Biodiversity monitoring using environmental DNA: Can it detect all fish species in a waterbody and is it cost effective for routine monitoring?

Lia Smith

*Edith Cowan University*

Follow this and additional works at: [https://ro.ecu.edu.au/theses](https://ro.ecu.edu.au/theses)

Part of the *Ecology and Evolutionary Biology Commons*, and the *Genetics and Genomics Commons*

**Recommended Citation**


This Thesis is posted at Research Online.

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.

- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author’s moral rights contained in Part IX of the Copyright Act 1968 (Cth).

- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
Biodiversity monitoring using environmental DNA: can it detect all fish species in a waterbody and is it cost effective for routine monitoring?

This thesis is presented in partial fulfilment of the degree of Master of Biological Science

Lia SMITH

Edith Cowan University
School of Science
2017
Abstract

The challenges associated with environmental monitoring such as the impact on the environment and the financial costs are problems we face when trying to conserve freshwater systems around the world. The need for precise and accurate results that are cost effective is important so that we can achieve our conservation goals.

The overall aim of this study was to explore Next-Generation - metabarcoding for the detection of feral and native freshwater fish species based on the DNA shed by individual organisms into the water column. Cytochrome c oxidase I (COI) primers were developed for this study using DNA from six freshwater species expected to be found in the waterbody. These primers, along with 16S rRNA (16S) primers, were assessed to ensure that the molecular method was robust and suitable for use in the field. Along with the cost effectiveness of the molecular method when compared to the more traditional surveying method of Fyke net surveying.

This study comprised development of field and lab protocols for the detection of freshwater fish species in a lentic system. Both the COI and 16S primer sets showed results that were comparable to previous Fyke net surveys, though both primer sets detected species that the other did not. Further qPCR analysis showed that there were differences in detection for both primers for each of the species. The molecular surveying of the waterbody has been proven sensitive enough to detect *Maccullochella peelii*. This species has a very low abundance in the waterbody (believed to be n=1) so these results suggest that this method can be used to target low abundance species.

The outcome of this study highlighted the need for multi-location sampling within a waterbody as increasing the number of locations sampled, led to an increase in the number of species detected. Along with the multi-location sampling, it was also important to sample throughout the year to account for seasonal variability. The eDNA study emphasized the importance of having knowledge of both the ecology and the biology of the species targeted so that a robust monitoring method can be implemented.
As well as comparing the apparent accuracy of Fyke netting and the eDNA approach in the study waterbody, a cost benefit analysis comparing the relative costs of multiplex DNA surveying, single species molecular surveying, and Fyke net surveying was undertaken. Molecular environmental surveying was found to be a cost effective method for monitoring, as the analysis suggested single species monitoring would break even after only 95 waterbodies were surveyed, and multiplex surveying would break even after 145 waterbodies, under the proposed scenario. The cost benefit analysis explored the costs associated with all three methods, including lab set up costs, along with the number of waterbodies that could be surveyed on both a weekly and yearly basis.
Copyright and Access Declaration

I certify that this thesis does not, to the best of my knowledge and belief:

i. incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education;

ii. contain any material previously published or written by another person except where due reference is made in the text of this thesis;

iii. or contain any defamatory material;

Signed…

Date 28/03/2017
USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.
Acknowledgements

I have so many people I NEED to thank; everyone I have met on this journey has helped me in some way.

To my supervisors, where would I have been without you?

Annette, thank you for having faith in me throughout this project, even when I had little faith in myself. Thank you for guiding me to be the confident and poised researcher, writer, and scientist. Thank you for taking a risk with an overeager, enthusiastic undergrad who just wanted to volunteer and do some more molecular research.

Quinton, you have made things easier with your explanations, help, and support and for that, I am eternally grateful. I am also grateful for your patience, your perseverance, and your dry sense of humour and your ‘can do attitude’ even when mine disappeared and I felt the end would never arrive.

Thank you for both teaching and pushing. I know that this project has been draining, overwhelming and frustrating, but I know that there are not many other people that I would like to see week in and week out so thank you. Thank you both for being there in the ups and in the downs. Thank you for pushing me to think outside of the boundaries and helping me to do that little bit better.

Mike, thank you for your support, your expertise, your knowledge, and your facilities. Without you, this project would not have gotten off the drawing board so I am forever grateful that I had this opportunity to work with you. I hope that the outcome of this project is something that you can implement and use in the future.

To my absolute favourite Science teacher of all time Dennis Haley, whose passion for science and not wearing shoes inspired me to study as much Science as I could, to my rock though the years at S.I.D.E Gerry Nolan, thank you both.
Another group of people who I must thank and who without them I would not have gotten this far, my family. My mum, who was more than just my mum, she has been my mum, my teacher (literally), my friend, and the woman who has kicked me up the butt multiple times and told me ‘if you don’t do it now, you will regret it later’. Thanks mum, for everything. To dad who showed me that art was not my thing (lol) and helped me copper plate things back in high school science, thank you. Also, thank you to my cousin Chris. Thank you for putting up with me that first year at Uni, I know it was hard having the straight-laced Nerd in your midst, but thank you. Thank you for always being there, 15+ years is a long time but somehow, I have not managed to shake you off just yet.

To my sestra, Tanya, thank you so much for all your support, without you I would not have had the ability to finish this degree. Thank you for putting a roof over my head, a smile on my face and love in my heart.

To the staff and my peers at ECU thank you for everything, without you this project would never have been completed. Charlie, Caitlin, Shanuka, Kate, Jess, Shanna, Mel, Casper, Mike and all the other post-grads who have listened to me complain, given me inspiration and helped me through the roughest of days, THANK YOU!!!!

David, your grammar support and encouragement has not gone unnoticed, thank you so much!

I could not have done this without you ALL!
Contents

Abstract ........................................................................................................................................ iii
Copyright and Access Declaration ...................................................................................... v
Acknowledgements .......................................................................................................... vii
Contents .................................................................................................................................... ix
List of Figures ....................................................................................................................... xi
List of Tables ......................................................................................................................... xii

1. Introduction ...................................................................................................................... 1
   1.1 Introduced fish species: the global context ............................................................. 1
   1.2 Introduced fish species: the local context for Perth, Western Australia ............... 3
   1.3 Surveying methods for detecting fish species .................................................... 5
       a) Traditional approach ....................................................................................... 6
       b) Molecular approach – single species detection .............................................. 7
       c) Molecular approach – multispecies detection via metabarcoding............... 8
       d) Comparison of molecular methods .................................................................. 9
   1.4 Importance of this project ..................................................................................... 10

2. Method Development .................................................................................................... 13
   2.1 Sample collection and storage ............................................................................. 13
   2.2 PCR, qPCR and Next Generation Sequencing (NGS) Primer design ................ 17
       a) in-silico development .................................................................................... 17
       b) Generation of reference DNA sequences from target species to support subsequent species identification from NGS data ...................................................... 18
   2.3 DNA extraction from environmental samples .................................................... 19
   2.4 DNA quantification to optimise input DNA into NGS ......................................... 19
   2.5 NGS preparation .................................................................................................. 20
       a) Fusion tagging .................................................................................................. 20
       b) Pooling and purification of samples for NGS .................................................. 21
       c) Dilution and pre-MiSeq quantification ............................................................ 22
   2.6 Next-Generation Sequencing ............................................................................ 22
   2.7 Bioinformatics and statistical analysis ................................................................ 22
   2.8 Application of developed method at the Department of Fisheries ..................... 23
List of Figures

Figure 2.1 The southwestern region of Western Australia, including the study site. Alexandria Lake (Canning Vale, Western Australia), showing the location of the three water collection sites of this study. Image from Google Maps. ................................. 13

Figure 3.1 Single source freshwater fish DNA qPCR amplification curves. qPCR was conducted on each primer set in triplicate to identify the number of cycles needed to detect each of the species using single source DNA. *Cyprinus carpio* (blue), *Gambusia holbrooki* (teal), *Leiopotherapon uniclor* (pink), *Pseudogobius olorum* (red), *Maccullochella peelii* (orange), and *Carassius auratus* (purple) 16S (a) and COI (b). .................................................. 25

Figure 3.2 Mean C_t values for six freshwater fish testing the 16S and COI primer sets for the number of cycles needed to detect single source DNA using qPCR. Open markers designate species not detected in environmental samples. ......................................................... 27

Figure 3.3 The total number of sequences detected at each location from 250 mL water collected, across five sampling periods, for both the 16S and COI markers, showing standard deviation for each location. ................................................................................. 30

Figure 3.4 Sequence counts for each location, surface, and depth, across the five sampling periods (16S (a) and COI (b)). Winter and winter 2 are repeat samplings one week apart, with rainfall occurring between weeks. ................................................................. 31

Figure 3.5 Water chemistry measurements (Dissolved Oxygen (mgL-1) (a), Conductivity (µS/cm) (b), pH (c) and (d) Water temperatures (°C)) taken over four seasons over a 12 month period at each sample. Winter values are the mean of two replicates measured 1 week apart. .............................................................................................................. 34

Figure 3.6 Freshwater fish species detected at each of the locations. The average number of identified unique sequences per location was 1375 with a standard deviation of ±344.2. Unique sequences are determined with 100% query coverage and a threshold of 98% maximum identity to allow for a 2% dissimilarity threshold. ....................... 36

Figure 3.7 Number of sequences detected for each species in each of the sampling periods. The mean number of identified unique sequences per season was 1556 with a standard deviation of ±897.8. Unique sequences are determined with 100% query coverage and a threshold of 98% maximum identity to allow for a 2% dissimilarity threshold. .... 37

Figure 5.1 Cumulative time to sample waterbodies for presence of fish species not including lab, or Fyke netting set up costs. A comparison of three methods of sampling showing the number of waterbodies sampled weekly for 40 weeks (~38-hour week). Molecular (multiplex) and molecular (single) both include field and laboratory. Horizontal lines shows the number of waterbodies that can be surveyed by each method in a year. Number of waterbodies rounded to the nearest 10. ...................... 59

Figure 5.2 Comparison between the three methods including start-up costs. Showing break-even intercepts of both molecular (multiplex) surveying and molecular (single target) surveying with traditional Fyke net surveying. ................................................................. 60

Figure 5.3 Cumulative costs for each method of detection over one year of surveying. Showing intercepts for each three methods including yearly cost and the number of waterbodies surveyed. Dollar values rounded to the nearest 1000 (AUD), number of waterbodies rounded to the nearest 10. ................................................................. 61

Figure 6.1 Invasion processes and management options (Invasive Species Council, 2010) .................................................................................................................. 65
List of Tables

Table 1.1 Non-native freshwater fish species established in Western Australia. Year of major/initial introduction or first record from the wild and reason for introduction. Table adapted from (Lintermans, 2004), with * indicating data from (Fletcher & Santoro, 2014)............................................................................................................ 4

Table 2.1 eDNA sample details. Including collection code, season, site location, surface, or depth along with where the DNA was extracted. (DoF – Department of Fisheries, ECU – Edith Cowan University, Cv – Canning Vale)..................................................... 14

Table 2.2 Primer pairs used for the detection of freshwater fish species in a waterbody located in Canning Vale. Universal fish primers for 16S were sourced from Dr. Michael Bunce (Curtin University (Unpublished)), COI primers were developed specifically for Western Australian freshwater fish................................................. 17

Table 3.1 16S Tukey HSD results. Sample size n = 3...................................................... 25

Table 3.2 NanoDrop results showing the concentration of DNA from each of the six single source freshwater fish species used during this project ........................................ 26

Table 3.3 All species detected from waterbody samples during NGS study. ................. 28

Table 3.4 Species detected by each of the markers and a comparison between molecular surveying results and traditional fyke net surveying................................................. 29

Table 5.1 Fyke netting maximum number of waterbodies that can be surveyed each week by a team of two working together in the field. Each person works the same number of hours (h).......................................................................................................................... 57

Table 5.2 Total number of waterbodies multiplexed surveyed each week, combining lab and field team hours........................................................................................................ 58

Table 5.3 Total number of waterbodies single target surveyed each week, combining lab and field team hours.................................................................................................. 58

Table 5.4 Weekly and yearly break down of cumulative hours and costs needed to survey waterbodies using all three methods, including weekly lab time. Based on a 38-hour week and a 1520-hour year................................................................. 62
1. Introduction

Globally, populations of wild flora and fauna are in decline due to anthropogenic disturbances (Brown & Paxton, 2009; Butchart et al., 2010). In the past five centuries, anthropogenic disturbances have started the sixth mass extinction, which may be comparable in rate and magnitude to the previous five mass extinctions recorded in Earth’s history (Barnosky et al., 2011). These extinction rates are impacting both human health and the sustainable future of the planet (Diaz et al., 2006). Overall, between 1970 and 2010 there was a biodiversity decline of 52%, with this loss being greatest in freshwater ecosystems (World Wildlife Fund, 2014). To prevent this decline from continuing, it is important to be able to identify accurately which species are present, along with species distribution, diversity and population sizes on a relevant ecological and political scale so that monitoring and management strategies can be implemented (Butchart et al., 2010; Novacek, 2001; Possingham et al., 2001).

Perhaps the biggest question we face when we wish to conserve biota is how can we manage and protect biodiversity when there are increasing financial pressures (such as the increasing cost of monitoring, along with the environmental and financial cost of invasive species impacts) when the need for precise and accurate results can be a costly endeavour? Alternatively, stated more succinctly, how can we maximise the return on financial investment so we can achieve our conservation goals? A useful case study of the increasing detrimental impacts of anthropogenic actions and the issue of how to invest funds in the most effective and efficient methods can be found in freshwater aquatic environments.

