highest values of ~0.8 on the cloglog scale, which is equivalent to a relative planktonic abundance 2.2 times higher than model average prediction, and lowest values of ~ −0.6 on the cloglog scale, equivalent to relative planktonic abundance of 0.54 times the model average. For DNA yield, the spatial effect shows difference in log(DNA yield) from the model average, with the range of values (-0.08 to 0.09) equivalent to a multiplicative change of 0.9 to 1.1 times the model average. There was, therefore, much greater spatial variability in detection likelihood than DNA yield.

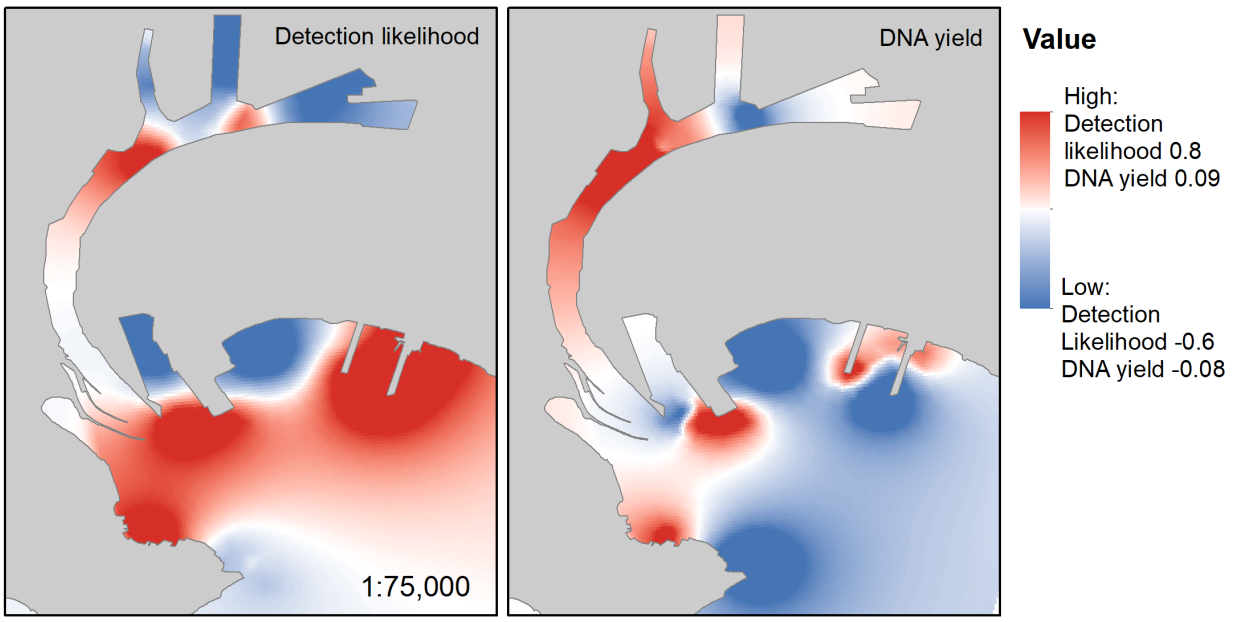

Figure 6. Spatial fields from zero added gamma model of *Hemigrapsus sanguineus* likelihood of detection (left) and DNA yield (right) in Melbourne. The spatial fields shows where each parameter is higher (red) or lower (blue) than model average prediction across the model area. The value shows the difference from model average on the cloglog scale for detection likelihood and log scale for DNA yield. See Appendix section 6.1.1 for further information on interpretation of these values.

Table 12. Spatial field parameters estimated by ZAG model for *Hemigrapsus sanguineus* detections in plankton samples from Melbourne.

Parameter:	Mean estimate (95% HDI)
Detection likelihood: range (km)	0.887 (0.083 – 2.11)
Detection likelihood: standard deviation	0.902 (0.177 – 1.90)
DNA yield: range (km)	1.17 (0.070 – 3.47)
DNA yield: standard deviation	0.222 (0.004 – 0.605)

3.3. Field specificity

In samples tested for this project, there were two detections by the Mros assay, in one sample each from Gove, NT, and Hay Point, Qld. There were no detections by the other eight assays, which were applied to 184 – 230 samples each (see Table 4 for source locations of samples tested for each species).

The Rhar, Cjap, Pper, and Pvir assays were applied by Wiltshire *et al.* (2020b) to 262 samples from WA, comprising 46 from Bunbury, 120 from Kwinana, 66 from Fremantle, and 15 from Geraldton. One detection by the Rhar assay was recorded in a sample from Fremantle by that project. The assays for *Perna canaliculus* and *Mytilopsis sallei* have been applied to over 2,000 samples each during surveillance around Australia since 2015 (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019c, 2020b), with no detections. The assays for *Arcuatula senhousia* and *Varicorbula gibba* have also been widely applied since 2015, including in areas with known populations of each of these pests. Detections have, however, also occurred in areas without known pest occurrence, including in some where environmental conditions are unlikely to be suitable for these pests to establish (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019c). There have been 141 detections of *Arcuatula senhousia* in 1337 samples, and 252 detections of *Varicorbula gibba* in 833 samples from regions where these pests are considered unlikely to establish populations. These areas include Gove, NT and Weipa, Hay Point, Gladstone and Brisbane, Qld, for both species, and, also Darwin, NT, and Geraldton, WA for *Varicorbula gibba.* Field specificity results for the 13 assays considered are summarised in Table 13.

Apparent specificity, i.e., the proportion of samples from areas free of each pest with no detection, was calculated for each assay. Apparent specificity is 100% for the nine assays with no detections in the sample set, while the Mros and Rhar assays have apparent specificity of 99.0 and 99.8%

respectively, the Asen assay has apparent specificity 89.5% and the Vgib assay 61.1% (Table 13).

3.4. Performance of assays applied to plankton

3.4.1. Detections in spiked samples

Target DNA was detected in 147–197 of the 200 samples that had pest tissue added for each species (Table 14). There were > 180 detections, providing > 90% apparent DSe, for the Dvex, Hsan, Mjap, Mros, Pcan, Cjap and Rhar assays, with the lowest apparent DSe being for the Pper and Pvir assays (both < 80%).

Table 13. Summary of samples tested to demonstrate field specificity of 13 assays. Samples included in the assessment are those from areas considered free of each pest. The total samples tested for each pest include testing of 9 species by the current project, plus results from previous molecular surveillance of relevant areas.

Species	Sample areas free of pest	Total samples tested (this project)	Detections	Apparent specificity
Arcuatula senhousia	Darwin, Gove, Weipa, Cairns, Hay Point, Gladstone, Brisbane, Newcastle, Botany Bay, Port Kembla, Bunbury, Geraldton	1337 (0)	141	89.5%
Charybdis japonica	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Fremantle, Kwinana, Bunbury, Geraldton, SE Tasmania	446 (184)	0	100%
Didemnum vexillum	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Bunbury, Fremantle, SE Tasmania	230 (230)	0	100%
Eriocheir sinensis	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Bunbury, Fremantle, SE Tasmania	230 (230)	0	100%
Hemigrapsus sanguineus	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Bunbury, Fremantle, SE Tasmania	200 (200)	0	100%
Maoricolpus roseus	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Melbourne, Bunbury, Fremantle	202 (202)	2	99.0%

Species	Sample areas free of pest	Total samples tested (this project)	Detections	Apparent specificity
Mya japonica	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Melbourne, Bunbury, Fremantle	202 (202)	0	100%
Mytilopsis sallei	Darwin, Gove, Weipa, Cairns, Hay Point, Gladstone, Brisbane, Newcastle, Botany Bay, Port Kembla, Melbourne, Adelaide, Bunbury, Fremantle, Kwinana, Geraldton, Devonport, Hobart	2170 (0)	0	100%
Perna canaliculus	Darwin, Gove, Weipa, Cairns, Hay Point, Gladstone, Brisbane, Newcastle, Botany Bay, Port Kembla, Melbourne, Adelaide, Bunbury, Fremantle, Kwinana, Geraldton, Devonport, Hobart	2170 (0)	0	100%
Perna perna	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Fremantle, Kwinana, Bunbury, Geraldton, SE Tasmania	446 (184)	0	100%
Perna viridis	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Fremantle, Kwinana, Bunbury, Geraldton, SE Tasmania	446 (184)	0	100%
Rhithropanopeus harrisii	Darwin, Gove, Weipa, Cairns, Hay Point, Brisbane, Newcastle, Botany Bay, Fremantle, Kwinana, Bunbury, Geraldton, SE Tasmania	446 (184)	1	99.8%
Varicorbula gibba	Darwin, Gove, Weipa, Cairns, Hay Point, Gladstone, Brisbane, Geraldton	632 (0)	246	61.1%

In the 40 samples without added pest tissue, there was a single detection for each of the Mros and Pcan assays, and there were three detections for the Asen assay, five for Vgib, 11 for Rhar and 38 for Pper. The *Maoricolpus roseus* detection was in a sample from the Adelaide coast, while the *P. canaliculus* detection was in a sample from Gladstone. The *A. senhousia* detections were in samples from Gladstone. These samples were surplus samples collected by Wiltshire *et al.* (2019a), and *A. senhousia* detections were also recorded in the Gladstone samples analysed for that project. *Varicorbula gibba* was detected in three Gladstone and two Adelaide samples, while *R. harrisii* was detected in three Gladstone and eight Adelaide samples. Detections of the

latter species occurred mainly in samples spiked with higher doses (1 or 2 g of sand), suggesting possible cross-contamination of the sand stock. The *P. perna* detections included 11 samples from Gladstone, 23 from Adelaide, three from Klein Point and one from Thevenard.

Table 14. Summary of detections of each species in samples with and without added pest tissue, and apparent assay performance calculated from these results.

Species	Detections in 200 samples with added pest	Detections in 40 samples without added pest	Apparent DSe	Apparent DSp
Arcuatula senhousia	170	3	85.0%	92.5%
Charybdis japonica	190	0	95.0%	100.0%
Didemnum vexillum	184	0	92.0%	100.0%
Eriocheir sinensis	166	0	83.0%	100.0%
Hemigrapsus sanguineus	192	0	96.0%	100.0%
Maoricolpus roseus	184	1	92.0%	97.5%
Mya japonica	189	0	94.5%	100.0%
Mytilopsis sallei	169	0	84.5%	100.0%
Perna canaliculus	181	1	90.5%	97.5%
Perna perna	156	38	78.0%	5.0%
Perna viridis	147	0	73.5%	100.0%
Rhithropanopeus harrisii	197	11	98.5%	72.5%
Varicorbula gibba	170	5	85.0%	87.5%

The high proportion of Pper detections in samples without added *P. perna* tissue prompted consideration that this assay is not specific against one or both of its congeners, *P. canaliculus* and *P. viridis*, both of which were added to these samples. Simpson *et al.* (2018) designed assays for each of the three *Perna* species, with their assays for Pper and Pvir implmented and tested here, while the Pcan assay implemented in the SARDI system, and assessed in this project, is the one designed previously by Bott and Giblot-Ducray (2011b). The forward and reverse primers of the Simpson *et al.* (2018) Pper and Pvir assays are the same (as are primers for their *P. canaliculus* assay that was not assessed here), and amplify all three *Perna* spp., although the probes were designed to be specific to each species. The Pper assay was applied to gDNA extracts from *P. canaliculus* and *P. viridis* specimens and to other bivalve and gastropod species added to the relevant sand stock, using the same specimens as used in the spiking, following the method used for specificity assessment (see section 2.1.3). The Pper assay detected *P. canaliculus* in all three gDNA extracts of this species with C_T of 23.7 – 27.3, while there was no detection in three tested extracts of *P. viridis* or other species (Table 15).

Rhithropanopeus harrisii is not closely related to any of the species added to the sand stocks, but to ensure the assay was not detecting DNA of any of these species, the Rhar assay was applied to the relevant gDNA extracts (i.e., from the same source specimens as used in spiking) with no detection recorded (Table 15). Aliquots of spiked sand with mass of 0.5, 1, 2 and 5 g that had not been added to plankton samples were also tested. The Rhar assay returned detections with C_T of 34.4 – 37.0 when applied to the spiked sand samples (Table 15), confirming that the sand had been contaminated with *R. harrisii* tissue or DNA.

Table 15. Additional qPCR test results for Pper and Rhar assays applied to gDNA extracts and spiked sand stock.

Sample tested	Sample details	Assay applied	C_T
Perna canaliculus	New Zealand	Pper	25.3
Perna canaliculus	New Zealand	Pper	23.7
Perna canaliculus	New Zealand	Pper	27.3
Perna viridis	Thailand	Pper	ND
Perna viridis	Thailand	Pper	ND
Perna viridis	Thailand	Pper	ND
Mytilopsis sallei	Darwin, NT	Pper	ND
Arcuatula senhousia	Perth, WA	Pper	ND
Mya japonica	Orford, Tas	Pper	ND
Maoricolpus roseus	D'Entrecasteaux Channel, Tas	Pper	ND
Varicorbula gibba	Melbourne, Vic	Rhar	ND
Charybdis japonica	Port Adelaide, SA	Rhar	ND
Eriocheir sinensis	Portugal	Rhar	ND
Maoricolpus roseus	D'Entrecasteaux Channel, Tas	Rhar	ND
Hemigrapsus sanguineus	Mount Martha, Vic	Rhar	ND
Perna canaliculus	New Zealand	Rhar	ND
Perna perna	South Africa	Rhar	ND
Perna viridis	Thailand	Rhar	ND
Mytilopsis sallei	Darwin, NT	Rhar	ND
Mya japonica	Orford, Tas	Rhar	ND
Didemnum vexillum	Canada	Rhar	ND
Sand stock E	0.5 g	Rhar	35.8
Sand stock E	0.5 g	Rhar	34.9
Sand stock E	0.5 g	Rhar	34.4
Sand stock E	1.0 g	Rhar	37.0
Sand stock E	1.0 g	Rhar	36.4
Sand stock E	2.0 g	Rhar	34.8
Sand stock E	2.0 g	Rhar	36.6
Sand stock E	5.0 g	Rhar	34.9

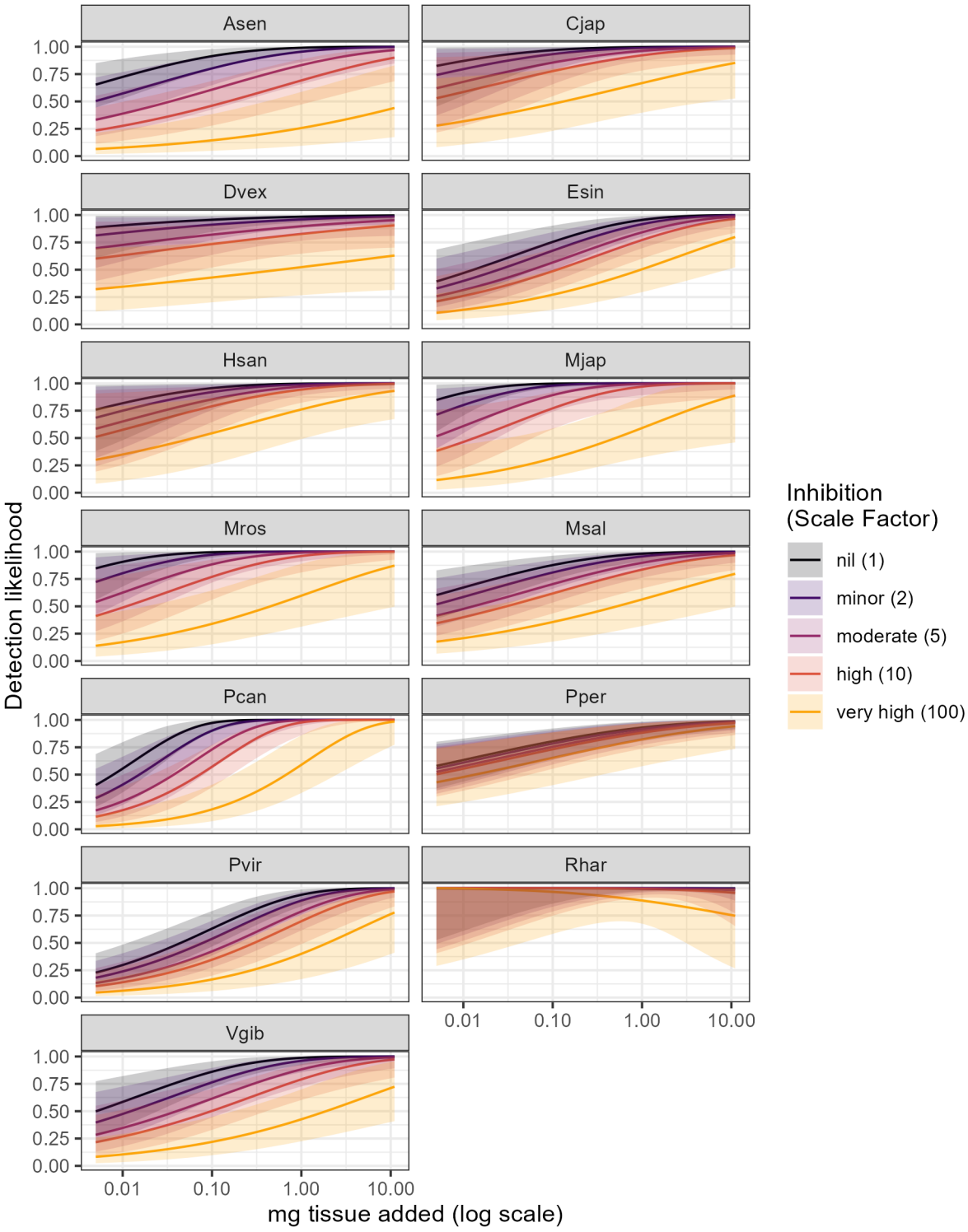

Figure 7. Predicted detection likelihood by each assay with tissue dose for five levels of inhibition as measured by scale factor.

