FRDC
FISHERIES RESEARCH & DEVELOPMENT CORPORATION

Developing a rapid molecular identification technique to improve egg production based fish biomass assessments

Richard J. Saunders, Shannon Kjeldsen, Roger Huerlimann, Thor Saunders, Shane Penny, Andrew Tobin and Dean Jerry

2019

FRDC Project No **2014/022**

ISBN **978-0-6485037-0-5**

Developing a rapid molecular identification technique to improve egg production based fish biomass assessments

FRDC Project No 2014/022

2019

Researcher Contact Details

Name: Richard Saunders
Address: 1 James Cook Dr
Douglas QLD 4814
Phone: +64 2 6423965
Fax: n/a
Email: richard.saunders@jcu.edu.au

FRDC Contact Details

Address: 25 Geils Court
Deakin ACT 2600
Phone: 02 6285 0400
Fax: 02 6285 0499
Email: frdc@frdc.com.au
Web: www.frdc.com.au

Contents

Tables

Figures

Acknowledgments

The authors would like to thank the crew of the JCU Research Vessel Kirby and JCU technical staff who assisted during the project Diana Pazmino-Jaramillo, Steve Moore and Heather Welladsen. Further, we would like to thank NT Fisheries Technical Staff Quentin Allsop, Wayne Baldwin and Nathan Crofts. Finally, we acknowledge the efforts of Jose Domingos and Julie Goldsbury who worked on an early version of the methods developed in this project.

Abbreviations

16S - 16S ribosomal RNA

BP – base pairs

CO1 / COI - Cytochrome Oxidase Subunit 1

CytB – Cytochrome B

DEPM – Daily Egg Production Method

DNA – Deoxyribonucleic acid

EtOH – Ethanol

GBR – Great Barrier Reef

JCU – James Cook University

NT – Northern Territory

NTC – non-template control

PCR – Polymerase chain reaction

PCR-AFLP – Polymerase chain reaction - Amplified Fragment Length Polymorphism

PCR-DGGE – Polymerase chain reaction - Denaturing Gradient Gel Electrophoresis

PCR-RAPD – Polymerase chain reaction - Random Amplified Polymorphic DNA

PCR-RFLP – Polymerase chain reaction - Restriction Fragment Length Polymorphisms

PCR-SSCP – Polymerase chain reaction - Single Strand Conformation Polymorphism

QDAF – Queensland Department of Agriculture and Fisheries

Qld – Queensland

RT – Room temperature

SA – South Australia

TE – Tris EDTA

TMAC - Tetramethylammonium chloride

Executive Summary

This project was a collaboration between James Cook University, Queensland Department for Agriculture and Fisheries and Northern Territory Department of Primary Industry and Resources with the aim of developing a molecular fish egg identification method and to assess its value for the application of the Daily Egg Production Method (DEPM) of biomass assessment.

Distribution and abundance of fish eggs as measured from plankton sampling can be a valuable tool to determine the distribution of spawning and for biomass assessment. Traditionally determination of distribution and abundance relies on morphological identification of eggs; however, this is very challenging because comparatively few fish species have had their eggs described and there are few distinguishing characteristics, particularly in the early stages of development. This has stymied the wider application of one fishery independent assessment method – the DEPM.

This project aimed to develop a rapid and affordable molecular method to identify fish eggs from plankton samples. The method selected was a multiplex bead array method where species-specific probes are developed and bound to beads which fluoresce when passed through a flow cytometer if bound to the target species DNA. This allows for identification of multiple species at one time as different probe-bead combinations can be used to identify different fish species. Furthermore, as molecular methods usually require preservation in ethanol which causes problems of egg staging, a critical component in modelling egg mortality and thus initial egg abundance, we also considered the impact of preservation method on egg staging and on DNA amplification.

Initially, species-specific probes were developed for a range of northern Australian fish species and tested against identified tissue for each of the species. Following this the bead array was developed and tested against a small number of wild collected fish eggs. Finally, the developed bead array method was applied to identify eggs of several species in a larger scale egg survey to determine its value as a method, particularly in the context of DEPM egg surveys. In addition, we also conducted an experiment investigating the impact of egg preservation on DNA amplification and, under controlled conditions, on the ability to stage eggs.

The method was able to identify eggs of several target species from wild collected plankton samples (confirmed by sequencing) with success rates varying between species from 50 to 100% identification success. We identified a series of issues that potentially limit the application of this method in the context of DEPM and egg identification. The principal issues were reliability of the chemistry and identified false negative results resultant from preservation of DNA. DNA amplification and egg staging were both impacted by preservation method with the best preservation method for egg staging being 5% formalin, however, this was the worst performer for DNA amplification. An ethanol based preservation method is essential for DNA based identification and this also allows for some egg staging, although egg staging prior to the presence of an embryo is very challenging, even for an experienced technician.

In the context of biomass assessment we caution against the use of this method as a principal egg identification method without addressing these issues. It does have value as a rapid and affordable method to confirm morphological identifications without the need to out-source sequencing.

In the context of DEPM, molecular methods for fish egg identification are important for verification of morphological identification methods but are imperfect as the sole egg identification method. The principal issues associated with molecular methods for identification of fish eggs are reliable preservation which maintains egg structures in order to age (stage) eggs at all developmental stages and determine the proportion of probable false negatives (where an egg of a target species is not able to be identified). The latter can be due to DNA amplification failure and/or failure of species-specific probes to bind to target DNA.

Keywords

DEPM, fish egg identification, Great Barrier Reef, plankton, ichthyoplankton, DNA, probes

1 Introduction

1.1 Why is fish egg identification important?

Ichthyoplankton surveys are a useful tool for the study of fish biology, fishery assessment and consequently fishery management. They can be used to define spawning areas (e.g. Aoyama et al. 2001), study recruitment processes (e.g. Murphy et al. 2012) and, with a good time series, track long term changes in fish populations (e.g. Ward et al. 2011). In particular, they provide essential data for egg production based biomass assessment of fishes (Lasker 1985). The most commonly used form of which is the Daily Egg Production Method (DEPM) which has now been applied to at least 18 species (Ward et al. 2011). DEPM is regularly used to assess spawning biomass of Australian Sardines (*Sardinops sagax*) in Australia (Ward et al. 2017) and the United States (Hill et al. 2018) and was central to management of the Northern Anchovy (*Engraulis mordax*) in the north Pacific (Stroatoudakis et al. 2006). It has been applied to numerous species, particularly small pelagic fishes (e.g. Australian Anchovy, *Engraulis australis*, Dimmlich et al. 2009; red bait, *Emmelichthys nitidis*, Neira & Lyle 2011), but also large benthic fish (e.g. Snapper, *Chrysophrys auratus,* Zeldis 1998; Jackson & Cheng 2001). DEPM is a particularly useful tool because it is a fishery independent assessment method, which avoids many of the problems associated with assessment models that rely on fishery dependent datasets. However, DEPM is not a panacea for fishery assessment because it relies on a detailed knowledge of spawning behaviour and output of a population, the distribution of spawning fish, the mortality of fish eggs, and crucially the distribution and abundance of fish eggs (Stratoudakis et al. 2006; Ward et al. 2011). The broader application of the DEPM for biomass assessment has been limited because of the challenges (financial and practical) of collecting these datasets. Fundamental to the success of the DEPM is the egg survey itself, which is dependent on accurate fish egg identification. Accurate fish egg identification, in particular, has stymied the wide application of the DEPM with particular problems identified in its application for Snapper (Oxley et al. 2017), *Trachurus* spp. (Neira et al. 2015) and Atlantic Cod, *Gadus morhua* (Fox et al. 2005).

1.2 Fish egg identification how and why is it difficult?

The focus of this project was on fish egg identification collected from ichthyoplankton samples. Fish egg identification has traditionally been done on morphological features of eggs. This approach can be straightforward where some species have eggs with easily visible and unique features, as is the case for Australian Sardine (*Sardinops sagax*) and Australian Anchovy (*Engraulis australis*) eggs in southern Australia. It can be much more challenging, especially in early developmental stages, where eggs can be indistinguishable between families, as is the case with Snapper and members of the Platycephalidae (Oxley et al. 2017). It is even more likely to be problematic with closely related species, as is the case amongst the *Trachurus* spp. (Neira et al. 2015). In some cases, when morphological identification approaches have been used, egg mis-identifications have occurred which can have serious implications for fishery assessment (e.g. Fox et al. 2005). Fish egg

identification is further stymied by the overarching lack of species-specific egg descriptions, thus there usually remains the possibility that species with undescribed eggs could be misidentified as eggs of a species with an available description. This problem escalates in regions where the fish fauna is particularly diverse. Thus, there is usually some uncertainty when using morphological methods to identify fish eggs, particularly in early developmental stages.

1.3 Molecular fish egg identification

Molecular methods are routinely applied to determine the identity of an unknown species, making use of a specific, informative gene region unique to species that can be assessed by a variety of methods. Most methods depend on a polymerase chain reaction (PCR) amplification of the target sequence, followed by a method that detects species-specific sequence variations. These methods include Restriction Fragment Length Polymorphisms (PCR-RFLP), Amplified Fragment Length Polymorphism (PCR-AFLP), Single Strand Conformation Polymorphism (PCR-SSCP), Random Amplified Polymorphic DNA (PCR-RAPD), Denaturing Gradient Gel Electrophoresis (PCR-DGGE), PCR using specific primers, PCR followed by sequencing, real-time PCR, or Microarrays (Teletcha 2009). More recently, next-generation sequencing technologies have added additional tools to the mix, including targeted amplicon sequencing. Each of these methods has a different combination of DNA quality requirements, development time, level of reproducibility, cost and speed of analysis (Teletcha et al., 2009). For example, PCR followed by sequencing produces a quick, reproducible, and relatively reliable species identification, at relatively high cost, but requires prior knowledge of the target sequence for identification.

Molecular identification of fish eggs presents some unique challenges. Specifically, the amount of DNA in a fish egg varies during development; indeed at the earliest developmental stage there is only a single cell. A further challenge is that fish eggs have a “shell”, known as the chorion, which has evolved as a protective barrier that physically shields the developing embryo, hardening immediately after fertilisation and allowing for transfer of some small molecules (Potts & Eddy 1973; Berios et al. 2011). This barrier must be broken for primers or probes to access the DNA of the egg.

A variety of approaches for molecular fish egg identification have been applied. These include sequencing methods (Dias et al. 2016), development of species-specific probes (Gleason & Burton 2012), or recently the application of *in-situ* hybridisation (see Oxley et al. 2017). Importantly, the approach used for identification of fish eggs depends on the needs of the study. In some cases, morphological methods may be sufficient to identify eggs to genus and molecular methods may only be needed to distinguish between two species (e.g., Neira et al. 2015). In other cases, it may be that morphological methods are insufficient to separate eggs at a much higher taxonomic level and thus eggs need to be screened to exclude a wider range of species (e.g., Oxley et al. 2017).

The feasibility of any molecular method to identify fish eggs is dependent on the preservation of DNA of the fish egg. With the exception of *in-situ* hybridization, most molecular methods used to identify fish eggs are primarily based on PCR amplification, followed by a specific method to detect the target species (Figure 1 and Teletcha 2009), and therefore PCR amplification success plays a

crucial part of detection rates. PCR failure can be attributed to poor preservation (DNA degradation), low DNA input (early stage eggs with limited cell numbers), failure in DNA extraction (small DNA input, dechorionation failure), presence of PCR inhibitors (e.g. algal metabolites), or primer mismatches ("universal primers" are not necessarily universal).

Preservation can be problematic because ichthyoplankton samples when collected at sea do not contain only fish eggs but a wide range of organisms that when preserved, usually in ethanol, react in different ways diluting the preservative and its efficacy. This can manifest as problem of false negatives when using molecular methods for egg identification.

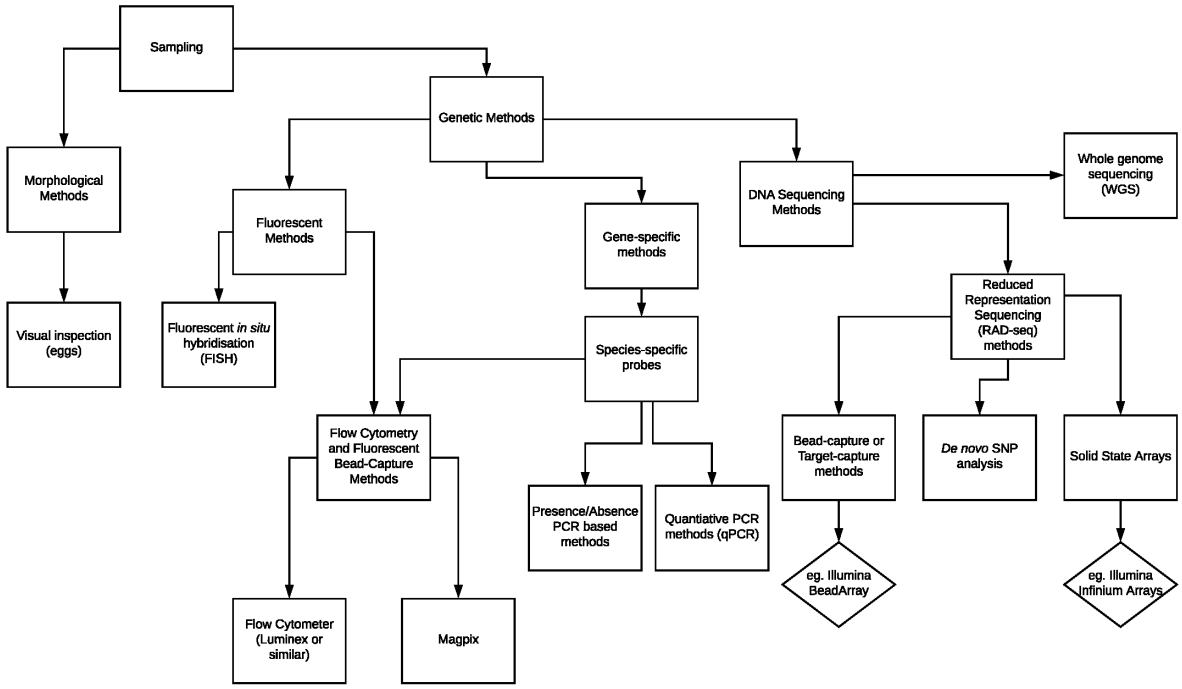


Figure 1. Flow chart of molecular egg identification approaches.

1.4 Egg staging (and estimating egg mortality)

The ability to stage (age) eggs is essential for developing egg mortality estimates that are critical to the success of a DEPM (Ward et al. 2011; 2018). As most successful DEPM surveys are based on formalin preserved eggs, egg staging is usually straightforward. Formalin is not a suitable preservative for use when using molecular methods to identify fish eggs as it renders DNA unsuitable for PCR amplification (although it remains theoretically possible to extract DNA). Thus, when using

molecular methods for identification, the most common approach has been to preserve plankton samples (fish eggs) in ethanol (70-100%), but this preservative impacts the morphology and renders the eggs relatively opaque making egg staging very difficult. Thus, when molecular methods are proposed to identify fish eggs for a DEPM survey an obstacle has been how to stage ethanol preserved eggs and thus estimate egg mortality. This directly informs the feasibility of the egg identification method for DEPM of biomass assessment, as failure of PCR means eggs cannot be identified and this would have a concomitant reduction in any biomass estimate, or at least increasing uncertainty around the estimates. Furthermore, the ability to stage eggs means egg mortality can be better estimated which is a key factor in the DEPM model.

Thus, to be suitable for DEPM any molecular fish egg identification method must be first based on good preservation of DNA, while maintaining the capacity to stage eggs, and finally, be able to be applied at a large scale and be cost effective.

1.5 Suspension bead arrays

Suspension bead arrays were designed initially for medical applications to rapidly screen for multiple genetic disorders. Likewise, a multiplex suspension bead array coupled with high-throughput flow cytometry offers a method to identify several species in one assay. The method involves developing species-specific probes that are attached to fluorescent beads. These fluoresce if the DNA of the target species has been bound to the bead via the probe and this fluorescence can be recognised by a flow cytometer. Multiplexing is achieved by using different coloured beads designed to target different species in a single assay. Systems are now capable of multiplexing hundreds of probes allowing rapid identification of many species rapidly (Gleason and Burton 2012).

This project involved binding commercially available fluorescent beads to genes of target fish species of interest. Fish DNA from eggs is mixed with the beads and if the DNA from the targeted fish species is present then it will hybridize to species-specific probes on the bead that is labelled with a fluorescent dye. Beads are then run through a flow cytometer, or equivalent method, and if fish DNA from the target species has hybridized to the bead a fluorescent signal is emitted and detected by the flow cytometer. This method allows both the qualitative identification of which fish species eggs are present and in what quantitative proportion they represent within the plankton tows.

1.6 Project progression

The initial stage of the project was the development of species-specific probes for the target species followed by testing of these probes in a multiplex bead array coupled with high throughput flow cytometry to identify wild collect fish eggs for a sub-set of eggs. Following the initial development of the techniques, we then proceeded to assess the suitability of the technology for DEPM. For this, a larger egg sampling regime in the Northern Territory and Queensland was implemented in an attempt to collect a significant number of eggs of the target species against which to test the egg identification methods. This was done in conjunction with a set of experiments investigating the effects of different preservation methods on egg identification success and egg staging.

Objectives

1. To develop a novel high-throughput, low cost DNA-based egg identification method for important fish species in northern Australia.

2. To assess the application of the technology developed for use in the daily egg production method (DEPM) for biomass estimation.