1.1 Introduced fish species: the global context

High levels of endemism are found in Mediterranean climate regions around the world, including southwestern Western Australia (Marr et al., 2010; Marr et al., 2013). Despite this, all Mediterranean climate regions around the world currently have more non-native fish species than endemics and despite being separated by huge distances and originally having few species in common, they now all contain similar species (Marr et al., 2010; Marr et al., 2013). One hundred and thirty-six species of fish from 26 families have been introduced into Mediterranean regions around the globe and recorded as established (Marr et al., 2013).
Introductions in these regions have been strongly driven by taxonomically biased anthropogenic interests such as recreational fishing, aquaculture, and ornamental pet species release, and illustrates the extent and importance of taxonomic homogenisation (Marr et al., 2010; Marr et al., 2013). The risk of further introductions continues due to increased interest in angling, poor public awareness of impacts of non-native introductions and difficulty in enforcing bans and penalties for non-native introductions (Beatty & Morgan, 2013; Beatty et al., 2011; Gill et al., 1999; Gozlan, 2008; Gozlan et al., 2010; Marr et al., 2013; Morgan, Allen, et al., 2011).

Homogenisation of taxa occurs when a few species dominate communities to the detriment of unique native species (Rahel, 2007) and an ever-increasing homogenisation of both flora and fauna on a global scale has been widely noted (Blair, 2001; Leprieur, Beauchard, Hugueny, et al., 2008; Lockwood & McKinney, 2001; Marchetti et al., 2001; Scott & Helfman, 2001). Within aquatic organisms, homogenisation is best documented in fish populations due to a small number of common species being introduced for reasons such as sport fishing, aquaculture, or ornamental releases (Rahel, 2007). For example, in areas close to reservoirs and where introduced fish species are abundant, homogenisation of native assemblages is strong (Hermoso et al., 2012), and a survey of non-native fish introductions into Europe, North America, Australia, and New Zealand, found native fish species populations were reduced or eliminated in 77% of cases (Ross, 1991).

It appears the introduced fish often meet the criteria for being considered an alien/invasive species (i.e. an alien species whose introduction and/or spread threaten biological diversity (Convention on Biodiversity, 2002)). Despite the well-documented detrimental effects of their release they will likely continue to be released for many reasons including accidentally through human transport, illegally through human intent, or legally via stocking for recreational fishing or as a biotic control (Cucherousset & Olden, 2011; Lintermans, 2004). Invasive species have been noted as being able to remodel ecosystems through habitat modification, introduce disease and parasites, and cause population fragmentation and localised extinction of endemic species along with the associated loss of genetic diversity (Allendorf & Lundquist, 2003; Corfield, 2008; Thurner, 2015; Townsend, 1996).
Ecological changes encourage the extension of geographic regions for some (often-introduced) species and reduction of others, which leads to biotic homogenisation. Introduction of disease and pathogens by food fish is widely documented (Austin & Austin, 2007; Brown & Gratzek, 1980; Corfield, 2008; Rimmer et al., 2015; Thurner, 2015) though there have been few studies showing impacts of invasions by ornamental fish (Ashburner, 1976; Langdon, 1988; Tompkins et al., 2015; Torchin et al., 2002).

1.2 Introduced fish species: the local context for Perth, Western Australia

Perth, Western Australia, is located in a Mediterranean climate region which is characterised by dry summers and mild, moist winters (Lionello et al., 2006; Peel et al., 2007). In Australia, 74% of the native species are endemic to the country and while no fish species has been listed as extinct since 1788 (Lintermans, 2013), it is not known if unknown species have been lost. Introduction and spread of freshwater fish is typically linked with a loss of native species (Clavero et al., 2004; Hermoso López et al., 2010; Leprieur, Beauchard, Blanchet, et al., 2008; Marr et al., 2010). While freshwater fish have not been comprehensively studied for the IUCN Red List, of the 169 species assessed up to 2003, 23% are threatened (World Wildlife Fund, 2014). Based on these figures, extinctions are predicted to occur in the next 30 years (Lintermans, 2013).

Negative impacts of introduced species manifest on many levels; genetic (gene transcription and hybridisation), individual (behaviour, morphology, vital rates), population (transmission of disease and parasites, demographic effects, distribution), community (extirpations, changes in composition, alterations to the food web) and ecosystem (biochemical cycles, energy fluxes between systems, ecological engineering) (Cucherousset & Olden, 2011; Marr et al., 2013; Smart et al., 2006). This is why introduction of feral species is considered one of the greatest threats to biodiversity conservation, second only to habitat destruction (Cambray & Pister, 2002; Leprieur et al., 2009).
Of the world’s 100 most invasive species, eight are fish (Lowe et al., 2000), and four of these: *Salmo trutta* (Brown Trout); *Cyprinus carpio* (Carp); *Oreochromis mossambicus* (Mozambique Tilapia); and, *Oncorhynchus mykiss* (Rainbow Trout), are found in Western Australia. Other commonly found invasive fish in Western Australia include *Gambusia holbrooki* (Mosquitofish), *Leiopotherapon unicolor* (Spangled Perch), *Carassius auratus* (Goldfish), and *Maccullochella peeli* (Murray Cod) (Table 1.1).

**Table 1.1** Non-native freshwater fish species established in Western Australia. Year of major/initial introduction or first record from the wild and reason for introduction. Table adapted from (Lintermans, 2004), with * indicating data from (Fletcher & Santoro, 2014).

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Common Name</th>
<th>Reason</th>
<th>Origin*</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyprinidae</td>
<td><em>Carassius auratus</em></td>
<td>Goldfish</td>
<td>Ornamental</td>
<td>Eurasia</td>
<td>1876</td>
</tr>
<tr>
<td></td>
<td><em>Cyprinus carpio</em></td>
<td>Carp</td>
<td>Aquaculture</td>
<td>Eurasia</td>
<td>1850-1880</td>
</tr>
<tr>
<td>Percidae</td>
<td><em>Perca fluviatilis</em></td>
<td>Redfin Perch</td>
<td>Acclimatisation</td>
<td>Europe</td>
<td>1862</td>
</tr>
<tr>
<td>Poecilidae</td>
<td><em>Gambusia holbrooki</em></td>
<td>Eastern Mosquitofish</td>
<td>Bio-control</td>
<td>South America</td>
<td>1925</td>
</tr>
<tr>
<td>Terapontidae</td>
<td><em>Leiopotherapon unicolor</em></td>
<td>Spangled Perch</td>
<td>Unknown</td>
<td>Gascoyne</td>
<td></td>
</tr>
<tr>
<td>Perciformes</td>
<td><em>Maccullochella peeli</em></td>
<td>Murray Cod</td>
<td>Recreational fishing</td>
<td>Eastern Australia</td>
<td>Late 1800s</td>
</tr>
</tbody>
</table>

Intentional translocation of feral fish species to Western Australia began in the 1870s and continues through to today (Lintermans, 2004; Townsend, 1991). Initially fish were translocated to Western Australia for sport angling and as a potential food source, beginning with *Perca fluviatilis* (Redfin Perch) in 1862 and *S. trutta* (Rainbow Trout) in 1864 (Coy, 1979; Lintermans, 2004). Twenty years later, *M. peeli* (Murray Cod), *Macquaria ambigu* (Golden Perch), *Bidyanus bidyanus* (Silver Perch), *Anguilla australis* (Short-finned eels), *Cy. carpio* (Carp), and *Tinca tinca* (Tench), were released into Western Australian waterways for the same purpose (Coy, 1979) and in the early 1900s, *Oncorhynchus mykiss* (Rainbow Trout) were introduced into rivers and lakes in the southwest of Western Australia (Coy, 1979).
From the 1930s until recent times, *G. holbrooki* (Mosquitofish) were released into many watercourses for controlling mosquito populations (Coy, 1979; Morgan et al., 2004). Ornamental escapes and liberation by fish owners have also occurred over the last four decades resulting in ornamental populations becoming established (Beatty & Morgan, 2013; Duffy et al., 2013). Species introduced due to this include *Cy. carpio* (Carp) and *O. mossambicus* (Mozambique Tilapia) and more recently popular aquarium species, *Xiphophorus helleri* (Swordtails), *Poecilia reticulata* (Guppies), *Phalloceros caudimaculatus* (One-spot Livebearers), and *Ca. auratus* (Goldfish) have been detected in Western Australian water bodies (Morgan et al., 2004). Other translocation events are due to escapes associated with expanding aquaculture practices and fishing (Lintermans, 2004): during fish stocking, or when living baitfish are discarded into waterways by anglers (Drake & Mandrak, 2014; Lintermans, 2004; Ludwig Jr & Leitch, 1996). *B. bidyanus* (Silver Perch) and *Pe. fluviatilis* (Redfin Perch) are the fish most commonly associated with this phase (Morgan et al., 2004). Irrespective of how these invasive species came to be in the waterbodies, they are a threat to native and endemic species in the region.

### 1.3 Surveying methods for detecting fish species

Surveying of fish is undertaken for many reasons including estimating total area of suitable habitat, total fish numbers, and species composition (including identifying introduced species) (Hankin & Reeves, 1988; Henry & Lyle, 2003). Deciding on the best survey method is the first requirement and typically the most difficult step when developing a sound monitoring and management program (Macnaughton et al., 2015). There are a variety of methods that can be used to sample fish in rivers, streams, and lakes, with perhaps the most common active methods being seine nets, scoop and dip nets, angling, visual census and electrofishing, while common passive methods include Fyke nets and gillnets (Department of Sustainability, 2011).

While these common methods are traditionally used, the ‘molecular revolution’ has opened up a pathway to using environmental DNA (eDNA). Environmental DNA is any genetic material obtained directly from environmental samples (e.g. freshwater without any obvious signs of the donating organism), and can be used to detect the presence of species in an environment (Thomsen & Willerslev, 2015).
The development of an eDNA-based sampling approach, and a comparison between it, and traditional methods, is the focus of this thesis. These two methods (traditional and molecular) are each explored below along with the more specific single species detection, prior to the specific aims of this study being discussed.

a) Traditional approach

When the aim of sampling is to detect all species in a waterbody, traditional methods should have the benefit of providing confidence that there will be no false positives. This is as long as the fish collected can be easily identified, however false negatives may be an issue (discussed below). Using traditional methods, phenotype plasticity and closely related species along with those that have similar appearances in various stages, may be misidentified (Butchart et al., 2010). If fish cannot be accurately identified they need to be humanely euthanized and brought back to the laboratory for identification (Daan, 2001); for example, Spangled Perch (*L. unicolor*) can be mistaken for Yellowtail Grunter (*Amniataba caudavittata*) as they look similar during different life stages (Morgan & Gill, 2006).

Generally, traditional methods such as Fyke netting and electrofishing are assumed beneficial, as there are few, if any, false positives when skilled staff are employed; however, there are many issues associated with these methods. There is a range of drawbacks associated with traditional methods of fish sampling. The gear used needs to be cleaned and disinfected before use at the next location to ensure that there is no contamination or translocation between sites and selectivity of the gear used for diversity analysis can lead to the misrepresentation of certain species of fish relative to their abundance in the environment (Murphy & Willis, 1996). Though by taking samples over a relatively long length of time (i.e. overnight), diurnal and mobility variations in fish behaviour are dampened (Hayes, 1989).

Catch rate variability is often high with Fyke nets and seine nets and therefore a relatively large number of sets are required to detect changes in relative abundance (Portt et al., 2006). Depending on the sampling method chosen, the surveying is relatively unaffected by turbidity or electrical conductivity and for the most part traditional surveying methods do not require a high level of technical expertise or specialist training (although experience and knowledge of field staff will influence catch efficiency) (Lake, 2013).
Trapping (and other ‘traditional’ methods such as electrofishing) may have a low chance of collecting data on target organisms, and can only be used as reliable indicators for species which have a moderate to high abundance (Magnuson et al., 1994). For rare or low abundance species probability of detection is much lower, which may lead to an incorrect conclusion that the species is absent (Gu & Swihart, 2004). This may also result from some fish species being trap-shy or developmental stages that may not allow fish to be easily surveyed (Ficetola et al., 2008; Jerde et al., 2011; Wilson et al., 2014). These traditional methods may also have ongoing organism-level sub lethal effects such as the potential for behavioural alterations, physiological and energetic costs, and associated reductions in feeding, growth, or reproduction of the fish being surveyed (Wilson et al., 2014).

There can be false negatives if the surveying does not include a wide range of habitats, as less mobile species are unlikely to be captured. Some methods such as Fyke and seine netting, and electrofishing can be less effective in deep water bodies, or waterbodies that have uneven or steeply sloped substrate (Lake, 2013). Predation can occur inside surveying nets affecting recorded capture rates of smaller fish (which may either be eaten or actively avoid entering nets containing predators) (Breen & Ruetz, 2006). Depending on the method chosen for biodiversity surveying, there is a bias towards highly mobile species, trap ‘happy’ species, cover-seeking species, and benthic species, along with larger species (Breen & Ruetz, 2006; Portt et al., 2006) and Fyke nets can also capture non-target species, such as birds and amphibians. Where there are rare or low abundance species, the only solution is to either increase sampling effort, which is time consuming and can be costly, or to change method of detection to something which has a greater detection probability (McDonald & Thompson, 2004).

b) Molecular approach – single species detection

Molecular surveying for single species detection has been used with success for detection of rare and low density species (Olson et al., 2012; Rees et al., 2014). This approach has growing interest and is important for ecosystem management. Using primers designed to target a single species allows for the rapid detection of target species (Rees et al., 2014) and could be used for on-site determination of the presence or absence of the target species, though stringent protocols need to be in place to ensure that the chance for false positive along with false negative results is reduced (Rees et al., 2014).
This method is good for confirmation, but does not give an overall understanding of the biodiversity of the lake system. Gaining an overall ‘snapshot’ of the species composition of the lake, means that it is likely that all species can be detected no matter the density. This would mean that the single species detection could be used as a monitoring tool once the alien or rare species had been detected via metabarcoding.

c) Molecular approach – multispecies detection via metabarcoding

As eDNA samples are usually composed of mixtures of DNA from tens, hundreds or even thousands of faunal and floral organisms (not to mention the incredibly large microbial diversity), these samples may allow for an in-depth view of the make-up of the ecosystem (Shokralla et al., 2012). Since the mid-1980s, eDNA has been used to detect and describe microbial communities in marine sediments, and during the mid-1990s, this technique was used to describe phytoplankton communities in the water column (Diaz-Ferguson & Moyer, 2014). Since these early studies, eDNA has successfully been sampled from marine environments, terrestrial sediments, ice cores, and fresh water lakes and rivers, revealing the biodiversity of the ecosystem, both modern and ancient (Ficetola et al., 2008; Matisoo-Smith et al., 2008; Parducci et al., 2013; Thomsen & Willerslev, 2015). Environmental DNA amplification was first applied to ancient sediments by Willerslev et al. (2003) as a method to assess diversity of macro-organism communities and revealed both extinct and extant mammals, birds and plants.