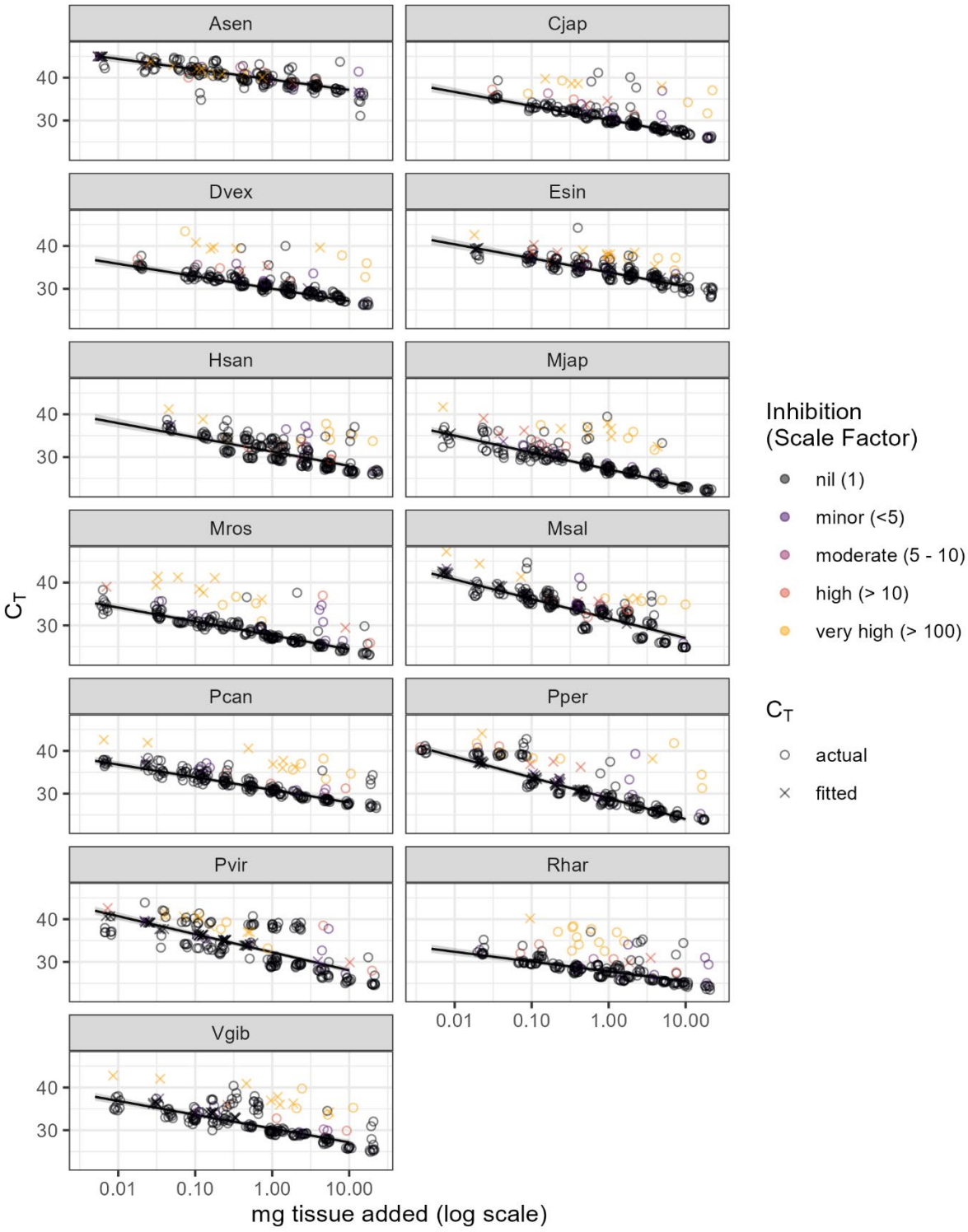

Figure 8. Predicted C_T value of each assay with tissue dose. Line shows prediction with no inhibition. Circles show C_T value of samples with detections and crosses show fitted C_T value for samples where no detection occurred, with shapes coloured by level of inhibition as measured by scale factor.

PCR inhibition, as measured by the scale factor, was recorded in some of the tested samples. Twenty samples, all from Gladstone, had scale factor > 5, with 13 of these showing very high inhibition (scale factor > 100). A further nine samples, comprising six from Gladstone and three from Adelaide, had minor inhibition (scale factor 2 – 5). Detections occurred in some samples with high inhibition, including one sample with scale factor > 3,000, but the 14 samples with the fewest detections (5 or fewer of 11 added species) included seven of those with scale factor > 100 and two more with scale factor > 10. The remaining samples with a low proportion of detections had scale factors between 1.4 and 3.5 and received lower doses (0.1 or 0.5 g) of spiked sand.

3.4.2. Tissue dose and inhibition effects on spiked sample detections

For each assay, detection likelihood increased with increasing tissue dose, except for the Rhar assay, which had detection likelihood of ~100% even at the lowest doses applied (Figure 7). Predicted detection likelihood for most assays was lower for the same tissue dose in the presence of inhibition, but detection likelihood approached 100% at higher tissue doses except at very high (scale factor = 100) levels of inhibition. The C_T value of detections decreased with increasing tissue dose for all assays, and increased with inhibition (Figure 8). The C_T value of detections by the Asen assay was generally higher than other assays for the equivalent tissue dose, while that of the Rhar assay was lower.

The lowest tissue dose providing predicted likelihood of detection ≥ 95% in the absence of inhibition was below 1 mg for each assay, and < 0.1 mg for seven assays (Table 16). The tissue dose required to provide ≥ 95% detection likelihood increased with increasing inhibition for all assays except Rhar, which maintained detection likelihood of > 95% across the range of tissue doses applied at even the highest sale factors observed.

3.4.3. Diagnostic performance from latent class modelling (LCM)

The LCM provided an estimate of DSe for each assay in the absence of inhibition, and an estimate of the scale factor effect on DSe for each (Table 17). In the absence of inhibition, DSe estimates were between 79.4% (Pvir) and 99.3% (Rhar). The scale factor effect varied between assays, with this effect for the Hsan, Msal, Pper and Rhar assays not considered statistically different from zero. The Asen, Mros and Mjap assays were the most affected by scale factor, but with a minimum coefficient of −0.47, indicating that the effect on DSe is not severe (Table 17; Figure 9).

Table 16. Minimum tissue dose (mg) required to provide predicted detection likelihood ≥ 95% for five levels of inhibition as measured by scale factor (scale factor used for prediction shown in brackets).

Assay	nil (1)	minor (2)	moderate (5)	high (10)	very high (100)
Asen	0.12	0.53	3.80	> 10	> 10
Cjap	0.03	0.08	0.36	1.08	> 10
Dvex	0.02	0.17	2.87	> 10	> 10
Esin	0.59	1.13	2.68	5.11	> 10
Hsan	0.05	0.11	0.31	0.69	9.68
Mjap	0.01	0.04	0.15	0.45	> 10
Mros	0.01	0.04	0.17	0.53	> 10
Msal	0.22	0.51	1.59	3.73	> 10
Pcan	0.06	0.12	0.29	0.56	5.25
Pper	0.88	1.19	1.80	2.44	6.83
Pvir	0.84	1.48	3.15	5.57	> 10
Rhar	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Vgib	0.22	0.53	1.72	4.21	> 10

The modelled response of DSe of each assay to PCR inhibition, as measured by the scale factor, is shown in Figure 9. DSe remains above ~75% for most assays to a scale factor of at least 10, the exceptions being the Pvir assay, which had lower DSe than other assays in the absence of inhibition, and the Asen assay, which had the largest response to scale factor. The Msal, Pper and Esin assays, which had DSe < 90% in the absence of inhibition, showed little to no decline in DSe with scale factor (Figure 9).

The DSp estimates for the assays ranged from 70.5% (Vgib) to 100.0% (Msal) (Table 17). The model estimates are influenced by the informative priors used, which were based on the field specificity results (section 3.3). The estimate of 93.4% for DSp of Pper therefore reflects the high specificity displayed by this assay in the field samples, despite the detections recorded in the spiked samples that did not have this species added. The fact that this assay displayed cross-reactivity with gDNA extracts of *P. canaliculus* but no detections in 446 field samples suggests that its DSp when applied to Australian samples (from which *P. canaliculus* is absent) is likely to be higher than the model estimate. DSp estimates for other assays were mostly > 99%. The assays for Mros, Mjap and Rhar, which returned detections in either some field samples or spiked samples without the relevant target species, had estimated DSp of between 97.8 and 98.8%. The Pcan assay, which had a single detection in a sample without tissue of this species added, had DSp of 99.9% given that no detections by this assay had occurred in > 2,000 previously tested field samples. The lowest DSp estimates were, unsurprisingly, for the two assays that had

displayed multiple field detections, with Vgib having the lowest DSp estimate (70.5%), and DSp of the Asen assay being 90.5%.

Table 17. Latent class model estimates of diagnostic sensitivity (DSe) and specificity (DSp), and coefficient for the scale factor effect on DSe on the complementary log-log scale. Each estimate shows the mean with the 95% HDI in brackets.

Assay	Modelled DSe	Modelled DSp	Scale factor effect on DSe
Asen	93.2% (89.3% – 96.6%)	90.5% (89.0% – 92.0%)	−0.47 (−0.69 – −0.28)*
Cjap	98.2% (96.2% – 99.7%)	99.8% (99.5% – 100.0%)	−0.34 (−0.53 – −0.18)*
Dvex	96.5% (93.8% – 98.9%)	99.7% (99.2% – 100.0%)	−0.37 (−0.56 – −0.19)*
Esin	88.0% (83.3% – 93.5%)	99.7% (99.2% – 100.0%)	−0.25 (−0.40 – −0.11)*
Hsan	97.8% (95.5% – 99.6%)	99.7% (99.2% – 100.0%)	−0.21 (−0.36 – −0.06)
Mros	99.0% (97.7% – 99.9%)	98.5% (97.3% – 99.6%)	−0.41 (−0.59 – −0.25)*
Mjap	97.8% (95.6% – 99.5%)	98.8% (97.7% – 99.8%)	−0.44 (−0.64 – −0.27)*
Msal	88.2% (83.4% – 92.8%)	100.0% (99.9% – 100.0%)	−0.22 (−0.37 – −0.06)
Pcan	94.8% (91.4% – 97.9%)	99.9% (99.8% – 100.0%)	−0.28 (−0.44 – −0.15)*
Pper	79.6% (73.7% – 85.8%)	93.4% (91.4% – 95.3%)	−0.10 (−0.22 – 0.00)
Pvir	79.4% (73.0% – 85.3%)	99.8% (99.5% – 100.0%)	−0.30 (−0.50 – −0.12)*
Rhar	99.3% (98.2% – 100.0%)	97.8% (96.6% – 98.9%)	−0.18 (−0.33 – −0.06)
Vgib	90.0% (85.6% – 94.2%)	70.5% (67.7% – 73.6%)	−0.28 (−0.44 – −0.12)*

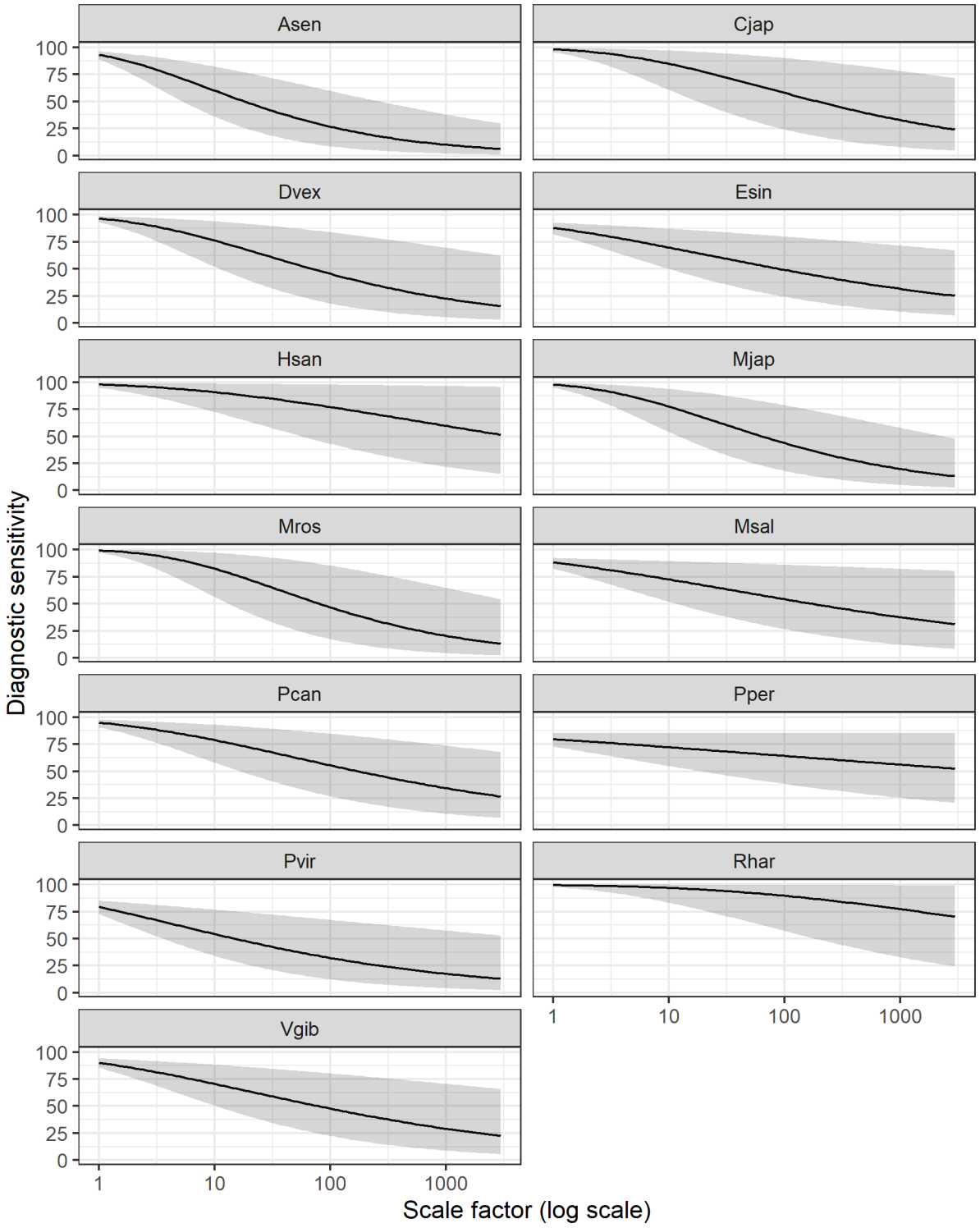

Figure 9. Modelled effect of PCR inhibition (measured by scale factor) on diagnostic sensitivity of each assay. The line shows mean model response with shaded area showing HDI of predictions.

3.5. High throughput sequencing results

Across the HTS runs, the average raw reads per library ranged from ~ 25,000 to > 80,000 (Table 18). Note that each library consisted of DNA from one plankton sample amplified with one primer set, with four primer sets applied to the runs for *A. senhousia* and *V. gibba* (which included testing for both species in each run) and two primer sets used otherwise (see section 2.5). All HTS runs were able to identify target sequence reads in at least some samples with qPCR detection by the relevant assay, with generally more target sequence reads in samples with lower C_T values (Figure 10). The targeted approach used, with application of selective primers, means that the total reads per library could be low in the absence of target species DNA, without this being an indication of poor amplification of sequencing. Total reads per library were typically higher in samples with qPCR detection (Figure 11), i.e., where target DNA, which would be amplified by the primers used in HTS, is expected to be present. Sequence detections and total reads by sample and primer set are shown in the Appendix (section 6.3) with results for each assay summarised below.

Table 18. Number of plankton samples and total libraries included in each HTS run showing number of samples with qPCR detection and average raw reads obtained per sample.