2 Methods

2.1 Target species

This project aimed to develop a bead array identification method targeting five important fishery species in northern Australia, specifically Goldstripe Sardinella (*Sardinella gibbosa*), Spotted Sardine (*Amblygaster sirm*), Black jewfish (*Protonibea diacanthus*), Spanish Mackerel (*Scomberomorus commerson*) and Grey Mackerel (*Scomberomorus semifasciatus*) (Table 1). In addition to this, because of the availability of target species eggs, we also included Mackerel tuna (*Euthynnus affinis*), Large Scale Grunter (*Terapon theraps*) and Eightband butterflyfish (*Chaetodon octofasciatus*) to better assess the practical application of the bead array methods for other species (Table 1).

Table 1 List of species for which species-specific probes were developed in the project and the rationale for their inclusion.

Scientific name	**Common name**	**Family**	**Australian distribution**	**Rationale**
Sardinella gibbosa	Goldstripe Sardinella	Clupeidae	Northern Australia - Qld, WA and NT	Targeted in NT small pelagic fishery
Amblygaster sirm	Spotted Sardine	Clupeidae	Northern Australia - Qld, WA and NT	Targeted NT small pelagic fishery
Protonibea diacanthus	Black Jewfish	Sciaenidae	Northern Australia - Qld, WA and NT	Requested inclusion by NT FRAB. Principally targeted in recreational fisheries in northern Australia (Qld, WA and NT)
Scomberomorus commerson	Spanish Mackerel	Scombridae	Northern Australia – NSW, Qld, WA and NT	Practical test species in need of alternative assessment methods.
Scomberomorus semifasciatus	Grey Mackerel	Scombridae	Northern Australia - Qld, WA and NT	Practical test species in need of alternative assessment methods.
Euthynnus affinis	Mackerel Tuna	Scombridae	Northern Australia – NSW, Qld, WA and NT	Opportunistic based on egg sample collection
Terapon theraps	Large Scale Grunter	Terapontidae	Northern Australia – Qld, NT and WA	Opportunistic based on egg sample collection
Chaetodon octofasciatus	Eightband	Chaetodontidae	Northern	Opportunistic based on egg

	Butterflyfish	Australia – Qld, NT and WA	sample collection

2.2 Ichthyoplankton sampling and sorting

2.2.1 Field sampling (2015)

In October 2015, an egg survey was done specifically targeting areas of the Great Barrier Reef immediately off Townsville where Spanish Mackerel have historically been targeted by commercial fishers and are known to spawn. The reefs targeted were John Brewer and Helix Reefs (Figure 2). In total, 24 plankton samples were collect by vertical tows using a twin-ring plankton net (30 cm diameter with 5:1 ratio and 300 µm mesh) (Figure 2). In addition, 24 hour plankton sampling was done by anchoring at a site (north-west corner of John Brewer Reef; Figure 2) that was identified as important for mackerel in a previous study (Tobin et al. 2015; FRDC Project No 2010/007). The site is also a known area for commercial fishers to target spawning Spanish Mackerel. Samples were passively collected by streaming a neuston net (50 x 100 x 300 cm with 300 µm mesh) in the current for 10 min from the stern of the anchored research vessel and sampling approximately every hour for 24 hours. This was done to maximise the opportunity to collect eggs of Spanish Mackerel.

All the plankton samples from the 2015 field sampling were sorted in the laboratory under a Nikon SMZ-18 dissecting microscope. Here a subset of possible eggs of Spanish Mackerel, Grey Mackerel, Gold stripe sardinellaand Spotted Sardine were selected (n=10) for sequencing and to be used to test the bead array. Eggs were selected based principally on egg size, as all these species have large eggs between 1.0 and 1.5mm in diameter. The 10 eggs were transferred to 1.5 mL microfuge tubes filled with ethanol and stored at room temperature until DNAextraction.

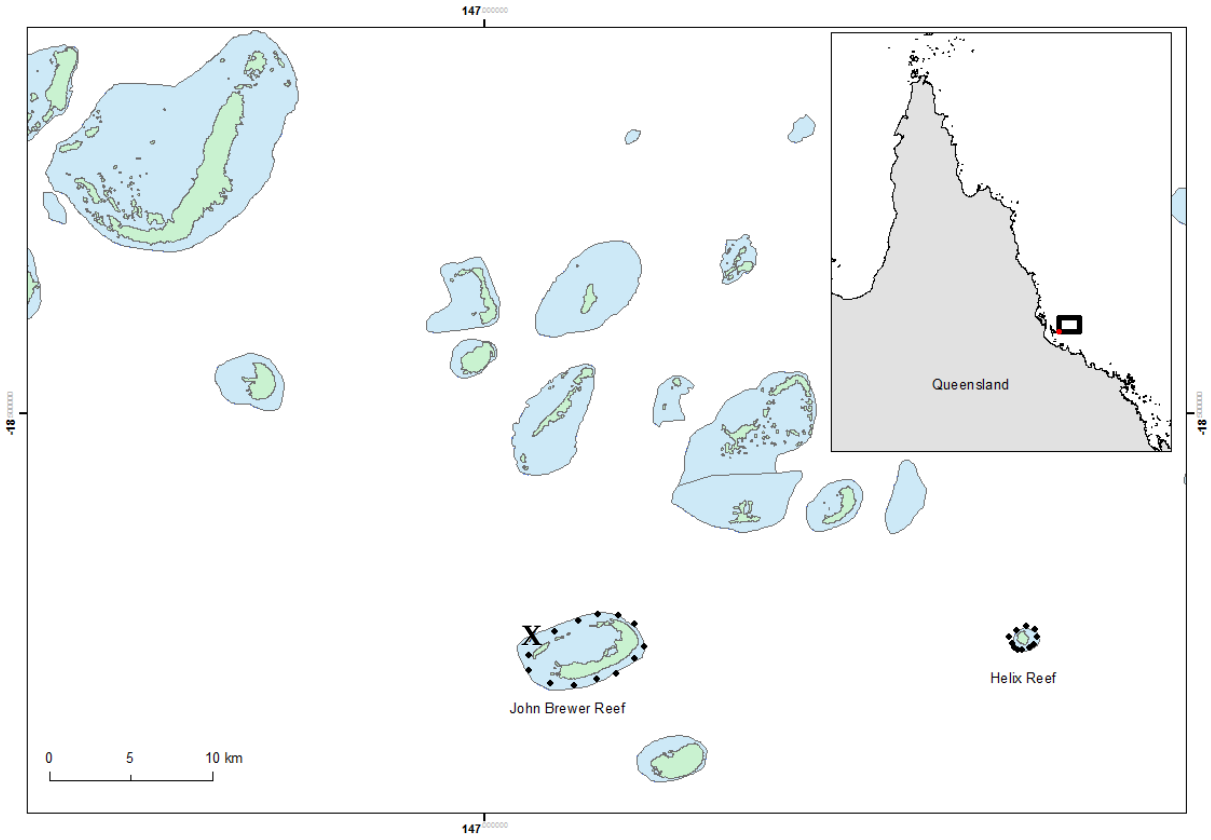


Figure 2. Map of Qld illustrating study region highlighted by a black square. Zoomed study area showing sites of plankton samples collected in October 2015 research survey. Black dots are twin-ring vertical tows, **x** is 24 hour neuston sampling site.

2.2.2 Field sampling (2016)

Great Barrier Reef

In Spring of 2016, plankton samples were collected from three reefs of the Great Barrier Reef off the Townsville Coast (Figure 3). In total 53 plankton samples were collected by vertical tow using a twin-ring plankton net (30 cm diameter with 5:1 ratio and 300 µm mesh). Spatial coverage was substantially increased over the 2015 sampling program to increase the probability of collecting eggs of the target species, particularly on the north-western corner of John Brewer Reef where spawning adults (females with hydrated oocytes and running males) were collected at the same time as the plankton samples. Samples were sorted in the same manner as above (2.2.1).

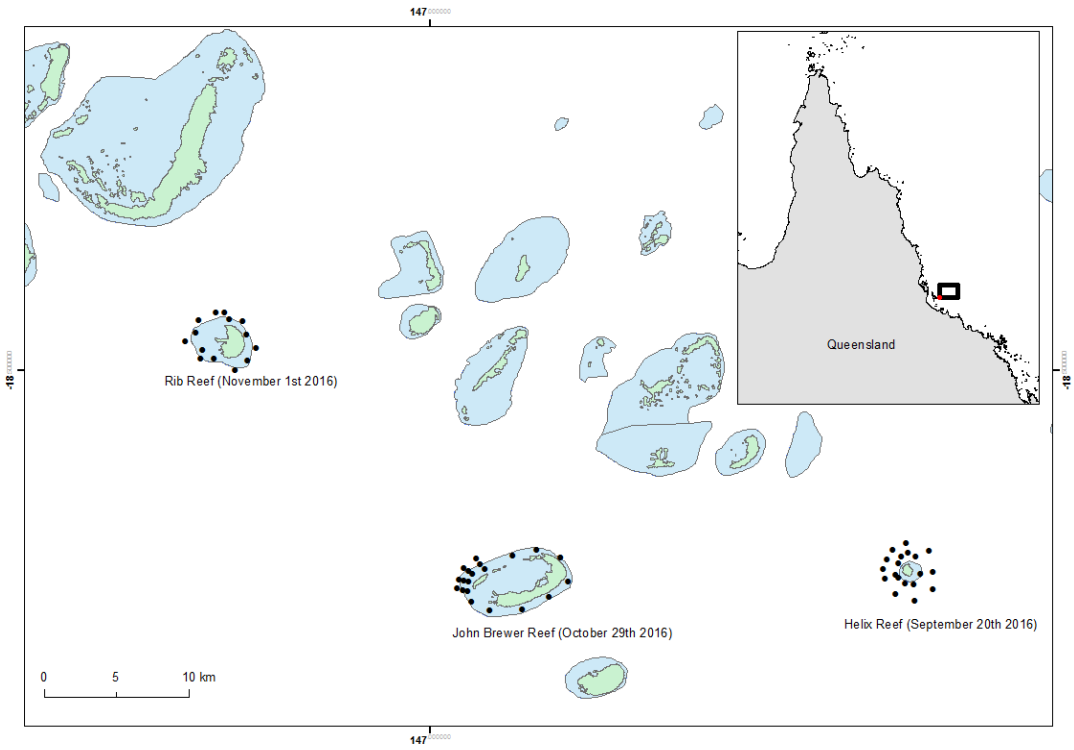


Figure 3. Map illustrating study region highlighted by a black square with zoomed study area showing sites of plankton samples collected in Spring 2016 plankton surveys. Black dots are the sites of the twin-ring vertical tows.

Northern Territory

In Spring 2016, eight ichthyoplankton samples were collected off the north-western coast of the Northern Territory (Figure 4). Samples were collected from the FRV Mallarra (NT Fisheries research vessel) using a twin ring 300mm x 1200mm x 300 µm nets towed horizontally for 10 min at 4 knots.

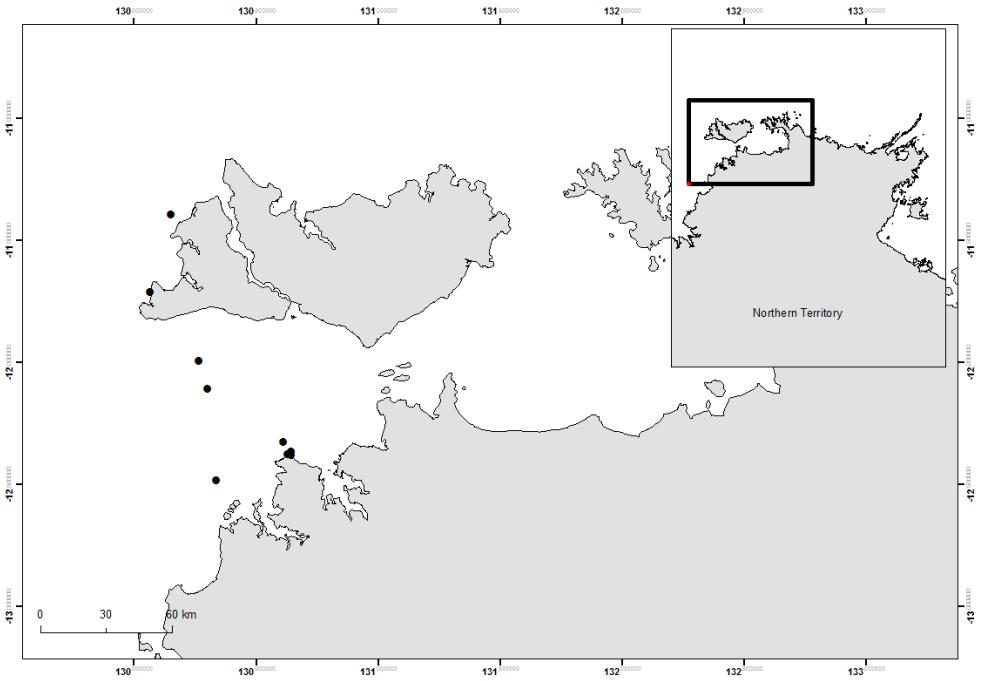


Figure 4. Map illustrating study region in the NT highlighted by a black square with zoomed study area showing sites of plankton samples collected (black dots).

All plankton samples for the 2016 field sampling were sorted under a Nikon SMZ-18 dissecting microscope in the laboratory and all fish eggs were removed and transferred to a 5 mL vial containing 100% EtOH. Subsequently, eggs were removed from these vials and digital images of each egg were taken, diameter was measured and other morphological characteristics, such as presence of an oil droplet and pigmentation, were noted. At that time, individual eggs were transferred to 96 well plate containing extraction buffer (Chemagic Arthropod DNA Extraction Kit LH, Chemagen) and stored at room temperature before DNA extraction.

2.3 Development of species-specific probes

Two mitochondrial genes, cytochrome c oxidase I (COI), and 16S rRNA, were chosen for probe design for all target species. Using two probes to identify each species increases the specificity of the assay. For initial testing, probes were designed for Spanish Mackerel, Grey Mackerel, Spotted Sardine, Goldstripe Sardinella and Black Jewfish. Probes were designed manually in Geneious (V8.1.9, http://www.geneious.com, Kearse et al., 2012) based on sequence alignments of the target species and multiple exclusion species, and probe specificity was verified using Primer BLAST through the National Centre for Biotechnology Information (NCBI) web site. Primer BLAST matches each primer against the NCBI database to find species that may also hybridize with the probe.

Another subset of species-specific oligonucleotide probes were also designed for Mackerel Tuna, Large Scale Grunter and Eightband Butterflyfish, because these species were found to be present in reasonable numbers in ichthyoplankton samples (based on genetic sequencing results; refer to Appendix 2). For these species, all publicly available sequences for COI and 16S on GenBank (NCBI) were collated and aligned using MUSCLE (Edgar, 2004) in Geneious (Kearse et al., 2012). Alignments

incorporated all target species (Table 1) to ensure probes were likely to exclude non-target species. All probes were between 15-30 bp long, ensuring each probe was placed within a conserved region on the target species gene of interest, and were a minimum of 2-3 bp different to all other non-target species (Table 2). All probes were rehydrated in molecular grade water.

Table 2. Probe sequences for each species. Note: Only a single gene could be targeted for some species due to the highly conserved nature of these genes across species.

Gene	Species	Probe ID	Probe	Included in final bead array 2016 samples
16S	Spanish Mackerel	16S-Scocom-1	/5mMC12/GTTTTAGGAGATAGTACTCCCGGTC	N
16S	Spanish Mackerel	16S-Scocom-2	/5mMC12/GCATGGTTCATTTGGTTTAGTCCTT	N
16S	Spanish Mackerel	16S-Scocom-3	/5mMC12/GGCAGCTCTAACGAACAGAATTTC	Y
COI	Spanish Mackerel	COI-Scocom-1	/5mMC12/GATTACCGGCAAGGGGTGGATAGAC	N
16S	Spanish Mackerel	16S-Scosem-1	/5mMC12/TTAGGAGGTCAAACTCCCTGTCCAC	N
16S	Grey Mackerel	16S-Scosem-2	/5mMC12/ACATGGGTCACTCGGTTTAGTCCTT	N
COI	Grey Mackerel	COI-Scosem-1	/5mMC12/TGAAGGGAAAAGATGGTTAGGTCAA	N
COI	Grey Mackerel	COI-Scosem-2	/5mMC12/GAATAAGTCAGTTTCCAAACCCTCC	N
16S	Goldstripe Sardinella	16S-Sargib-1	/5mMC12/TCGCGTTATTTGGTGGCCCGT	N
16S	Goldstripe Sardinella	16S-Sargib-2	/5mMC12/TGGCCCGTCGTCGGGGTGCTCTT	N
COI	Goldstripe Sardinella	COI-Sargib-1	/5mMC12/TGCGTGGGCTAAGTTGCCCGCTAGG	N
COI	Goldstripe Sardinella	COI-Sargib-2	/5mMC12/CGCCGTAACGATAACATTGTAGATT	N
16S	Spotted Sardine	16S-Ambsir-1	/5mMC12/TCCAGGCTTTATGTCCCAGTCCTCG	N
16S	Spotted Sardine	16S-Ambsir-2	/5mMC12/GGAGAAGTTCAGTTTGGCTGGTTTC	N
16S	Spotted Sardine	16S-Ambsir-3	/5mMC12/CAAACTGAACTTCTAACCGC	Y
COI	Spotted Sardine	COI-Ambsir-1	/5mMC12/ATCATTAGGGGTACCAGTCAGTTTC	N
COI	Spotted Sardine	COI-Ambsir-2	/5mMC12/AAGGCGTGTGCGGTGACGATGACGT	N
16S	Black Jewfish	16S-Protnib-1	/5mMC12/TTATGGGAGCAGGTGGTACTCTCAT	N
16S	Black Jewfish	16S-Protnib-2	/5mMC12/CATTTAGTTCAGCCCCTTGGTCAGG	N
COI	Black Jewfish	COI-Protdia-1	/5mMC12/CGCCAATTATTAAGGGCACAAGTCA	N
COI	Black Jewfish	COI-Protdia-2	/5mMC12/TAAAGATTTGATCGTCTCCGAGGAG	N
16S	Mackerel Tuna	Euthaff-1	/5mMC12/CTCTAATAAGCAGAATTTCT	Y
COI	Mackerel tuna	Euthaff-2	/5mMC12/TCTACTATCCCTCCCAGTCCTT	N
COI	Eightband Butterflyfish	Chaeoct-1	/5mMC12/GCCTCCCGCTATGTCCCAATAT	Y
16S	Large Scale Grunter	Terathe-1	/5mMC12/CTAAACCCTATGAACCCTGCCCTA	Y

2.4 DNA extraction

For adult samples, tissue and fin clips were preserved in 96% ethanol and stored at 4°C. Initially, a subset of samples were extracted using a modified CTAB / Chloroform-Isoamyl method (Adamkewicz & Harasewych 1996) and further purified using a magnetic bead clean-up step using Sera-Mag Speedbeads Carboxilate - Modified Magnetic Particles (Thermo Scientific, California, USA).