There is great potential for improving biodiversity monitoring using eDNA, especially if monitoring using eDNA is sensitive enough to detect a single organism in a waterbody. Low abundance organisms can be difficult to detect using traditional methods, and it is hoped that eDNA monitoring can overcome this traditional method limitation, though for this project, DNA sequences are not being used as a measure of abundance.
Metabarcoding of macro organisms in aquatic environments is in its infancy and the studies that have been completed in natural environments have emphasised the challenges posed by this method (Balasingham, 2016; Deiner et al., 2015; Evans et al., 2017; Miya et al., 2015; Thomsen, Kielgast, Iversen, Wiuf, et al., 2012; Tucker et al., 2016). Many studies have looked into the uses of molecular surveying in aquatic environments, with studies conducted in both marine (Foote et al., 2012; Kelly et al., 2014; Landi et al., 2014; Radulovici et al., 2010; Thomsen, Kielgast, Iversen, Moller, et al., 2012; Thomsen & Willerslev, 2015) and freshwater systems (Deiner et al., 2016; Hänfling et al., 2016).

Environmental DNA surveying has been applied to common invasive species of amphibians and fish (Ficetola et al., 2008; Goldberg et al., 2011; Goldberg et al., 2013; Jerde et al., 2013; Jerde et al., 2011), as well as with native fish species such as the Mekong giant catfish (Pangasianodon gigas) (Bellemain et al., 2016), spotted gar (Lepisosteus oculatus) (Boothroyd et al., 2016), and Macquarie perch (Macquaria australasica) (Bylemans et al., 2016), along with herpetological surveys in North America (Lacoursière-Roussel et al., 2016).

Environmental DNA has been shown to be a promising and effective tool for detection of some invasive species, yet its viability has not yet been assessed for most communities and ecosystems (Dougherty et al., 2016). This study will evaluate the effectiveness of eDNA in identifying fish species found in a small lake in Western Australia, with particular focus on the non-native and invasive species such as M. peelii, G. holbrooki, and L. unicolor. This approach will be tested to see if it can be reliably used as a standalone survey to replace approaches that are more traditional.

d) Comparison of molecular methods

While there are benefits from both molecular methods the approaches and field utilisation differ. For single source DNA the water is collected in the field and then the DNA of the target species is extracted. The DNA then undergoes Sanger sequencing using species specific primers and is BLASTed to confirm that the DNA collected is from the target species. For the multiplex metabarcoding method, the water from the field also has the DNA extracted, but using universal primers it undergoes quantification in a quantitative polymerase chain reaction (qPCR). After this stage the quantified DNA is then prepared for metabarcoding using Next-Generation sequencing.
After sequencing, the results are BLASTed to identify species present. The sensitivity of each method relies on the concentration of DNA present along with the sensitivity of the primers. For the single source method low density species are likely to be detected as the primers are developed to specifically target that species and the DNA is not likely to be ‘swamped-out’ due to other species. With the metabarcoding method the ability to detect low density species depends on primer binding affinity, concentration of DNA compared to other species along with specificity of the primers used.

Both methods have the ability to detect species that are likely to be present in low densities, though there are limitations with both methods. This means that careful consideration needs to be put in place when deciding which to use in the field.

**1.4 Importance of this project**

This project, along with others that tackle the field of biodiversity in new and novel ways, illustrates the unique insights that eDNA identification can bring to the field. Identification of fish assemblages can be expensive and time consuming. For example, in Western Australia, the Department of Fisheries surveyors travel to each site multiple times for overnight deployment of nets, with setting and collection taking time each visit (M. Snow pers. comm., 2013). Furthermore, additional time is taken with gear handling, fish handling, and disinfection of traps and gear on collection (M. Snow pers. Comm., 2013). Yet as previously discussed, results of these surveys may not present accurate results due to limitations of the surveying method (Connolly, 1994). Typically, traditional monitoring techniques have been highly invasive to species and/or the ecosystem under study, while many marine based studies have relied on damaging techniques (Jones, 1992) to achieve desired results.

Contemporary research indicates that eDNA is useful for detecting rare or cryptic species, and especially those species that would be difficult to find using traditional methods due to either low density or trap shyness (Shaw et al., 2016; Simpfendorfer et al., 2016; Thomsen et al., 2016). This means that proof of concept projects such as this one are important for determining the presence or absence of endangered and low density species or for the monitoring of an advancing biological invasion (Jones, 2013).
The monitoring of invasion fronts has already proven valuable for detecting single species invasions, such as the Asian Carp in the Great Lakes in the United States of America (Jerde et al., 2013). Jerde et al. (2013) reported that a commercial angler caught the invasive carp species only a few kilometres from where the nearest positive eDNA detection had occurred only a few months earlier. This shows that eDNA can be used as a powerful tool for early detection of invasions and for allocating resources where they are most beneficial. Extending on from the successes of Jerde et al. (2013) where a single species was targeted, this thesis shows multispecies detection can also be achieved using eDNA methods.

One reason why eDNA is so powerful for contemporary conservation and invasive species detection is because decay of eDNA in freshwater beyond the limits of detection occurs over a scale of only days to weeks (Dejean et al., 2011; Pilliod et al., 2014; Thomsen, Kielgast, Iversen, Wiuf, et al., 2012). This means that a positive detection is likely to be associated with the recent presence of a species, while older signals from transient or deceased organisms are not likely to be detected over repeat sampling periods. This is particularly useful in closed systems, though less so in river systems where long distance transportation of eDNA has been reported (Deiner & Altermatt, 2014; Dejean et al., 2011; Staley et al., 2013).

Freshwater eDNA research is not only important for invasive species detection but can also be useful in the discovery of endangered species. Research in the field of endangered species has shown positive detection of both amphibians (Dejean et al., 2012; Goldberg et al., 2011; Pilliod et al., 2014; Pilliod, Goldberg, Laramie, et al., 2013; Rees et al., 2014) and freshwater fish (Bellemain et al., 2016; Boothroyd et al., 2016; Bylemans et al., 2016) along with quantification of species present (Pilliod, Goldberg, Arkle, et al., 2013).
The following aims, and sub-aims, test if multiplexed eDNA analysis of freshwater lake samples can be used as a viable alternative to more traditional approaches.

Aim -

- To develop a cost effective and efficient, environmental DNA method for the detection of all species of finfish present in a freshwater lentic system.

Sub-aims -

1. To develop and test COI primer sets, along with testing existing 16S primer sets for fish species to reliably detect native Western Australian and non-native fish species in an eDNA water sample (PCR, qPCR and next generation sequencing).

2. To optimise field and laboratory protocols for determining freshwater fish species presence in freshwater waterbodies and to compare results with those of traditional net-based sampling.

3. To compare the financial costs of a commonly employed traditional method of fish sampling (specifically Fyke netting) to eDNA analysis.

In order to complete these aims COI were developed using Geneious software (Biomatters) and both primer sets were tested in the lab before being used in on the field samples. Water samples (250 mL) were collected at three locations and two depths from the waterbody to test if there was any change in detection of species within either the water column or location around the lake. Samples were collected five times over a year to account for seasonal variability in results. Results were then compared with previous year’s Fyke net surveying results, to see if molecular surveying produced similar results. Finally, the costs, financial and work hours, involved in molecular multiplex surveying, molecular single target surveying and Fyke net surveying, to see if molecular surveying had a cost effective role in future management programs.
2. Method Development

Materials and methods developed for, and used, during this study are described in this chapter, including sample collection and storage, primer design, DNA extraction and quantification, Next Generation sequencing (NGS), and bioinformatics.

2.1 Sample collection and storage

Water samples for eDNA extraction were collected from three locations at Alexandria Boulevard Lake, approximately 25 km south-southeast of the Perth Central Business District (Figure 2.1). Sample collection occurred once during each season (to test for seasonal variability) in 2015, except in winter when two sampling days occurred one week apart (to test for repeatability of results (Table 2.1)).

Figure 2.1 The southwestern region of Western Australian, including the study site. Alexandria Lake (Canning Vale, Western Australia), showing the location of the three water collection sites of this study. Image from Google Maps.
Table 2.1 eDNA sample details. Including collection code, season, site location, surface, or depth along with where the DNA was extracted. (DoF – Department of Fisheries, ECU – Edith Cowan University, Cv – Canning Vale)

<table>
<thead>
<tr>
<th>Collection number</th>
<th>Season</th>
<th>Date</th>
<th>Site</th>
<th>Surface/Depth</th>
<th>Extraction location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cv.Su.1A</td>
<td>Summer</td>
<td></td>
<td>Southern</td>
<td>Surface</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Su.1B</td>
<td>Summer</td>
<td></td>
<td>Southern</td>
<td>Depth</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Su.2A</td>
<td>Summer</td>
<td>27/2/15</td>
<td>Central</td>
<td>Surface</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Su.2B</td>
<td>Summer</td>
<td></td>
<td>Central</td>
<td>Depth</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Su.3A</td>
<td>Summer</td>
<td></td>
<td>Northern</td>
<td>Surface</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Su.3B</td>
<td>Summer</td>
<td></td>
<td>Northern</td>
<td>Depth</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Au.1A</td>
<td>Autumn</td>
<td></td>
<td>Southern</td>
<td>Surface</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Au.1B</td>
<td>Autumn</td>
<td></td>
<td>Southern</td>
<td>Depth</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Au.2A</td>
<td>Autumn</td>
<td>28/5/15</td>
<td>Central</td>
<td>Surface</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Au.2B</td>
<td>Autumn</td>
<td></td>
<td>Central</td>
<td>Depth</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Au.3A</td>
<td>Autumn</td>
<td></td>
<td>Northern</td>
<td>Surface</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Au.3B</td>
<td>Autumn</td>
<td></td>
<td>Northern</td>
<td>Depth</td>
<td>ECU</td>
</tr>
<tr>
<td>Cv.Wi1.1A</td>
<td>Winter (1)</td>
<td>10/8/15</td>
<td>Southern</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv.Wi1.1B</td>
<td>Winter (1)</td>
<td></td>
<td>Southern</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv.Wi1.2A</td>
<td>Winter (1)</td>
<td></td>
<td>Central</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv.Wi1.2B</td>
<td>Winter (1)</td>
<td></td>
<td>Central</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv.Wi1.3A</td>
<td>Winter (1)</td>
<td></td>
<td>Northern</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv.Wi1.3B</td>
<td>Winter (1)</td>
<td></td>
<td>Northern</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Collection number</td>
<td>Season</td>
<td>Date</td>
<td>Site</td>
<td>Surface/Depth</td>
<td>Extraction location</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------</td>
<td>------------</td>
<td>--------</td>
<td>---------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Cv.Wi2.1A</td>
<td>Winter (2)</td>
<td></td>
<td>Southern</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Wi2.1B</td>
<td>Winter (2)</td>
<td></td>
<td>Southern</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Wi2.2A</td>
<td>Winter (2)</td>
<td>18/8/15</td>
<td>Central</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Wi2.2B</td>
<td>Winter (2)</td>
<td></td>
<td>Central</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Wi2.3A</td>
<td>Winter (2)</td>
<td></td>
<td>Northern</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Wi2.3B</td>
<td>Winter (2)</td>
<td></td>
<td>Northern</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv.Sp.1A</td>
<td>Spring</td>
<td></td>
<td>Southern</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Sp.1B</td>
<td>Spring</td>
<td></td>
<td>Southern</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Sp.2A</td>
<td>Spring</td>
<td>25/11/15</td>
<td>Central</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Sp.2B</td>
<td>Spring</td>
<td></td>
<td>Central</td>
<td>Depth</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Sp.3A</td>
<td>Spring</td>
<td></td>
<td>Northern</td>
<td>Surface</td>
<td>DoF</td>
</tr>
<tr>
<td>Cv. Sp3B</td>
<td>Spring</td>
<td></td>
<td>Northern</td>
<td>Depth</td>
<td>DoF</td>
</tr>
</tbody>
</table>
Initial samples were collected in 500 mL, sterile, gamma radiated bottles (Thermo Fisher Scientific Australia Pty Ltd), which were labelled with site location (1 (southern), 2 (central), or 3 (northern)) and as surface (A) or depth sample (B). Water chemistry including, pH, conductivity, temperature, dissolved $O_2$ and salinity levels were recorded with each sample using a Thermo Scientific Orion 5-Star Plus pH/ORP/ISE/Cond/DO Portable Meter. A Kemmerer bottle was used to collect the depth water samples from approximately 3.1 meters below the surface. At the northern location the sampling location was not deep enough for effective use of the Kemmerer bottle so water was collected by wading out to waist depth and then collecting water from as close to the bottom as possible. Surface samples were collected directly into sample bottles. The Kemmerer bottle was washed with a 10% bleach solution and rinsed with MilliQ water between sites. Samples were then transported back to ECU on ice where they were stored at -20°C until processed.