Target species	HTS run #	Samples	Libraries	qPCR positives	Mean raw reads per library
Hemigrapsus sanguineus	1	46	94	16	80,890
Maoricolpus roseus	2	120	241	91	24,935
Mya japonica	3	43	87	19	72,352
Arcuatula senhousia	4	93	376	31	26,458
Arcuatula senhousia	5	31	128	18	70,699
Varicorbula gibba	4	93	376	62	26,458
Varicorbula gibba	5	31	128	21	70,699

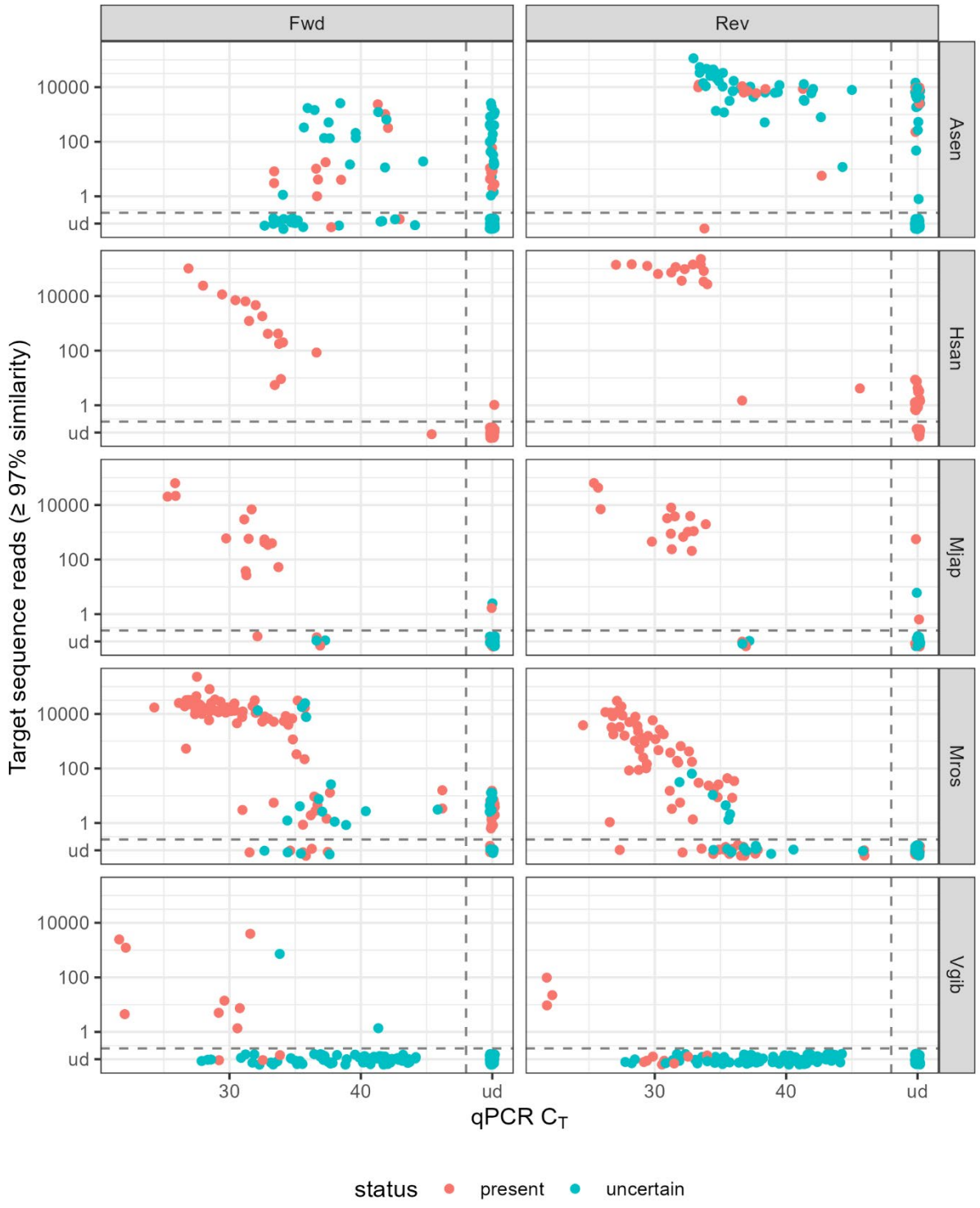

Figure 10. Number of target sequence reads identified using HTS with assay forward (Fwd) and reverse (Rev) primer sets against qPCR cycle threshold (CT) value. 'ud' = undetected. Dashed lines separate points with detection by qPCR/HTS from those without. Status indicates where species is expected to be present or where qPCR detections are uncertain. See Table 2 for assay name definitions.

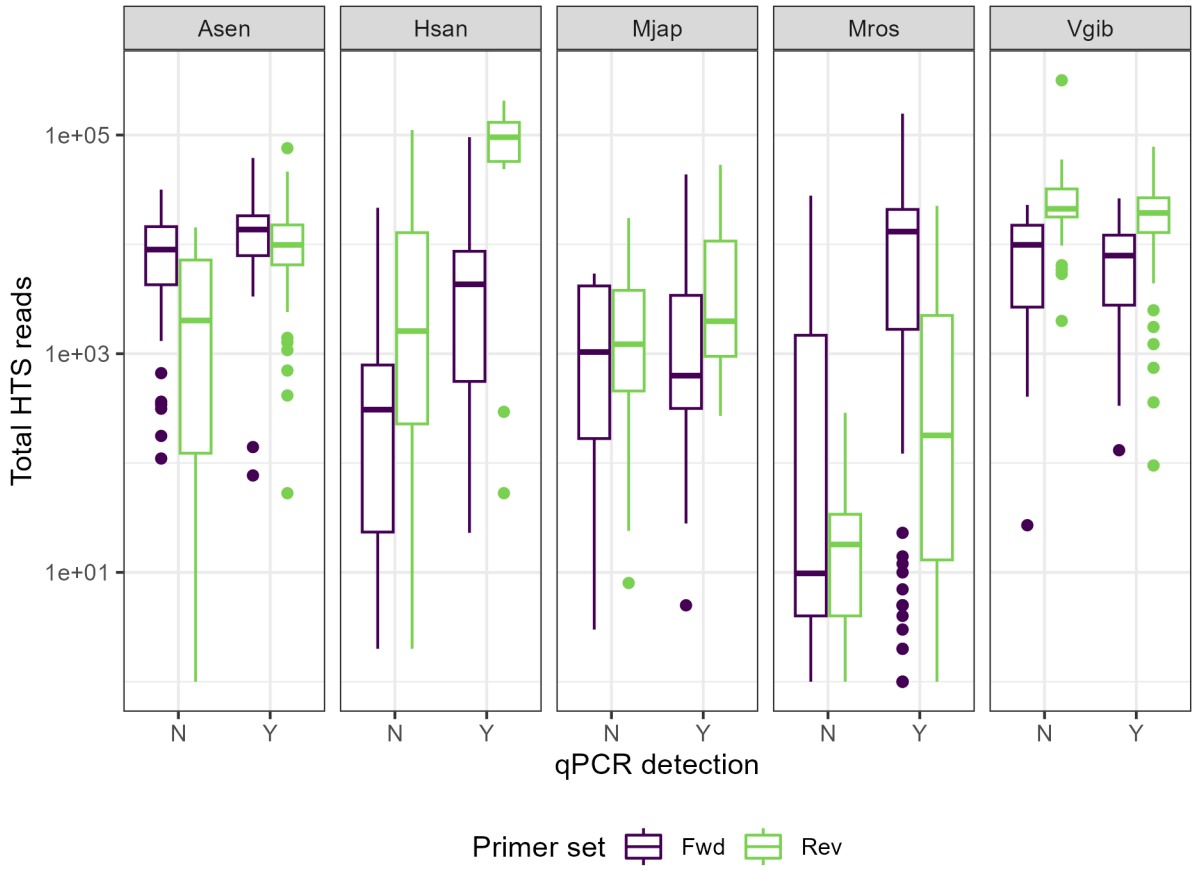

Figure 11. Total reads per HTS in libraries prepared with forward (Fwd) and reverse (Rev) primer sets in samples with and without qPCR detection by the relevant species assay. See Table 2 for assay name definitions. Note that total reads are shown on a log scale.

3.5.1. *Hemigrapsus sanguineus*

The 46 samples from Melbourne used for *Hemigrapsus sanguineus* HTS testing included 16 with Hsan qPCR detection, with C_T values between 27.0 and 45.5. HTS identified OTUs with ≥ 97% sequence similarity to *H. sanguineus* in both the forward and reverse primer sets for 15 of the 16 samples, being all those with C_T < 36.8 (Figure 10; Table 19; Appendix section 6.3), confirming the qPCR detections. The single sample with high (45.5) C_T Hsan detection had target sequences in amplicons from the reverse primer but not the forward primer, which was likely due to the low concentration of target DNA in this sample. Low numbers (≤ 8 reads) of *H. sanguineus* sequences were identified in reverse primer amplicons for 19 of the 30 samples without Hsan qPCR detection, with one of these samples also having one target sequence in forward primer amplicons. This suggests that very low levels of target DNA could be present in some samples without qPCR detection, although it is likely some reads may be due to indexing read errors

leading to sequences being incorrectly assigned to samples. One OTU was identified that had 92% similarity to *H. sanguineus* but was identical in the assay region, and so could cross-react with the assay. This OTU was rare, being present as three reads across two samples, and could be due to sequencing error. Both samples containing the potential confounding sequence had Hsan qPCR detections and target sequences present in HTS.

Table 19. Correspondence between samples with qPCR and HTS target detection for *Hemigrapsus sanguineus* and average target sequence reads returned by forward (Fwd) and reverse (Rev) primer sets in samples with HTS detection. 'ud' = undetected.

qPCR detection	No. samples	HTS – both primers	HTS – Fwd only	HTS – Rev only	HTS ud	Av Fwd reads	Av Rev reads
N	30	1	0	18	11	1	2.22
Y	16	15	0	1	0	10,416	90,810

3.5.2. *Maoricolpus roseus*

The 120 samples used for *Maoricolpus roseus* HTS testing included 91 with Mros qPCR detection, with C_T values between 24.5 and 40.5. Most qPCR detections were confirmed by HTS (Figure 10; Table 20; Appendix section 6.3). Six samples with C_T values between 35.3 and 37.7 did not have any OTUs with matches to *M. roseus* identified in either forward or reverse primer sets, but these samples returned 0 – 2 total reads with each primer set indicating a possible low abundance of target DNA. Other samples with similar or higher C_T value, however, detected OTUs containing the target sequence, although only in forward primer amplicons. In total, 26 samples with qPCR detection had *Maoricolpus roseus* sequences identified only in forward primer amplicons, and four had target sequences only in reverse primer amplicons. The samples with Mros qPCR detection from Hay Point (C_T 40.5), Gove (C_T 37.7), Klein Point (C_T 36.9) and Adelaide (C_T 37.9) were amongst those samples with target sequences in only the forward primer amplicons. Twenty two of 29 samples without qPCR detection had some reads (≤ 22) identified as *Maoricolpus roseus* in forward primer amplicons, indicating possible presence of target DNA or index read errors. Contamination during library preparation is also a potential source of detection in the sample with 22 reads, with other qPCR negative samples returning ≤ 9 reads. These samples were all from either Portland or Tasmania and were collected in the same sampling events as samples that had qPCR detections. Two OTUs, comprising < 0.12% of the total reads, were identified that were 96 and 97% similar to *Maoricolpus roseus* but 98% similar to each other. These OTUs would be detected by the Mros assay and could represent a confounding taxon, although it is also possible that *Maoricolpus roseus* has > 3% divergence in

COI and that these OTUs are of a *Maoricolpus roseus* haplotype. The potential confounding OTUs occurred in 35 samples, which were from Tasmania (33 samples) and Portland (2), and which predominantly had Mros qPCR detections (33 samples) and/or target sequences present (34 samples).

Table 20. Correspondence between samples with qPCR and HTS target detection for *Maoricolpus roseus* and average target sequence reads returned by forward (Fwd) and reverse (Rev) primer sets in samples with HTS detection. 'ud' = undetected.

qPCR detection	No. samples	HTS – both primers	HTS – Fwd only	HTS – Rev only	HTS ud	Av Fwd reads	Av Rev reads
N	29	0	22	0	7	4.73	0
Y	91	55	26	4	6	4768	885.9

3.5.3. *Mya japonica*

The *Mya japonica* HTS runs were affected by primer dimerisation, with not all samples able to be successfully cleaned of primer dimers and sequenced. The 43 samples that were successfully sequenced included 19 with Mjap detection and HTS identified target OTUs in 15 of these (Figure 10; Table 21; Appendix section 6.3). *Mya japonica* OTUs were recorded using both forward and reverse primer pairs for 14 of these 15 samples, and in the reverse primer run only for the remaining sample. The four samples with Mjap qPCR detection without *Mya japonica* OTUs in HTS all had C_T values > 36, indicating HTS lacked sensitivity to detect the target at low abundance. These samples included two from locations without known *Mya japonica* occurrence: Port Giles, collected during 2021-22 surveillance (Wiltshire *et al.* 2022) and D'Entrecasteaux Channel, collected during field testing of the Mros and Mjap assays (Giblot-Ducray *et al.* 2022). The other two samples were from the Orford - Triabunna region where *M. japonica* is established. Although presence of target DNA was not confirmed by HTS, it appears unlikely that the two qPCR detections from locations without *Mya japonica* are cross-reactions because the BLAST results did not identify any sequences across the set of samples tested that would confound the assay. Giblot-Ducray *et al.* (2022) note that the D'Entrecasteaux Channel sample was inadvertently processed with samples from Orford – Triabunna and may have thereby been contaminated with *Mya japonica* DNA. Three samples with Mjap qPCR detection that were collected in Blackman Bay in 2017 and tested by Giblot-Ducray *et al.* (2022), and individual samples with Mjap detection from Port Lincoln and Klein Point from 2021-22 surveillance (Wiltshire *et al.* 2022) were unable to be sequenced, but it is likely that HTS would not have been able to detect target sequences in these samples in any case due to each having C_T > 36. Two

samples without Mjap qPCR detection each returned two target sequence reads in the forward primer run, with one of these samples having a single target read in the reverse primer set, with these reads potentially due to indexing errors. One sample without Mjap qPCR detection returned a moderate number (539) of reads in the reverse primer set. This sample was from the Orford – Triabunna region and hence possibly represents a qPCR false negative in a sample with target DNA occurrence. Alternatively, contamination or incorrect sample selection may have occurred at the library preparation stage, with the repeated preparation required for this run increasing these risks.

Table 21. Correspondence between samples with qPCR and HTS target detection for *Mya japonica* and average target sequence reads returned by forward (Fwd) and reverse (Rev) primer sets in samples with HTS detection. 'ud' = undetected.

qPCR detection	No. samples	HTS – both primers	HTS – Fwd only	HTS – Rev only	HTS ud	Av Fwd reads	Av Rev reads
N	24	1	1	2	20	2	180
Y	19	14	0	1	4	3117	3435

3.5.4. *Arcuatula senhousia*

HTS testing for *Arcuatula senhousia* included 124 samples across the two runs, of which 49 had Asen qPCR detection. OTUs with ≥ 97% similarity to *A. senhousia* were identified in both HTS forward and reverse primer sets of 50 tested samples, with 25 of these corresponding to Asen qPCR detections and the remaining 25 in samples without Asen qPCR detections (Figure 10; Table 22; Appendix section 6.3). A further 23 of the samples with qPCR detection had OTUS with ≥ 97% similarity to *A. senhousia* identified from the reverse primer set, with a single qPCR positive sample having no target OTUs identified. Target OTUs were also identified by the reverse primer set in 11 additional qPCR negative samples and by the forward primer set in two, with no target OTUs identified in the remaining 37 qPCR negative samples. In contrast to the runs for *H. sanguineus, Maoricolpus roseus* and *Mya japonica*, where the number of target sequence reads in qPCR negative samples was low (mostly < 20 reads), target OTU read numbers for *A. senhousia* in qPCR negative samples were in the same range (up to ~10,000 in reverse primer sets and >2,000 in forward primer sets) as in several samples with qPCR detection (Figure 10; Appendix section 6.3). Samples with *A. senhousia* OTUs detected by HTS included 8 of 9 samples from Hobart, a location where there have been no Asen detections recorded in qPCR testing of 155 samples collected from 2015-18 (Deveney *et al.* 2017; Wiltshire *et al.* 2019a), and where no specimens were found in 2017-18 traditional surveillance (Wiltshire *et al.* 2019a). In

combination with the generally poor correspondence between qPCR results and HTS target detection, this suggests that at least some identified sequences may not be *A. senhousia* despite ≥ 97% similarity to this species. The identity of the organism producing these sequences is, however, unknown. The 97% similarity cut-off is often applied for assignment to species, but is not always appropriate due to variation in sequence divergence across taxa (Porter and Hajibabaei 2020). Amplicons produced by the primer sets for HTS were in the order of 300 bp, which is a relatively short fragment length for determination of species identity, especially for 28S, which is typically more conserved than COI (Andújar *et al.* 2018). In addition, 28S divergence in Mytilidae is particularly low (Colgan and da Costa 2013; Kartavtsev *et al.* 2018).