Large-scale fish egg extraction was carried out on a Zephyr® Automated Liquid Handling System (Perkin Elmer) using a Chemagic Arthropod DNA Extraction Kit LH (Chemagen). Before extraction, eggs were placed into a 96 deep well plate with lysis buffer 1a, frozen at -80°C for 10 min, and thawed at 65°C for 5 min. The freeze thaw cycle was conducted three times to rupture the chorion, before proteinase K was added and the samples were incubated overnight at 55°C.

2.5 DNA amplification for sequencing

The same 16S primers (without biotinylated ends) and PCR conditions were used to amplify the 16S region (Tables 3, 4 & 5) of all tested eggs. PCR products were cleaned using Sephadex before sequencing. Paired end sequencing was carried out by the Australian Genome Research Facility. Sequence alignment and quality control was done in Geneious (Kearse et al., 2012), and the final sequences were annotated using BLAST against the NCBI nucleotide database (https://blast.ncbi.nlm.nih.gov/).

2.6 DNA amplification with the addition of biotinylated end for the bead array

Two genes were used for species identification; COI and 16S. Two sets of universal primers were designed to amplify the respective genes across all fish species (Tables 1 & 3). Each forward primer incorporated a biotinylated attachment site at the 5’ end to facilitate fluorescent probe attachment prior to screening.

Each sample was amplified for a single gene (Table 3), for 35 cycles (Tables 4 & 5), using 1 µl of template DNA for a 40 µl reaction volume. Following PCR amplification, 1 µl of each reaction was run on a 0.8% agarose gel, pre-stained with GelGreen, to confirm amplification. Each PCR batch incorporated one negative control sample.

Following PCR amplification, amplification success was recorded for future analysis, and each sample was purified using a magnetic bead clean approach (as described in https://genome.med.harvard.edu/documents/sequencing/Agencourt_AMPure_Protocol.pdf), to remove any substances which may inhibit downstream reactions. Each sample was then quantified approximately using a NanoDrop Microvolume Spectrophotometer, to assess the purity of each sample.

Table 3. Universal primer sets used in the study.

Gene	Forward sequence (5’-3’)	Reverse Sequence (5’-3’)	Reference

COI	TCAACCAACCACAAAGACATTGGCAC	TAGACTTCTGGGTGGCCAAAGAATCA	Ward et al. 2005
16S	GCCTGTTTACCAAAAACATCGC	CCGGTCTGACTCAGATCACGT	Palumbi 1991 (the forward primer was modified to better include fish)

Table 4. General PCR recipe

Total volume:	**20µl**
5x MyTaq Buffer	4
Forward Primer	0.4
Reverse Primer	0.4
MyTaq	0.2
H2O	15
Template	0.5

Table 5. PCR conditions for COI and 16S gene

Gene	**Step**	**Temperature**	**Time**	**Cycles**
COI	Initial denaturation	95	1:00	1
	Denaturation	95	0:15	35
	Annealing	50	0:15	
	Extension	72	0:20	
	Final extension	72	5:00	1
16S	Initial denaturation	95	1:00	1
	Denaturation	95	0:15	35
	Annealing	60	0:15	
	Extension	72	0:20	
	Final extension	72	5:00	1

2.7 Coupling of species-specific probe to bead

Each probe was coupled with a specific magnetic microparticle (Bio-Plex Pro Magnetic COOH Beads, order number MC10023, MC10026, MC10037, MC10044, MC10052, supplied by Bio Rad Laboratories Australia), with each magnetic microparticle being associated with a unique fluorescent dye combination (see Bio-Plex Pro Magnetic COOH Bead documentation). Beads were coupled with each species-specific probe (Figure 5), as per the protocol described in Appendix 1. Beads were eluted into 20 µL of 1xTE, or 20 µL of water, to give a final concentration of ~53,000 beads/µl, and stored at 4 °C until needed (up to 4 months).

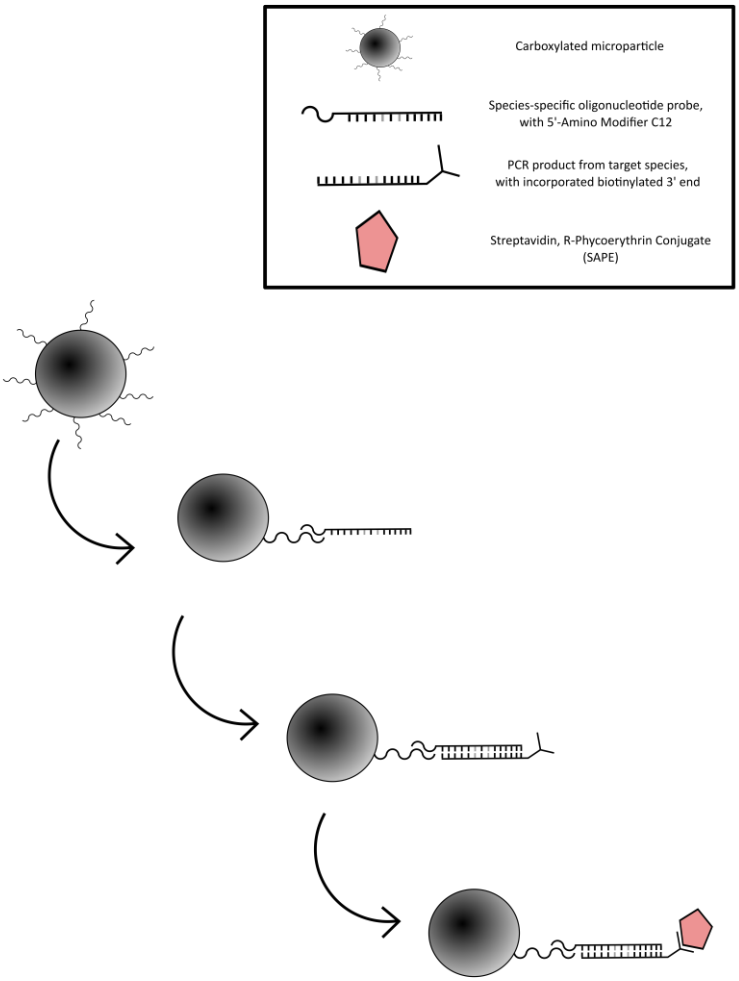


Figure 5. Schematic representation of the bead binding via the species-specific probe to DNA of target.

2.8 Species-specific fluorescent probe approach

2.8.1 Hybridisation of probe to sample and attachment of fluorescent marker (SPE)

To dilute working stocks of coupled beads to an appropriate concentration for processing on the Magpix platform, coupled bead stocks were diluted in 1.5xTMAC solution (1:150). This dilution equates to approximately 354 beads/µL.

For hybridisation of template DNA to species-specific probes, 1 µL of cleaned PCR product was combined with 12.5 µL of working bead solution, and incubated at 57°C for 30 min, following an initial denaturation at 95 °C for 2 min. Following hybridisation, a fluorescent marker dye was added, and incubated for a further 10 min at 57 °C (streptavidin-R-phycoerythrin (SPE), order number S-866, Life Technologies Australia PTY LTD).

2.8.2 Initial testing of probes

The 16S and COI gene fragments were amplified with universal primers from DNA extracted from Spanish Mackerel, Grey Mackerel, Spotted Mackerel, Goldstripe Sardinella, Spotted Sardine and Black Jewfish fin-clips. Each gene was tested in a multiplex bead assay containing beads with probes for each species. Workflow is detailed in Appendix 1.

2.8.3 High through-put testing: Preparation and running on Magpix platform

Prior to running on the Magpix platform, each sample was placed on a magnetic rack to allow removal of TMAC residue. Each sample was washed twice with Magpix driver fluid, and finally eluted into 150 µL of driver fluid. Each plate was allowed to mix on a shaker for 5-10 min prior to being placed on the Magpix, to ensure beads were fully resuspended in solution.

A total of 50 µL of coupled, hybridised bead solution was used to record median fluorescence (MFI), with a minimum bead count set to 30. For each plate of samples, a minimum of two wells containing negative controls were included (coupled beads in water), to obtain an accurate baseline fluorescence value. Non-target species, which had previously been identified through sequencing, were also included on each plate to determine if probes were specific to their target species.

2.8.4 Data analysis

Any wells reading below 30 beads were discarded, and an average MFI value was calculated across all replicates for each bead. Average MFI values of all negative control samples (no DNA), were used as a baseline, with any samples significantly deviating from this value being deemed 'likely positive'. Known target species were also compared with known non-target species, to determine if non-target binding was occurring. If non-target binding was identified, probe sequences were aligned with the non-target species, and location of all sample wells was noted, to determine if the non-target fluorescence was likely a result of high similarity between gene regions, or if it was likely to be technical error resulting from bead carry-over between wells. An average MFI value of all non-target samples was calculated, along with standard error and standard deviations, and compared to average MFI values for target samples. A target sample was considered a 'strong positive' if it deviated significantly from the median fluorescence of both the negative control, and the non-target species, with a threshold of 1.5x greater than the baseline (of both negative control and non-target samples) being deemed 'strong positive'.

2.9 Fish egg staging trial when put into different preservatives

Fish egg staging was assessed using eggs spawned from Barramundi (*Lates* calcarifer) in a hatchery facility at James Cook University following the procedures in Thépot & Jerry (2015). Eggs were removed at different developmental stages and preserved in one of five different ways: 5% formalin, 100% Ethanol, 70% Ethanol, RNA L8R and the final portion was soaked for 30 sec in 5% formalin, washed in seawater and finally preserved in 100% Ethanol. A minimum of 10 eggs were preserved in each solution for each egg stage. The eggs were stored in the dark at 4 °C. Eggs were examined

under a Nikon SMZ-18 dissecting microscope, features of each egg stage were examined and digital images recorded. The morphological features that define each egg stage (see Thepot & Jerry 2015) were examined and the capacity for each stage to be determined was determined subjectively.

2.10 DNA amplification in different preservatives

To measure egg preservation success from wild plankton samples, four plankton samples were collected at Helix Reef on the GBR using a Neuston Net (Section 2.2.2). The net was deployed from the RV Kirby while at anchor and streamed in the current for 15 min (site: -18°37.7', 147°17.3'). On retrieval, each sample was split into four roughly equal parts and each part preserved in one of four target preservation solutions outlined above. Plankton samples were stored in the dark at 4 °C until laboratory processing. In the laboratory the first 20 fish eggs encountered in each sample were removed and placed individually into a round bottom 96 well plate as follows: one egg in each well containing 300 µl of lysis buffer 1A, with the exception of wells 1A and 2A used as negative and positive control, respectively. Before starting the DNA extraction protocol, three cycles of freeze and thaw were carried out at -80 and 65 °C in order to break the chorion and ease access to DNA. Following the last thaw cycle, 10 µl of Proteinase K were added to each well and mixed by vortex for 5 sec. Samples were lysed overnight at 55 °C using a dry incubator. DNA extraction of lysed samples was performed in the Zephyr Robotics system (Caliper, Life Sciences) following the Chemagic Arthtopod DNA extraction Kit LH (Chemogen) protocol. Extracted DNA was eluted in 40 µl of elution buffer and stored at -20 °C.

The 16S gene was amplified by PCR using MyTaq DNA polymerase (Bioline). PCR was performed using the same universal 16S primers (Table 3). Reactions were carried out in a total volume of 10 µl containing 2.0 µl 5X MyTaq Reaction Buffer, 0.2 µl of each of the Forward and Reverse primers, each at 10 µM, 0.1 µl MyTaq DNA polymerase and 1 µl of undiluted DNA with PCR cycling conditions as per Table 5. Successful amplification was verified by running the PCR amplicon on an agarose gel (1.5% agarose in 1X TBE buffer) and only a single band of the expected size was obtained.

3 Results

3.1 Initial testing of probes

For each of the five initial species of interest, Spanish Mackerel, Grey Mackerel, Spotted Sardine, Goldstripe Sardinella and Black Jewfish, two probes targeting two different genes (16S and COI) were designed (where possible), and tested for species specificity. Flow cytometry results show that at least one probe per species successfully hybridizes to these species with high specificity, resulting in an increased fluorescence.

The Spanish Mackerel samples showed a positive result with the species-specific probe (Figure 6 A,B), and a negative result with the probe specific *to* Grey Mackerel (Figure 6 C,D). Similarly, the Grey Mackerel samples showed a positive result with the species-specific probe (Figure 6 C,D), and a

negative result with the probe specific to Spanish Mackerel (Figure 6 A,B). No positive results were observed for the negative control species Spotted Mackerel (Figure 6 A-D). A similar negative result was obtained with COI from Black Jewfish samples (data not shown).

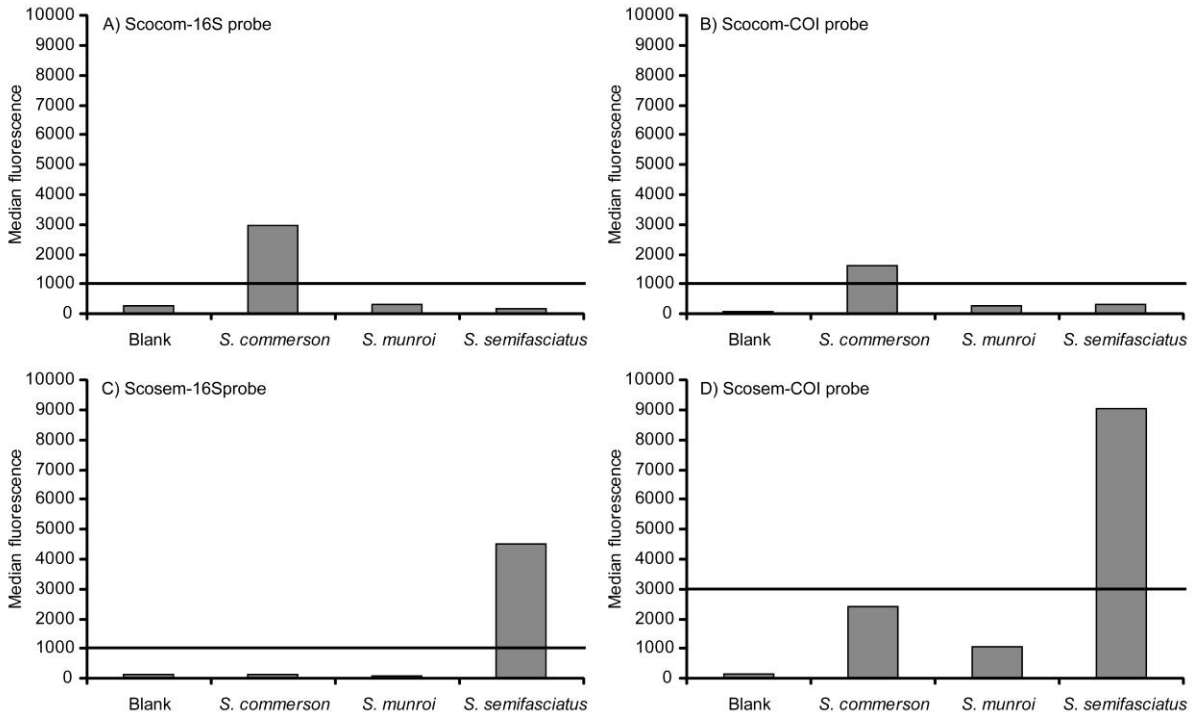


Figure 6. Median fluorescence for four probes A) Spanish Mackerel (*Scomberomorus commerson*) (Scocom) 16S, B) Scocom COI, C) Grey Mackerel (*S. semifasciatus*) (Scosem) 16S, and D) Scosem COI tested with Spanish Mackerel, Spotted Mackerel and Grey Mackerel16S and COI fragments amplified from fin-clip DNA. The horizontal line represents the threshold for a positive assignment to the respective species.

The *A. sirm* specific 16S probe successfully identified the *A. sirm* sample, while being negative for the other two species (Figure 7 A). In contrast, the COI universal primers did not amplify the *A. sirm* sample, and the *A. sirm* specific COI probe exhibited unspecific binding (Figure 7 B). The *P. diacanthus* specific 16S probe did not show a positive result for the species (Figure 7 C); however, the COI probe exhibited a strong, species-specific result and therefore successfully identified the correct species (Figure 7 D). The *S. gibbosa* 16S probe successfully identified the correct species, to the exclusion of the other species (Figure 7 E), while the COI probe only showed a minimal response (Figure 7 F). Full probe details, including those developed for the additional species included in the 2016 field samples (*Euthynnus affinis*, *Terapon theraps* and *Chaetodon octofasciatus*) are provided in Supplementary Material S1.