Two of the initial bottles cracked while in the freezer so subsequent samples were collected in Nalgene Sterile Wide-Mouth HDPE Bottles. A designated field blank was included for each time point. This consisted of a sample bottle filled with MilliQ that was opened at the central site, exposed to the air for approximately one minute, closed, and then submerged in the waterbody before being stored with the rest of the samples from that time point (Furlan et al., 2016; Jerde et al., 2013). These blanks were processed along with the water samples, though they were not used in the final analysis as MilliQ water at Edith Cowan University was found to contain freshwater fish amplicon contamination. As the blanks were MilliQ water sourced from ECU it is believed that there were airborne freshwater fish amplicons that had been spread through the building through the air conditioning system. The amplicons detected were likely generated by the single source DNA experiments that had been completed in the early stages of the project, this meant that in subsequent field samples UV treated MilliQ water was used to rinse the Kemmerer between sites. Samples from winter 1, winter 2 and spring were then processed at the Department of Fisheries (DoF) (Table 2.1).
2.2 PCR, qPCR and Next Generation Sequencing (NGS) Primer design

a) in-silico development

For this study, 16S ribosomal RNA (16S) and Cytochrome c oxidase I (COI) gene regions were analysed (Table 2.2).

Table 2.2 Primer pairs used for the detection of freshwater fish species in a waterbody located in Canning Vale. Universal fish primers for 16S were sourced from Dr. Michael Bunce (Curtin University (Unpublished)), COI primers were developed specifically for Western Australian freshwater fish.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Marker</th>
<th>Sequence (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>121F</td>
<td>COI</td>
<td>NATYGGNGGNTTTYGGNAAYTGRY</td>
</tr>
<tr>
<td>324R</td>
<td>COI</td>
<td>GATGCTCCGGCGTGKGCTAA</td>
</tr>
<tr>
<td>Fish16sF/D</td>
<td>16S</td>
<td>GACCCCTATGGAGCTTTAGAC</td>
</tr>
<tr>
<td>16s2R–degenerate</td>
<td>16S</td>
<td>CGCTGTATCCCTADRGTAACT</td>
</tr>
</tbody>
</table>

To develop COI primers, firstly six sequences (Gambusia holbrooki, Maccullochella peeli, Carassius auratus, Pseudogobius olorum, Leiopotherapon unicolor, and Cyprinus carpio) obtained from DNA extracted from fish muscle were aligned using ClustalW (Geneious software (Biomatters), and alignments were examined for conserved regions across the sequences using the Primer3 design tool (Untergasser et al., 2012). These conserved regions were used to determine regions where primers may be placed, following design parameters (Product Size = 70-150bp; Consensus = 75%; Number of pairs = 5), and each of the potential primer pairs were then evaluated against the alignment using Geneious (Biomatters).

Once primer pairs satisfying all the above parameters were found, degenerate bases were introduced to ensure that at least a proportion of each primer mix shared 100% identity with the target DNA of all of the six species. The primers were evaluated against the DNA sequences of thirty-three relevant endemic and invasive species retrieved from GenBank, accessed via the National Centre for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/) (Appendix 3). The available relevant freshwater fish sequences were aligned and tested using the primer pairs, which were further degenerated as needed until the DNA of all species would be potentially amplified.
The suitable primer pairs were further examined to ensure that they would have similar melting temperatures \( T_m \), would not form hairpins or dimers, and still had sufficient guanine and cytosine content after degeneracy to be able to identify organisms to a species level, specifically near the 3’ ends of each of the primers in the set. The finalised primer pair was tested against DNA extracted from species (Appendix 3) that were likely to be in the waterbody to ensure that species from the non-target taxonomic groups were not detected.

b) Generation of reference DNA sequences from target species to support subsequent species identification from NGS data

To test the primers developed in the previous step and to develop the reference library, DNA was extracted from tissue samples of six different species *G. holbrooki* (Mosquitofish), *Cy. carpio* (Carp), *L. unicolor* (Spangled Perch), *C. auratus* (Goldfish), *M. peelii* (Murray Cod), and *Ps. olorum* (Swan River Goby), using the Qiagen Blood and tissue kit. The extracted DNA was then analysed using a Bio-Rad NanoDrop to determine the concentrations of DNA extracted. A region of both the COI and 16S mitochondrial genes approximately 170 bp in length were amplified using the primers previously developed. PCR of tissue samples was performed on a Bio-Rad IQ5 multicolor real-time qPCR using 2 µl of DNA. The following components were used for each 25 µL reaction: 14.4 µL H₂O, 2.5 µL Taq Gold buffer (Applied Biosystems [ABI], USA), 2 µL 25 mM MgCl₂ (ABI, USA), 1 µL 10 mg/mL BSA (Fisher Biotech, Australia), 0.25 µL 25mM dNTPs (Astral Scientific, Australia), 1 µL each of forward and reverse primers (10 µM) (Integrated DNA Technologies, Australia), 0.6 µL of 1/10,000 SYBR Green dye (Life Technologies, USA), 0.05 U/µL (0.25 µL) of Taq polymerase Gold (ABI, USA). The PCR conditions consisted of an initial denaturing step at 95°C for 10 minutes followed by 50 cycles of 95°C (30s), primer annealing at 54°C (30s) and 72°C for (45s), with a final extension at 72°C (10 minutes).
PCR products were sent to the Australian Genome Research Facility (AGRF) Perth node for forward and reverse sequencing using both the 16S and COI primers. Sequence reads obtained from each of the above species were imported into Geneious (Biomatters) and processed to give one clean consensus sequence for each species, which was then BLASTed (Basic Local Alignment Search Tool) against all known freshwater fish species on NCBI’s GenBank to ensure that there was no contamination and that the fish species were correctly identified using these primers.

2.3 DNA extraction from environmental samples

Ultra violet treated ultra-pure water was used as extraction and PCR controls, and was treated in the same manner as the environmental water samples, lab samples were sampled prior to environmental samples being opened. The lab controls and environmental samples were filtered using a vacuum pump attached to a filter tower in which was placed a sterile GN-6 Metricel® S-Pack Membrane Disc Filter, 0.45 µm, 47 mm, grid (Life Sciences). After filtering, the water filter from each location was cut in half then cut into smaller pieces, with the two halves separated for processing. eDNA was extracted from both halves of a filter using the tissue protocol of a DNeasy Blood and Tissue Kit (QIAGEN), with the two halves filtered through the same column before the elution step to create a single sample with all lab work was conducted using barrier pipette tips and commercially available nuclease free water. Elution was completed three times for each sample, each time with 50 µL of Buffer AE, which had been heated to 55°C, to ensure maximum DNA yield.

2.4 DNA quantification to optimise input DNA into NGS

DNA quantification was completed using quantitative PCR (qPCR) and each of the COI and 16S primer sets. Each qPCR was carried out for three dilutions of DNA (neat, 1/10 and 1/100), and included extraction controls. The qPCR chemistry was the same as had been used in section 2.2 (b).
Each qPCR was then run on an Applied Biosystems step-ONE qPCR thermocycler under the following conditions; 5 minutes at 95°C, followed by 45 cycles of denaturation, primer dependant annealing and extension; 95°C for 30 seconds, 54°C for 30 seconds and 72°C for 45 seconds. This was followed by a final extension of 10 minutes at 72°C and a one-second melt curve. The melt curve was used to determine whether amplicons detected were those being targeted, primer dimer, or some other form of contamination.

qPCR results were then inspected to determine if the amplification resulted in a good sigmoidal curve, cycle threshold ($C_t$) value, and a clear melt curve peak. The $C_t$ value is the number of cycles that the reaction undertakes before the fluorescent signal exceeds the defined threshold level of fluorescence. During this study, a threshold of 500 was used for consistency.

2.5 NGS preparation

Next Generation Sequencing data for this project was obtained using an Illumina MiSeq sequencer. Both the final sequencing and the library build stage of the project were performed in the Trace and Environmental DNA (TrEnD) labs at the Curtin University of Technology under the supervision of Professor Michael Bunce. Steps involved in the library build were (a) fusion tagging, (b) pooling and purification of samples for NGS, (c) dilution and pre-MiSeq quantification, before the final sequencing occurred.

a) Fusion tagging

The samples that showed good amplification in the previous stage, including any blanks that showed amplification, were then fusion tagged. Fusion tagging attaches the Multiplex Identifier (MID) tags to samples so that they can be easily identified after sequencing. These tags allow up to 96 samples to be multiplexed and then samples from a source, or in the case of this project, waterbody location, and time point, can be later identified and analysed. Fusion tagging results were then viewed to ensure that amplification occurred and that $C_t$ values were not too high. The melt curve was also checked.
The following components were used for the fusion reaction for a total volume of 25 μl: 15.6 μl H₂O, 2.5 μL Taq Gold buffer (Applied Biosystems [ABI], USA), 2 μL 25 mM MgCl₂ (ABI, USA), 1μL 10mg/mL BSA (Fisher Biotech, Australia), 0.25 μL 25 mM dNTPs (Astral Scientific, Australia), 0.5 μL each of forward and reverse primers (10 μM) (Integrated DNA Technologies, Australia) with the reverse primer added separately as the reverse primers are the unique identifiers, 0.6 μL of 1/10,000 SYBR Green dye (Life Technologies, USA), 0.05 U/μL (0.25 μL) of Taq polymerase Gold (ABI, USA). The master mix components were added in the order listed above.

Each tagging qPCR was then run on an Applied Biosystems step-ONE qPCR thermocycler for 5 minutes at 95°C, followed by 45 cycles of denaturation, primer dependant annealing and extension; 95°C for 30 seconds, 54°C for 30 seconds and 72°C for 45 seconds. This was followed by a final extension of 10 minutes at 72°C and a one-second melt curve. The melt curve was used to determine whether amplicons detected were those being targeted, primer dimer, or some other form of contamination.

b) Pooling and purification of samples for NGS

After successful amplicon tagging, 16S and COI PCR amplicons were then pooled in accordance with their fusion tagging Cₜ values. Samples, which amplified well and had low Cₜ values, were pooled together (up to four samples per pool) and samples that had higher Cₜ values were pooled (with only one or two samples per pool). Amplicon pools were then purified following the Agencourt AMPure XP bead PCR purification process to remove excess dNTPs and remaining primers, as well as dimers, salts and other impurities (Beckman Coulter, 2000-2013).

Four microliters of each purified pool were run on a 2% agarose gel pre-stained with ethidium bromide to confirm. This confirmed DNA was the expected length and for visual confirmation of relative concentrations of pools (based on intensity of bands). The final pooling of amplicons occurred in accordance with the visualised concentrations on the gel; darker bands require less DNA in the final pool. By combining pools in this way, a final equimolar DNA library was created. The final library was then purified twice more following the AMPure process to ensure there was no remaining primer or contaminates.
c) Dilution and pre-MiSeq quantification

The next stage in library preparation ensures that the bead to template ratio for emulsion PCR (emPCR) is optimised by quantifying the template. This quantification is a series of dilutions of the library run alongside a set of known standard synthetic oligonucleotides of a known molarity. This allows a standard curve to be created and then calculations are used to determine DNA concentration in library dilutions. Once DNA concentrations are known, the optimum amount of DNA can be added to the emPCR. The amount determined in calculations is used to determine the prime bead to template ratio. A prime template to bead ratio is crucial as a correct ratio prevents more than one template copy from binding to the sequencing beads during emPCR. If multiple templates bind with the bead, the sequencing bead will give a mixed read during sequencing.

Dilutions of 1:10, 1:100, 1:1000, 1:5000, 1:25000, 1:125000, and 1:625000 were tried against standard curve. After the optimal amount of amplicon library had been determined for the emPCR the DNA library was sequenced.

2.6 Next-Generation Sequencing

The prepared library was sequenced on a single end 300-cycle version 2 reagent kit and a Nano flow cell, using an Illumina MiSeq.

2.7 Bioinformatics and statistical analysis

Raw data was imported into Geneious (Biomatters) where it was separated into time point and location. Sequences were checked to ensure they contained the correct P7 adapter, which was then trimmed off the sequence (along with the primers), and singletons were removed. The sequences were dereplicated using Geneious (Biomatters); one read became the representative read for each unique sequence. Singletons were then removed. As there were rare species or species present in low numbers i.e. *M. peelii* it was decided to keep sequences that had two or more sequences detected to ensure there were no species missed. Remaining sequences were BLASTed using Geneious (Biomatters) against the GenBank (NCBI) database along with a local database that contained the unpublished *Ps. olorum* sequences. Extracted DNA was deemed to show a positive presence of species if there was 100% query coverage. A threshold of 98% maximum identity allowed for a 2% dissimilarity threshold to allow for intraspecific diversity (Mächler et al., 2014).
Sequence counts were analysed using SPSS (SPSS, 2011). One-way ANOVA was used to test significance between locations, as well as surface and depth, this analysis was also used to test for significance between seasons. Multidimensional scaling (MDS) was used to test if any of the water chemistry, season, or location variables affected the number of DNA sequences detected.

2.8 Application of developed method at the Department of Fisheries

The extractions performed at the Department of Fisheries Laboratories followed the previous protocol developed at Edith Cowan University, with the following exceptions; extractions were done in a lamina flow hood (airflow switched off) and a one-time elution of 100 µL of Buffer AE was used.