Table 22. Correspondence between samples with qPCR and HTS target detection for *Arcuatula senhousia* and average target sequence reads returned by forward (Fwd) and reverse (Rev) primer sets in samples with HTS detection. 'ud' = undetected.

qPCR detection	No. samples	HTS – both primers	HTS – Fwd only	HTS – Rev only	HTS ud	Av Fwd reads	Av Rev reads
N	75	25	2	11	37	370.0	4760
Y	49	25	0	23	1	1227	8987

Examination of OTUs with ≥ 97% similarity to *A. senhousia* across runs showed that the most abundant OTU detected in the forward primer set was the same in Fremantle, Kwinana and Melbourne, where *A. senhousia* is established and recently recorded (Wiltshire *et al.* 2019a; Wiltshire *et al.* 2020b), but this OTU was not detected in other locations (Darwin, Cairns, Gladstone, Weipa, Hay Point, Thevenard, Hobart, Devonport). The OTUs with ≥ 97% similarity to *A. senhousia* in forward primer sets for Darwin, Cairns, Hobart, Gladstone, Weipa, and Hay Point samples were predominantly absent from other sample locations, being recorded as ≤ 3 reads in 1 – 2 samples each, which were potentially due to indexing error. The same OTUs were recorded in Hobart as in Devonport but these OTUs were not detected in other locations, while there were no OTUs assigned to *A. senhousia* with ≥ 97% similarity in the forward primer set for Thevenard samples aside from one read in a single sample, which may have been due to indexing error. Sequences identified in the forward primer set for Thevenard had 95 – 96% similarity to *A. senhousia,* which was the closest match to these sequences in Genbank, and matched the assay region. OTUs with ≥ 97% similarity to *A. senhousia* in the reverse primer sets were shared across sample locations, with the most common OTU being detected in Thevenard, Fremantle, Kwinana, Devonport, Darwin, Cairns, Hobart, Gladstone, Weipa and Hay Point, and another common OTU being shared between Fremantle, Kwinana and Gove. OTUs in reverse primer sets were different between Melbourne and other locations, although sequences in each case were

≥ 97% similar to *A. senhousia* and near identical to one another. The differences, particularly in the forward primer OTUs, between Fremantle, Kwinana and Melbourne and tropical locations suggests that one or more taxa may occur in the tropical locations that are closely related to *A. senhousia*, and have highly similar 28S sequences, but do not have sequences available in Genbank. The sequences occurring in Thevenard samples similarly appear to represent a taxon that is closely related to *A. senhousia,* but which is different from taxa present in the tropical locations.

In addition to the OTUs with < 97% similarity to *A. senhousia* in Thevenard and the OTUs from tropical locations that would be detected by the Asen assay, potentially confounding sequences with 74 – 90% similarity to *A. senhousia* were identified in samples from several locations. These sequences could not be ascribed to species but appear to represent at least five different Mytilidae taxa. In each case, sequences provided a 100% match to the assay region, hence, these non-target sequences would also be efficiently detected by the Asen assay.

3.5.5. *Varicorbula gibba*

Across the two HTS runs for *V. gibba*, 124 samples were tested, including 83 with Vgib qPCR detection. OTUs with ≥ 97% similarity to the target were identified in the forward primer sets in ten of the 83 qPCR positive samples, and in the reverse primer sets for three of the 83 qPCR positive samples. Except for one sample from Thevenard, which had C_T of 41.3 and a single target sequence read detected, samples returning *V. gibba* OTUs from the forward primer sets had C_T < 34. The three samples with target OTUs detected by the reverse primer sets all had $C_T \leq 22$, indicating a possible lack of HTS sensitivity for *V. gibba* detection. The samples with target OTUs in both primer sets were all from Hobart, where *V. gibba* is established and was detected in traditional surveillance in 2017-18 (Wiltshire *et al.* 2019a), with a further four Hobart samples among those having target sequences in the forward primer sets only. The remaining three samples with *V. gibba* sequences in the forward primer sets were one from each of Thevenard, Melbourne and Portland. The Thevenard sample had a single target sequence read, which is likely due to indexing error given the lack of target sequence detection in other high C_T samples. *Varicorbula gibba* is established in Melbourne and confirmed to occur in 2017 (Wiltshire *et al.* 2019a), and has been recorded in Portland, albeit in a single sample from 1996 (Parry *et al.* 1997). The Portland HTS sample had 611 reads of the same *V. gibba* OTU as detected in the Melbourne and Hobart samples, supporting that Portland Vgib HTS and qPCR detections could represent true *V. gibba* occurrence.

Table 23. Correspondence between samples with qPCR and HTS target detection for *Varicorbula gibba* and average target sequence reads returned by forward (Fwd) and reverse (Rev) primer sets in samples with HTS detection. 'ud' = undetected.

qPCR detection	No. samples	HTS – both primers	HTS – Fwd only	HTS – Rev only	HTS ud	Av Fwd reads	Av Rev reads
N	41	0	0	0	41	0	0
Y	83	3	7	0	73	955.3	14.7

HTS identified sequences with < 97% similarity to *V. gibba,* but for which *V. gibba* was the closest match in Genbank, in samples from Darwin, Gove, and Weipa, where *V. gibba* is not recorded and is unlikely to occur given its thermal tolerance (Summerson *et al.* 2016). These < 97% similar sequences were present in seven samples with Vgib qPCR detection, but also occurred, with read number ≤ 29, in four samples that were qPCR negative.

A lack of *V. gibba* OTUs in HTS for most samples with qPCR detection could be due to the qPCR detections being of non-target species DNA, particularly in samples with C_T < 34. While few target sequences were identified by HTS, especially in reverse primer sets, many non-target sequences were amplified (Figure 11; Appendix section 6.3). Detections of potentially confounding sequences occurred in nearly all samples with Vgib qPCR detection, including in samples where *V. gibba* OTUs were identified. Most of the potentially confounding sequences could not be ascribed to species and likely represent DNA from taxa without publicly available sequences. The nearest matches to the potentially confounding sequences appear to be bivalve molluscs of the families Arcidae, Teredinidae, and Mytilidae, the gastropod family Cerithiidae, and the polychaete family Phyllodocidae. Where taxonomy could be attributed to potentially confounding sequences, the assigned identities represented several species that are not recorded in Australia, including *Apachecorbula muriatica*, a deep-water Red Sea species. It is likely these sequences belong to Australian taxa without sequences in Genbank, but that are related, and have highly similar (98 – 99%) 28S sequences, to the identified species. As noted for *Arcuatula* senhousia, the 28S barcode is typically more conserved than COI (Andújar *et al.* 2018) and related species sequences may therefore have > 97% similarity to one another.

Further investigation determined that mismatches between the Vgib assay and 28S sequences of Australian taxa, including *Tegillarca granosa* (Arcidae) and *Cerithium coralium* (Cerithiidae) are unlikely to be sufficient to prevent the assay detecting these sequences, with the sequence occurring in *Tegillarca granosa* also present in other Arcidae. The 28S sequence of the *Anadara trapezia* gDNA extract that was detected by the Vgib assay is very similar to the

Tegillarca granosa sequence in GenBank, particularly in the Vgib assay regions. *Tegillarca granosa* and *Anadara trapezia* sequences each provide a 100% match to the Vgib assay probe region, one mismatch to the reverse primer, and six or seven mismatches to the forward primer region respectively. There were, however, only one or two mismatches respectively in the first 13 bp of the 3' end of the Vgib forward primer with remaining five mismatches at the 5' end (positions 14 – 18 of the 18 bp total primer length). The position of the mismatches means that amplification of these sequences is possible, though efficiency would be lower than for target DNA, leading to higher C_T for equivalent DNA concentration. Arcidae bivalves may therefore be responsible for cross-reactions occurring in some locations, but there are likely to be additional taxa contributing to non-specificity of this assay based on the HTS data.

4. DISCUSSION

Important steps for the validation of 13 qPCR assays for detection of marine pest species have been completed, providing data to adequately design surveys and to appropriately interpret their results for these pests. The assays for the marine pests *Hemigrapsus sanguineus, Eriocheir sinensis, Rhithropanopeus harrisii, Charybdis japonica*, *Didemnum vexillum*, *Perna viridis, P. canaliculus, Mytilopsis sallei, Mya japonica* and *Maoricolpus roseus* demonstrated performance suitable for application to routine surveillance, i.e., DSe ≥ 70% and DSp ≥ 98%. The assay for *P. perna* is also suitable for surveillance, but with some caveats, discussed below, given that it also detects *P. canaliculus* DNA. The assays for *Arcuatula senhousia* and *Varicorbula gibba*, while having good DSe, are not suitably specific, and require re-design before application to Australian samples.

Data on diagnostic performance are now available for all assays that have been implemented in the SARDI testing system. DSe data are used for calculation of the number of samples required to achieve a target survey confidence (likelihood of detection in at least one sample of a survey) for a given pest population size or density. The survey design tool developed by Wiltshire (2021) calculates the number of samples required for the plankton sampling method, based on a selected planktonic concentration and survey confidence, considering the DSe of assays to be applied in the survey. For a given target pest abundance, a greater number of samples would need to be collected and analysed to achieve the same survey confidence for an assay with lower than higher DSe, e.g. 43 samples to detect a pest with planktonic concentration of 0.0075 m^{-3} with 80% confidence for DSe = 60%, compared with 37 samples for DSe = 70 % and 32 for DSe = 80% (survey design tool) The DSe data provided by this project, in combination with existing data for previously validated assays, therefore allow this tool to be used to design molecular surveillance for all 19 species for which qPCR assays are available.

While the collection and analysis of a greater number of samples can permit detection by an assay with lower DSe, this would necessarily incur extra costs, and the number of samples to collect may become infeasible. It is therefore highly desirable for assays applied to routine surveillance to have DSe ≥ 70%. The diagnostic sensitivity of each of the assessed assays was within or higher than the range (73 – 91%) for DSe of the assays the key species of ballast water concern assessed by Wiltshire *et al.* (2019b). These additional species could therefore be included in surveillance without requiring the collection of more samples than are needed to achieve adequate confidence of detection for the priority ballast water species.

The DSe values estimated by this study show the average detection likelihood across the range of tissue doses that were used for each species. Likelihood of detection in a specific sample will, however, depend on the quantity of target DNA present. Detection likelihood for most assays increased with increasing tissue dose, and C_T value decreased, demonstrating higher levels of target DNA yield as expected (see Results section 3.4.2). The range of tissue doses was chosen to be environmentally relevant based on available data, but ideally samples from locations with known occurrence of each pest should be tested to verify assay performance in environmental samples and to determine the typical DNA yield that occurs in samples from a location with target species presence. The C_T values and DNA yields obtained in the spiked samples were within the ranges found during surveillance for marine pests where these occur (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019c, 2020b), providing confidence that the doses used, and therefore the DSe estimates, are relevant. We used a range of tissue doses in the samples tested for DSe estimation, but if future investigation shows that some of the doses were either too high or too low, the data can be re-analysed to derive the DSe for each assay at more appropriate doses.

PCR inhibition was evident in some of the plankton samples used for the spiking experiment. Substances that may cause inhibition include structural proteins, enzymes, alcohols, complex polysaccharides, humates, calcium, urea, chlorides, and detergents (Bessetti 2007). Given the wide range of possible causes, it is typically not possible to predict when inhibition will occur or to determine the specific cause of inhibition in a sample. PCR inhibition is a recognised issue for environmental molecular surveillance (Goldberg *et al.* 2016), but qPCR assays vary in their response to inhibitors, complicating the identification of inhibition using internal controls and of predicting the impacts of inhibition on assay performance (Lance and Guan 2020). The assays we assessed showed differing responses to inhibition, which was measured using a scale factor calculated from DNA yield of an internal control. This demonstrated that the assays responded differently to inhibition relative to each other, and relative to the assay used for the internal control. Scale factor may also be influenced by extraction efficiency. Scale factor is therefore an imperfect measure of inhibition, but it was still shown to be useful at identifying samples where detection likelihood may be reduced. DSe decreased and the C_T value of detections increased with increasing scale factor for each assay, although the magnitude of the effect varied across assays. Detections occurred in some samples 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, at least for some assays, even in the absence of inhibition.

Future surveillance should continue to assess inhibition to identify sample sets where detection may be compromised.

Hemigrapsus sanguineus DNA was successfully detected in plankton samples collected in Melbourne, providing additional confidence in performance of this assay and preliminary data on seasonality of detection for this species. This species was not detected in samples collected in July – August 2017 or June 2018, but was detected in samples from November 2017, March 2018, and February 2022. This result suggests that sampling in spring – summer is better for detection of this species than autumn – winter, although this finding should be considered preliminary given limited data availability. In North America *H. sanguineus* spawns over a 5- to 6-month period spanning summer (Brousseau and McSweeney 2016), providing support that summer is likely to be a suitable season for detecting this species. Sampling twice yearly, in late winter and late summer, using the plankton sampling and qPCR testing method, allows detection across seven species of domestic ballast water concern, because species that have low detection likelihood (due to low planktonic abundance) in summer have high detection likelihood in winter and *vice versa* (Wiltshire 2021). Application of this seasonal sampling strategy appears suitable for detection of *H. sanguineus*, although further data would be needed to confirm seasonality in planktonic abundance and consequently seasonal detectability of this species in Australia.

Specificity was high for most of the assays assessed but was lower for some assays than the > 99% DSp of the previously assessed assays (Wiltshire *et al.* 2019b; Wiltshire *et al.* 2021b). DSp is not used to calculate required sample number in the survey design tool, with the calculation used in that tool assuming perfect (100%) DSp (Wiltshire 2021). It is desirable for assays used in surveillance to have DSp as close to 100% as possible, because lower DSp increases the likelihood of false positive detections caused by non-target DNA, which can complicate the interpretation of surveillance results (Darling *et al.* 2020; Sepulveda *et al.* 2020; Wiltshire 2023). The likelihood of a false positive also increases with increasing sample number, but using a sample set of 35 as applied in recent surveillance (Wiltshire *et al.* 2022), the risk of a false positive remains low for the assays with DSp > 98%. DSp can be used, along with DSe, to interpret survey results by allowing estimation of true prevalence, which can provide a maximum plausible estimate of likelihood of occurrence for species with no detections, or where detections occur that may be false positives (Low-Choy 2013; Stanaway 2015; Wiltshire 2023). Where data on sample volumes are also available, e.g., from a flow meter fitted to the plankton net used for sampling, the DSe and DSp data allow the application of modelling approaches that can provide estimates of planktonic pest concentration while accounting for differences in assay diagnostic performance

across surveyed species (Wiltshire 2023). These modelling approaches can account for DSp < 100% where false positives occur randomly and sporadically, but model estimates are compromised where false positives occur systematically, e.g., due to assay cross-reactivity with DNA of non-target species (Wiltshire 2023).

False positives in qPCR testing may occur due to sample contamination with target DNA, or due to a non-specific reaction with DNA of a different species. For the spiking experiment, it was necessary to sample tissue from multiple pest species, and time and equipment constraints meant that pest tissue samples were freeze-dried together, which may have resulted in cross-contamination between samples. The detections by the Rhar assay in samples not spiked with *R. harrisii* suggested that cross-contamination occurred, which was supported by further testing that showed no detection using gDNA of other target species added to the samples, while detections occurred in sand that had not been added to plankton. The Rhar assay is therefore likely to have higher DSp than the model estimate.

The Pper assay, however, cross-reacts with DNA of *Perna canaliculus*, which resulted in a high proportion of false positive detections for Pper in the spiked samples that did not contain *P. perna* tissue. In field samples, however, no issues with the Pper assay were identified, and it appears unlikely that this assay cross-reacts with DNA of native Australian species. *Perna canaliculus* is exotic to Australia, therefore the DSp of the Pper assay applied to Australian plankton samples is likely to be close to 100%. If Pper detections occur, further investigation may be needed to determine whether the detected species is *P. perna* or *P. canaliculus*, but, where both assays are applied, a Pper detection without Pcan detection would indicate presence of *P. perna* DNA. A detection by both assays with a lower C_T for Pcan than Pper would suggest *P. canaliculus* presence. The Pper assay should therefore not be applied to samples without also testing using the Pcan assay, because Pper detections would otherwise be ambiguous. Sequencing approaches could alternatively or additionally be applied to confirm which species DNA is present. Where uncertainty remains, further investigation would be needed to rule out co-occurrence of both species, but, while many parts of Australia provide suitable habitat for establishment of both, simultaneous introduction would be unlikely given their disparate ranges. Should *P. canaliculus* become established in Australia, however, suitability of the current Pper assay for Australian surveillance would need to be re-assessed.

The Asen assay has provided detections in samples from several tropical locations where *Arcuatula senhousia* is unlikely to occur, but, in surveillance of temperate regions from 2017 to

2020, this assay only returned detections in areas with known pest presence (Deveney *et al.* 2017; Wiltshire *et al.* 2019a; Wiltshire *et al.* 2019c, 2020a, b). It therefore appeared likely that this assay cross-reacts with a species found only in tropical regions but was specific when applied to temperate areas. In 2021-22 surveillance, however, the Asen assay returned detections in samples from Thevenard, a location without known occurrence of *A. senhousia*, although other surveillance that may have detected the species at this location is lacking (Wiltshire *et al.* 2022). The Vgib assay, in contrast, has returned detections widely across temperate areas beyond the known distribution of *Varicorbula gibba*, as well as in tropical areas that would be unsuitable for occurrence of this pest (Deveney *et al.* 2017; Wiltshire *et al.* 2017; SARDI data). A lack of other surveillance means there is a lack of certainty about the absence of *V. gibba* from all temperate locations with detections, but the pattern of detections suggests that cross-reactions of this assays with non-target DNA may be occurring in both temperate and tropical locations.