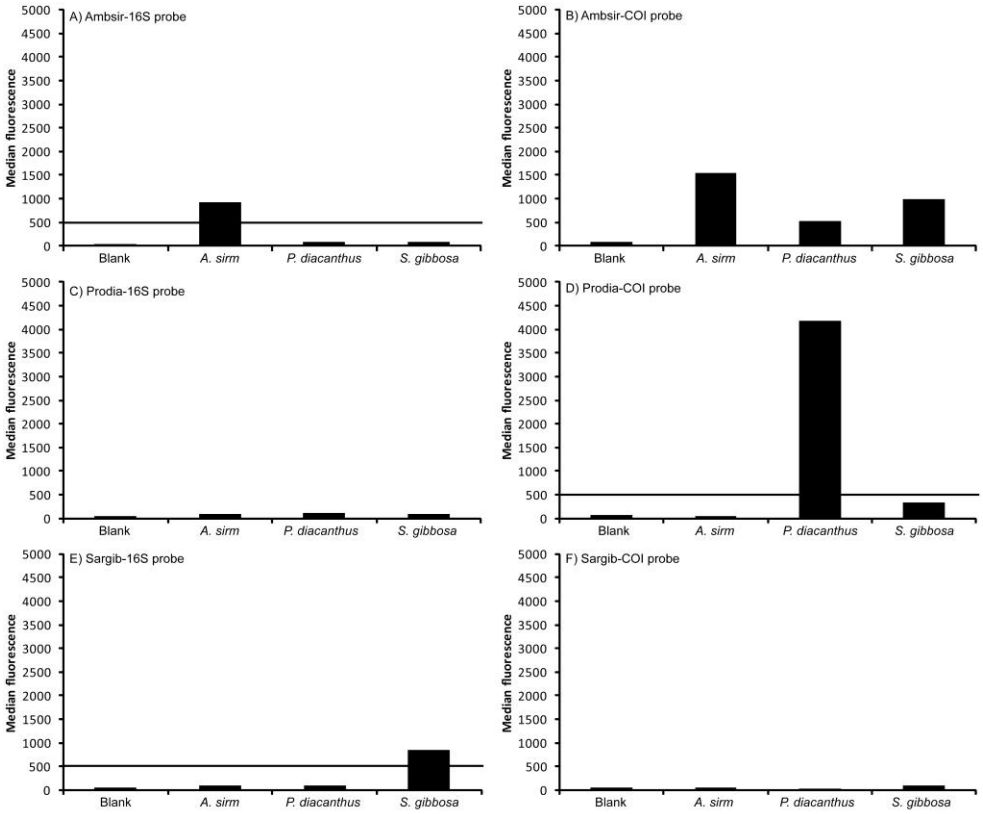


Figure 7. Median fluorescence for six probes A) Spotted Sardine (*Ambligaster sirm*) (Ambsir) 16S, B) Ambsir COI, C) Black Jewfish (*Protonibea diacanthus*) (Prodia) 16S, D) Prodia COI, E) Goldstripe Sardinella (*Sardinella gibbosa)* (Sargib) 16S, and F) Sargib COI. 16S and COI fragments for hybridization were amplified from fin-clip DNA of the respective fish species. The horizontal line represents the threshold for a positive assignment to the respective species.

3.2 Species identification from field egg surveys

3.2.1 Bead array egg identification (2015)

The samples for identification by bead array consisted of a negative control (non-template control, NTC), DNA extracted from fin-clips of Spanish Mackerel (PC-SC) and Spotted Sardine (PC-AS) as positive controls, and DNA extracted from 10 selected fish eggs (see Figure 8). The DNA was amplified using a universal 16S primer pair. Species-specific 16S probes for Spanish Mackerel (Scocom) and Spotted Sardine (Ambsir) were coupled to two different bead lots, enabling simultaneous screening for both species (i.e. multiplexing). The three control samples and 10 unknown samples were individually hybridized to the bead mixture containing both probes. After the flow cytometric analysis, median fluorescence values were calculated for all samples for each probe. The threshold for positive identification was set using the NTC, negative controls and positive controls as guidance. Samples were deemed positive if they were clearly above the threshold. Figure 8 illustrates the successful identification of egg number EB18 as Spanish Mackerel and egg numbers EB05 and EB37 as Spotted Sardine using the threshold level. Correct identification was confirmed by

Sanger sequencing (Table 1; sequencing results). Egg identification results between the multiplex bead array and the sequencing were identical.

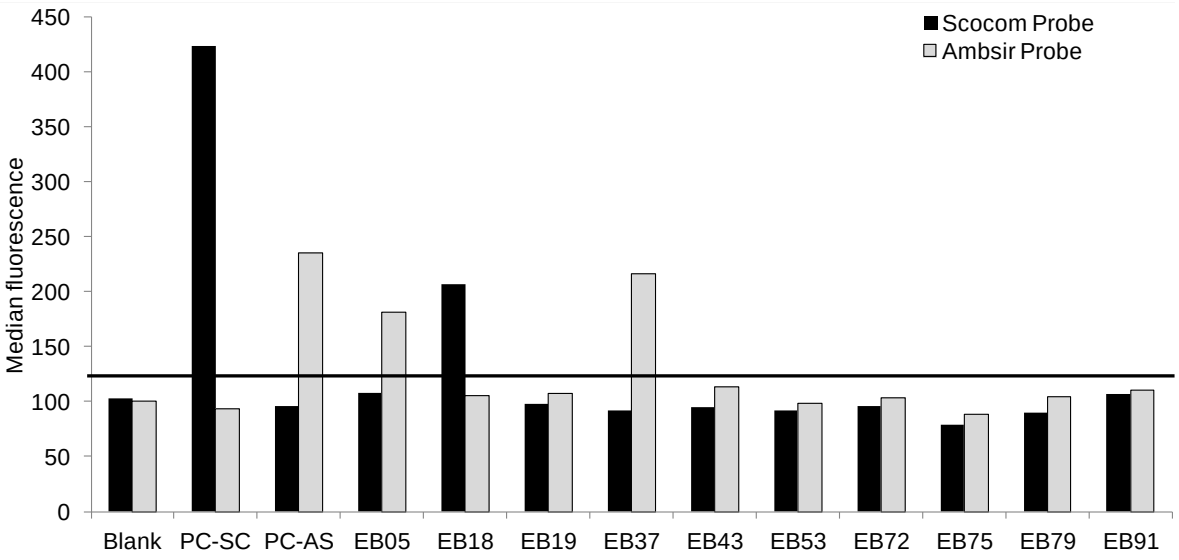


Figure 8. Results from bead array identification of 10 eggs collected during the October 2015 plankton survey. Fluorescence above the threshold line indicates a positive detection. Control tissue from each target species (PC-SC and PC-AS) were used to demonstrate the array was working effectively.

Table 6. Sequencing results for identification of 10 unknown eggs collected during the October 2015 egg survey using 16S probe.

Sample ID	Spanish Mackerel	Spotted Sardine	% pairwise identity	Query coverage	Best hit (Scientific name)	Best hit (common name)	Taxonomic level of assignment
NTC	negative	negative	-	-	-	-	-
PC-SC	positive	negative	-	-	-	-	-
PC-AS	negative	positive	-	-	-	-	-
EB05	negative	positive	99	96	*Amblygaster sirm*	Spotted Sardine	Species
EB18	positive	negative	100	100	*Scomberomorus commerson*	Spanish Mackerel	Species
EB19	negative	negative	100	99	*Strongylura incisa*	Reef longtom	Species
EB37	negative	positive	97	96	*Amblygaster sirm*	Spotted Sardine	Species
EB43	negative	negative	100	100	*Scarus globiceps*	Violetline Parrotfish	Species
EB58	negative	negative	100	100	*Synodus kaianus*	Black Lizardifsh	Species
EB72	negative	negative	100	100	*Choerodon fasciatus*	Harlequin Tuskfish	Species
EB575	negative	negative	100	100	*Fistularia petimba*	Rough Flutemouth	Species

EB79	negative	negative	100	100	*Choerodon jordani*	Dagger Tuskfish	Species
EB91	negative	negative	100	97	*Synodus variegatus*	Variegated Lizardfish	Species

3.2.2 Amplification success of 16S and COI, and incorporation of biotinylated attachment site for bead array testing (2016 egg survey)

To understand the proportion of false negatives in the bead array resultant from failed DNA amplification we present the following results for the 2016 field collected egg samples. Both universal primers (16S and COI) appeared to amplify well across all samples, with a PCR success rate of 79% across all plates but amplification success was not consistent with rates ranging from 74% to 95% across sample locations (Table 7). Weaker bands were observed in many Spotted Sardine samples, compared to other species when amplified using the COI primer set.

Table 7. 2016 egg collections. *Helix Experiment eggs were preserved in a range of solutions, numbers in brackets are for those eggs preserved in 100% EtOH.

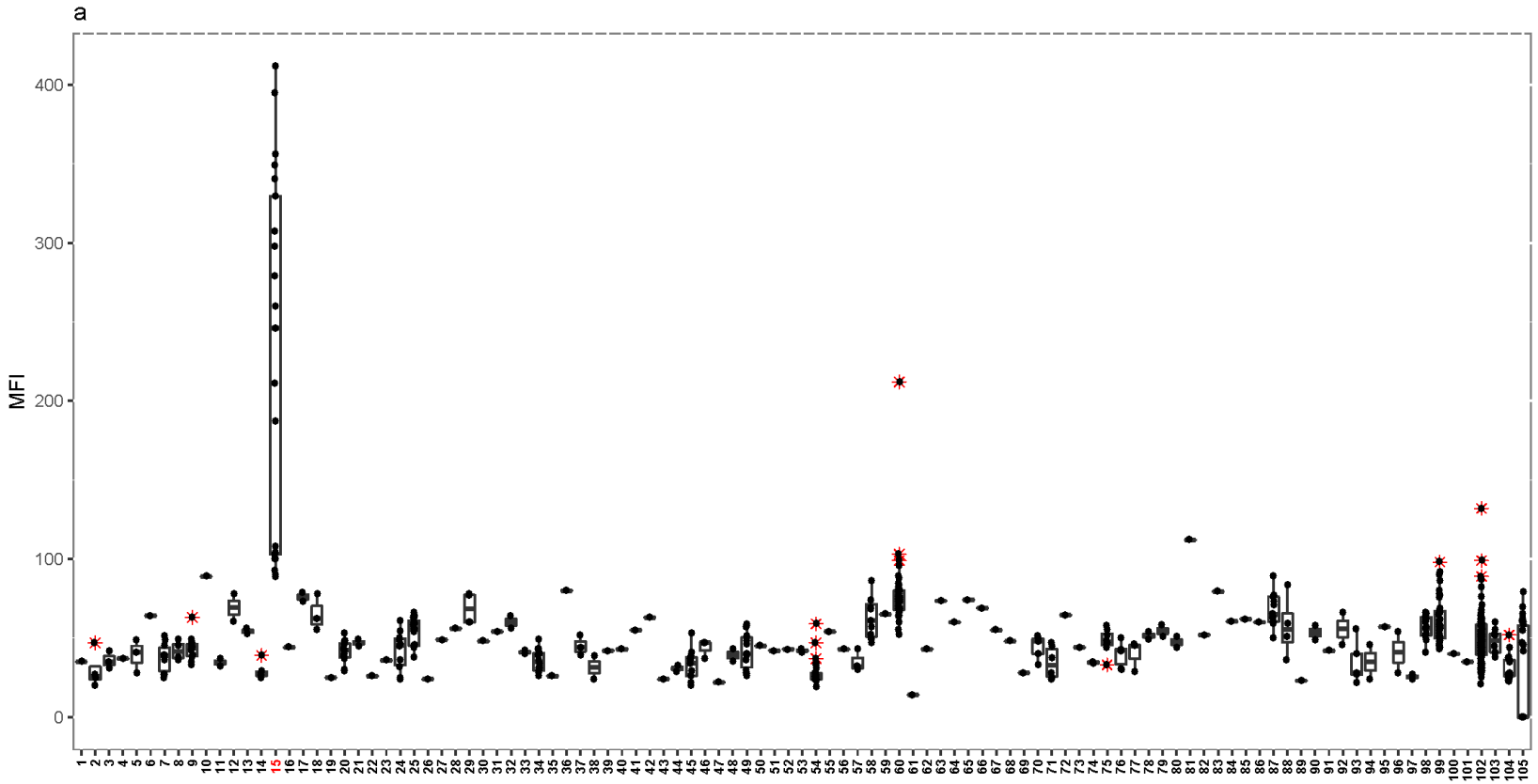
Location	**Total samples**	**Number of fish eggs**	**Number of eggs with successful DNA amplification**	**% amplification success**	**Spherical eggs between 1.0 and 1.5mm diameter (size range for Spanish Mackerel)**
Brewer	20	439	296	67.4	12
Helix Experiment	4	160 (40) *	150 (38)*	93.8 (95.0)*	21
Helix Survey	17	146	116	79.5	2
Rib	14	310	230	74.2	9
NT	8	309	286	92.6	8

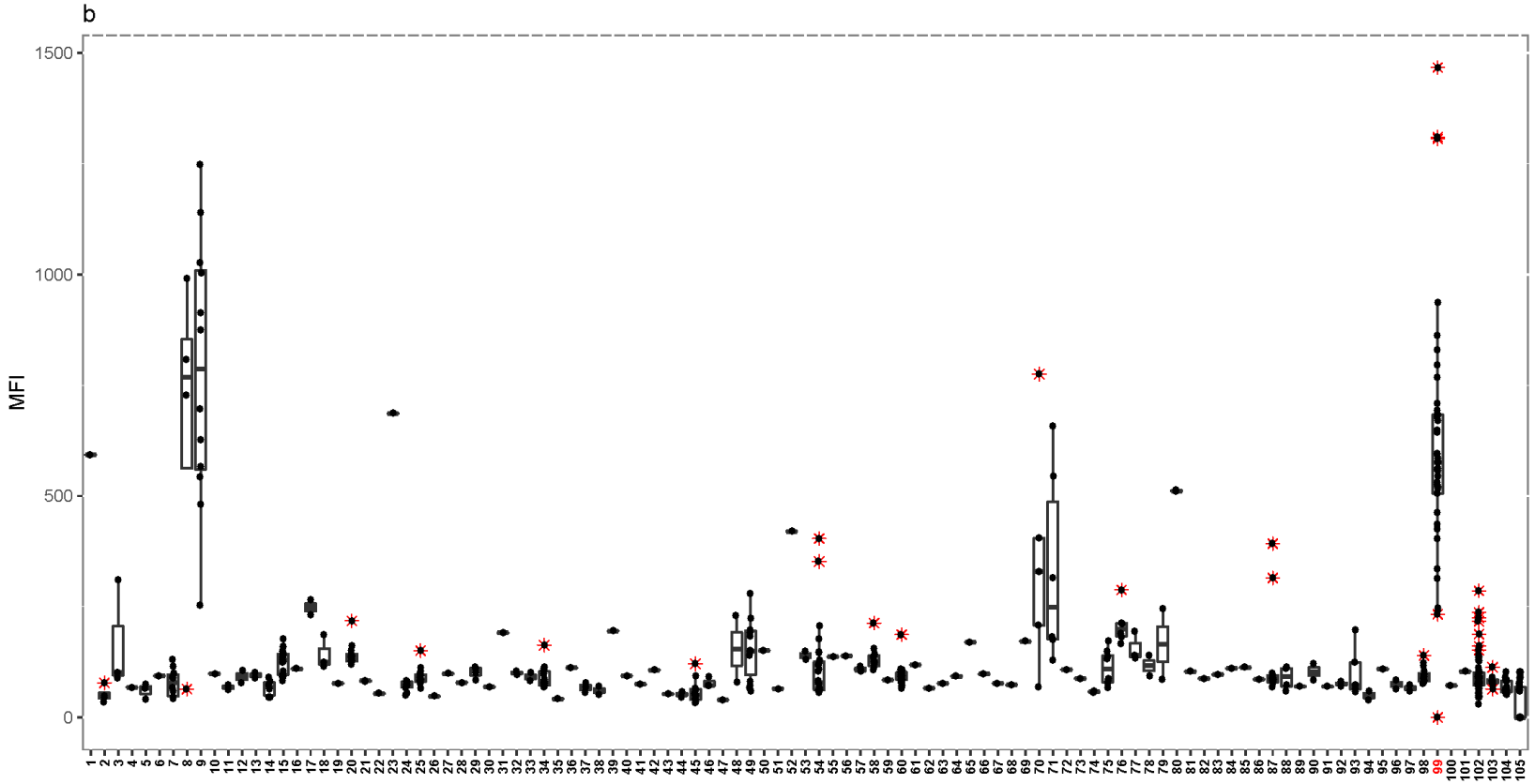
3.2.3 Bead array egg identification (2016 survey)