The water samples were filtered in a laminar flow hood that had undergone 30 minutes of UV treatment before use, then when the hoods were ready for use, the flow was switched off to ensure that there was less chance of airborne contamination being introduced to the work area. All equipment including pipettes, filter housings, scalpels, falcon tubes, and waste containers were also subjected to UV treatment prior to use. The qPCR MasterMix was then loaded on the plate in a clean room. The DNA was then loaded onto the plate in another UV treated hood in the PCR room, once again with the flow switched off.
3. Results

A biodiversity assay was completed over a year to assess variability of results between location and time points along with the practicality of using the developed method in the field. Data from the primer development, along with data from field surveying (location and season, including water chemistry, and absence and presence of a species) were analysed to determine primer performance along with the accuracy of the metabarcoding method. Water chemistry along with sequence counts for location and seasonal data was analysed using mixed model testing, Binary Logistic Regression, and GLM, though as there was a low number of data points these analyses had a low power of analysis which meant that there was no way to extract any significance between locations or seasons.

3.1 Initial Evaluation of COI and 16S primers using single source DNA and conventional PCR

Both COI and 16S primers resulted in generation of a clean single PCR product of the expected size (~150 bp) using DNA extracted from *Cy. carpio*, *G. holbrooki*, *L. unicolor*, *P. olorum*, *M. peelii*, and *Ca. auratus*.

3.2 Evaluation of primer performance using single source DNA in qPCR

The primer performance of the multiplex primer sets (16S and COI) was evaluated by amplifying single source DNA using qPCR (Figure 3.1). The 16S curves had the classic sigmoidal shape and a C_t range of 8.8 cycles (22.13 - 30.93) (Figure 3.1(a)). A one-way ANOVA with Tukey post hoc test showed that there was a significant difference in amplification between species (Table 3.1) and the six freshwater fish species were separated into three subsets. These subsets grouped species with similar mean C_t values together based on the Tukey results. Subset 1, *(G. holbrooki, L. unicolor, and M. peelii)*, subset 2 *(L. unicolor, M. peelii, and Ps. olorum)* and subset 3 *(Ca. auratus and Cy. carpio)*. These subsets indicate the cycles that the DNA was detected on from the species that was detected first, *L. unicolor* to the species that was detected last, *Cy. Carpio*. There was clear separation between subsets one and three with *Gambusia holbrooki* amplifying much more efficiently than *Cyprinus carpio/Carassius auratus*. Subset two had overlap with subset one showing that there was not much difference between these species.
Figure 3.1 Single source freshwater fish DNA qPCR amplification curves. qPCR was conducted on each primer set in triplicate to identify the number of cycles needed to detect each of the species using single source DNA. *Cyprinus carpio* (blue), *Gambusia holbrooki* (teal), *Leiopotherapon uniclor* (pink), *Pseudogobius olorum* (red), *Maccullochella peelii* (orange), and *Carassius auratus* (purple) 16S (a) and COI (b).

Table 3.1 16S Tukey HSD results. Sample size n = 3

<table>
<thead>
<tr>
<th></th>
<th>G. holbrooki</th>
<th>L. uniclor</th>
<th>M. peelii</th>
<th>P. olorum</th>
<th>Ca. auratus</th>
<th>Cy. carpio</th>
</tr>
</thead>
</table>

The COI curves had a slight bump on the top of the sigmoidal shape for *G. holbrooki*, *Pseudogobius olorum*, and *Ca. auratus* though this did not seem to affect the amplification efficacy (Figure 3.1). The C<sub>t</sub> range of 17 cycles (19 – 36) was broader than that of the 16S primers though this did not seem to affect the results from the single source testing or field-testing (Figure 3.1(b)). A one-way ANOVA with Tukey post hoc test showed that there was no significant difference in the efficacies of amplification between species with all six species grouped in the same subset.
DNA from six single source freshwater fish species used to test the primers was analysed on a BioRad NanoDrop One to determine the concentration of DNA in the samples (Table 3.2). Both *L. unicolor* and *Ca. auratus* showed high levels of DNA (~15 ng/µl) while the other four species reported between 3.9 (M. peelii) and 6.8 ng/µl (Ps. olorum). Despite these variations in the DNA concentrations, this did not seem to affect the qPCR results (Figure 3.1).

Using the 16S primers *C. auratus* which had the highest DNA concentration was detected second last in the qPCR and *C. auratus*, which had the second highest DNA concentration was detected the earliest. *M. peelii* which had the lowest concentration of DNA was detected midrange using the 16S primers. For the COI primers, *Ca. auratus* was detected the earliest of the six species and *M. peelii* which had the lowest concentration was detected the second earliest. For the COI primers *L. unicolor* which had the second highest concentration of DNA was detected the latest in the qPCR experiments.

**Table 3.2** NanoDrop results showing the concentration of DNA from each of the six single source freshwater fish species used during this project.

<table>
<thead>
<tr>
<th>Species</th>
<th>ng/µl of DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>L. unicolor</em></td>
<td>15.1</td>
</tr>
<tr>
<td><em>G. holbrooki</em></td>
<td>4.3</td>
</tr>
<tr>
<td><em>Ps. olorum</em></td>
<td>6.8</td>
</tr>
<tr>
<td><em>M. peelii</em></td>
<td>3.9</td>
</tr>
<tr>
<td><em>Cy. carpio</em></td>
<td>6.7</td>
</tr>
<tr>
<td><em>Ca. auratus</em></td>
<td>15.5</td>
</tr>
</tbody>
</table>

Each species, other than *M. peelii*, showed a large difference in mean C_t values between 16S and COI (Figure 3.2). The COI primers had lower C_t values than the 16S primers for *Cy. carpio* and *Ca. auratus*. However, *G. holbrooki*, *Ps. olorum*, and *L. unicolor* had lower C_t values when using 16S compared to COI.
3.3 DNA detection results from environmental samples

Field samples were analysed for two genetic markers (COI and 16S) using the NGS-metabarcoding approach resulting in the detection of 17 unique species: eleven freshwater fish, two marine fish, two species of duck, along with cow and pig (Table 3.3). The two species of marine fish, *Salmo salar*, and *Seriola lalandi* are likely to be contamination. This contamination is likely to have occurred either during the extraction process or during qPCR. The samples that showed this contamination went through these processes at the laboratories at the Department of Fisheries (DoF) where marine fish DNA was also processed; the *S. salar* may have been misidentified *Oncorhynchus mykiss*, which is more likely as *S. salar* is not found in Western Australia (Fiske, 2006; Llewellyn, 2015; Yurtseva et al., 2014). Two of the freshwater fish, *Cyprinus multitaeniata* and *Amniataba percoides*, could be false positives from sequencing error.

![Figure 3.2](image-url)  
**Figure 3.2** Mean Cₜ values for six freshwater fish testing the 16S and COI primer sets for the number of cycles needed to detect single source DNA using qPCR. Open markers designate species not detected in environmental samples.
Table 3.3 All species detected from waterbody samples during NGS study.

<table>
<thead>
<tr>
<th>Freshwater fish</th>
<th>Marine fish</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Amniataba percoides</em> (Barred Grunter)</td>
<td><em>Salmo salar</em> (Atlantic salmon)</td>
</tr>
<tr>
<td><em>Cyprinus carpio</em> (Carp)</td>
<td><em>Serio lalandi</em> (Yellowtail amberjack)</td>
</tr>
<tr>
<td><em>Cyprinus multitaeniata</em></td>
<td></td>
</tr>
<tr>
<td><em>Gambusia holbrooki</em> (Mosquitofish)</td>
<td></td>
</tr>
<tr>
<td><em>Leiopotherapon unicolor</em> (Spangled Perch)</td>
<td></td>
</tr>
<tr>
<td><em>Maccullochella peelii</em> (Murray Cod)</td>
<td></td>
</tr>
<tr>
<td><em>Macquaria novemaculeata</em>^ (Australian bass)</td>
<td></td>
</tr>
<tr>
<td><em>Nannoperca vittata</em> (Western Pygmy Perch)</td>
<td></td>
</tr>
<tr>
<td><em>Perca fluviatilis</em> (Redfin Perch)</td>
<td><em>Gallinula tenebrosa</em> (Dusky Moorhen)</td>
</tr>
<tr>
<td><em>Poecilia sphenops</em> (Molly)</td>
<td><em>Bos taurus</em> (Cow)</td>
</tr>
<tr>
<td><em>Pseudogobius olorum</em> (Swan River Goby)</td>
<td><em>Sus scrofa</em> (Pig)</td>
</tr>
</tbody>
</table>

*possible sequencing error ^ species not expected to be found

Using 16S over the course of the year, seven species were detected (Table 3.4), G. holbrooki, L. unicolor, M. peelii, Pe. fluviatilis, Po. sphenops, Ps. olorum, and N. vittata. This was one more than was detected using COI (Table 3.4). COI detected L. unicolor, M. peelii, Pe. fluviatilis, P. olorum, N. vittata, and C. carpio. The species detected were similar using both markers, though Po. sphenops and G. holbrooki were only detected with 16S and C. carpio were only detected using COI. These detections by specific primers are likely to have occurred due to the C_{t} values previous described. Po. sphenops and G. holbrooki had low C_{t} values using 16S while the C_{t} values for Cy. carpio were much lower using the COI primers than the 16S primers. The molecular surveying results and the traditional (though dated) surveying results (Table 3.4) were comparable with all species detected using Fyke net surveying. The two species not detected using Fyke net surveying may have been introduced to the lake after the last survey was completed in 2014.
Table 3.4 Species detected by each of the markers and a comparison between molecular surveying results and traditional fyke net surveying.

<table>
<thead>
<tr>
<th></th>
<th>16S</th>
<th>COI</th>
<th>Traditional Surveying (up to 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. carpio</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>G. holbrooki</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>L. unicolor</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>M. peelii</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>N. vittata</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pe. fluviatilis</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Po. sphenops</td>
<td>✓</td>
<td>×</td>
<td>-</td>
</tr>
<tr>
<td>Ps. olorum</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

3.4 Total freshwater fish sequence counts from each location and season

A comparison of sequence counts was conducted between markers as well as season and location. This was completed to see if there was any variation between the markers and to see if a season or location represented a better time or place to sample from than another.

DNA counts (Figure 3.4) between the 16S and COI markers were analysed using a one-way ANOVA and there was a significant difference ($p = 0.005$) between total sampling counts for each marker. There was no significant difference between seasons for 16S ($p = 0.44$) or COI ($p = 0.057$) at $p = 0.05$ in the total number of sequences extracted between seasons. There is a wide variation in the number of sequences detected at each location with both COI and 16S.

The southern depth site had the greatest number of sequences detected using the 16S marker which was mirrored by the COI results (Figure 3.3). There were consistent high numbers detected by 16S at all six locations. The central location had similar results between the surface and depth locations, while the northern depth location had a slightly higher number of detected sequences than the surface location.
The number of sequences detected using the COI marker was lower across all locations than the 16S marker. The northern surface location had the lowest number of sequences detected using 16S and the central surface location had the lowest numbers using COI. The southern location had the largest number of detected unique sequences followed by the northern then central locations. The average number of identified unique sequences per location was 2751 with a standard deviation of ±373.5. Unique sequences were determined with 100% query coverage and a threshold of 98% maximum identity to allow for a 2% dissimilarity threshold.

Using the 16S primers, winter 2 had the highest number of sequences detected (Figure 3.4), this may have been due to the rainfall event that occurred between the winter 1 and winter 2 sampling periods, or may have just been due to the cool conditions encouraging a greater mobility in the fish species. These three sampling periods resulted in a higher number of sequences being detected for both 16S and COI. Spring had the best results using COI, but showed a similar trend to 16S with the same three sampling results yielding better results than summer and autumn. For both markers, summer and autumn had the worst results with much fewer sequences being detected (Figure 3.4).

There was variation between marker, season, and location, though trends were similar between markers. For both markers, the same two sampling periods (summer and autumn) were the poorest and this is likely due to the variables addressed above. Results were similar across locations; the southern location had the best total number of sequences detected using both makers, though 16S was more consistent across the locations.

**Figure 3.3** The total number of sequences detected at each location from 250 mL water collected, across five sampling periods, for both the 16S and COI markers, showing standard deviation for each location.
Figure 3.4 Sequence counts for each location, surface, and depth, across the five sampling periods (16S (a) and COI (b)). Winter and winter 2 are repeat samplings one week apart, with rainfall occurring between weeks.
3.5 Water Chemistry values for each location and season

Water chemistry was expected to play a role in the number of sequences detected. Dissolved oxygen (DO), conductivity, temperature, and pH were recorded at each of the locations over the year and these results were used to see if the waterbody was comparable to other lentic systems following these parameters. The waterbody had a mean dissolved oxygen level of 7.7 mg/L (±1.3 mg/L standard error), conductivity of 610.8 µS/cm (±101.8 µS/cm standard error), temperature of 19.1˚C (±3.18 ˚C standard error), and pH of 7.4 (±1.24 standard error). These parameters were all considered to be within normal ranges for a lentic system in Australia though fluctuations did occur across seasons (Kay et al., 2001). These variables were used in MDS analysis but it was determined that the power of analysis was too low to determine if there was one or more variables that affected either the species detected or the number of DNA sequences detected.

Dissolved oxygen (DO) ranged from 12.4 mgL⁻¹ at the central surface location in winter 2 to a low of 1.8 mgL⁻¹ at the northern depth location in winter 1, however over the year the DO was on average between 6.8 (winter 2) and 8.0 mgL⁻¹ (autumn) (Figure 3.5 (a)). The central location had consistently high levels of dissolved oxygen while the northern location had the lowest levels of dissolved oxygen over the year. Dissolved oxygen was similar at different water depths for each season, ranging between a 0.0 and 1.5 mgL⁻¹ difference, with the exception of the deeper water at the northern site. Summer and autumn had a much lower DO at the southern sub-surface site than near the surface and there was the least amount of variation in DO between sites in spring. Little variation in DO was also observed between seasons, except winter when the DO at the central site was greater than in the other seasons, whereas it was lower in the other two sites.