An approach using HTS was applied to further assess specificity of the newly developed assays for *H. sanguineus, Maoricolpus roseus* and *Mya japonica*, and to assess detections of *A. senhousia* and *V. gibba.* Detections by the Hsan, Mros and Mjap assays were confirmed by the identification of target species OTUs by HTS in relevant samples. For *H. sanguineus*, confirmed detections included those obtained from testing archived DNA from 2017-18 samples from Melbourne. For *Maoricolpus roseus*, target species OTUs were identified in two samples, which were also qPCR positive, from locations without species occurrence, demonstrating that these detections were due to either transient DNA presence or sample contamination. The number of target sequence reads from HTS was ≤ 22 in each case, and qPCR detections had C_T values ≥ 36.9, demonstrating a low level of target DNA occurrence in each case. HTS runs for *Mya japonica* were complicated by primer dimerisation and lower HTS sensitivity than for the other two species. Not all samples of interest could be tested and samples with qPCR detections at C_T > 36, including those from locations without recorded species occurrence, could not be confirmed. There was, however, no evidence to suggest cross-reactivity of the Mjap assay with non-target sequences. Detections by the Mjap and Mros assays in un-spiked samples from locations outside known species ranges occurred in 6 of 682 tested with these assays across this project and Wiltshire *et al.* (2022), indicating < 1% likelihood of detection due to either contamination or transient DNA presence.

A small number of potentially confounding sequences were identified for *H. sanguineus* and *Maoricolpus roseus*, but these sequences are unlikely to be problematic for application of the assays. The organisms which have either potentially confounding sequence have not been

identified. For *Maoricolpus roseus*, these sequences were ≤ 4% divergent from the target, and are potentially a divergent haplotype of this species. Some molluscan species show intraspecies divergence in COI (up to ~10 %) that is greater than what is typical for other taxa (Layton *et al.* 2014; Sun *et al.* 2016). The divergent sequences were found exclusively in samples from locations where *Maoricolpus roseus* occurs, and in only one sample where no sequences with ≥ 97% similarity to *Maoricolpus roseus* were identified. Sequencing of additional *Maoricolpus roseus* specimens would assist in determining whether these sequences are a haplotype of this species, but the sequences do not pose a concern for application of the assay based on current data. For *H. sanguineus*, the potentially confounding sequences were represented by only three reads in two samples and may represent sequencing artefacts. A lack of Hsan detection in un-spiked plankton samples from locations other than Melbourne tested by this project and by Wiltshire *et al.* (2022) supports that this assay is suitably specific for use in Australia.

Primer re-design is possible to provide specificity against the divergent sequences for *Maoricolpus roseus* and *H. sanguineus*, but without identification of the species associated with divergent sequences, re-designed assays could have reduced sensitivity to detect the targets. Re-design could also affect assay performance and specificity against other taxa and therefore re-designed assays would require validation steps to be repeated. Given that the potentially confounding sequences occurred in low abundance in samples from areas with known target species occurrence, and field specificity testing of this assay over a broad geographic range did not identify any issues with either assay, re-design is unlikely to be worthwhile.

HTS results for *H. sanguineus, Mya japonica* and *Maoricolpus roseus* identified target sequences in some samples without qPCR detection, with HTS target sequence detections being more common in qPCR negative samples than for the other two species. In all cases, however, qPCR-negative samples with HTS detection were from locations with known species occurrence and collected in the same sampling events as samples with qPCR detections. The samples may, therefore, have contained low levels of target DNA, although the frequency of HTS reads in these samples, even in the case of *Maoricolpus roseus,* was within the range expected for HTS indexing errors (Wright and Vetsigian 2016). Results from testing of environmental samples, and diagnostic sensitivity determined from the spiking experiment, demonstrated that each of these qPCR assays has suitably high sensitivity for routine surveillance, although non-detection may occasionally occur, particularly in samples with low target DNA abundance.

In HTS testing for *V. gibba*, target species OTUs were identified in samples from Hobart and Melbourne, where this species is known to occur (Wiltshire *et al.* 2019a), and in a sample from Portland. OTUs with < 97% similarity to *V.* gibba, but for which *V.* gibba was the nearest match in Genbank, were identified in samples from Gove, Darwin, and Weipa. Some species show > 3% intra-species divergence in gene regions, but the 28S region used for the Vgib assay typically has low divergence in bivalves (Mazón-Suástegui *et al.* 2016; Guo *et al.* 2021). Given that the tropical locations where these sequences occurred are outside the recorded temperature tolerance range of *V. gibba*, it is likely that the sequences detected belong to a related corbulid mollusc that does not have sequences in Genbank.

In addition to the sequences with high similarity to *V. gibba,* HTS data identified many OTUs with close matches to all assay regions, that were otherwise different from the relevant Vgib assay gene region. These OTUs would be detected by the Vgib assay, though with lesser efficiency and hence higher C_T, than target sequences. Few of these confounding OTUs could be ascribed to species, but the nearest matches to each belong to several different taxa, including bivalves in the Arcidae. Non-specific Vgib detection was also recorded in gDNA testing of the ark cockle *Anadara trapezia* (Arcidae), with sequencing showing close matches to the Vgib assay regions and strong similarity to potentially confounding sequences identified by HTS. This demonstrates that the assay target regions, although chosen to be specific based on sequence data available at the time of development, were serendipitously similar to sequences in other taxa, including some that are not closely related. Sequences from these taxa did not provide 100% match to the Vgib assay, but the mismatches are unlikely to be sufficient to prevent cross-reactions, although detections would occur with lower efficiency and hence sensitivity than for target DNA. The potential for broad cross-reactivity of the current Vgib assay means that this assay should be re-designed.

The status of *V. gibba* in most locations with qPCR detection remains unclear, because, although cross-reactivity of the assay is likely, the low sensitivity of the HTS method for *V. gibba* means that the presence of target sequences cannot be confidently ruled out. The data support, however, that, in addition to Melbourne and Hobart, *V. gibba* may occur in Portland, where *V. gibba* sequences were identified that were the same as those in Melbourne. An individual *V. gibba* was detected in 1997 surveillance of Portland Harbor (Parry *et al.* 1997) although none were recorded in 2011 traditional surveys (Werner *et al.* 2012). In 2021-22 molecular surveillance, Vgib qPCR detections occurred in 2 of 35 winter samples and 3 of 35 summer samples (Wiltshire *et al.* 2022).

The relatively low number of qPCR detections suggests that, if present, *V. gibba* is not abundant in Portland, and may therefore be difficult to detect with traditional surveillance.

For *A. senhousia,* sequences with > 97% similarity to this species were widely detected, including in samples without qPCR detection, but it is unclear whether the default 97% threshold is appropriate for assigning sequences to species in this case. Divergence of 28S is < 1% between several congeneric mytilid species, and there is < 2% divergence between some mytilid genera (Colgan and da Costa 2013; Kartavtsev *et al.* 2018), therefore sequences with 3% divergence in 28S could represent different species or even different genera within the Mytilidae. Examination of OTUs, which are groupings of sequences with ≥ 99% similarity to one another, demonstrated that OTUs identified in forward primer sets from tropical locations were different from those in Melbourne, Fremantle, and Kwinana, where *A. senhousia* is recorded. The OTUs in forward primer sets for Thevenard had < 97% similarity to *A. senhousia*. These data suggest that different taxa occur in the tropical locations and in Thevenard that are related to *A. senhousia* and have similar 28S sequences, including identical assay regions, but do not have publicly available sequences. It is unclear what taxa may have these sequences, and they may comprise one or more undescribed species. *Arcuatula senhousia* is unlikely to be the species detected in Thevenard given the divergent 28S sequences identified in samples from this location. It is difficult, however, to confidently rule out *A. senhousia* occurrence in other locations, because there are few 28S sequence data for relevant mytilids that would assist in species discrimination.

In addition to sequences of taxa that are likely to be close relatives of *A. senhousia* that would amplify with the Asen assay, HTS identified other sequences that were dissimilar to *A. senhousia,* but which contained regions perfectly matching the assay. These sequences appear to represent at least five mytilid taxa, which could not be ascribed to species, that would efficiently cross-react with the Asen assay. These confounding sequences were not widespread or abundant in the tested samples, but in combination with uncertainty regarding identity of the taxa detected by the Asen assay in Thevenard and in tropical locations, the HTS results show that re-design of this assay is warranted.

Testing of gDNA from a wide range of bivalve species from around Australia did not identify any species that cross-react with the Asen assays, therefore, the identity of cross-reacting taxa for this assay remains unresolved. The low-level detection by Vgib of *Anadara trapezia* gDNA was likely a cross-reaction, given that the Vgib probe and reverse primer do not confer specificity against the *Anadara trapezia* sequence, and mismatches in the forward primer are probably

insufficient to prevent amplification. HTS data also demonstrated that several species in the genera *Anadara* and *Tegillarca* (Arcidae) have 28S sequences that are expected to be detected with the Vgib assay, although not all species with confounding sequences found by BLAST searches of GenBank occur in Australia. Other taxa may also cross-react with the Vgib assay based on the HTS results.

The HTS data of 28S sequences could be used to assist re-design for the Asen and Vgib assays. Modification of the 28S assays for either Asen or Vgib to achieve suitable specificity, however, is likely to be infeasible given low inter-species divergence for this region in bivalves, a lack of sequence data for Australian taxa, and the common occurrence of sequences with high similarity to those of the assay target in environmental samples. Candidate assays targeting 28S for *Mytella strigata* that were specific *in silico* returned non-specific detections when applied to plankton (Wiltshire *et al.* 2021b), further supporting that this gene target is unlikely to be suitable for species-specific detection in bivalves. Low inter-species divergence in 28S also limits the utility of HTS to confirm detections because sequencing errors, generated during PCR amplification and HTS, may cause similar levels (1 – 2%) of sequence divergence to that between some species (Salk *et al.* 2018), and sequences cannot therefore be confidently assigned at the species level even where Genbank sequence data are available. The Asen and Vgib assays should therefore be re-designed to target a different gene region, with COI a potential candidate given the availability of sequence data for this barcode and its generally good performance for species-specific identification (Andújar *et al.* 2018). Suitability of COI for *A. senhousia* and *V. gibba* assay design would need to be investigated, however, because this barcode is not suitable for diagnostic identification in some taxa (Deagle *et al.* 2014). Re-testing DNA, where available, from locations with uncertain occurrence of *A. senhousia* or *V. gibba* with re-designed assays could assist in clarifying the status of these species in these locations.

Testing target gDNA extracts using most assays returned C_T values in the range expected for assays with good analytical performance, which was supported by high DSe estimates from the spiking experiment. All *A. senhousia* specimens, however, returned high C_T values (>28) in comparison to the archived *A. senhousia* DNA that was used during Asen assay development (C_T = 21.0). The Asen assay also demonstrated high C_T values, mostly > 35, relative to other assays in spiked plankton samples. The Asen assay had DSe > 93% in the spiked samples despite the high C_T results, but this assay has low analytical efficiency in comparison to the other assessed assays. Re-design of this assay should also aim to improve its efficiency, which is likely to also improve DSe. DSe of the Asen assay was more severely affected by PCR inhibition than most

other assays, possibly because detections by this assay comprised relatively high C_T results. For most assays, inhibition increased C_T values, but detections still occurred, at least at higher tissue doses, while, for Asen, an increase in C_T beyond normal detections would provide non-detection within 45 PCR cycles. Some individual gDNA extracts returned high C_T values for other assays, but this is likely to be due to varying quality of the gDNA subsamples rather than problems with assay performance, given that extracts from other target specimens were detected with low C_T, as were DNA extracts from spiked plankton samples.

The assays examined in this project can be considered operationally validated because their performance in plankton samples has been characterised, but additional data are needed for some applications. The assays are likely to be suitable for testing other environmental sample types, e.g., settlement plates, but diagnostic performance would need to be investigated separately for these sample types. The efficacy of other sample types at capturing DNA of each target pest would also need to be assessed. The ability of plankton tows to effectively sample the pests that are exotic to Australia has not yet been assessed, but performance of plankton tows as a sample type has been confirmed across multiple taxonomic groups, providing confidence in this method. The collection of plankton samples from relevant overseas locations to demonstrate detection of target species in samples from areas where each occurs was not practical in this project. Plankton samples from relevant areas, however, should be tested if the opportunity arises to confirm efficacy of this sample type for these species.

Optimising the timing of surveillance using plankton tows requires knowledge of seasonality of planktonic pest DNA occurrence (Wiltshire 2021), which is lacking for the majority species considered here. The best time to sample can be informed by spawning and larval period seasonality where known, but the timing of reproduction can shift in introduced ranges, and between years for some species, so this aspect requires ongoing investigation. In the absence of data on seasonality, samples from at least two opposing seasons should be collected and tested to improve detection likelihood across species.

5. CONCLUSIONS

Data on diagnostic performance is now available for all 19 assays that have been implemented in the SARDI testing system, allowing surveillance to be designed and appropriately interpreted. The assays for 11 of the 13 marine pests assessed in the current project: *Hemigrapsus sanguineus, Eriocheir sinensis, Rhithropanopeus harrisii, Charybdis japonica*, *Didemnum vexillum*, *Perna perna, P. viridis, P. canaliculus, Mytilopsis sallei, Mya japonica* and *Maoricolpus roseus* demonstrated performance suitable for application to routine surveillance, albeit with some caveats for *P. perna.*

For these 11 assays, DSe in plankton samples without PCR inhibition ranged from > 79 to 99%. Inhibition decreased DSe for most assays, but the effect was generally minor, and detections occurred in some samples with very high inhibition (scale factor > 100). DSp estimates for these 11 assays were 93.4 – 100.0%. The lowest DSp estimate was for the *P. perna* assay, which also detects *P. canaliculus*. DSp of this assay is likely to be close to 100% in samples from Australia where *P. canaliculus* is currently absent. The Pper assay should not, however, be applied to samples that are not also tested using the Pcan assay, otherwise, the identity of detections would be ambiguous, or sequencing would be required to confirm which species was detected. Should *P. canaliculus* become established within Australia, utility of the current Pper assay would need to be re-assessed. Specificity of the assays for *H. sanguineus, Maoricolpus roseus* and *Mya japonica* was further confirmed by HTS testing.

Inclusion of these 11 species in surveys would not require collection of more samples than are needed for detection of the species of ballast water concern. The understanding of DSe and DSp for these species assists with interpretation of survey results and allows estimation of true prevalence where detections occur, and of the maximum plausible likelihood of occurrence in the case of non-detection.

The other two assays assessed, i.e., those for *Arcuatula senhousia* and *Varicorbula gibba*, had good DSe of 93.2 and 90.0%, but low estimated DSp of 90.5 and 70.5% respectively. HTS data indicate that non-target sequences are being detected by both assays and the Vgib assay was also shown to detect *Anadara trapezia* gDNA. The HTS data further suggest that the 28S region is not sufficiently diagnostic for species-specific detection of these taxa, and the lack of divergence in 28S between several bivalves limits the utility of HTS as a confirmatory testing method. These two assays should therefore be re-designed using a different genetic marker.

Hemigrapsus sanguineus DNA was successfully detected by qPCR in plankton samples collected in Melbourne, including in archived DNA from samples collected in November 2017 and March 2018. Spring – summer appears to be the best time for detection of this species, although further investigation is needed to confirm seasonality. Plankton samples from areas where the pests that are exotic to Australia occur should be obtained and tested if the opportunity arises to confirm the efficacy of this sample type for detection of these pests. Samples from the established range should ideally be collected across multiple seasons to provide preliminary data on seasonality of detection likelihood. Further investigation is needed to characterise assay performance in other sample types and the efficacy of these sample types at capturing target DNA.

Surveillance using the assays assessed in this project can assist in detecting incursions of the currently exotic species and contribute to improving knowledge of the distribution of the pests that occur in Australia, informing management steps that will aid in preventing their further spread.

6. APPENDIX

6.1. Supplementary modelling methods and results

6.1.1. Use and interpretation of link functions

Several link functions are available for GLM (or GLMM) of binary data, with each link function providing a monotonic transform of probability or proportion data from the (0,1) interval to the real number line $(-\infty, \infty)$ (Fox 2015; Damisa *et al.* 2017). The logistic, or log-odds, link is the most commonly applied option due to ease of interpretation and good performance in many cases, but does not always provide a good fit, particularly where the data are asymmetric around a probability of 0.5 (Damisa *et al.* 2017). The complementary log-log (cloglog) link is suitable in many cases for asymmetric data (Fox 2015; Damisa *et al.* 2017) and has direct interpretation in terms of underlying abundance in modelling detection probabilities (Royle and Dorazio 2009). Specifically, where detection relies on capturing a randomly distributed target within a sample, the likelihood of detection (π) is given by:

$$\pi = 1 - exp(-\lambda)$$

Where λ is the expected number of targets per sample, which is a function of target abundance (Royle and Dorazio 2009).