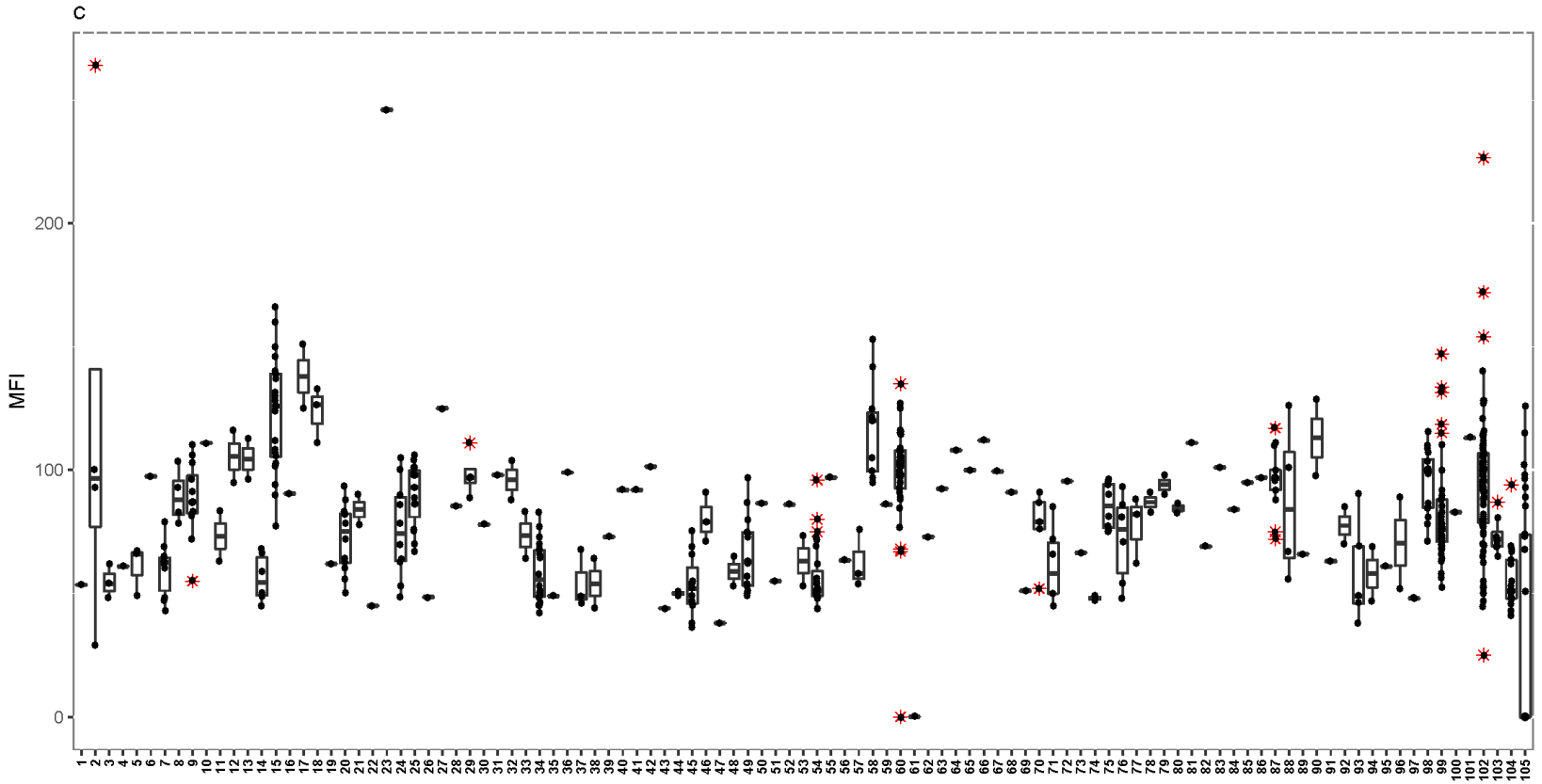
Following initial optimisation, a final panel of five species-specific probes (Table 2) were selected for use to identify target species in ichthyoplankton samples. Using the species identification results obtained through sequencing (Appendix 2), a total of 576 fish eggs (six plates) were processed using the Magpix platform (Figure 9a – e; comprising a total of 86 target species, 456 non-target species and 34 negative controls). All remaining samples were omitted from this testing, as sequencing results indicated that no target species were present within these plates.

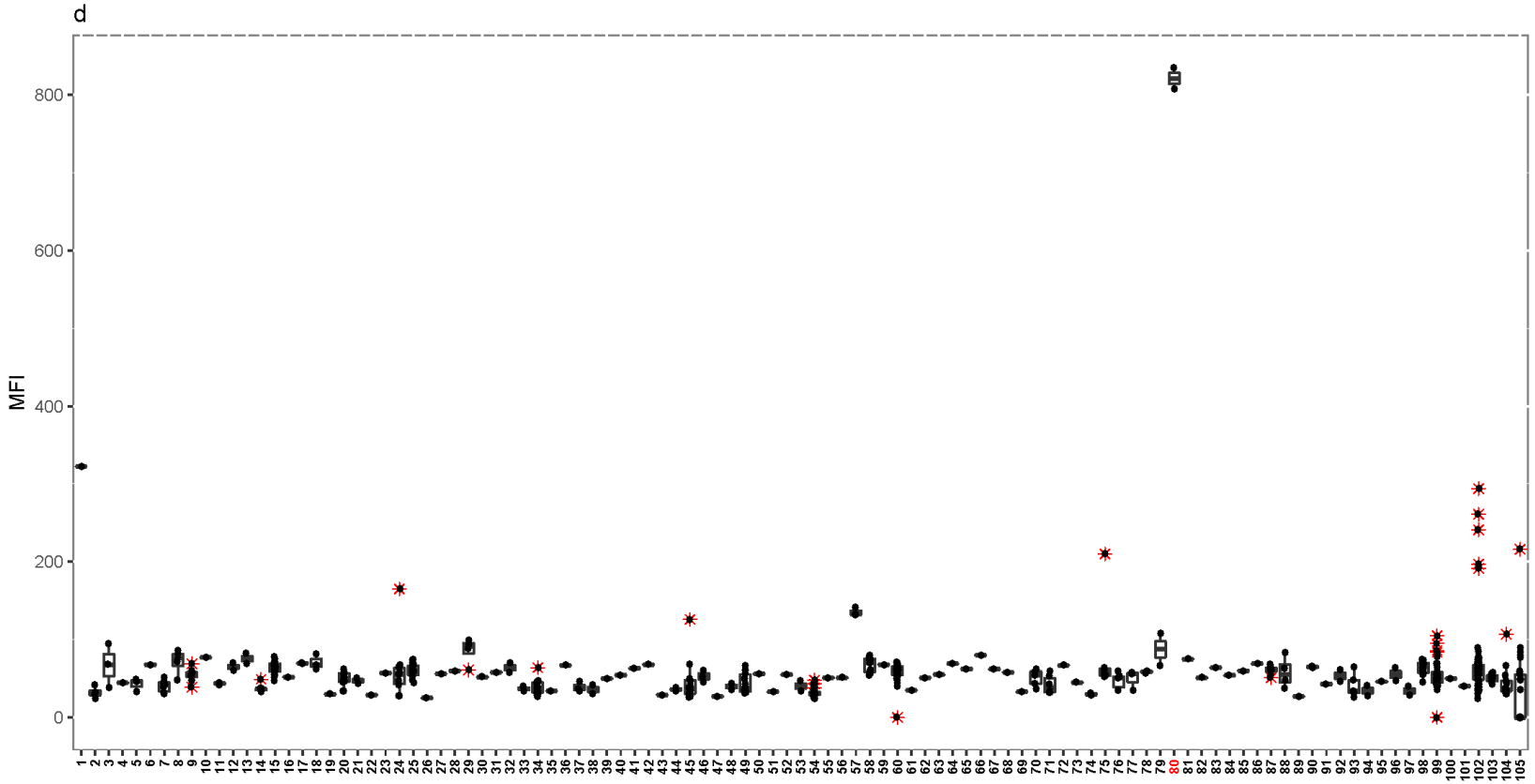
Probes designed for Eightband Butterflyfish(Figure 9a), Large Scale Grunter (Figure 9b), Spanish Mackerel (Figure 9d) and Mackerel tuna (Figure 9e) all displayed markedly higher fluorescence values than non-target species, and negative controls. However, a large degree of variation (ranging from MFI = 50–1,250) was observed within these species, indicating that probe hybridisation may have failed in some samples. Non-target binding to Spangled Emperor (*Lethrinus nebulosus*) samples

was also observed for the Mackerel Tuna probe, and when this probe was aligned to publicly available 16S sequences for this species, only a single nucleotide differentiated these species, meaning that this probe may be useful in identifying both Spangled Emperor and Mackerel Tuna samples, but not to differentiate the two species. Non-target binding was also observed in the Large Scale Grunter probe, which gave positive results for Yellowtail Fusilier, *Caesio caerulaurea*, *Pterocaesio pisang*, and Bigtail Caesio *Pterocaesio marri*. All of these species belong to the order Perciformes, and only contained 2 bp differences within the probe binding region. Only 50% of Spotted Sardine samples showed 'likely positive' results, but were not as clearly differentiated from baseline fluorescence as other target species (Figure 9c). Other probes also failed in some samples (Table 11).


a
MFI
400
300
200
100
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105


b
MFI
1500
1000
500
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105


C
MFI
200
100
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105


d
MFI
800
600
400
200
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105

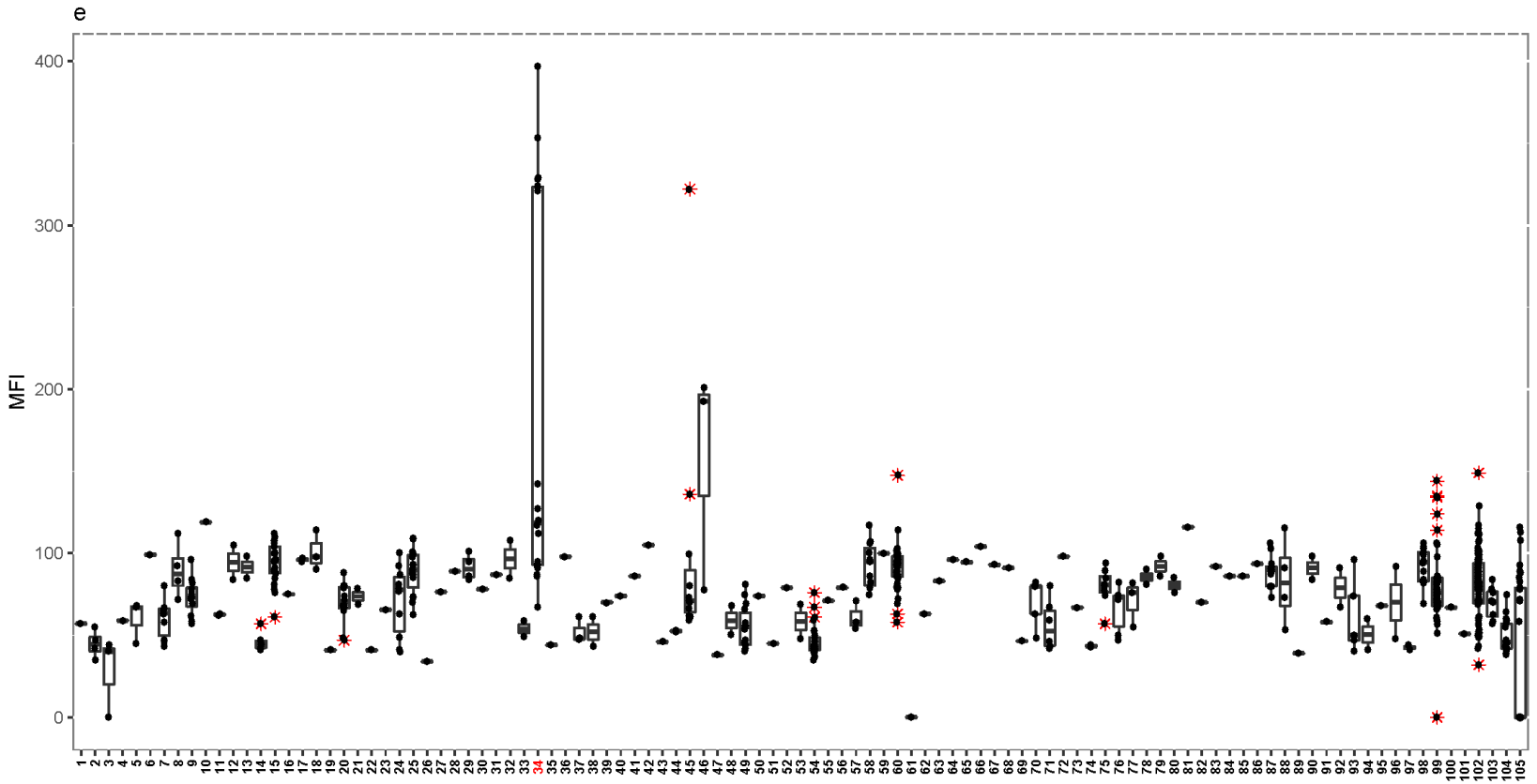


Figure 9 a-e. Fluorescence values for each egg with target species highlighted in red; a) *C. octofasciatus*, b) *T. theraps*, c) *A. sirm*, d) *S. commerson* and e) *E. affinis.* (Species names for the numbers listed on the x-axis are provided in the Appendix 3.

Table 8. Comparison of identification success between sequencing and the bead array (on Magpix platform) for 2016 samples.

Species	Number in 2016 Samples (from sequencing)	Number successfully identified using Bead Array	Bead Array (% successfully identified)
Spanish Mackerel	2	2	100.00%
Grey Mackerel	0	-	-
Spotted Sardine	4	2	50.00%
Goldstripe Sardinella	0	-	-
Black Jewfish	0	-	-
Mackerel Tuna	18	10	55.56%
Large Scale Grunter	41	40	97.56%
Eightband Butterflyfish	21	21	100.00%

3.5 Fish egg staging in different preservatives

There were visible differences among the different preservation methods (Table 9) and an experienced technician should be able to distinguish most egg stage in F+100% preservation from very early in development but in 100% and 70% ethanol preserved eggs, reliable staging does not appear possible until the point at which optic vesicles are visible as they become very opaque (Tables 9, 10). Egg staging using RNAlater was challenging because the preservative deforms the egg markedly (Table 9; however, many of the important features in early-mid development (e.g., visibility of the blastodermal cap) are readily distinguishable, as the egg does not become opaque. At the stage when somites develop and pigmentation develops, (i.e., essentially when an embryo is visible) egg staging is relatively straightforward using all methods.

Table 9. The impact of preservation on the opacity and shape of barramundi eggs. 100% EtOH = 100% ethanol, 70% EtOH = 70% ethanol, RNAL8R = RNA L8R, F+100% = eggs washed in 5% formalin and preserved in 100% ethanol, Formalin 5% = Formalin 5%.

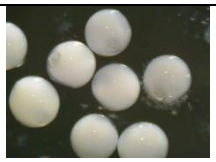
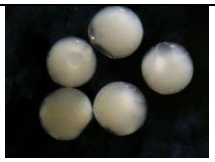
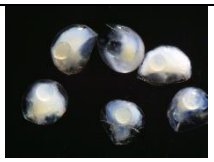
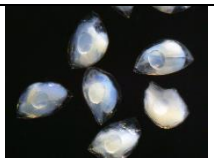
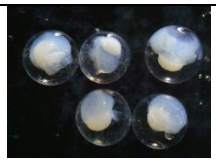
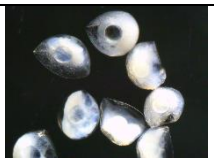
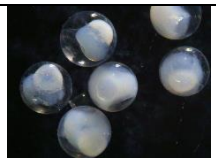
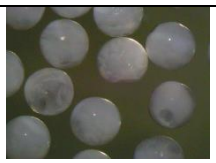
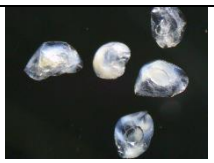
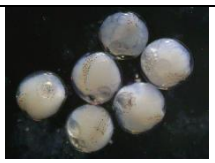
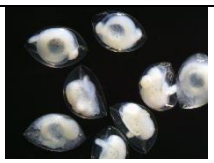
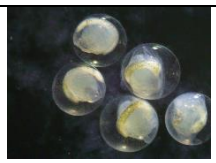
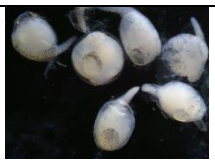
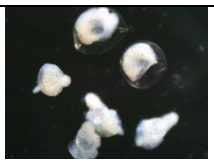
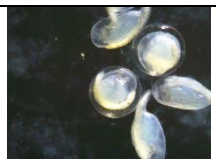
Stage	100% EtOH	70% EtOH	RNAL8R	F+100%	Formalin 5%
128 cells					
256 cells					
germ ring					
optic vesicles					
15 somites					
hatch					

Table 10. The impact of different preservatives on the subjective ability to stage barramundi eggs. 100% EtOH = 100% ethanol, 70% EtOH = 70% ethanol, RNAL8R = RNA L8R, F+100% = eggs washed in 5% formalin and preserved in 100% ethanol, Formalin 5% = Formalin 5%.

Stage	100% EtOH	70% EtOH	RNAL8R	F+100%	Formalin 5%
128 cells	N	N	Y	Y	Y
256 cells	N	N	Y	Y	Y
Germ ring	N	N	Y	Y	Y
Optic vesicles	Y (~10% of cases)	N	Y	Y	Y
15 somites	Y	Y	Y	Y	Y
Hatch	Y	Y	Y	Y	Y

3.6 PCR amplification success in different preservatives

PCR success was uniformly high across all preservation methods (Table 11). The poorest performance was from eggs preserved in 70% ethanol.

Table 11. Rate of PCR success using a variety of preservation methods on wild collected plankton samples. 100% EtOH = 100% ethanol, 70% EtOH = 70% ethanol, RNAL8R = RNA L8R, F+100% = eggs washed in 5% formalin and preserved in 100% ethanol.