Conductivity ranged from 757 µS/cm at the northern surface location in summer to 440.0 µS/cm at the northern depth location in winter, however over the year the conductivity was on average between 515.1 µS/cm (winter 2) and 734.1 µS/cm (summer) (Figure 3.5 (b)). Conductivity was similar at water depths for each season, with the exception of the central and northern sites in summer and the southern site in autumn. Summer had a much higher conductivity at central sub-surface, and northern surface, than central surface, and northern sub-surface. While in autumn, the southern sub-surface site had a much higher conductivity than near the surface and these two seasons had the highest overall conductivity for all sites. This is likely due to the higher evaporation levels
concentrating the naturally occurring salts in the water. There was the least amount of variation between sites in spring and little variation in conductivity was observed between seasons, except winter. In winter, conductivity at the northern site was lower than in the other seasons, where it was greater in the other two sites.

pH ranged from 8.4 at the northern surface location in summer to 6.5 at the southern depth location in winter, however over the year the pH was on average between 7.6 (summer) and 7.1 (winter 2) (Figure 3.5 (c)). pH was different at water depths for each season, with the exception of the surface water at the central site. Autumn had a much higher pH at the central surface site, than the sub-surface. There was the least amount of variation of pH between sites in spring, however, autumn showed the highest amount of variation which is probably due to both the water temperatures during this season and the high level of conductivity recorded in autumn. There was variation in pH observed between seasons. In autumn, pH at central surface and northern sub-surface increased while at the other sites, the pH deceased, while in winter, sites central surface along with both northern sites all recorded an increase in pH with the other seasons recording a decrease. There was little variation in pH in spring, though there was a decrease reported in pH recorded at the southern site, while the other five sites all recorded an increase in pH.

Temperature ranged from 28.6 °C at the central surface location in summer to 14.2 °C at the central depth location in winter 1, however, over the year the temperature was on average between 27.2 °C (summer) and 14.4 °C (winter 1) (Figure 3.5 (d)). Temperature had a large variation between water depths for each of seasons. Winter had the least variation while for summer the southern site and the central site in spring had the highest variation in temperature. These high variations are likely due to higher air temperatures. There was the least amount of variation between sites in winter and there was a high variation in temperature observed between seasons with the greatest decrease in temperatures being observed between summer and autumn, the greatest increase in temperature was between winter and spring. Between autumn and winter, sites southern sub-surface and central surface showed a slight decrease in temperature, while the other sites all recorded an increase.
Figure 3.5 Water chemistry measurements (Dissolved Oxygen (mgL\(^{-1}\)) (a), Conductivity (µS/cm) (b), pH (c) and (d) Water temperatures (°C)) taken over four seasons over a 12 month period at each sample. Winter values are the mean of two replicates measured 1 week apart.
3.6 Location – surface and depth detection of freshwater fish species

*Pseudogobius olorum*, *L. unicolor*, *N. vittata*, and *C. carpio* were detected at all three sampling sites and DNA for these species was detected at both the surface and at depth (*Figure 3.6*). *M. peelii* was detected at the southern location, both at the surface and at depth, along with at the central surface location. *Perca fluviatilis* was only detected at the southern location at both the surface and depth. *Poecilia sphenops* was also detected at each of the sampling sites and at both depths except for central depth but the number of copies of sequences was very low. These two species, one an ornamental (*Po. sphenops* (Molly)) and the other a pest species (*Pe. fluviatilis* (Redfin Perch)) are not native to the area and may not have survived long in the system. As with the water chemistry variables location (surface and depth) was analysed using one-way ANOVA and MDS to determine if the depth where the water was collected from was important.

The combined southern location (surface and depth) had a positive detection of all eight freshwater species (*Figure 3.6*). The central location showed positive detection of seven of the freshwater species, *Pe. fluviatilis* was not detected at this location. The last location, northern, had a positive detection of six of the species, with neither *Pe. fluviatilis* or *M. peelii* being detected here. A one-way ANOVA was used to determine if there was any difference between the number of sequences detected at the surface and sub-surface locations. The results of this analysis showed that there was a difference (*p*=0.028) with more sequences being detected at the sub-surface locations.

Using 16S, the central depth location only had 37.5% of the fish species detected, while the southern surface location had 100% of the species detected. The COI had three locations where only 50% of the species were detected, central depth and both northern locations.
Figure 3.6 Freshwater fish species detected at each of the locations. The average number of identified unique sequences per location was 1375 with a standard deviation of ±344.2. Unique sequences are determined with 100% query coverage and a threshold of 98% maximum identity to allow for a 2% dissimilarity threshold.

3.7 Seasonal detection of freshwater fish species

As with the water chemistry variables and location (surface and depth) seasonal detection of species was analysed using MDS to determine if the season when the water was collected from was important. Across the five sampling periods, there was some variation between the species detected (Figure 3.7). Two fish that were consistently detected using both markers were _Ps. olorum_ and _L. unicolor_. These two species have been trapped there in large numbers and are thought to dominate the waterbody community (M. Snow, pers.com. 2014).

_Nannoperca vittata_ was detected in the cooler winter and spring months. This is when it is known to breed and it was consistently detected using both the markers. There were a greater number of sequences detected in winter before the rainfall event than in the spring sampling period. _Maccullochella peelii_ along with the other species were sporadically detected. In the case of _M. peelii_, it was detected only during two seasons, summer, and the first winter sampling period. This may be due to the low number (one) in the waterbody and the locations from where the water was sampled.
Figure 3.7 Number of sequences detected for each species in each of the sampling periods. The mean number of identified unique sequences per season was 1556 with a standard deviation of ±897.8. Unique sequences are determined with 100% query coverage and a threshold of 98% maximum identity to allow for a 2% dissimilarity threshold.

Winter 2 had the highest number of unique fish sequences detected; this may have been due to the rainfall event, which occurred during the week between the two winter sampling periods. Summer and winter were the two seasons with the highest percentage of species detected using COI, 62.5%, and autumn only had a detection of 25% of the fish species. Using 16S, summer had the highest detection rate of 87.5% of the species detected and autumn had the lowest detection rate of 37.5% species detected.
4. Discussion

This study investigated whether next-generation sequencing of eDNA extracted from water samples could be used in metagenomic fish biodiversity analysis of an urban, lentic system by developing a method to detect native and invasive freshwater fish (Appendix 1). In order to assess whether this approach is robust and suitable to use in the field, two different primer sets (for the 16S and COI gene regions) were tested and the potential impacts of water chemistry, location, and timing of sample collection considered.

4.1 Species detected using eDNA monitoring

The results of the this eDNA study showed that pooling water samples collected from multiple locations around the waterbody and across all seasons detected all fish species previously been detected using Fyke netting (Table 3.4). This suggests that the molecular surveying method outlined and tested in this thesis can be as accurate as the more tradition methods of surveying, such as Fyke netting. In addition to the species detected using both methods *P. fluviatilis* and *P. sphenops* were also detected (Table 3.4). The last Fyke net surveying data came from 2014 and it is possible that these species are recent additions to this ecosystem which may be why they are not present in any of the Fyke net data. *Cyprinus carpio* and *Carassius auratus* have been listed together as these species can hybridise (Hume et al., 1983; Panicz et al., 2013; Qin et al., 2016; Taylor & Mahon, 1977) and so it is not certain whether these results are from pure genetic individuals or hybrids.

4.2 16S and COI single source and field results

Results from the eDNA surveying in this project highlight the importance of testing primers before utilising them for monitoring. The COI and 16S gene regions are both routinely used for a wide-variety of PCR-based investigations (Aylagas et al., 2016; Gariepy et al., 2007; Medlin et al., 2002; Powers, 2004), and the COI gene region is considered to be the ‘barcoding gene’ in animals (Hebert et al., 2003; Moritz & Cicero, 2004; Ratnasingham & Hebert, 2007), so it was important to test whether these would prove either individually, or in combination, to be robust and reliable for eDNA monitoring. As this method utilises shorter fragments ~150 bp, testing of primers was important, as the short fragments may not include enough discriminatory power individually.
In the eDNA results, the two primer sets consistently detected the same five fish species (Table 3.4); however, the 16S primers also detected *Po. sphenops* in two sampling periods, along with *G. holbrooki* in three sampling periods, where COI did not. The COI primers, however, detected *Cy. carpio* in four of the five sampling periods where 16S did not. An explanation of why the primers performed differently may have been identified by using single source DNA, as it was found that *Cy. carpio* and *Ca. auratus* were the two single source species that had the highest C\textsubscript{t} values for values for 16S (Figure 3.1). This late detection may have occurred because there was not a strong affinity between the 16S primers and the DNA of these species, which may suggest that for the field samples, the DNA from these species may be ‘swamped out’ (other species that have a much lower C\textsubscript{t} value being preferentially amplified (Spoto & Corradini, 2012). On the other hand, *G. holbrooki*, which were not detected by the COI primers in the eDNA samples (and had a high C\textsubscript{t} value using single source DNA) were the species with the lowest C\textsubscript{t} values using the 16S primers.

*Pseudogobius olorum* and *M. peelii* exhibited similar C\textsubscript{t} values with both markers and it is therefore likely that either primer set would have detected these two species, though in the field *M. peelii* was only detected using COI (Figure 3.1). *Maccullochella peelii* was detected mid-range with 16S, and if there are low numbers, such as in this project where only one organism may be present, of DNA available in the eDNA sample, then it is likely that this species was ‘lost’ in the field samples (Hajibabaei et al., 2011; Spoto & Corradini, 2012). The failure to detect this low abundance species using 16S may be due to the bias associated with binding PCR primers to target-template DNA in the extracted DNA (Hajibabaei et al., 2011). Primers that have a higher affinity to the target DNA, or species that have a higher abundance (i.e. *P. olorum* or *L. unicolor*) may capture more primer molecules during the DNA annealing stage of PCR (Hajibabaei et al., 2011). This means that species that have a lower primer binding affinity and/or abundance may not yield amplicons (Hajibabaei et al., 2011).
If a target marker has too few copies present in a sample, then a false negative result can occur. The low copy numbers may be due to a low abundance of the target species, degradation following collection, low efficiency of amplification or poor laboratory techniques, along with where the water was collected i.e. patchy distribution (Darling & Mahon, 2011; Schultz & Lance, 2015). These can be controlled using appropriate sampling methods by adjusting parameters of the sampling protocol such as sample volume and location, number of samples, and in the lab, the number of PCR replicates (Schultz & Lance, 2015). If a target marker is detected in a sample this may indicate the presence of the target species, though this requires the knowledge of the target species and their preferred habitat (Schultz & Lance, 2015) or it may be a false positive. False positives which can occur for two reasons, either because a non-target species is mistaken for the genetic marker of the target species, or if the eDNA samples are contaminated (Ramsey et al., 2015; Taberlet et al., 1999) also need to be taken into consideration.

In the case of this study, Cyprinus multitaeniata and Amniataba percoides are both likely to be incorrect assignment; Cy. multitaeniata has only one base pair different from Cy. carpio in the target amplified COI region, and A. percoides has only one base pair different from L unicolor. Macquaria novemaculeata (Australian bass) was detected during winter but not in any of the other periods, suggesting that this is likely to be a false positive as well, though the detected sequence had 10 copies and was a 100% match over 200 base pairs to sequences available on NCBI. Macquaria novemaculeata may not have been a true positive as there have been no reported sightings of this species in the waterbody and this species has a distribution from Tin Can Bay, Qld to Wilsons Promontory, Vic and has been translocated widely to impoundments in southeast Qld (Atlas of Living Australia; Harris, 1988; Reinfelds et al., 2013).

A robust study design may overcome the potential limitations of this method, though this project has shown that even with a volume of water sampled, it was still able to detect a low abundance species such as M. peelii. The number of samples needed, or primer design, needs to be taken into consideration (Moyer et al., 2014). For projects, such as this one, where multiple species are being detected, it is likely robust results will be achieved by analysing samples with multiple primers instead of trying to design primers that will accurately detect large numbers of organisms with a variety of densities.
Filtration of more litres of water may have overcome the false negative issues outlined above, though this may lead to its own set of problems. For example, small micron filters clog after a few hundred millilitres of freshwater from lakes or rivers have been processed, meaning that filters must be changed often with large amounts of water. Extra filtration can be time consuming: during the project where there were high levels of algae present in the samples and filtration of 250 mL could take around 5 hours to complete.

Multiple smaller filtrations would mean that the concentration of the target DNA would be too low to amplify well, which means there would be added steps to concentrate the DNA, also multiple filtrations would increase the cost of this step, so it is unlikely that this would be a cost effective solution.

Contamination and the subsequent risk of the generation of false positives is another possible limitation when using this method. This method uses massive parallel sequencing and even minuscule contamination can be detected (Cannon et al., 2016). The detection of rare species may be overcome with an increase in qPCR cycles, but this does lead to an increase in amplification, which in turn increases the likelihood of contamination. It is critical to have stringent lab and collection practices in place to reduce the possibility of cross contamination and obtain true results (Cannon et al., 2016).

For species that may be present in lower abundance, or may not shed DNA at the same rates as other species, singletons may be more important. There were single sequences (singletons) detected in the data and it was unclear whether these were true detection of species or artefact. This project used a platform that has an increased read platform (Illumina MiSeq) and therefore these singletons are likely to be artefacts and can be readily discarded when doing biodiversity assays (Edgar, 2013).