The cloglog link is (Fox 2015):

$$\eta = log(-log(1 - \pi))$$

Where η is the linear predictor in GL(M)M. The inverse link function of the cloglog transform is therefore (Fox 2015):

$$\pi = 1 - exp(-exp(\eta))$$

The exponent of the linear predictor, η, of a GL(M)M using the cloglog link can therefore be interpreted as λ, i.e., the relative number of targets per unit sample volume (Royle and Dorazio 2009).

In our ZAG models, which included sample set as a categorical fixed effect, the linear predictor for sample set *i*, η_i, is given by:

$$\eta_i = \beta_0 + \beta_i$$

Where β_0 is the intercept and β_i is the coefficient for sample set *i*. The difference between the coefficients for two sample sets, $\beta_1 - \beta_2$, assessed using linear combinations, is therefore equal to $\eta_1 - \eta_2$, and the exponent of this difference can be interpreted as the multiplicative difference in abundance, λ, as follows:

$$exp(\eta_1 - \eta_2) = exp(\eta_1)/exp(\eta_2) = \lambda_1/\lambda_2$$

The log link used in the Gamma component of the ZAG model similarly results in coefficients that indicate multiplicative difference when exponentiated. For the Gamma component, the mean for sample set *i*, $\mu_i = exp(\eta_1)$, and the exponent of the difference between sample sets is therefore:

$$exp(\eta_1 - \eta_2) = exp(\eta_1)/exp(\eta_2) = \mu_1/\mu_2$$

We applied the cloglog link in ZAG modelling due to the direct link between planktonic abundance and likelihood of capture in a plankton sample, as per Wiltshire (2021). Note that this is the planktonic concentration of discrete DNA-containing particles, which is different from the total DNA yield. DNA-containing particles may be spores, eggs, sperm, larvae, or cells which will contain varying amounts of DNA, affecting the resulting DNA yield in a sample. The cloglog link was also regarded as suitable for modelling detection likelihood in constructed samples to which pest tissue was added because detection in these samples will depend on abundance of target DNA, which is related to the added tissue dose.

6.1.2. Supplementary ZAG model results

Table 24. Coefficients of ZAG model and predictions for each model component. Sample set was applied as a fixed factor with the Feb 2022 sample set as the reference level. Note that, with no detections in the Jul 2017 or Jun 2018 sample sets, no data were available to estimate coefficients for DNA yield and estimates for these sample sets therefore reflect the prior used. N/A = not available.

Coefficient	**Mean estimate (95% HDI)**	**Prediction (95% HDI)**
Detection likelihood intercept (Feb 2022)	-0.78 (-1.95 – 0.35)	0.35 (0.10 – 0.71)
Detection likelihood Set: Jul 2017	-15.26 (-29.62 – 8.86)	0.00 (0.00 – 1.00)
Detection likelihood Set: Nov 2017	-1.21 (-2.50 – 0.26)	0.13 (0.028 – 0.32)
Detection likelihood Set: Mar 2018	-1.71 (-3.03 – -0.13)	0.072 (0.011 – 0.24)
Detection likelihood Set: Jun 2018	-15.25 (-29.62 – 8.88)	0.00 (0.00 – 1.00)
DNA yield intercept (Feb 2022)	6.20 (5.17 – 7.29)	391.7 (111.7 – 1,126)
DNA yield Set: Jul 2017	0.00 (-62.03 – 61.89)	N/A
DNA yield Set: Nov 2017	-2.93 (-4.61 – -1.28)	16.78 (3.54 – 69.85)
DNA yield Set: Mar 2018	-2.32 (-4.25 – -0.53)	23.72 (3.06 – 145.9)
DNA yield Set: Jun 2018	0.00 (-62.03 – 61.89)	N/A

6.2. JAGS model code

```
model{
  for(i in 1:n){
    # se = test DSe
    # sp = test DSp
    # Assay se accounting for inhibition
      cloglog(se[i,1]) <- max(min(12, a[species[i]] + bSF[species[i]]
* lnSF[i]),-12)
      se[i,2] <- 1 # reference test
      sp[i,1] <- SP[species[i]]
      sp[i,2] <- 1 # reference test
    for(k in 1:2){
      s1[i,k] <- se[i,k]^x[i,k]*((1-se[i,k])^(1-x[i,k]))
      s2[i,k] <- sp[i,k]^(1-x[i,k])*((1-sp[i,k])^x[i,k])
      }
      # probability of detection
      eta[i] <- prod(s1[i,1:2])
      theta[i] <-prod(s2[i,1:2])
      prob[i]   <-   max(min(1-1e-9,   (pi[species[i]]*eta[i]   +   (1-
pi[species[i]])*theta[i])),1e-9)
      # Likelihood using zero trick
      z[i] ~ dpois( - log(prob[i]))
  }
  # Priors
  for(j in 1:K) {
    # prevalence per species
    pi[j] <- 200/240
    # se intercept
    a[j] <- max(min(12, cloglog(se.int[j])), -12)
    se.int[j] ~ dbeta(1,1)T(1-SP[j], )
    # scale factor effect
    bSF[j] ~ dnorm(0,0.154)T(,0) #95% prob in range -5 to 0
    # specificity
    SP[j] ~ dbeta(spA[j],spB[j])
    }
}
```

6.3. High throughput sequencing target detections by sample

The mean total reads per sample with each primer set are shown in Table 25 for qPCR negative and positive samples. Note that the targeted HTS approach used means that samples without target DNA present could potentially return no or very few total reads, which is not indicative of method failure given primer sets used were selective. The number of reads identified as the relevant target (with ≥ 97% similarity) from HTS results is shown in Tables 26-30, along with information on sample location and target species qPCR result (C_T value or non-detection). The number of reads with < 97% similarity to the target but with sequences identified by BLAST as having matches to probe and complementary primer sequences is also shown, along with the total reads for each library (note, each library is amplicons from a sample – primer set combination). The sequences with BLAST matches were further investigated to determine their potential to cross-react with the relevant assay based on the number and location of mismatches.

Table 25. Mean total reads per sample with each primer set showing mean ± standard deviation and range (minimum to maximum) in brackets.

Primer set	Mean reads in qPCR negative samples	Mean reads in qPCR positive samples
Hsan Fwd	1,437 ± 4,037 (0 – 21,609)	12,676 ± 24,225 (23 – 95,606)
Hsan Rev	13,519 ± 28,276 (0 – 110,932)	96,334 ± 57,833 (53 – 206,181)
Mros Fwd	3,352 ± 7,658 (0 – 27,890)	14,937 ± 20,827 (0 – 156,860)
Mros Rev	29 ± 74 (0 – 288)	2,161 ± 4,834 (0 – 22,487)
Mjap Fwd	1,942 ± 2,009 (0 – 5,408)	7,194 ± 13,228 (5 – 43,609)
Mjap Rev	2,902 ± 3,828 (8 – 17,422)	10,411 ± 16,906 (270 – 53,381)
Asen Fwd	10,442 ± 7,846 (0 – 31,674)	14,287 ± 9,855 (0 – 61,653)
Asen Rev	3,380 ± 4,015 (0 – 14,301)	13,719 ± 14,081 (0 – 75,808)
Vgib Fwd	9,393 ± 6,735 (27 – 22,959)	8,493 ± 6,303 (131 – 26,346)
Vgib Rev	31,410 ± 47,550 (1,994 – 316,915)	24,115 ± 18,073 (95 – 78,107)

Table 26. Location, qPCR result and number of reads assigned to *Hemigrapsus sanguineus* in forward (Fwd) and reverse (Rev) primer libraries by sample, plus number of non-target reads identified by BLAST as having matches to probe and complementary primer for each library and total reads per library. ND = not detected by qPCR.

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Melbourne	ND			2	2		2
Melbourne	ND			59	5		653
Melbourne	ND			3	2		111
Melbourne	ND			316			13,808
Melbourne	ND			672	2		38,222
Melbourne	ND			1,047	1		30,577
Melbourne	33.71	7		23	53,843		54,551
Melbourne	ND			0	2		406
Melbourne	ND			26			0
Melbourne	ND			13	2		110,932
Melbourne	ND			9			367
Melbourne	45.51			6,733	4		294
Melbourne	33.89	148		737	68,732		74,171
Melbourne	ND			433			0
Melbourne	33.83	611		22,567	46,646		58,271
Melbourne	ND			3			30
Melbourne	31.25	4,585		4,862	100,331	1	103,680
Melbourne	28.14	35,116		35,219	203,842		206,181
Melbourne	33.48	6		117	105,333		106,692
Melbourne	ND	1		717	2		106,022
Melbourne	29.34	9,359		18,285	121,831	2	124,869
Melbourne	27.01	93,570		95,606	145,617		147,318
Melbourne	30.27	5,633		5,904	76,934		78,471
Melbourne	32.16	4,087		4,166	47,954		48,752
Melbourne	33.70	132		242	150,389		153,635

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Melbourne	32.90	405		4,463	168,049		170,959
Melbourne	ND			16	2		41,928
Melbourne	ND			228	1		255
Melbourne	ND			17	1		22
Melbourne	ND			154	9		8,995
Melbourne	ND			4,398			12,462
Melbourne	36.77	133		133	3		53
Melbourne	ND			1,168			4
Melbourne	ND			143			163
Melbourne	ND			2,649	1		21,639
Melbourne	ND			163	6		2,158
Melbourne	ND			300	1		553
Melbourne	ND			429	1		1,885
Melbourne	ND			1,564	1		1,712
Melbourne	ND			5,824	3		799
Melbourne	ND			0			1,515
Melbourne	ND			468	1		31
Melbourne	ND			677			3,561
Melbourne	ND			21,609	1		6,763
Melbourne	32.30	1,370		1,492	123,837		125,392
Melbourne	31.45	1,079		2,261	85,706		88,058

Table 27. Location, qPCR result and number of reads assigned to *Maoricolpus roseus* in forward (Fwd) and reverse (Rev) primer libraries by sample, plus number of non-target reads identified by BLAST as having matches to probe and complementary primer for each library and total reads per library. ND = not detected by qPCR.

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Adelaide	37.89	1		9,254			13
Gove	37.66	22	1	122			1
Hay Point	40.54	2		1,818			1
Klein Point	36.89	11		8,838			3
Portland	37.66			0			2
Portland	ND			1			1
Portland	35.53	4		4	1		2
Portland	37.09	4		7			0
Portland	ND			6			0
Portland	ND	4		4			0
Portland	38.78	1		1			0
Portland	35.66	5,086	3,208	8,531			0
Portland	ND	5		8			0
Portland	ND	5		7			0
Portland	ND	8		25			2
Portland	35.38			0			0
Portland	34.59	1		1			1
Portland	ND			4			19
Portland	35.52	11,620		11,823	4		7
Portland	ND			0			34
Portland	31.96	17,064		17,263	31	1	32
Portland	34.53			0	11		17
Portland	ND	4		5			5
Portland	32.83			0	66		77
Portland	ND	4		5			0

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Portland	35.80	17,316		17,406	2		7
Portland	ND	10		12			0
Portland	ND	11		16			1
Blackman Bay	30.43	7,048	960	19,173	295		347
Prince of Wales Bay	34.76			987	30		51
Prince of Wales Bay	ND	3		1,536			90
Prince of Wales Bay	ND	11		23			8
Derwent	31.61			0	154	22	179
Derwent	31.97	23,294		23,410			2
Derwent	36.08			0			0
Derwent	36.44	3		3			0
Derwent	34.12	5,993		6,795	24		35
Pipeclay Lagoon	26.67	540	59	741	2,248	93	2,350
Pipeclay Lagoon	28.64	72,232	11,542	85,302	1,280	127	1,423
Pipeclay Lagoon	27.48	23,958	1,485	28,820	4,557	86	4,651
Dunalley	34.38	3,339		18,131			0
Blackman Bay	31.13	8,176	991	11,607	4		4
Blackman Bay	29.57	12,063	1,260	18,218	127	33	164
Pipeclay Lagoon	ND			0			31
Derwent	ND	2		4			0
Pipeclay Lagoon	33.54	6		856			0
Pipeclay Lagoon	ND	1		5,052			18
Pipeclay Lagoon	31.18	3		9,830	20		496
Pipeclay Lagoon	35.04	209		24,935			0
Pittwater	ND	2		3			0
Pittwater	ND	3		748			0
Pittwater	ND	5		3,305			0
Pittwater	ND	1		6,025			18

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Pipeclay Lagoon	28.86	16,852	496	32,252	900	19	948
Pipeclay Lagoon	29.94	18,636	919	42,264	1,042	7	1,055
Pipeclay Lagoon	24.46	20,968	311	20,785	4,495	12	4
Pittwater	ND	2		33,269			7
Pittwater	32.01	8,013		40,184	6		4,515
Pittwater	35.65			1			2
D'Entrecasteaux	29.32	13,303	5	13,385	126	2	128
D'Entrecasteaux	29.36	19,178	777	20,178	833	14	849
D'Entrecasteaux	26.71	22,591	124	22,894	1		1
D'Entrecasteaux	29.43	20,101		20,240	2,126		2,128
D'Entrecasteaux	28.06	7,569	37	7,675	59	2	61
D'Entrecasteaux	28.55	15,751	217	16,066	2,562	55	2,619
D'Entrecasteaux	29.11	17,257	1,506	19,177	254	17	272
D'Entrecasteaux	28.42	8,484	914	9,814	7,125	240	7,377
D'Entrecasteaux	28.82	37,483	2,031	40,107	70		70
D'Entrecasteaux	27.53	17,301	522	18,127			0
D'Entrecasteaux	26.95	22,386	1,258	23,847	1,305	65	1,374
D'Entrecasteaux	26.69	20,622	160	20,882	16,964	40	17,102
D'Entrecasteaux	28.66	12,727	1,021	13,951	346	30	383
D'Entrecasteaux	27.88	18,327	821	19,499	7,199	89	7,301
D'Entrecasteaux	28.67	16,198	435	16,802	3,226	46	3,407
Orford - Triabunna	34.92	7,751		8,958			0
Orford - Triabunna	32.65	7,041		7,154	127		127
Orford - Triabunna	35.67	7,720		8,578	53		59
Orford - Triabunna	35.04	35,730		35,932			0
Orford - Triabunna	31.04	5,553	255	5,891	403	5	417
Orford - Triabunna	32.53	8,091		8,156	301	45	393
Orford - Triabunna	ND	22		27,890			288

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Orford - Triabunna	30.55	21,673	4,965	27,270	1,281		1,283
Orford - Triabunna	ND	2		1,849			245
Orford - Triabunna	26.12	27,815	1,166	55,520	12,685	211	12,960
Orford - Triabunna	27.43	13,593	29	20,067	20,471	54	20,581
Orford - Triabunna	29.66	18,792	1,128	25,805	5,040	44	5,294
Orford - Triabunna	30.37	14,951		19,856	3,193		3,224
Orford - Triabunna	35.76	255		292	10		10
Orford - Triabunna	ND	2		8334			172
D'Entrecasteaux	27.07	18,078	2,006	20,843	20,422	633	22,487
D'Entrecasteaux	27.82	20,721	1,087	22,048	1,429	65	1,558
D'Entrecasteaux	26.69	27,906	583	29,015	12,698	18	14,334
D'Entrecasteaux	28.60	26,645	459	27,449	10,117	74	10,976
D'Entrecasteaux	27.49	30,273		30,445	15,643	25	17,642
D'Entrecasteaux	27.51	152,573	2,949	156,860	13,015	121	13,590
D'Entrecasteaux	26.91	20,840	150	21,104	9,135	10	9,212
D'Entrecasteaux	ND	9		727			0
D'Entrecasteaux	35.51	1		5			0
D'Entrecasteaux	ND	1		1			0
D'Entrecasteaux	36.62	9		10			8
D'Entrecasteaux	34.47	12,451		12,520	14		164
D'Entrecasteaux	35.65	16,957		17,319			33
D'Entrecasteaux	35.46			0			0
D'Entrecasteaux	36.64	12		14			0
Orford - Triabunna	37.65	9		12			0
Orford - Triabunna	ND	4		5			0
Orford - Triabunna	ND	1		743			0
Orford - Triabunna	37.61			0			0
Orford - Triabunna	37.57	2		2			0

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Orford - Triabunna	ND			0			0
Orford - Triabunna	31.86	12,562		13,086	767		769
Orford - Triabunna	36.76	5		5			0
Orford - Triabunna	34.74	1,837		1,847	13		13
Orford - Triabunna	33.33	4,718		4,777	22		22
Orford - Triabunna	33.05	10,029		10,077	1		1
Orford - Triabunna	30.59	14,946	1,254	20,477	1,678		1,679
Orford - Triabunna	31.74	13,268	2	13,407	175		180
Orford - Triabunna	36.11	2		2	52		52
Orford - Triabunna	ND			1			0
Orford - Triabunna	ND	4		4			0

Table 28. Location, qPCR result and number of reads assigned to *Mya japonica* in forward (Fwd) and reverse (Rev) primer libraries by sample, plus number of non-target reads identified by BLAST as having matches to probe and complementary primer for each library and total reads per library. ND = not detected by qPCR.