Preservation	Average PCR success % ± SE
70% EtOH	92 ± 7.5
100% EtOH	95 ± 2.9
F +100%	95 ± 2.9
RNA L8R	95 ± 3.2

4 Discussion

The project successfully identified target species from wild collected plankton samples when they were present. The method was effective when tested using fin clips (100% successful) and results were excellent during the testing phase when low numbers of eggs were assessed (100% successful). However, when scaled up to screen for multiple species in wild collected plankton samples containing eggs of a diverse range of species, the ability of the bead array method to successfully identify to a species level was challenged (ranging from 50% to 100% success depending on species) and three major concerns were identified.

The first project concern was that low numbers of eggs of the target species were collected. This was identified as a risk to the project in the initial Risk Analyses, but it remains disappointing and surprising not to have collected significantly more eggs of the target species, especially for Spanish Mackerel for which we had reasonable information to target spawning times and aggregations. At the time the egg survey was done on the Great Barrier Reef, particularly in the area surrounding John Brewer Reef, we also targeted adult Spanish mackerel to confirm that spawning was indeed occurring at the time of the plankton sampling. We were unsuccessful in capturing any female Spanish Mackerel in 2015. Commercial catch data also indicated that the catch was relatively low during the period we targeted our survey in 2015 suggesting the possibility that fish had not yet aggregated to spawn (http://qfish.fisheries.qld.gov.au/). The following year we expanded our survey to include Rib Reef and collected additional samples around Helix and Brewer Reefs but again egg numbers for Spanish Mackerel remained very low. This issue was overcome by developing the bead array method for some additional species which were present in reasonable numbers in the samples and this allowed for a better assessment of the identification method.

The second concern was the presence of false negative results. False negative results are when an egg of a target species is present, but has not been able to be identified. There are several different sources of false negatives. If DNA amplification fails, likely due to poor preservation of the egg, the egg cannot be identified. In controlled conditions using eggs sourced from aquaculture broodstock this occurs rarely, but in wild collected plankton samples it was typically much higher (Table 12). This particular issue is not solely a problem using the bead array but rather is likely to occur for all molecular identification methods. Indeed, while not all previous studies identifying fish eggs report amplification failure rates, those that do range from 4% to 10% (Table 12). The amplification failure can be either due to problems with the DNA extraction (e.g., failed DNA extraction or degradation), or problems with the PCR reaction (e.g., low input DNA, inhibition, or “universal primers” not being completely universal).

Table 12. Range of published methods for fish egg identification identifying DNA amplification failure rates.

Species	**Method**	**Amplification failure rate (gene)**	**Wild or captive spawned eggs**	**Preservation**	**Publication**
Various (eggs and larvae)	PCR – sequencing	51.4% CO1 61% 16S	Wild	100% ethanol	Ahern et al. 2018
Snapper	Real time PCR (Silicone membrane)	9.5% (CO1 & 16S)	Wild	Chilled then frozen	Dias et al. 2016
	Real time PCR (Quick extraction)	1.5% (CO1 & 16S)			
Gadidae	TaqMan multiplex PCR	6% (Cyt B)	Wild	Chilled then 100% ethanol	Fox et al. 2005
Various species	Multiplex bead array (PCR universal primers)	~10% (CO1)	Wild	Ethanol 100% (stored 5-10 years)	Gleason & Burton 2012
Trachurus trachurus	PCR – FRAP	80% (Cyt B) 0% (Cyt B)	Captive spawned	Formalin Ethanol	Karaiskou et al 2007
Various	PCR-RFLP	20% Rockling (Cyt B) 13.6% Gadidae (Cyt B)	Wild	Formalin	Lelièvre et al. 2010
Trachurus spp.	PCR/sequencing	4% range 0-6.6% depending on egg stage (CO1)	Wild	Ethanol	Neira et al. 2015

Another source of a false negative that is harder to identify is when the species-specific probe fails to bind the fluorescent bead to the DNA of the target. This appears to have occurred occasionally and is particularly evident in examining the Mackerel Tuna egg data. In this case, DNA of eggs was successfully amplified using a universal primer, but only 55% of these eggs were successfully identified using the bead array in comparison to sequencing (Table 8).

In the context of a DEPM assessment, these results are of considerable concern as they would alter two important parameters in the model. First, they would result in a lower number of eggs identified, which would have a concomitant reduction in the biomass estimate. Second, it would also have an impact on estimates of egg mortality in the DEPM, the direction of which would be dependent on whether false negatives were more or less likely at different egg developmental stages. Intuitively it is probable that false negatives are more likely in earlier developmental stages when less DNA is present, which would skew results to a lower estimate of egg mortality and again ultimately a lower biomass estimate. However, Neira et al. (2015) reported no PCR amplification failures for early stage eggs, but counterintuitively this increased to 6.7% for later stage eggs. Oxley et al. 2017 note that the success of in situ hybridisation was independent of egg stage, although this conclusion appears to be based on eggs sourced from aquaculture broodstock. In our study, we did not have sufficient data to assess the effect of egg stage on amplification success. In our case we have attempted to quantify false positive and false negative identifications by contrasting the bead array approach with traditional sequencing methods but the equivalent is rarely done in the case of morphological egg identifications. It is noteworthy that when checks have been done morphological identifications have also been found to be imperfect (e.g., Fox et al. 2005).

The impacts of false negatives on DEPM biomass estimates are complex, and depending on their distribution across egg stages, could either increase or decrease biomass estimates. Thus, it is clear that minimising false negatives in egg identification is imperative in providing precise biomass estimates using the DEPM. The amplification failure rates observed across species in this study were 21%. This is at the upper end of failure rates reported in the literature for fish eggs, which range from 4% (Neira et al. 2015) to 20% (Lelièvre et al. 2010) for ethanol preserved eggs and as high as 80% (Karaiskou et al. 2007) in formalin preserved eggs (Table 9), although it is noteworthy that not all studies report this data. In the case of Oxley et al. (2017), using *in-situ* hybridisation, DNA amplification is not part of the method, thus a measure of preservation success is not available. The level of uncertainty around false negative results means that molecular methods are compromised for egg identification in the context of DEPM. It is important to note however that false negative results for morphological identifications are rarely quantified and may well be a significant problem in itself.

The third major concern was with the reliability of the chemistry. The approach described above requires a number of successful reactions and creates a “daisy chain” of oligonucleotide, and antibody-like molecules (Supplementary Material S1, Figure 5). An accurate identification of the target species requires each step in this process to be successful, with little opportunity to accurately assess the success of each intermediate step (or at which point the process failed). Despite the successful identification of 80.6% of target samples, there was a high level of variability observed between individual samples, even when the assay appeared to give some positive results. This can be seen in Figure 9a &e, where it should be noted, that while a subset of the target species were identified successfully, several of the samples returned a negative result. The failure of some samples could be a result of any of the issues outlined in Table 13, but it is difficult (and costly) to accurately determine why individual samples failed. We were able to successfully develop this

method and the Magpix platform allowed for high throughput processing of many samples at a time, but the reliability of the method was not suitably high enough to allow for efficient and accurate processing of plankton samples on a large scale.

Table 13. Issues and solutions for the bead array method (on Magpix platform).

Steps in which the process could fail	Possible Issue	Solution
PCR failure (template)	Complete amplification failure	Ensure the primers are appropriate to the target species, and always use positive controls. Ensure that PCR conditions are appropriate.
PCR failure (streptavidin attachment site not attached)	Successful amplification of target gene, but failure of primers to incorporate carboxylate modified attached site to PCR product	This can be assessed by using streptavidin-coated magnetic beads to clean up PCR product (Sigma Aldrich Product code 11641778001). Only DNA fragments with the streptavidin attachment site will be retained.
Failure of probe binding to bead	Failure to attach species specific probe to fluorescent magnetic bead	Ensure that probes incorporate an amino-modified C12 to the 5' end of the probe, which will hybridise to the carboxylate coated magnetic beads. This process requires a catalyst to the reaction (EDC solution), if this catalyst is not freshly prepared, then hybridisation will not occur efficiently, or will fail completely.
Failure of probe/bead combination to bind to the target species PCR product	Failure of probe binding to the PCR product, or non-specific binding, either as a result of inappropriate hybridisation temperature, or the probe being inappropriately designed (too many, or too few nucleotide mismatches)	For each set of species-specific probes, a hybridisation temperature gradient should be performed to assess the optimal temperature. Probes should also be thoroughly checked to ensure that they are specific to the target species only.
Failure of the SPE to bind to the attachment site	Failure of SPE binding to streptavidin attachment site on the end of the PCR product, either as a result of inappropriate hybridisation temperature, or denaturation of SPE reagent	Ensure that the SPE reagent is within its 'used by' date, and keep cold at all times, avoid exposure to light wherever possible.
Failure of the SPE to fluoresce	Failure of SPE to fluoresce, likely as a result of previous over excitation of SPE molecule, and overexposure to light.	Ensure that once SPE is added and hybridised to bead, that the assay is exposed to as little light as possible. Keep covered with foil at all times until the plate

		is to be read.
Loss of beads	Loss of magnetic beads during one of the several wash steps, as a result of plates/samples not being placed on an appropriately string magnetic plate holder	Ensure that an appropriate, strong rare earth magnetic is used for all wash steps. Decanting appears to be more successful than removing washes by pipetting.
Failure of Magpix to read beads	Failure of Magpix (or flow cytometer) to accurately read beads, either due to inability to draw beads through the machine, or an error in software protocol (reading an incorrect bead region)	Software protocols can be rerun to cross check bead regions, and once the assay is ready to be run on the Magpix platform, it appears to be stable so long as it is not exposed to light. Plates may be stored in the fridge and rerun at a later date if the Magpix platform malfunctions.
Cross contamination between wells, leading to false positives	Inability to determine a true positive result as a result of cross contamination of beads into nearby wells	Low bead counts are indicative of cross contamination – a minimum threshold of beads should be implemented to identify possible cross contamination

Other concerns include that all molecular identification methods rely on accurately identified reference material and good coverage of the sequences of target genes of species that co-occur with a study region. This data is essential for the development of genuinely species-specific probes. Thus, confidence in the species-specificity of these probes is largely dependent on how well studied the fauna of a particular region is. For example, in the work by Gleason & Burton (2012) in which a bead array was developed for a large number of species in California, the coverage of sequences of fish fauna in the region for the target genes was close to 100%. At the time we developed the probes for this study, coverage in Australia was considerably lower, and for the Great Barrier Reef region the proportion of fish species that occur within the region with sequences uploaded into GenBank was 40% for 16S and 60% for C01. With many of these sequences coming from specimens taken in areas outside the Great Barrier Reef. It is worth noting that while sequencing was good for accurate identification of the target species, unknown eggs species were only able to be determined in 50.5% of cases, which roughly aligns to the coverage of GBR fish species in GenBank. There is a need for a concerted effort to improve this coverage for projects such as this, but also for other molecular methods such as biodiversity surveys using environmental DNA which relies on the same data sources.

Finally, rupturing the chorion of the fish egg to access the genetic material for molecular identification of fish eggs is a known issue. It stymied the work of Oxley et al. (2017), where to get *in-situ* hybridisation to work the probe needed to be injected into the fish eggs individually which impacted the efficiency of the method. Other studies have physically broken the chorion using toothpicks (Neira et al. 2015) and in our study we overcame the issue by freezing and thawing the eggs once they were loaded into plates to physically rupture the chorion; however, this process may

have contributed to our relatively low DNA amplification success and add to the issue of false negatives. If a chemical method could be developed to remove the chorion that did not damage DNA this it would be a considerable step forward for molecular egg identification, however, the thorough attempts by Oxley et al. (2017) to chemically remove the chorion of snapper failed and demonstrate the hardiness of this barrier.

The bead array method developed in this study has numerous benefits. The ability to multiplex means multiple species can be identified in one pass, as it is PCR based it is effective on early stage eggs, can be applied to all life stages (not just eggs), it is affordable when hundreds of eggs of a target (or multiple target) species need to be identified. Furthermore, false positives were not identified (when compared to sequencing results). However, some issues remain, the plankton still requires sorting, it is destructive to the eggs, it is only cost effective when doing hundreds of eggs, the chemistry has proven temperamental, and confidence is only as good as the reference material and online resources.

5 Conclusion

The most suitable approach for fish egg identification will depend on the accuracy and precision required for a particular need. The bead array approach in this project was excellent in as much as false positive results can be eliminated if the species specific probes are genuinely species specifid. This is a considerable advance over morphological methods where false positives can easily go un-noticed. However, in the context of DEPM,the rate of false negatives (i.e. the inability to identify some eggs) of the target species is important. Thus, using the bead array method to identify eggs would most likely produce a lower biomass estimate. While this would be a conservative approach for fishery management, the loss of eggs due to DNA amplification failure and the occasional failure of the binding of the beads is a significant flaw and would likely viewed as unacceptable for fishing industry stakeholders.

6 References

Ahern, A. L. M., Gomez-Gutierrez, J., Aburto-Oropeza, O., Saldierna-Martinez, R. J., Johnson, A. F., Harada, A. E., Sanchez-Uvera, A. R., Erisman, B., Arvizu, D. I. C. & Burton, R. S. (2018). DNA sequencing of fish eggs and larvae reveals high species diversity and seasonal changes in spawning activity in the southeastern Gulf of California. Marine Ecology Progress Series. 592, 159-79.

Aoyama, J., Ishikawa, S., Otake, T., Mochioka, N., Suzuki, Y., Watanabe, S., Shinoda, A., Inoue, J., Lokman, P.M., Inagaki, T., Oya, M., Hasumoto, H., Kubokawa, K., Lee, T.W., Fricke, H. & Tsukamoto, K. (2001). Molecular approach to species identification of eggs with respect to determination of the spawning site of the Japanese eel *Anguilla japonica.* Fisheries Science, 67, 761-763.

Berois, N., Arezo, M. J., & Papa, N. G. (2011). Gamete interactions in teleost fish: The egg envelope. Basic studies and perspectives as environmental biomonitor. Biological Research, 44, 119–124.

Gleason, L. U., & Burton, R. S. (2012). High-throughput molecular identification of fish eggs using multiplex suspension bead arrays. Molecular Ecology Resources, 12, 57–66.

Dias, P. J., Wakefield, C. B., Fairclough, D. V., Jackson, G., Travers, M. J., & Snow, M. (2016). Real-time PCR validation of visually identified Snapper *Chrysophrys auratus* (Sparidae) eggs. Journal of Fish Biology, 88, 811–819.

Dimmlich, W. F., T. M. Ward, & W. G. Breed. (2009). Spawning dynamics and biomass estimates of an anchovy *Engraulis australis* population in contrasting gulf and shelf environments. Journal of Fish Biology, 75, 1560–1576.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research, 32(5), 1792-1797.

Fox, C. J., Taylor, M. I., Pereya, R., Villasana, M. I., & Rico, C. (2005). Taq-Man DNA technology confirms likely overestimation of cod (*Gadus morhua* L.) egg abundance in the Irish Sea: Implications for the assessment of the cod stock and mapping of spawning areas using egg-based methods. Molecular Ecology, 14, 879–884.

Hill, K.T., Crone, P.R., Zwolinski, J.P. (2017). Assessment of the Pacific sardine resource in 2018 for U.S. Management in 2018-19. NOAA Technical Memorandum NMFS. https://doi.org/10.7289/V5/TM-SWFSC-600

Jackson, G., & Cheng, Y. W. (2001). Parameter estimation with egg production to estimate snapper, Pagrus auratus, biomass in Shark Bay, Western Australia. Journal of Agricultural, Biological, and Environmental Statistics, 6, 243–257.

Karaiskou N, Triantafyllidis A, Alvarez P, Lopes P, Garcia-Vazquez E, Triantaphyllidis C. (2007). Horse mackerel egg identification using DNA methodology. Marine Ecology – An Evolutionary Perspective. 28, 429-34.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647-1649.

Lasker, R., 1985. An Egg Production Method for Estimating Spawning Biomass of Pelagic Fish: Application to the Northern Anchovy, *Engraulis mordax*. NOAA Technical Report NMFS 36.

Lelievre S, Verrez-Bagnis V, Jerome M, Vaz S. PCR-RFLP analyses of formalin-fixed fish eggs for the mapping of spawning areas in the Eastern Channel and Southern North Sea. (2011). Journal of Plankton Research, 32, 1527-39.

Murphy, H.M., Jenkins, G.P., Hamer, H.A. & Swearer, S.E. (2012). Interannual variation in larval survival of snapper (*Chrysophrys auratus*, Sparidae) is linked to diet breadth and prey availability. *Canadian Journal of Fisheries and Aquatic Sciences,* 69, 1340-1351.

Neira, F.J. & Lyle, J.M. (2011). DEPM-based spawning biomass of *Emmelichthys nitidus* (Emmelichthyidae) to underpin a developing mid-water trawl fishery in south-eastern Australia. Fisheries Research, 110, 236-243.

Neira, F. J., Perry, R. A., Burridge, C. P., Lyle, J. M., & Keane, J. P. (2015). Molecular discrimination of shelf-spawned eggs of co-occurring *Trachurus* spp. (Carangidae) in southeastern Australia: A key step to future egg-based biomass estimates. ICES Journal of Marine Science, 72, 614–624.

Oxley, P.A., Catalano, S.R., Wos-Oxley, M.L., Westlake, E.L., Grammer, G.L. & Steer, M.A. (2017). Using in situ hybridisation to expand the daily egg production method to new fish species. Molecular Ecology Resources, 17, 1108-1121.

Palumbi, Steve. "Simple fool's guide to PCR." (1991).

Pereira, F., Carneiro, J. & Amorim, A. (2008). Identification of species with DNA-based technology: current progress and challenges. Recent Patents on DNA & Gene Sequences, 2, 187-200.

Potts, W. T. W., & Eddy, F. B. (1973). The permeability to water of the eggs of certain marine teleosts. Journal of Comparative Physiology, 82, 305–315.