This assumption was tested to ensure that there was no loss of diversity if singletons were removed and as the number of species detected was not altered, it was decided to remove the singletons as a form of quality control (Flynn et al., 2015). The results from this study and the Flynn et al. (2015) study showed that retention of singletons did not increase species detection. However, depending on the needs of the project, the trade-off between genuine OTU numbers and the ability to detect genuinely rare species needs to be taken into consideration.
4.3 Comparison of sequence counts between location, depths and seasons

Currently there is little known about the impact of habitat on the number of sequence counts obtained in an eDNA sample (Hänfling et al., 2016; Port et al., 2016), however, there are some trends detected in this study that warrant discussion and investigation. Understanding the link between sequence counts and location/season may help inform investigators when developing their study design.

The southern depth location had the highest number of sequences detected by both the 16S and COI primers (Figure 3.3) and the largest number of unique sequences. This result may be due to the presence of a fountain providing an underwater structure encouraging fish to congregate. In addition to this, the re-circulation of water through the fountain would mix the waterbody water in this area (Akinwole et al., 2014; Kostic, 1999), thus creating a more homogenous mixture than other areas of the waterbody. This would likely increase both the total amount of DNA (i.e. number of sequences) and increase the chance of detection of a wider range of species (the southern depth location had the most species detected (seven)). When using the 16S primers all six sampling locations detected consistently high numbers of DNA copies. At all locations, 16S primers detected a higher number of sequences than COI primers.

The sequence copies for both 16S and COI were low in both the central and northern locations. The northern surface location (which had the lowest number of sequences of all sites detected using 16S) may have had low copy numbers due to particle settling rates as the area is undisturbed by fountains or significant water movement and therefore the DNA may drop to the sediment faster than in other locations (Turner et al., 2015). In situations such as this, sediment sampling may prove to be a better method of detection – this warrants further investigation. The central surface location (which had the lowest numbers using COI lacks any structure, vegetation and the low numbers may simply be an indicator of a lower number of fish, or it could possibly be related to water chemistry in this particular area not being optimal for DNA preservation.

Not only did location influence copy numbers, so too did season. The cooler sampling periods yielded the higher number of sequences detected using both primer sets (Figure 3.7). The increased sequence counts in the cooler seasons may be simply due to temperature (Moyer et al. (2014) showed that DNA was 1.67 times less likely to be detected for every 1°C water temperature increase), but could also be related to water chemistry or increased mobility of the fish (possibly related to breeding events) at these times of year.
Surprisingly, summer had the greatest number of species detected (seven) and autumn had the least number of species detected (three) despite the inverse pattern seen in the number of counts. This high number of species detected may have been due to the higher level of activity exhibited by the fish species leading to more DNA being shed, along with less non-target DNA being present.

In this study, summer and spring’s average water temperature were high (Figure 3.5), which likely affected the degradation rates of the DNA, as the overall sequence counts were low for these seasons, this may have been due to the water chemistry (Lacoursière-Roussel et al., 2016; Robson et al., 2016; Strickler et al., 2015), extraction practices or due to a change in mobility of fish in the warmer weather (Becker & Genoway, 1979; Rice et al., 1983). Several studies have assumed that temperature is the main influence on DNA survival (Hofreiter et al., 2001; Lindahl, 1993; Willerslev et al., 2004) since temperature has been shown to exhibit a strong influence on DNA degradation. Water temperatures of 5°C lead to much less DNA degradation than water temperatures of 20° and 35°C (Strickler et al., 2015).

Autumn and winter had the coolest average temperatures (Figure 3.5) and winter 2 had the highest amount of DNA detected which may have been due to the rainfall event that happened between the winters 1 and 2 sampling events. This influx of fresh water may have reduced the degradation of the DNA along with encouraging fish to be more active (Pilliod, Goldberg, Arkle, et al., 2013; Pilliod, Goldberg, Laramie, et al., 2013). Despite the cooler temperatures, autumn had one of the lowest sequence counts of all of the sampling periods: only summer was lower. This suggests that there are other factors than temperature influencing detection of DNA (Corinaldesi et al., 2008).

4.4 Water chemistry effects on eDNA

Studies have indicated abiotic environmental characteristics influence DNA counts through a variety of mechanisms. These include the indirect impact of higher temperatures which increase enzyme kinetics and microbial metabolism (Hofreiter et al., 2001; Mann et al., 2009; Moyer et al., 2014; Okabe et al., 2007; Walters et al., 2009) along with other experimental studies showing how light exposure has been shown to affect the rate of DNA degradation (Dick et al., 2010; Green et al., 2011). Both indicate continuing uncertainty around the effects of environmental variables on eDNA.
The water chemistry averages were within the normal ranges for a lentic system in Australia, (Kay et al., 2001). However, the temperature ranges, especially in summer and spring, were much were higher than average (Figure 3.5). Conductivity ranges in summer and winter 2 were the most variable, while all other water chemistry recordings showed little variation across the five sampling periods.

While temperatures greater than 50˚C can degrade DNA by denaturation, it is more likely that moderate temperature does not directly relate to eDNA degradation, and moderate temperatures similar to the ones recorded during this study are more likely to encourage microbial metabolism and exonuclease activity (Hofreiter et al., 2001; Pote et al., 2009; Zhu, 2006). The microbial activity in the water body directly contributes to enzymatic hydrolysis by producing exogenous nucleases that break down DNA into its components (Lindahl, 1993). DNA has been shown to be broken down by chemical hydrolysis, either through exposure to acid or by enzymatic hydrolysis (Strickler et al., 2015). Water with a higher acidity has been shown to have a higher degradation rate than alkaline water (Strickler et al., 2015) but as the average pH across all of the sampling seasons was ~7 (neutral) this is unlikely to have affected the overall DNA counts.

Water conductivity was highest in summer and autumn (Figure 3.5) and these two seasons had the lowest overall DNA counts for both markers. For this study, spring had consistently high DNA counts for both the 16S and COI markers and this may be due to the low conductivity, and moderate temperatures (25˚C).

This study has identified several key issues concerning eDNA and water chemistry. Further research into the types of genetic materials that organisms most commonly release into their environment would be valuable, and would support the design of better DNA collection and extraction methods. Increased quantification of production rates and degradation of DNA in the environment is needed, as are studies in a wide variety of situations such as; lentic vs. lotic water bodies, water bodies that are either slightly acidic or basic along with a comparison of eDNA results from water compared to sediment results.
Lacoursière-Roussel et al. (2016) suggest that the concentration of DNA found in environmental samples is a variable which depends on both the rate of shedding experienced by the target organism along with the rate of degradation in the field, these factors coupled with complex interactions between metabolism, ecology, and environmental conditions (Barnes et al., 2014; Strickler et al., 2015). This suggests that a comparison of taxonomic group shedding rates would also allow researchers to decide if results from studies such as this one would be transferable across species.

Laboratory based experiments (Pilliod et al., 2014) suggest that DNA persists in aquatic environments for between eight and eighteen days, depending on temperature and light conditions while Dejean et al. (2011) determined that eDNA detection in freshwater systems is determined by the length of DNA extracted, 300-400 bp lengths could be detected after one week under controlled conditions, while short lengths are more slowly degraded and therefore more readily detected in environmental samples. In tropical environments, it has been found that there is a greater amount of DNA shedding at higher temperatures than at moderate temperatures (Robson et al., 2016), though this study did suggest that there was no real difference in degradation rates of DNA between the moderate and high temperatures (~35°C). The temperatures in this project were not recorded daily, though recorded water temperatures did not exceed 29°C.

These experiments are useful in helping to design experiments for eDNA collection though they were run under northern hemisphere conditions and may not accurately reflect the conditions and results found in the southern hemisphere as temperature along with water chemistry are likely to be different due to environmental differences. Water chemistry characteristics, along with other biotic influences, and whether they work synergistically or antagonistically to preserve or degrade DNA are important factors that need to be considered. Further studies exploring these would increase the interpretability of eDNA surveillance results.

4.5 Location and seasonality

Both location of the water sample collected, along with the season collected was shown to affect the results of this study. This indicates that to have a robust method for environmental surveying, these factors need to be taken into consideration. A knowledge of fish ecology and habitat preference is another way to overcome this limitation.
Across the six sampling locations (Figure 3.6), there was a range of four to seven species detected. This means there was no location where all eight species were detected over the course of the project. Repeated detection of fish species seems to be dependent on the ecology and habitat preferences of the fish species.

*Leiopotherapon unicolor* and *Ps. olorum* are both highly abundant in this area of Western Australia (Beuma, 1979; Beumer, 1979; Neira et al., 1992). *Leiopotherapon unicolor* is a highly mobile species (Humphries & Walker, 2013) and *P. olorum* is a bottom dweller that has been observed performing aquatic surface respiration (Larson, 2001), which, along with their abundance in this region, are the likely reasons these two species were detected at all locations. *Cyprinus carpio* are also a highly mobile species (Koehn, 2004) and as this species has a high spawn rate, with females laying up to 300,000 eggs in a single spawn (Hicks et al., 2012), it is no surprise that this species was detected at all locations. *Nannoperca vittata* were also detected at each of the locations, though this is likely due to a spawning event. This is a relatively small fish (growing to ~8 cm) (Axelrod et al., 1986) and it may not shed DNA at the same rates as the larger fish in the waterbody, which would mean that outside of spawning events, they could be much harder to detect.

The ornamental species, *Poecilia sphenops*, were detected in all but one of the locations (central depth) (Figure 3.6). This is an ornamental species, and it may have been released into the waterbody, and then preyed upon, as it was only detected twice throughout the project and not detected during the previous Fyke net trapping data. *Gambusia holbrooki* were detected at both of the northern locations along with the southern depth and central surface locations. These fish are typically found in shallow vegetated areas (Schaefer et al., 1994), which describes the habitat found at the northern location. They also prefer standing to slow moving water (Pyke, 2005), which is typical of the other two locations this species was detected. *Maccullochella peeli* was detected at both southern locations along with the central surface location, this species tend to stay in their home territories, that is, specific areas of a waterbody or river (Kearney & Kildea, 2001) which may be why they were not detected in other sampling locations.
Finally, *Perca fluviatilis* was only detected at the southern location. This species is restricted to the south west corner of Western Australia, though it has been found in 10 river systems including the Swan River (Morgan et al., 2003). This species has broad environmental and habitat tolerance and is successful in systems where there is an absence of predators (Morgan et al., 2003). As *M. peelii* is a predatory and territorial species, the *P. fluviatilis* may have been preyed upon, which may be why it was not detected in any of the other locations, or it may not have been detected due to low abundance in the waterbody, though it had not been detected during the previous Fyke net surveying.

Across the five sampling periods, there was a range in the number of species detected (Figure 3.6). *Leiopotherapon unicolor* and *Pseudogobius olorum* were the only species that was consistently detected during each of the sampling periods, this is likely due to the abundance and high mobility of these species. Other species such as, *M. peelii* have a very low abundance in the waterbody, there is only likely to be one present, and are generally thought to be sedentary from late summer to late winter (Koehn, 2009), which is possibly why they were only detected during the summer and winter 1 sampling periods. Seasonality in detection of species is likely due to seasonal changes in activity (de Souza et al., 2016; Furlan et al., 2016). For example, *Nannoperca vittata* were detected in both winter sampling periods and in spring and these fish breed from July to November (Morgan, Beatty, et al., 2011). Detection during these periods is likely due to the large amounts of DNA are released into the waterbody during the spawning event (Bylemans et al., 2016; Erickson et al., 2016). Spawning events are often characterised by higher concentrations of nuclear than mitochondrial eDNA, while outside the reproductive period it is likely that both forms of eDNA are present in equal amounts (Bylemans et al., 2016). This is contrary to a study on bighead carp (Erickson et al., 2016) that suggested that there was no relationship between spawning events and eDNA detected though the study does state that they may have over saturated the sample with DNA or they may not have collected water from the plume during the spawning event.

*Cyprinus carpio* is a highly abundant species in the waterbody, and can been seen surfacing around the waterbody. They were consistently detected in all sampling periods other than autumn. Both *Perca fluviatilis* and *Poecilia sphenops* were species that were only identified during the first two sampling periods, with *P. fluviatilis* being detected in summer and *P. sphenops* in summer and autumn.
*Perca fluviatilis* have been spotted in this region (Hourston et al., 2014) and it may have been due to low densities that this species was not detected in subsequent sampling periods. *Maccullochella peelii* have been known to feed on *P. fluviatilis* which may be a reason that redfin was only identified during the one sampling period, though there is no way to verify this.

### 4.6 Importance of multiple location surveying

Understanding the distribution, population dynamics, and abundance of species traditionally requires the collection and identification of organisms at study locations via methods such as trapping (Beggel et al., 2015; Moyer et al., 2014; Suter et al., 2015; Thomas et al., 2016; Viard et al., 2016). Disciplines such as phylogenetics, conservation biology, and ecology rely on species detection and morphology has been useful in the past (Moyer et al., 2014). However, reliable detection is gaining importance to ensure that species in low abundances - either rare species or early invasion fronts - are taken care of as needed (Ficetola et al., 2008; Goldberg et al., 2011; Goldberg et al., 2013; Jerde et al., 2013; Jerde et al., 2011; Lodge et al., 2012; Pilliod et al., 2014; Pilliod, Goldberg, Arkle, et al., 2013; Pilliod, Goldberg, Laramie, et al., 2013).

Over the course of this project, no single location or season showed a presence of all species at the same time. This highlights a limitation that exists with this technique, but this can be overcome with robust sampling design. Taking multiple samples from each waterbody and combining them, creates a better ‘snapshot’ and if the survey is targeting a specific species, then knowledge of the organism’s ecology and biology could allow for targeted surveying with greater accuracy.