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Port Giles	37.33			630			843
Prince of Wales Bay	ND			3,677			7,670
Prince of Wales Bay	ND			5,369			3,611
Prince of Wales Bay	ND			1,235			3,774
Pipeclay Lagoon	ND			1,719			183
Dunalley	ND	2		7			1,297
Blackman Bay	ND		1	960			301
Blackman Bay	ND			5,408			759
D'Entrecasteaux	ND			2,495			665

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
D'Entrecasteaux	ND			4,393			223
D'Entrecasteaux	ND			71			24
D'Entrecasteaux	ND		1	36			926
D'Entrecasteaux	ND			48	8		858
Orford - Triabunna	32.66	317		372	3,012		3,458
Orford - Triabunna	31.05	32		168	5,352	2	5,570
Orford - Triabunna	33.72	41		6,059	2,155		16,710
Orford - Triabunna	ND	2	18	4,044	1		425
Orford - Triabunna	36.79		2	159			1578
Orford - Triabunna	ND			4,296			5,603
Orford - Triabunna	31.62	8,992	6	18,272	5,942	2	14,053
Orford - Triabunna	32.65	642	80	1,376	1,127		1,349
Orford - Triabunna	36.78		1	536			31,383
Orford - Triabunna	ND			4,449	539		17,422
D'Entrecasteaux	ND			844		1	5,439
D'Entrecasteaux	ND			1,036			3,265
D'Entrecasteaux	ND			738		1	6,598
D'Entrecasteaux	ND			395			2,739
D'Entrecasteaux	ND			528			3,860
D'Entrecasteaux	ND			4,840			1,148
D'Entrecasteaux	ND			0		2	2,387
D'Entrecasteaux	36.62			268			389
D'Entrecasteaux	ND			3			467
D'Entrecasteaux	ND			15			8
Orford - Triabunna	32.93	374	2	378	293		294
Orford - Triabunna	31.31	25		28	1,288		1,294
Orford - Triabunna	31.09	1,914	4	1,927	2,707	1	2,712

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Orford - Triabunna	25.83	43,557	40	43,609	53,293	7	53,381
Orford - Triabunna	31.37	533		541	268		270
Orford - Triabunna	33.16	538	1	745	740		740
Orford - Triabunna	29.85	929		1,168	369	2	1,981
Orford - Triabunna	26.01	32,886	84	33,732	8,153	1	8,176
Orford - Triabunna	25.47	26,377	42	26,721	52,553	8	52,568
Orford - Triabunna	32.23			5	1,049	1	1,060

Table 29. Location, qPCR result and number of reads assigned to *Arcuatula senhousia* in forward (Fwd) and reverse (Rev) primer libraries by sample, plus number of non-target reads identified by BLAST as having matches to probe and complementary primer for each library and total reads per library. ND = not detected by qPCR.

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Brisbane	ND			14,149	1		532
Brisbane	ND			12,834	71		690
Brisbane	ND			18,240	1,864	4	2,325
Brisbane	44.88	19	445	18,484	12,009	53	15,392
Bunbury	ND			3,784			141
Bunbury	ND			11,834			525
Bunbury	ND			9,125			77
Cairns	36.50	1,240	149	19,905	8,359	3	8,660
Darwin	ND	320	13	1,943	2,297	1	2,588
Darwin	38.56	3,214	299	13,945	6,745	4	6,838
Darwin	ND	8		6,951	2,737	5	2,925
Darwin	ND	2		110	8,555	15	8,884
Devonport	42.06	220	5	16,889	11,097	15	11,700

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Devonport	ND	10	1	17,500	6,658	9	7,237
Devonport	ND			1,307	171	1	329
Devonport	42.81			4,857	5		415
Fremantle	33.74			0			0
Fremantle	ND	8	3,492	8,022	7,760	8	8,309
Fremantle	ND	7	2,661	6,334	6,806	6	7,058
Fremantle	33.49	6	4,542	11,602	14,192	13	14,945
Fremantle	33.49	2	5,868	17,536	10,393	21	11,196
Fremantle	ND	59	4,286	11,889	14,086	14	14,301
Geraldton	ND			3,317			2,956
Geraldton	ND			14,935			2,156
Geraldton	ND			13,121			2,947
Gladstone	35.61	272	83	17,729	2,304	3	2,403
Gladstone	37.37	366	69	18,708	8,388	9	8,618
Gladstone	39.21	11	2	4,910	9,849	12	10,574
Gladstone	39.57	153	22	23,518	9,178	7	9,772
Gladstone	ND	24	69	17,772	9,977	14	11,062
Gladstone	ND			7,879	4,906	4	5,474
Gladstone	39.51	90	56	12,858	5,971	2	6,596
Gladstone	ND	304	356	25,169	12,267	11	12,826
Gladstone	ND			0	9,349	12	9,647
Gladstone	35.88	1,272	143	13,058	11,399	11	11,878
Gladstone	ND	255	238	24,939	8,082	7	8,301
Gladstone	37.69	116	12	26,053	6,932	6	7,500
Gladstone	ND	11	9	16,266	5,832	6	6,704
Gladstone	ND	136	122	19,974	7,816	9	8,804
Gladstone	37.02	144	75	5,141	9,328	12	9,784
Gove	ND		196	31,674	2,724	12	2,821

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Gove	41.31		39	27,420	2,339	16	2,480
Gove	ND		70	30,343			0
Gove	ND		16	14,733	648	5	695
Gove	34.77		39	13,614	1,208	3	1,270
Gove	41.47		608	25,841	3,914	13	3,975
Hay Point	ND			9,939			14
Hay Point	38.28			8,703	691		702
Hay Point	ND			24,885			139
Hay Point	ND			13,821	229	1	297
Hay Point	42.62			12,331	576	3	1,079
Hay Point	44.10			12,721	18		53
Hobart	ND	262	21	12,275	7,281	79	7,855
Hobart	ND	884	20	11,618	8,274	19	8,642
Hobart	ND	123	3	6,737	7,273	35	7,905
Hobart	ND	832	29	8,798	7,994	84	9,521
Hobart	ND	49	1	4,147	4,717	39	6,620
Hobart	ND	318	3	4,898	9,259	53	11,360
Hobart	ND			2,435			430
Hobart	ND	21		7,980	1,978	4	2,191
Hobart	ND	21		6,452			0
Klein Point	ND			364			9
Kwinana	ND		1,806	14,658	9,069	11	10,709
Kwinana	ND	2	1,497	7,797	6,729	16	8,173
Kwinana	ND		820	17,319	5,413	5	7,150
Kwinana	37.85		656	7,915	4,811	10	6,213
Kwinana	36.85	1	1,334	13,046	4,928	8	5,738
Kwinana	37.20	14	832	6,513	5,003	4	5,754
Kwinana	36.74	11	942	16,207	6,636	9	7,725

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Kwinana	ND	4	131	664	5,089	8	5,742
Kwinana	ND	4	8	10,647	2,945	1	4,232
Kwinana	36.79	4	1,023	12,728	7,989	17	8,955
Kwinana	38.58	4	180	6,828	6,584	6	7,295
Melbourne	41.21	2,935	4	5,160	12,057	5	12,065
Melbourne	41.83	1,473	2	4,122	10,409	4	10,430
Newcastle	ND			11,570			310
Newcastle	ND			12,930			1,039
Newcastle	ND			7,505			48
Port Adelaide	ND			177			175
Port Adelaide	ND			5,064			43
Port Botany	ND			15,641			33
Port Botany	ND			3,271			109
Port Botany	ND			8,738			220
Port Kembla	ND			23,314			1,316
Port Kembla	ND			11,667			36
Port Kembla	ND			314			61
Port Lincoln	ND			3,688			3
Port Lincoln	ND			9,575			3
Port Lincoln	ND			4,193			3
Port Lincoln	ND			11,349			3
Port Lincoln	ND		1	2,698			1
Port Lincoln	ND	1		8,562			3
Port Lincoln	ND			1,856			1
Port Lincoln	ND			6,064			0
Port Lincoln	ND			6,568			785
Port Lincoln	ND			3,528			828
Port Lincoln	ND			8,191			1,049

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Portland	ND			1,734			32
Portland	ND			342			46
Thevenard	35.16		349	24,292	10,402	4	10,416
Thevenard	35.22		106	7,814	1,392	1	1,397
Thevenard	34.05		271	22,247	30,562	21	30,624
Thevenard	34.27		700	61,653	35,053	18	35,095
Thevenard	33.88	1	322	22,059	13,717	13	13,740
Thevenard	33.52		114	3,335	28,886	11	28,914
Thevenard	34.86		212	15,155	14,450	15	14,471
Thevenard	35.78		178	15,476	8,635	5	8,646
Thevenard	34.93		262	14,394	21,666	9	21,691
Thevenard	33.55		195	13,724	19,365	11	19,385
Thevenard	32.83		1	140	75,754	17	75,808
Thevenard	33.23		203	14,437	46,268	13	46,301
Thevenard	34.36		191	12,989	38,061	8	38,077
Thevenard	33.86		271	22,190	32,768	22	32,800
Thevenard	34.72		302	14,987	24,095	19	24,118
Thevenard	35.10		252	18,184	30,720	18	30,751
Weipa	ND	2,804	659	12,222	9,552	6	9,951
Weipa	41.90	855	264	7,605	9,610	5	10,001
Weipa	41.87	9	4	77	7,080	8	7,248
Weipa	41.45	1,298	404	10,945	12,555	11	12,740
Weipa	ND	966	797	4,518	4,013	9	4,306
Weipa	ND	2,644	847	12,380	7,397	10	7,653
Whyalla	ND			30,250			2,013
Whyalla	ND			25,502			0
Whyalla	ND			20,193			2,091

Table 30. Location, qPCR result and number of reads assigned to *Varicorbula gibba* in forward (Fwd) and reverse (Rev) primer libraries by sample, plus number of non-target reads identified by BLAST as having matches to probe and complementary primer for each library and total reads per library. ND = not detected by qPCR.

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Brisbane	38.92			9,855			6,400
Brisbane	33.44			13,180			25,144
Brisbane	35.24			10,370			24,008
Brisbane	ND			5,044			30,504
Bunbury	40.80			10,042			6,516
Bunbury	42.66			8,576			12,720
Bunbury	42.43			17,981			6,356
Cairns	ND			6,006			5,367
Darwin	32.34		1	17,605			7,880
Darwin	33.23		3	15,443			14,227
Darwin	ND		30	13,614			19,362
Darwin	30.85		21	10,591			19,149
Devonport	ND			7,197			1,994
Devonport	41.90			5,572			20,109
Devonport	44.14			8,749			16,408
Devonport	ND			14,959			20,133
Fremantle	ND			19,855			32,823
Fremantle	ND			12,361			35,776
Fremantle	37.36			10,234			26,557
Fremantle	38.94			7,658			18,929
Fremantle	ND			4,893			26,508
Fremantle	36.81		1	17,182			22,547
Geraldton	40.53		1	16,739			20,739
Geraldton	40.87			15,949			16,469
Geraldton	40.47			10,011			17,878

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Gladstone	36.47			21,866			12,559
Gladstone	34.59			26,346			21,136
Gladstone	37.20			4,035			16,646
Gladstone	ND			15,335			17,098
Gladstone	ND			4,019			19,378
Gladstone	ND			17,808			50,466
Gladstone	37.85			5,217			9,652
Gladstone	ND		1	17,483			13,184
Gladstone	41.47			16,074			18,717
Gladstone	34.73			613			95
Gladstone	33.65		1	17,167			23,144
Gladstone	ND			21,302			27,505
Gladstone	ND			11,177			24,663
Gladstone	ND			17,256			20,135
Gladstone	ND			22,959			22,535
Gove	ND		3	10,832			12,115
Gove	28.43		18	18,129			15,017
Gove	28.34		2	22,359			26,903
Gove	ND		2	9,900			21,155
Gove	ND		2	17,348			10,873
Gove	27.85		80	9,051			21,996
Hay Point	39.96			4,361			8,145
Hay Point	ND			13,809			9,749
Hay Point	42.34			17,633			22,399
Hay Point	ND			15,275			6,449
Hay Point	41.44			17,063			14,145
Hay Point	ND			2,668			5,859
Hobart	30.82	5		19,679			2,497

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Hobart	29.71	11	1	15,647			12,399
Hobart	32.68		1	1,940			11,151
Hobart	29.32	6		9,950			12,946
Hobart	30.66	1		5,042			16,671
Hobart	29.26		1	9,750			22,111
Hobart	21.94	3		2,829	67	2	10,565
Hobart	21.71	2,469	8	8,208	13	2	16,600
Hobart	22.02	1,204	1	1,487	24		15,981
Klein Point	41.94			1,616			28,825
Kwinana	ND		1	5,891			15,203
Kwinana	40.18			8,168			26,741
Kwinana	39.56			7,896			6,722
Kwinana	ND			11,857			20,456
Kwinana	ND			7,690			20,160
Kwinana	ND			13,927			18,844
Kwinana	ND			6,811			19,247
Kwinana	37.65			12,946			359
Kwinana	37.83			9,238			23,083
Kwinana	40.26		1	7,204			23,250
Kwinana	ND			10,759			21,685
Melbourne	31.59	3,633	26	5,241			36,174
Melbourne	33.88			2,545			78,107
Newcastle	40.79			7,944			19,093
Newcastle	43.24			2,732			17,416
Newcastle	42.00			7,570			26,418
Port Adelaide	41.62		1	7,285			743
Port Adelaide	43.05			1,022		1	75,191
Port Botany	40.98			3,297			15,897

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Port Botany	40.99			8,386			15,446
Port Botany	43.64			11,073			18,904
Port Kembla	42.22			11,214			19,394
Port Kembla	42.76			2,372			24,565
Port Kembla	42.99			7,571			13,700
Port Lincoln	35.64			5,023			30,480
Port Lincoln	35.31		1	10,169			1,223
Port Lincoln	35.15			11,750			8,854
Port Lincoln	ND		1	1,946			24,292
Port Lincoln	40.15		2	1,172			52,837
Port Lincoln	40.63		2	3,034			43,107
Port Lincoln	39.83		2	131			63,823
Port Lincoln	ND		3	772			41,220
Port Lincoln	40.26			334			63,126
Port Lincoln	43.43		2	678			34,662
Port Lincoln	ND			630			45,624
Portland	33.88	611	1	3,418			52,833
Portland	38.20			6,479			65,393
Thevenard	42.10		1	1,156			45,742
Thevenard	ND			27			53,639
Thevenard	ND		19	710			32,128
Thevenard	43.38		1	2,338		1	61,252
Thevenard	41.76		10	730			45,913
Thevenard	43.51		65	1,568			55,401
Thevenard	ND		2	710		1	42,323
Thevenard	ND		1	404		1	31,478
Thevenard	ND		2	1,702		2	316,915
Thevenard	41.74			406		1	51,753

Location	qPCR C_T	Fwd primer target species (≥ 97% match)	Fwd primer BLAST match to probe + rev	Fwd library total reads	Rev primer target species (≥ 97% match)	Rev primer BLAST match to probe + fwd	Rev library total reads
Thevenard	41.96		33	2,423			52,255
Thevenard	41.42		17	1,543			26,283
Thevenard	ND			983		1	59,742
Thevenard	41.29	1	63	1,643			59,166
Thevenard	ND			1,234			32,783
Thevenard	44.04		191	819		1	4,400
Weipa	31.79			6,372			20,315
Weipa	ND			8,287			19,203
Weipa	ND			15,928			17,806
Weipa	ND		3	13,743			21,413
Weipa	31.89			6,742			21,593
Weipa	31.41		1	4,788			1,755
Whyalla	36.92			20,621			11,463
Whyalla	37.13			13,555			22,688
Whyalla	37.00			12,541			21,698

7. REFERENCES

Andersen J, Kallenbach E, Thaulow J, Hesselsoe M, Bekkevold D, Hansen B, Jacobsen L, Olesen C, Moller P, Knudsen S (2018) Development of species-specific eDNA-based test systems for monitoring of non-indigenous species in Danish marine waters. NIVA Denmark Water Research, Oslo.

Andújar C, Arribas P, Yu DW, Vogler AP, Emerson BC (2018) Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology* 27, 3968-3975.

Arab A (2015) Spatial and Spatio-Temporal models for modeling epidemiological data with excess zeros. *Int J Environ Res Public Health* 12, 10536-48.

Baillie S, McGowan C, May-Mcnally S, Leggatt R, Sutherland B, Robinson S (2019) Environmental DNA and its applications to Fisheries and Oceans Canada: National needs and priorities. Fisheries and Oceans Canada, Ottawa.

Bakka H, Rue H, Fuglstad GA, Riebler A, Bolin D, Illian J, Krainski E, Simpson D, Lindgren F (2018) Spatial modeling with R-INLA: A review. *Wiley Interdisciplinary Reviews: Computational Statistics* 10, e1443.

Bakka H, Vanhatalo J, Illian JB, Simpson D, Rue H (2019) Non-stationary Gaussian models with physical barriers. *Spatial statistics* 29, 268-288.