Stratoudakis, Y., Bernal, M., Ganias, K., & Uriarte, A. (2006). The daily egg production method: Recent advances, current applications and future challenges. Fish and Fisheries, 7, 35–57.

Teletchea, F. (2009). Molecular identification methods of fish species: reassessment and possible applications. Reviews in Fish Biology and Fisheries, 19, 265-293.

Ward, T. M., Burch, P., McLeay, L. J., & Ivey, A. R. (2011). Use of the Daily Egg Production Method for stock assessment of sardine, *Sardinops sagax*, lessons learned over a decade of application off southern Australia. Reviews in Fisheries Science, 19, 1–20.

Ward, T.M., Smart, J. & Ivey, A. (2017). Stock assessment of Australian Sardine (*Sardinops sagax*) Report to PIRSA Fisheries and Aquaculture. Sardine Publication No. F2017/000765-6. SARDI Research Report Series No. 971.

Ward, T.M., Grammer, G.L., Ivey, A.R., Smart, J.J. & Keane, J.P. (2018). Spawning biomass of Jack Mackerel (*Trachurus declivis*) and Sardine (*Sardinops sagax*) between western Kangaroo Island, South Australia and south-western Tasmania. Report to the Australian Fisheries Management Authority. SARDI Publication No. 2018/000174-1. SARDI Research Report Series No. 983.

Ward R.D., Zemlak T.S., Innes B.H., Last P.R. & Hebert P.D.N. (2005) DNA barcoding of Australia's fish species. Philosophical Transactions of the Royal Society of London B. Biological Sciences 360, 1847–1857.

Zeldis, J. (1998). A daily egg production method estimate of snapper biomass in Hauraki Gulf, New Zealand. ICES Journal of Marine Science, 55, 522–534.

7 Implications

This project identified complications in the practical application of molecular identification of fish eggs, especially using the bead array approach. In particular, it was clear that the online resources offered remarkably poor species coverage in the Great Barrier Reef Region. To increase confidence in probes and their specificity there is a need to improve and curate these resources.

Furthermore, the identification of fish eggs using molecular methods is stymied by the practicalities of preservation of field collected ichthyoplankton samples with substantial numbers of eggs unable to be identified due to PCR failures. These failures could result from either poor preservation or that "universal fish primers" are not truly universal. These issues can be overcome and rates of egg loss could be incorporated into any models of abundance but these would need significant resources to develop.

8 Recommendations

- The bead array method does have value for identifying multiple species in one pass but given the difficulties with the reliability of the chemistry the method would need to be applied regularly and on a large scale to be efficient.
- For application to DEPM, the authors consider that this approach is not suitable and that pursuit of alternative molecular egg identification methods such as *in-situ* hybridisation would be more appropriate, although this has its own hurdles (e.g. dechorionation) before it can be applied on scale.

9 Further development

- Development of approaches to dechorionation
- Preservation approaches that allow for accurate staging of early developmental stages of fish eggs
- Improved coverage of fish fauna for CO1, 16S and 12S in Genbank, particularly for commercial and non-commercial species in Australia

10 Extension and Adoption

Previous extension activities

- Presentation to stakeholders: Saunders, R.J., Penny, S. & Jerry, D.J. (2014). Molecular identification of fish eggs. Invited presentation. International technical workshop and stakeholder forum on small pelagic fisheries. Adelaide, South Australia. 14-18 July 2014.
- Presentation to Northern Territory RAC (2017)

Planned extension activities

- Presentation to NT and Qld fishery managers
- Report circulated to Northern Territory and Queensland fishery managers
- Report circulated to Australian agencies using DEPM (e.g. SARDI & IMAS)

11 Project coverage

Newspaper article: “Casting a beady eye at fish eggs. Cairns Post. 5th December 2014.

Appendices

Appendix A1.

Method optimization

Carbodiimide coupling of amine-modified oligonucleotides to carboxylated microspheres

IMPORTANT: Microspheres should be protected from prolonged exposure to light throughout this procedure.

1) Take stock uncoupled microsphere vials out of fridge and warm to room temperature
2) Take the EDC powder out of the -20C freezer and bring to room temperature in a desiccator
3) Weight in two times 10 mg EDC in a 1.5 mL tube and place into desiccator
4) Mix uncoupled microsphere stock vials by gently inverting by hand for 1 min
5) Transfer 85 µL uncoupled microsphere stock (~ 10^6 beads) into a 1.5 mL microfuge tube
6) Centrifuge microspheres for 2 minutes at 8000 x g
7) Remove the supernatant and resuspend the pellet in 10 µL of 0.1 M MES pH 4.5 by vortexing.
8) Add 1 µL (0.1 nmole) of the 0.1 mM probe to the resuspended microsphere and mix by vortexing
9) Add 1 mL of dH_20 to the 10 mg of EDC
10) Add 2.5 µL of freshly made 10 mg/mL EDC to the microspheres (final conc. 0.5 µg/µL) and vortex
11) Incubate for 30 min at room temperature in the dark
12) Repeat steps 9 to 11
13) Add 0.5 mL of 0.02% Tween-20 to the coupled microspheres.
14) Centrifuge microspheres for 2 minutes at 8000 x g
15) Remove supernatant and resuspend the coupled microspheres in 0.5 mL of 0.1% SDS and vortex
16) Centrifuge microspheres for 2 minutes at 8000 x g
17) Remove supernatant and resuspend the coupled microspheres in 20 µL TE, pH 8.0 and vortex

DNA hybridization to microspheres

IMPORTANT: Microspheres should be protected from prolonged exposure to light throughout this procedure.

1) Resuspend the microspheres by vortexing for approximately 20 seconds
2) Prepare a Working Microsphere Mixture by diluting 3 µL coupled microsphere in 597 µL 1.5X TMAC Hybridization Solution <!>TOXIC<!>
3) Mix the Working Microsphere Mixture by vortexing
4) Add 46 µL of Working Microsphere Mixture to a 200 µL PCR tube per sample
5) Add 4 µL of biotin-labelled DNA per sample
6) Mix reaction wells by vortexing
7) Incubate the samples in a thermocycler with the following program:

    95°C for 2 min

    57°C for 120 min

8) Prepare fresh Reporter Mix by diluting 1 µL streptavidin-R-phycoerythrin with 249 µL of 1X TMAC Hybridization Solution <!>TOXIC<!>
9) Add 15 µL SPE to each sample, mix well, and incubate 10 min at 57°C
10) Transfer samples to 1.5 mL microfuge tubes and centrifuge at 8,000 x g for 2 min
11) Remove supernatant and discard in amber bottle; add 400 µL Isoflow to each sample
12) Repeat steps 10 and 11.

For validation, samples were run on a Fortessa Flow Cytometer using Forward and Side scatter to generally identify single beads, using APC and PE-Cy7 channels to identify bead Ids, and PE to detect the presence of the target species.

Methods

In initial optimisation, a subset of positive control samples were purified using streptavidin coated magnetic microparticles, to confirm that the PCR amplification was successfully adding a biotinylated end to each fragment. This process only allows DNA fragments with a biotinylated end to attach to the beads, washing any unmodified fragments away. To assess if inhibitors were present in PCR products, and if this was affecting probe binding efficiency, eight samples of each streptavidin bead cleaned, carboxylate modified bead cleaned, and uncleaned PCR products, were processed on the Magpix platform.

To examine if different bead concentrations effected the florescence readings, two additional final bead concentrations (177 beads/µl and 88.5 beads/µl) were tested.

To determine which hybridisation temperature was optimal for each species specific probe, a gradient hybridisation was conducted between 47°C and 57°C, at ~1°C increments, using eight replicates for each positive control sample species.

Results

Clean up trials

To confirm that a biotinylated attachment site was successfully incorporated, a subset of 8 samples were amplified for each gene of interest, and subsequently cleaned using streptavidin coated magnetic beads. Following cleaning and visualisation on an agarose gel, strong bands were observed in both clean and uncleaned product, indicating that the PCR had successfully incorporated a biotinylated end. However, no discernible difference was observed when these samples were processed through on the Magpix platform, with no significant difference being observed in median fluorescence between target samples (both clean and unclean), and the negative control (Figure 9a & b).

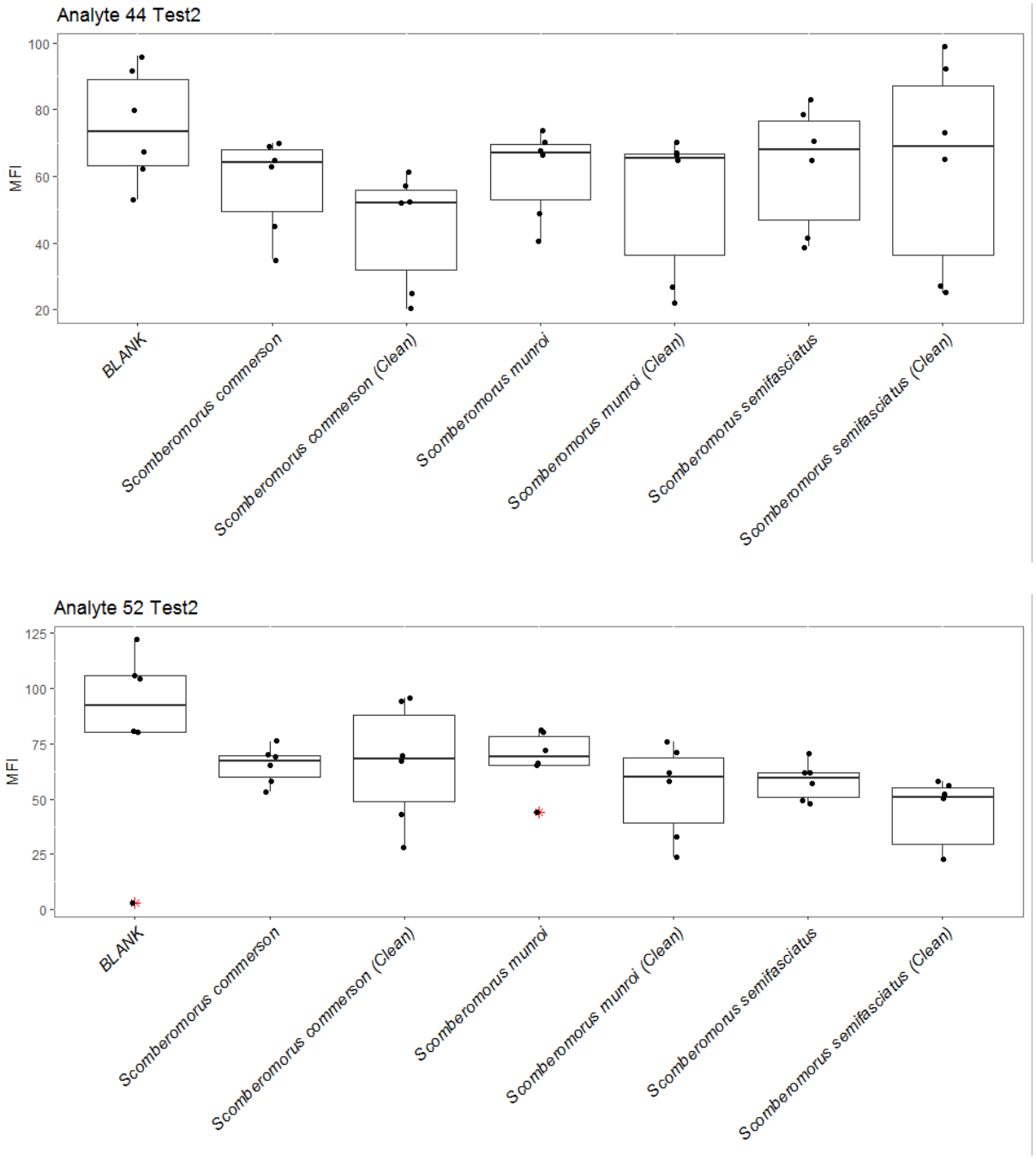
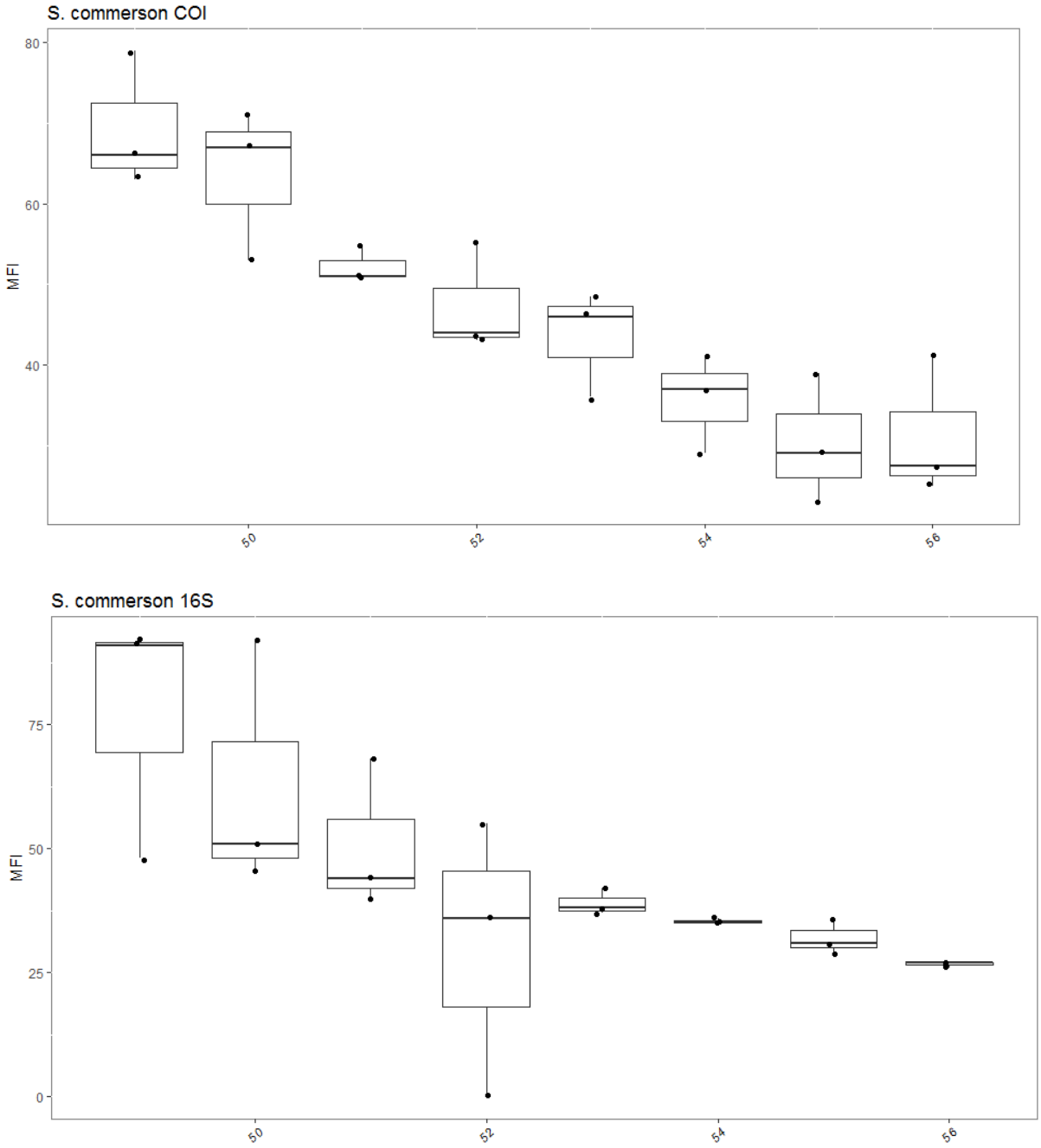


Figure 9a & b. Comparson of median flourescence values between samples cleaned using streptavidin coated magnetic beads, and uncleaned samples.

Optimisation of species specific probes, and Hybridisation of probe to sample and attachment of fluorescent marker (SPE)

Temperature trials

To determine appropriate hybridisation temperatures for the target species probes, samples from both *S. semifasciatus*, and *S. commerson* (both 16S and COI) were hybridised to their respective probes at temperatures ranging from 49°C – 56°C in ~1°C increments. Generally, median fluorescence increased with lower hybridisation temperatures (Figure 10a - d), likely as a result of more non-specific binding. No clear trend was observed when changing hybridisation temperatures of the *S. semifasciatus* COI probe (Figure 10c), with all temperature treatments having a high median fluorescence value (~ 7 fold higher than all other probes).



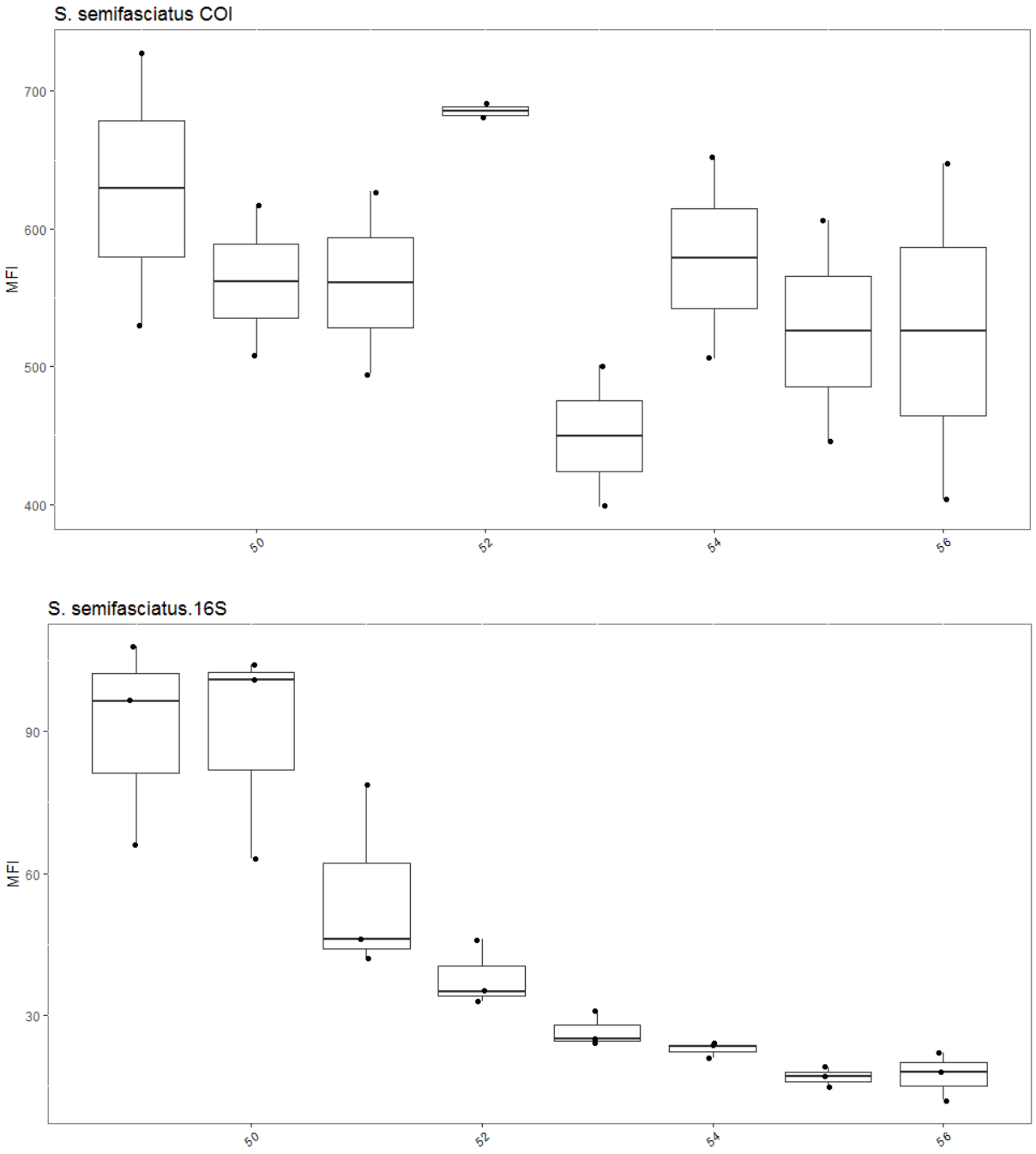


Figure 10 a-d. Mean fluorescence values of samples hybridised to probes at different temperatures (x-axis, °C), ranging from 49-57°C for two target species; *S. commerson* and *S. semifasciatus.*

Dilution of bead stocks (counts)

Initial trials indicated that >1,000 beads were sampled from each well when 50 µL was set as the collection volume, with no dilution. The Magpix platform recommends a minimum of 50, and a maximum of 250 bead to be read at each well. To examine if bead counts significantly affected

median fluorescence readings, three dilution factors were compared to fluorescence values. The three dilutions did not return discrete bead count readings across all samples, with a continuous being observed instead. No significant differences in median fluorescence were observed when bead counts were within between 1-1,000 beads/50ul, however to ensure confidence in each reading, fluorescence values were only retained if >40 beads/events were recorded for each sample.

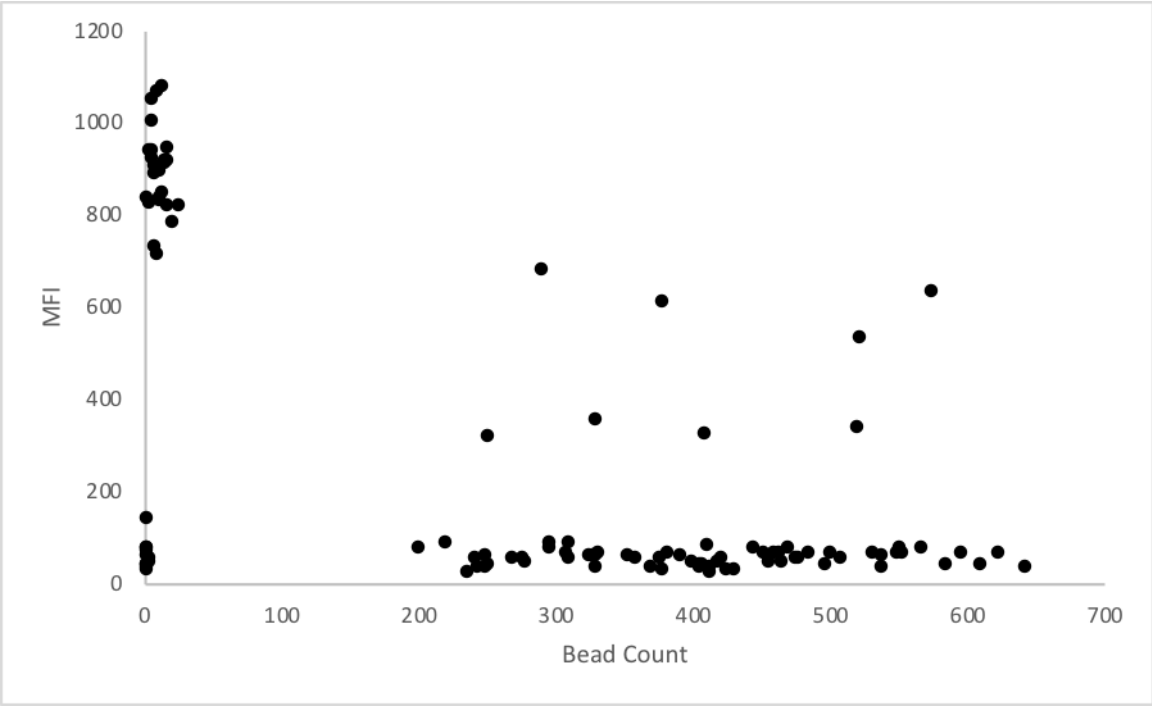


Figure 11. Scatter plot of mean fluorescence values correlated with a range of bead counts

Appendix 2.

Sequencing results 2016 egg survey. Pairwise match was set for species at <99% species, 97% to 99% for Genus, 95% to 97% for Family and 75% to 95% Order. Below 75% was undetermined.

Species	Species	Genus	Family	Order
Abudefduf vaigiensis				1
Acanthurus leucosternon		4		
Acanthurus lineatus	5			
Acanthurus nigrofuscus	1			
Alepes kleinii			2	
Amblygaster sirm		5		
Anampses twistii		1		
Arnoglossus polyspilus				1
Auxis rochei	3			
Bolbometopon muricatum	14			
Bothus myriaster	4			
Brachyhypopomus occidentalis				1
Bregmaceros nectabanus				10
Caesio caerulaurea	26			
Caesio cuning	32	1		
Callionymus enneactis				1
Carangoides armatus		1		
Carangoides malabaricus		5		
Cephalopholis boenak	2			
Cephalopholis miniata	2			
Cephalopholis sonnerati	6			
Cetoscarus bicolor	9			
Chaetodon baronessa	6			
Chaetodon lunulatus	2			
Chaetodon octofasciatus	17	5		
Chaetodon rainfordi	1			
Chaetodontoplus mesoleucus		2		
Cheilinus oxycephalus	4			
Cheilinus trilobatus	2			
Chelidonichthys capensis		1		
Chelmon rostratus		3		
Chilara taylori				1
Chlorurus microrhinos	5			

Chlorurus sordidus	25	2			
Chlorurus spilurus	1				
Choerodon frenatus	2				
Choerodon sugillatum	2				
Coilia mystus					24
Coilia reynaldi					2
Coris gaimard	1				
Cromileptes altivelis	2				
Ctenochaetus binotatus	4				
Decapterus macarellus	1				
Decapterus macrosoma	20				
Decapterus maruadsi	42	1			
Dendrochirus brachypterus	1				
Dendrophysa russelii					1
Deveximentum megalolepis	6				
Diagramma labiosum	4				
Diagramma picta	1				
Dussumieria elopsoides					2
Ellochelon vaigiensis	4				
Encrasicholina heteroloba					2
Engraulis australis					1
Engyprosopon maldivensis	4				
Engyprosopon sp. BD-2002				2	2
Epibulus insidiator	8				
Epinephelus areolatus	1				
Epinephelus malabaricus	2				
Epinephelus ongus	1				
Equulites elongatus		3			
Equulites leuciscus			25	1	
Euthynnus affinis	20				
Euthynnus lineatus	1				
Fistularia commersonii	1				
Gerres oyena					1
Gomphosus varius	3				
Gymnothorax minor	1				
Gymnothorax pictus					3
Gymnothorax reticularis		2			
Halichoeres hortulanus	1				
Halichoeres melanurus					1
Halichoeres nebulosus	1				
Halichoeres nigrescens	4				

Halichoeres trimaculatus	1				
Hemigymnus fasciatus	1				
Hemigymnus melapterus	8				
Hemiodopsis gracilis	1				
Hipposcarus longiceps	11				
Hyperoglyphe antarctica					10
Iniistius aneitensis					36
Labrichthys unilineatus	2				
Labroides dimidiatus	1				
Lactoria diaphana					1
Lates calcarifer	1				
Leptojulis cyanopleura	1				
Lethrinus laticaudis				1	2
Lethrinus lentjan	2				
Lethrinus nebulosus		1	1	2	12
Lethrinus obsoletus	5				
Lethrinus sp. Kakeroma	1				
Lophonectes gallus					2
Lutjanus argentimaculatus	1				
Lutjanus carponotatus	9				
Lutjanus erythropterus	2	13			
Lutjanus fulviflamma	4				
Lutjanus malabaricus	1	2			
Lutjanus quinquelineatus	1				
Lutjanus rivulatus					3
Lutjanus russellii	1				
Lutjanus vitta		1		2	
Macquaria ambigua					1
Monotaxis grandoculis					28
Naso brevirostris	1				
Nemipterus bipunctatus	3				
Nemipterus furcosus	5				
Nemipterus peronii	8				
Notesthes robusta					2
Novaculichthys taeniourus	1				
Nuchequula decora	1				
Ostracion rhinorhynchos					1
Oxycheilinus digramma	9				
Parapercis haackei					1
Parastromateus niger				12	
Pegasus volitans	1				

Pentapodus aureofasciatus					42
Pentapodus nagasakiensis	1				
Pentapodus setosus		4			
Plectorhinchus gibbosus	1				
Plectorhinchus lineatus	1				
Plectropomus leopardus	4				
Pomadasys kaakan	1				
Psammoperca waigiensis		1			
Pseudanthias hypselosoma		1			
Pseudocheilinus hexataenia		2			
Pseudorhombus pentophthalmus					1
Pterocaesio digramma	1				
Pterocaesio marri	29				
Pterocaesio pisang	11				
Pterocaesio tile	1				
Rogadius patriciae	2	1			
Sander lucioperca					2
Sardinella fimbriata	8				
Sardinella hualiensis				1	
Sardinops melanostictus	3				
Sardinops sagax	10				
Sargocentron rubrum	1				
Saurida undosquamis			3	3	
Saurida wanieso				2	
Scarus aff. ghobban SY	1				
Scarus festivus	2				
Scarus flavipectoralis	25				
Scarus forsteni	1				
Scarus frenatus	2				
Scarus globiceps	23	2			
Scarus niger	28	1			
Scarus psittacus	18				
Scarus schlegeli	4				
Scarus spinus	5				
Scolopsis bilineata	1				1
Scolopsis monogramma			1	2	1
Scomberoides commersonnianus			1	1	
Scomberomorus commerson	3				
Scomberomorus munroi	1				
Scopeloberyx robustus					4
Scorpaena scrofa					1

Scorpaenodes scaber					1
Scorpaenopsis cirrosa					1
Sebastes alutus					1
Selenotoca multifasciata	13				
Sillago asiatica				4	
Sorsogona tuberculata	1				
Sphyraena barracuda					3
Stethojulis bandanensis	3				
Stethojulis strigiventer		2			
Stolephorus chinensis					3
Stolephorus waitei					2
Symphorichthys spilurus	2				
Symphorus nematophorus	8				
Synchiropus splendidus					2
Synodus kaianus	1				
Synodus lacertinus					2
Synodus variegatus	1				1
teleost environmental sample	32	5			4
Terapon theraps	41				
Thalassoma lunare	1				
Trachinocephalus myops	1				
Upeneus japonicus				8	
Upeneus tragula		7		18	
Uranoscopus scaber					1

Appendix 3.

Species ID allocations for each species identified through sequencing of planktonic samples (refer to Figure 9 in body text).

Sample	*Sample_ID*
Acanthurus lineatus	1
Amblygaster sirm	2
Auxis rochei	3
Bolbometopon muricatum	4
Bothus myriaster	5
Brachyhypopomus occidentalis	6
Bregmaceros nectabanus	7
Caesio caerulaurea	8
Caesio cuning	9
Callionymus enneactis	10
Carangoides malabaricus	11
Cephalopholis boenak	12
Cephalopholis miniata	13
Cephalopholis sonnerati	14
Chaetodon octofasciatus	15
Chaetodon rainfordi	16
Chaetodontoplus mesoleucus	17
Chelmon rostratus	18
Chilara taylori	19
Chlorurus sordidus	20
Choerodon frenatus	21

Choerodon sugillatum	22
Ctenochaetus binotatus	23
Decapterus macrosoma	24
Decapterus maruadsi	25
Dendrochirus brachypterus	26
Diagramma labiosum	27
Dussumieria elopsoides	28
Ellochelon vaigiensis	29
Encrasicholina heteroloba	30
Epinephelus areolatus	31
Equulites elongatus	32
Equulites leuciscus	33
Euthynnus affinis	34
Fistularia commersonii	35
Gerres oyena	36
Gymnothorax pictus	37
Gymnothorax reticularis	38
Halichoeres melanurus	39
Hemigymnus melapterus	40
Hemiodopsis gracilis	41
Hyperoglyphe antarctica	42
Lactoria diaphana	43
Lethrinus laticaudis	44
Lethrinus nebulosus	45
Lethrinus obsoletus	46
Lophonectes gallus	47
Lutjanus carponotatus	48
Lutjanus erythropterus	49

Lutjanus fulviflamma	50
Lutjanus malabaricus	51
Lutjanus quinquelineatus	52
Lutjanus rivulatus	53
Monotaxis grandoculis	54
Nuchequula decora	55
Ostracion rhinorhynchos	56
Oxycheilinus digramma	57
Parastromateus niger	58
Pegasus volitans	59
Pentapodus aureofasciatus	60
Pentapodus nagasakiensis	61
Pentapodus setosus	62
Plectorhinchus gibbosus	63
Plectropomus leopardus	64
Pomadasys kaakan	65
Psammoperca waigiensis	66
Pseudanthias hypselosoma	67
Pseudocheilinus hexataenia	68
Pterocaesio digramma	69
Pterocaesio marri	70
Pterocaesio pisang	71
Sardinella hualiensis	72
Sargocentron rubrum	73
Saurida wanieso	74
Scarus flavipectoralis	75
Scarus globiceps	76
Scarus niger	77

Scarus psittacus	78
Scarus spinus	79
Scomberomorus commerson	80
Scomberomorus munroi	81
Scopeloberyx robustus	82
Scorpaena scrofa	83
Scorpaenodes scaber	84
Scorpaenopsis cirrosa	85
Sebastes alutus	86
Selenotoca multifasciata	87
Sillago asiatica	88
Sorsogona tuberculata	89
Sphyraena barracuda	90
Stolephorus chinensis	91
Stolephorus waitei	92
Symphorus nematophorus	93
Synchiropus splendidus	94
Synodus kaianus	95
Synodus lacertinus	96
Synodus variegatus	97
teleost environmental sample	98
Terapon theraps	99
Thalassoma lunare	100
Trachinocephalus myops	101
Unknown	102
Upeneus japonicus	103
Upeneus tragula	104
BLANK	105

FRDC FINAL REPORT CHECKLIST

Project Title:	Developing a rapid molecular identification technique to improve egg production based fish biomass assessments		
Principal Investigators:	Richard J. Saunders, Shannon Kjeldsen, Roger Huerlimann, Thor Saunders, Shane Penny, Andrew Tobin and Dean Jerry		
Project Number:	2014-022		
Description:	Brief one/two paragraph overview of what the project did and achieved.		
Published Date:	2/7/19	**Year:**	2019
ISBN:	978-0-6485037-0-5	**ISSN:**	n/a
Key Words:	Molecular identification, fish eggs, Spanish Mackerel, Grey Mackerel, Goldstripe Sardinella, Spotted Sardine, bead array, flow cytometer		

Please use this checklist to self-assess your report before submitting to FRDC. Checklist should accompany the report.

	Is it included (Y/N)	**Comments**
Foreword (optional)	N	
Acknowledgments	Y	
Abbreviations	Y	
Executive Summary	Y	
– **What the report is about**	Y	
– **Background – why project was undertaken**	Y	
– **Aims/objectives – what you wanted to achieve at the beginning**	Y	
– **Methodology – outline how you did the project**	Y	
– **Results/key findings – this should outline what you found or key results**	Y	
– **Implications for relevant stakeholders**	Y	
– **Recommendations**	Y	
Introduction	Y	
Objectives	Y	
Methodology	Y	
Results	Y	
Discussion	Y	
Conclusion	Y	
Implications	Y	

Recommendations	Y	
Further development	Y	
Extension and Adoption	Y	
Project coverage	N	
Glossary	N	
Project materials developed	N	
Appendices	Y	