Bax N, Dunstan P, Gunasekera R, Patil J, Sutton C (2006) Evaluation of National Control Plan management options for the North Pacific Seastar *Asterias amurensis*. National Heritage Trust.

Bax NJ, McEnnulty FR, Gowlett-Holmes KL (2003) Distribution and biology of the introduced gastropod, *Maoricolpus roseus* (Quoy and Gamard, 1834) (Caenogastropoda: Turritellidae) in Australia. Centre for Research on Introduced Marine Pests Technical Report Number 25. CSIRO Marine Research, Hobart.

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

Bott NJ (2015) Detection of Invasive Species. In 'Springer Handbook of Marine Biotechnology'. (Ed. S-K Kim) pp. 441-453. (Springer Berlin Heidelberg: Berlin, Heidelberg)

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. *Preventive Veterinary Medicine* 68, 145-63.

Brousseau DJ, McSweeney L (2016) A comparison of reproductive patterns and adult dispersal in sympatric introduced and native marine crabs: implications for species characteristics of invaders. *Biological Invasions* 18, 1275-1286.

Buxton AS, Groombridge JJ, Zakaria NB, Griffiths RA (2017) Seasonal variation in environmental DNA in relation to population size and environmental factors. *Sci Rep* 7, 46294.

Camacho-Sanchez M, Burraco P, Gomez-Mestre I, Leonard JA (2013) Preservation of RNA and DNA from mammal samples under field conditions. *Molecular Ecology Resources* 13, 663-673.

Cavieres J, Nicolis O (2018) Using a spatio-temporal Bayesian approach to estimate the relative abundance index of yellow squat lobster (*Cervimunida johni*) off Chile. *Fisheries Research* 208, 97-104.

Cheung A, Dufour S, Jones G, Kostoulas P, Stevenson MA, Singanallur NB, Firestone SM (2021) Bayesian latent class analysis when the reference test is imperfect. *Revue scientifique et technique (International Office of Epizootics)* 40, 271-286.

Clouthier SC, McClure C, Schroeder T, Anderson ED (2021) Bayesian latent class model estimates of diagnostic accuracy for three test methods designed to detect spring viremia of carp virus. *Preventive Veterinary Medicine* 190, 105338.

Colgan DJ, da Costa P (2013) Invasive and non-invasive lineages in *Xenostrobus* (Bivalvia: Mytilidae). *Molluscan Research* 33, 272-280.

Commonwealth of Australia (2023) Carpet sea squirt (*Didemnum vexillum*). Accessed: 6 July 2023. URL: https://www.outbreak.gov.au/current-outbreaks/carpet-sea-squirt

Damisa S, Musa F, Sani S (2017) On the Comparison of Some Link Functions In Binary Response Analysis. *Biomedical Statistics and Informatics* 2, 145-149.

Dann A, Ellard K, Grove SJ, Willan RC (2020) Genetic confirmation of *Mya japonica* Jay, 1857 (Bivalvia: Myidae) in Tasmania, Australia: first record of any species of *Mya* in the southern hemisphere. *BioInvasions Records* 9, 103-108.

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, Jerde CL, Sepulveda AJ (2020) What do you mean by false positive. *Environmental DNA* 3, 879-883.

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

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

DAWE (2021) Asian Shore Crab | Marine Pests Accessed: 27 May 2021. URL: https://www.marinepests.gov.au/pests/identify/asian-shore-crab

de Souza LS, Godwin JC, Renshaw MA, Larson E (2016) Environmental DNA (eDNA) detection probability is influenced by seasonal activity of organisms. *PLoS One* 11, e0165273.

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. *Biology Letters* 10.

Deveney MR, Giblot-Ducray D, Wiltshire KH (2017) Marine pests testing methods: contributing to establishment of an Australian Testing Centre for Marine Pests. South Australian Research and Development Institute (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.

Fox J (2015) Applied regression analysis and generalized linear models. (Sage Publications)

Fuglstad G-A, Simpson D, Lindgren F, Rue H (2019) Constructing priors that penalize the complexity of Gaussian random fields. *Journal of the American Statistical Association* 114, 445-452.

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.

Giblot-Ducray D, Herdina, Ellard K, Deveney MR, Wiltshire KH (2022) Development of TaqMan MGB quantitative PCR assays for the invasive species *Maoricolpus roseus* (Quoy & Gaimard, 1834) and *Mya japonica* (Jay, 1857). Report to the Department of Agriculture, Fisheries and Forestry. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2022/000222-1. SARDI Research Report Series No. 1149., Adelaide.

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.

Gómez-Rubio V (2020) Bayesian inference with INLA. (CRC Press)

Guo Z, Wang Z, Hou X (2021) Comparative analysis of the nrDNA repeat unit of Manila clam *Ruditapes philippinarum* and Quahog *Mercenaria mercenaria*. *Fishes* 6, 42-50.

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.

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.

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, Campbell ML, Schaffelke B (2007) Introductions of seaweeds: accidental transfer pathways and mechanisms. *Botanica Marina* 50, 326-337.

Hewitt MJ, Hourston M, McDonald JI (2018) A long way from home: Biosecurity lessons learnt from the impact of La Nina on the transportation and establishment of tropical portunid species. *PLoS One* 13, e0202766.

Hijmans RJ (2021) raster: Geographic Data Analysis and Modeling. R package version 3.4-13. https://CRAN.R-project.org/package=raster.

Hinlo R, Furlan E, Suitor L, Gleeson D (2017) Environmental DNA monitoring and management of invasive fish: comparison of eDNA and fyke netting. *Management of Biological Invasions* 8, 89-100.

Hourston M, McDonald J, Hewitt M (2015) Public engagement for the detection of the introduced marine species *Charybdis japonica* in Western Australia. *Management of Biological Invasions* 6, 243-252.

Izquierdo F, Menezes R, Wise L, Teles-Machado A, Garrido S (2022) Bayesian spatio-temporal CPUE standardization: Case study of European sardine (*Sardina pilchardus*) along the western coast of Portugal. *Fisheries Management and Ecology* 29, 670-680.

Johnson WO, Jones G, Gardner IA (2019) Gold standards are out and Bayes is in: Implementing the cure for imperfect reference tests in diagnostic accuracy studies. *Preventive Veterinary Medicine* 167, 113-127.

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.

Kean JM, Burnip GM, Pathan A (2015) Detection survey design for decision making during biosecurity incursions. In 'Biosecurity surveillance: quantitative approaches'. (Eds F Jarrad, S Low Choy and K Mengersen) pp. 238-252. (CAB International: Oxfordshire, England)

Knudsen SW, Møller PR (2020) Development of species-specific eDNA-based test systems for monitoring of non-indigenous Decapoda in Danish marine waters. NIVA-rapport 8257772798.

Kralik P, Ricchi M (2017) A basic guide to real time PCR in microbial diagnostics: definitions, parameters, and everything. *Fronteirs in Microbiology* 8, 1-9.

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

Lance RF, Guan X (2020) Variation in inhibitor effects on qPCR assays and implications for eDNA surveys. *Canadian Journal of Fisheries and Aquatic Sciences* 77, 23-33.

Layton KK, Martel AL, Hebert PD (2014) Patterns of DNA barcode variation in Canadian marine molluscs. *PLoS One* 9, e95003.

Lefever S, Pattyn F, Hellemans J, Vandesompele J (2013) Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clinical Chemistry* 59, 1470-80.

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.

Lindgren F, Rue H (2015) Bayesian spatial modelling with R-INLA. *Journal of Statistical Software* 63, 1-25.

Lindgren F, Rue H, Lindström J (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73, 423-498.

Low-Choy S (2013) Bayesian modelling of surveillance and proof of freedom: the mathematical, logical & psychological challenges.

Mansournia MA, Heinze G, Geroldinger A, Greenland S (2017) Separation in logistic regression: causes, consequences, and control. *American Journal of Epidemiology* 187, 864-870.

Martins TG, Simpson D, Lindgren F, Rue H (2013) Bayesian computing with INLA: new features. *Computational Statistics & Data Analysis* 67, 68-83.

Mazón-Suástegui JM, Fernández NT, Valencia IL, Cruz-Hernández P, Latisnere-Barragán H (2016) 28S rDNA as an alternative marker for commercially important oyster identification. *Food Control* 66, 205-214.

McDonald J (2012) Detection of the tropical mussel species *Perna viridis* in temperate Western Australia: possible association between spawning and a marine heat pulse. *Aquatic Invasions* 7, 483-490.

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)

Mölder F, Jablonski K, Letcher B, Hall M, Tomkins-Tinch C, Sochat V, Forster J, Lee S, Twardziok S, Kanitz A, Wilm A, Holtgrewe M, Rahmann S, Nahnsen S, Köster J (2021) Sustainable data analysis with Snakemake. *F1000Research* 10, 33.

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.

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.

Paradinas I, Conesa D, Pennino MG, Muñoz F, Fernández AM, López-Quílez A, Bellido JM (2015) Bayesian spatio-temporal approach to identifying fish nurseries by validating persistence areas. *Marine Ecology Progress Series* 528, 245-255.

Parry GD, Currie DR, Crookes DP (1997) Exotic marine pests in Portland Harbour and environs. Marine and Freshwater Resources Institute, Queenscliff.

Plummer M (2017) JAGS Version 4.3.0 user manual.

Porter TM, Hajibabaei M (2020) Putting COI metabarcoding in context: the utility of exact sequence variants (ESVs) in biodiversity analysis. *Frontiers in Ecology and Evolution* 8, 1-15.

Quiroz ZC, Prates MO, Rue H (2015) A Bayesian approach to estimate the biomass of anchovies off the coast of Peru. *Biometrics* 71, 208-217.

R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rahman AKMA, Smit S, Devleesschauwer B, Kostoulas P, Abatih E, Saegerman C, Shamsuddin M, Berkvens D, Dhand NK, Ward MP (2019) Bayesian evaluation of three serological tests for the diagnosis of bovine brucellosis in Bangladesh. *Epidemiology and Infection* 147, e73.

Royle JA, Dorazio RM (2009) Occupancy and abundance. In 'Hierarchical modeling and inference in ecology'. (Eds JA Royle and RM Dorazio) pp. 127-157. (Academic Press: San Diego)

Royle JA, Link WA (2006) Generalized site occupancy models allowing for false positive and false negative errors. *Ecology* 87, 835-841.

Royle JA, Nichols JD (2003) Estimating abundance from repeated presence–absence data or point counts. *Ecology* 84, 777-790.

Rue H, Held L (2005) Gaussian Markov random fields: theory and applications. (CRC press)
Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71, 319-392.
Rue H, Riebler A, Sørbye SH, Illian JB, Simpson DP, Lindgren FK (2017) Bayesian computing with INLA: A review. *Annual Review of Statistics and Its Application* 4, 395-421.
Salk JJ, Schmitt MW, Loeb LA (2018) Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nature Reviews Genetics* 19, 269-285.
Schaffelke B, Hewitt CL (2007) Impacts of introduced seaweeds. *Botanica Marina* 50, 397-417.
Sepulveda AJ, Nelson NM, Jerde CL, Luikart G (2020) Are environmental DNA methods ready for aquatic invasive species management? *Trends in Ecology and Evolution* 35, 668-678.
Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH (2017) Penalising model component complexity: a principled, practical approach to constructing priors. *Statistical Science* 32.
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.
So KYK, Fong JJ, Lam IPY, Dudgeon D (2020) Pitfalls during *in silico* prediction of primer specificity for eDNA surveillance. *Ecosphere* 11.
Stafford H, Willan RC, Neil KM (2007) The invasive Asian Green Mussel, *Perna viridis* (Linnaeus, 1758)(Bivalvia: Mytilidae), breeds in Trinity Inlet, tropical northern Australia. *Molluscan Research* 27, 105-109.
Stanaway M (2015) Evidence of absence for invasive species: roles for hierarchical Bayesian approaches in regulation. In 'Biosecurity Surveillance: Quantitative Approaches'. (Eds F Jarrad, S Low Choy and K Mengersen) pp. 265. (CAB International: Oxfordshire, England)
Stoeckle MY, Soboleva L, Charlop-Powers Z (2017) Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PLoS One* 12, e0175186.
Summerson R, Zhao S, Arthur T, Robinson A, McCrudden R (2016) Ballast Water Risk Assessment. Exploring new methods for estimating risk: Using satellite sea surface temperature data; incorporating vessel voyage data. CEBRA, Melbourne.
Sun Se, Li Q, Kong L, Yu H, Zheng X, Yu R, Dai L, Sun Y, Chen J, Liu J, Ni L, Feng Y, Yu Z, Zou S, Lin J (2016) DNA barcoding reveal patterns of species diversity among northwestern Pacific molluscs. *Scientific Reports* 6, 1-17.
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.
Vilela R, Burger C, Diederichs A, Bachl FE, Szostek L, Freund A, Braasch A, Bellebaum J, Beckers B, Piper W, Nehls G (2021) Use of an INLA latent gaussian modeling approach to assess bird population changes due to the development of offshore wind farms. *Frontiers in Marine Science* 8.
Wang C, Lin X, Nelson KP (2020) Bayesian hierarchical latent class models for estimating diagnostic accuracy. *Statistical Methods in Medical Research* 29, 1112-1128.

Wells FE (2017) If the Asian green mussel, *Perna viridis* (Linnaeus, 1758), poses the greatest invasive marine species threat to Australia, why has it not invaded? *Molluscan Research* 37, 167-174.

Werner G, Heislers S, White C, Blake S, Hirst A (2012) Portland Harbour marine pest surveys - Final Report. Department of Primary Industries, Fisheries Victoria Technical Report Series No. 166, Queenscliff, Victoria.

Wilcox TM, McKelvey KS, Young MK, Jane SF, Lowe WH, Whiteley AR, Schwartz MK (2013) Robust detection of rare species using environmental DNA: the importance of primer specificity. *PLoS One* 8, e59520.

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), SARDI Publication No. F2021/000152-1. SARDI Research Report Series No. 1095., Adelaide.

Wiltshire KH (2023) Statistical modelling approaches for the interpretation of molecular surveillance data. Report to the Department of Agriculture, Fisheries and Forestry. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2023/000144-1. SARDI Research Report Series No. 1177., Adelaide.

Wiltshire KH, Giblot-Ducray D, Deveney MR (2017) Applying molecular techniques to surveys for introduced marine pests at ports in Spencer Gulf, South Australia. Report for PIRSA Biosecurity SA. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2017/000202-1. SARDI Research Report Series No. 947, Adelaide.

Wiltshire KH, Giblot-Ducray D, Deveney MR (2021a) Design of a TaqMan quantitative PCR assay for the invasive crab, *Hemigrapsus sanguineus* (DeHaan, 1853). Report to the Department of Agriculture, Water and the Environment. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2021/000236-1. SARDI Research Report Series No. 1102, Adelaide.

Wiltshire KH, Giblot-Ducray D, Deveney MR (2021b) Development and validation of a quantitative PCR assay for the invasive mussel, *Mytella strigata* (Hanley, 1843). Report to the Department of Agriculture, Water and the Environment. South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2021/000305-1. SARDI Research Report Series No. 1107, Adelaide.

Wiltshire KH, Giblot-Ducray D, Deveney MR (2022) Molecular surveys for priority marine pests in 2021-22 for Thevenard, Port Lincoln, Port Giles, Klein Point, Port Adelaide, South Australia, and Portland, Victoria. Report to the Department of Agriculture, Fisheries and Forestry South Australian Research and Development Institute (Aquatic Sciences), SARDI Publication No. F2022/000236-1. SARDI Research Report Series No. 1147., Adelaide.

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), SARDI Publication No. F2019/000040-1. SARDI Research Report Series No. 1019., Adelaide.

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), SARDI Publication No. F2019/000337-1. SARDI Research Report Series No. 1036., Adelaide.

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), SARDI Publication No. F2019/000406-1. SARDI Research Report Series No. 1044., Adelaide.

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), SARDI Publication No. F2020/000289-1. SARDI Research Report Series No. 1068., Adelaide.

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.

Wright ES, Vetsigian KH (2016) Quality filtering of Illumina index reads mitigates sample cross-talk. *BMC Genomics* 17, 876.

Wright ES, Yilmaz LS, Ram S, Gasser JM, Harrington GW, Noguera DR (2014) Exploiting extension bias in polymerase chain reaction to improve primer specificity in ensembles of nearly identical DNA templates. *Environmental Microbiology* 16, 1354-65.

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, 1-17.

Zhao S, Caley P, Arthur T, Robinson A, Gregg M (2012) Ballast water risk table reconstruction. ACERA Project 1104E Report 1. Australian Centre of Excellence for Risk Analysis, Melbourne.

Zuur AF, Ieno EN (2018) Spatial, temporal and spatial-temporal ecological data analysis with R-INLA. Volume II: GAM and zero-inflated models. (Highland Statistics Ltd: Newburgh, UK)

Zuur AF, Ieno EN, Saveliev AA (2017) Spatial, temporal and spatial-temporal ecological data analysis with R-INLA. Volume I: Using GLM and GLMM. (Highland Statistics Ltd: Newburgh, UK)