Collection of water from various locations around the waterbody has shown to be important, especially for species that may be present in low numbers, or that may be less mobile. Species density is very important for accurate detection and in a study published in 2014 (Moyer et al., 2014), multiple water samples (1 L samples) from the ponds with high-density numbers of fish gave a 95-100% probability for the confirmation of the presence or absence of a target species. From a single water sample (1 L sample), this probability of detection dropped to 55% (Moyer et al., 2014). For ponds that contained medium and low-density fish numbers the chance of detecting the target species dropped to 7% and 3% respectively if a single sample was taken, and it was shown that 42–73 L and 100 L respectively would be required to achieve a 95 – 100% probability (Moyer et al., 2014).
This study showed similar results; water from a single location did not detect all species, and without multiple sampling, low-density species such as the \textit{M. peelii} were unlikely to be detected. For this study, using COI, the probability of detection across the waterbody was lowest for \textit{M. peelii} where the chances of detection ranged between 0.00-5.84\% and \textit{Ps. olorum} had the highest probability of detection with a detection range of 27.45-49.72\%. Using 16S, the probability of detection was least for \textit{Poecilia sphenops} with a detection range of 0.00-0.32\% across the waterbody, \textit{Ps. olorum} were the species most likely to be detected in this waterbody with a detection range of 40.51-91.93\%. The sensitivity of this method needs to be taken into account and can be increased by testing greater volumes of water, or by developing a primer set or sets with an increased eDNA sensitivity.

Without currents and water flow, DNA does not seem to be evenly distributed throughout the system, and without a robust sampling design, it would be easy to have a false negative on some species. Moyer et al. (2014) had results that differed from this study as they found that DNA was most readily detected from the surface of their systems, while in this study, DNA was more readily detected from the water-collected sub-surface (p = 0.028) (\textbf{Figure 3.4}) (~3.1 m). In the Moyer et al. (2014) study, African jewelfish (\textit{Hemichromis lifalili}), which are benthic dwellers, were most readily detected on the surface of the pond. The ponds utilised by Moyer et al. (2014) were spring fed and this may have been the reason for the differences between the two studies as it is not known if the study waterbody is spring fed. However, the two studies have shown that even bottom dwelling species, such as the \textit{P. olorum} (Gill et al., 1996), can be detected from surface water.

The results of this project, primarily in regards to the \textit{M. peelii}, supported multiple location, time, and depth surveying to detect species that are present in low numbers, though there is still a high level of stochastic variability that needs to be addressed. This project has highlighted the importance of the need to understand the target species ecology along with its biology.
4.7 Early detection and what this project means for the future

Environmental DNA metabarcoding has a role in environmental surveying; it has the potential to detect early invasion fronts of non-native species as well as targeting rare and low-density species. It has the added benefits of being able to be used in habitats where traditional methods cannot be easily implemented, due to either logistics or limited effectiveness in the habitat. With further modification, primers used and developed for this project could be used for further freshwater finfish studies around the world, allowing for a global freshwater biodiversity assessment, not just a local or national assay.

When using the molecular approach it is important to take limitations highlighted in this project into consideration (e.g. how many sites and how much water needs to be collected to get accurate results) and plan accordingly. As this is often the case with traditional approaches as well, these considerations are minor, and with the results showing that species in low abundance can be detected using the molecular method certainly has the potential to be the surveying way of the future.

Early detection of invasive organisms, whether they are aquatic, terrestrial, or air borne is crucial as it allows for the monitoring of, and, if possible, the removal of the invasive species before it dominates an ecosystem (Crooks et al., 1999; Larson et al., 2011; Poon et al., 2007; Pyšek & Richardson, 2010; Sousa et al., 2014). Traditionally, early stage invader detection has been near impossible without species density exceeding certain thresholds (Harvey et al., 2009; Hulme, 2006). These detection thresholds are dependent on the methods used for monitoring and may be reached only once the species has become well established in the ecosystem (Myers et al., 2000). At that point, the cost for control or removal can become prohibitively high (Dejean et al., 2012) and complete eradication of the invasive species may be impossible to achieve. Without complete eradication of the invader, endemic species recovery may be compromised along with the environment (Myers et al., 2000). If there are multiple invasion fronts it can be difficult to prioritise which invasion fronts are important, and early detection may help to decide what course of action needs to be taken (Cruz et al., 2008; Dawson et al., 2015; Howald et al., 2007; Pimentel et al., 2005).
By improving the methods for probability of detection, along with reducing the impact on the ecosystem being monitored, funding and person-hours can be spent on eradication programs. While much literature has been focussed on post detection control and eradication, along with the prevention of invasions, few have looked at the role of early detection (Mehta et al., 2007). The use of eDNA may increase the chances of accurate detection at an invasion front, which in turn may reduce the damage to the ecosystem, along with making control of the invasion less expensive and it may increase the effectiveness of the eradication efforts (Mehta et al., 2007).

Until recently, metagenomics and metagenetics have only been used in analysis of microorganism biodiversity; however, these approaches are now being applied to macro-organisms (Bohmann et al., 2014). As sequencing technology has improved, the ability to detect eDNA in nearly all environments - terrestrial, marine, freshwater, and airborne - has improved drastically, availing researchers of a new detection tool for a wide variety of applications (Bohmann et al., 2014; Foote et al., 2012; Moyer et al., 2014; Taberlet et al., 2012). These applications include, for example, conservation biology and informing policy decisions (Schnell et al., 2012), population genetics (Zhu et al., 2011), along with species detection and biomass (Thomsen, Kielgast, Iversen, Wiuf, et al., 2012) and increasing our understanding of ecosystems through invasive species or wildlife disease detection (Dejean et al., 2012; Folloni et al., 2012; Walker et al., 2007).

4.8 Troubleshooting

This project has faced issues and challenges and within this section are the methods that have been explored to overcome the challenges. One of the challenges faced in this project is the contamination of the blank field samples. This contamination, once sequenced was shown to originate from the same target fish species as those expected to be found at the sample location.

It has also been proposed that any sample which contains very little DNA is at risk of being contaminated with exogenous DNA from the same species and that different forms of contamination have to be treated using different forms of decontamination (Champlot et al., 2010). Thomsen, Kielgast, Iversen, Moller, et al. (2012) found that the detection threshold below which DNA could not be detected was, on average, 25 DNA molecules per 400 mL of water. This would also support the previous suggestion that not enough water was filtered from each sample.
Low copy numbers were another potential issue identified during this study. Thomsen, Kielgast, Iversen, Moller, et al. (2012) found that for 400 mL of water filtered yielded concentrations of between 48 and 214 DNA molecules so filtering 250 mL as per this study may not have been enough to get a high concentration of DNA. The water samples from both autumn and summer were then blasted against the local database to identify which species were present in the lake. There were a low number of unique sequences for all the samples sequenced which was likely due to small amount (250 mL) of water filtered.

After summer and autumn eDNA samples had been extracted at the Edith Cowan University labs it was discovered that there was amplicon contamination from the single source freshwater fish species that had been used to test the COI and 16S primers. This contamination was detected in the lab blanks that were analysed alongside the environmental samples. Contamination of samples when using degraded DNA (such as DNA found in water) is not uncommon (Lusk, 2014) and the problem arises because there is an inverse relationship between the concentration of the DNA and the length of the target fragment (Champlot et al., 2010; Lusk, 2014). Common problems that most often occur in the literature include sample contamination, laboratory surface contamination, carry over contamination and contamination of reagents (Champlot et al., 2010) and it is likely that in this case, the contamination was spread through the air conditioning system.

A recent study suggests that while negative control libraries that have been prepared from ‘blank’ samples recovered the highest frequency of contaminants, low frequency contaminates were also not well controlled (Lusk, 2014). Negative controls processed in parallel are essential for detecting contaminated samples, but these negative controls are limited in their ability to recover low-frequency contamination (Lusk, 2014). It has been suggested that there is not a single method of decontamination for all possible contamination sources and the most common methods of decontamination are not stringent enough to decontaminate short DNA fragments found in low concentration (Champlot et al., 2010). Contamination is often underestimated and it has been suggested that the greatest danger may not be the contamination itself but rather in ignoring or neglecting it completely (Champlot et al., 2010).
In this study, despite rigorous cleaning and lab practices the lack of an isolated room meant that there was amplicon contamination found in DNA free nuclease water showing that at the very least, the blanks were contaminated. One of the biggest problems in this project is that there was not a ‘blanket’ contamination i.e. not all the species that were found in the blanks were also found in each of the environmental samples. For contamination by exogenous species, which may occur during collection and processing, Champlot et al. (2010) suggest that a minimum of gloves up to a full body protection suit should be worn in an attempt to reduce contamination and cross contamination between sites. Contamination of laboratory surfaces and instruments such as centrifuges and pipettes can occur at any stage of the processing workflow; this contamination can be reduced using clean rooms and decontamination processes. These processes include UV irradiation and the use of bleach and DNAerase where necessary. Before the first samples were processed and before subsequent processing, the laboratory surfaces were cleaned using bleach, ethanol, or DNAerase as necessary. All equipment that was brought into the lab was also cleaned and any that could be autoclaved or undergo UV treatment were treated in this way before being used.

Contamination of reagents, whether by carry over or during reagent production is challenging to overcome as decontamination must be highly effective without adversely affecting the efficacy or sensitivity of the qPCR (Champlot et al., 2010) and during this project reagents were changed out to reduce contamination. It is expected that there were amplicons spread throughout the building in the air-conditioning system. Due to the airborne contamination even with changing the reagents the amplicons were present once the reagent containers were opened. After UV irradiation, the nuclease free water was found to be contamination free, and so the water was irradiated before being used in the lab. Commercial DNA reagents have been found to contain contamination by human DNA along with DNA from domestic animals such as cows, pigs and chickens (Leonard et al., 2007). There have been many methods for dealing with reagent contamination as outlined by Champlot et al. (2010) but these methods have shown inconsistent results and it was shown that these methods were also ineffective for removal of DNA with a molecular mass of less than 200 bp which is the size of the DNA targeted in this project.
4.9 Assessment of methods used

Completing the second round of extractions at the DoF facilities allowed us to test the robustness of this project for use by agencies for both monitoring and detection of low density species. One of the questions that was tested during this project was if there were any issues processing the samples in a lab where other fish, mainly marine, were also processed. This was an important question to answer as concerns over contamination would make this approach unfeasible for a department such as DoF. Initial tests showed that the workflow detailed above removed all contamination when using the 16S primers, and the COI primers showed late amplification (~42 – 45 cycles) of contaminants.

For further studies, several changes to the workflow are recommended. The first of these changes is to ensure that all time points, i.e. summer and winter, were processed separately, from filtering and extraction to dilution and possibly even qPCR. This would be to ensure that there is no cross contamination between time points (seasons). Secondly, if possible the qPCR plates would be loaded in a room separately from the post-PCR workspace. This would ensure that there is no amplicon contamination on the plates. One of the things learnt while doing the extractions and qPCR at the labs at ECU is that air-conditioning systems can transport DNA and amplicons throughout a building, which means that without a purpose built, separate laboratory for steps such as filtering/extraction along with pre- and post-PCR it is very hard to control contamination.

With the 50-cycle protocol that was used during this project, it is easy to pick up background contamination from air-borne DNA and amplicons. Using a 50-cycle protocol means that the method is sensitive for picking up rare species, or fish with low numbers but it does also mean that the method is much more likely to amplify the contamination. For future studies a 40-cycle protocol should be employed, as there is less likelihood that background contamination from amplicons will be amplified during the process. This may mean that fish that are present in low numbers may not be identified, but will make the method more robust.

This method of surveying for freshwater fish has been shown to be both efficient and sensitive in the detection of not only high density species i.e. *Cy. carpio* but also species expected to be present in very low densities i.e. *M. peelii*. This suggests that with some fine tuning in respect to contamination control and the number of qPCR cycles needed to detect the low density species this method can be used in a practical.
5. eDNA, is it worth it?

This chapter looks at costs involved with the traditional Fyke net surveying approach used by the Department of Fisheries Western Australia, and costs identified with molecular surveying. Three scenarios will be examined; Fyke net surveying using data supplied by the Department of Fisheries, molecular multiplex surveying of potentially all fish present using data obtained during this project and molecular surveying of a single target species. The comparison of the three scenarios lead to discussion as to whether the way of the future is through the molecular approach.

5.1 Method

There are many costs involved with the three scenarios analysed. For the following scenarios, several assumptions were made (Appendix 4). Time needed for molecular surveying at the waterbody was based on the time taken during this project, while the Fyke net surveying times were estimated from literature and anecdotal evidence from staff at the Department of Fisheries, Western Australia. The time needed at each waterbody was much greater in the traditional Fyke net surveying as a return trip is needed to collect the nets and score the fish. This return trip is one of the big restrictions of Fyke netting method, significantly reducing the number of waterbodies that can be surveyed compared to molecular surveying.

The costs involved with transport include the hiring of a four-wheel drive through a specific Australian hire company (Appendix 4), which may be reduced if a fleet car is used instead. The national average for fuel costs in Australia was used to give the average fuel costs for the week (1-18 October 2016), based on waterbodies being 30 km apart and a fuel economy of 6.9 L/100 km. (Appendix 4). The distance of 30 km, may be an overestimate for city surveillance, but the figure was averaged between city and country distances. It was estimated that to travel 30 km would take 45 minutes. This estimation was used for all scenarios. All of the assumptions used for this analysis are outlined in Appendix 4.

These figures were used for all three scenarios as there was no difference in the distances between waterbodies; the only difference was the number of waterbodies that could be surveyed each week. The salary for the staff was converted to an hourly rate by dividing the weekly salary by 38 hours. This hourly rate was used for both the field staff, as well as the laboratory staff for the molecular surveying techniques.
For molecular surveying, laboratory time is required as well as field time. Laboratory time includes filtering of the water, DNA extraction, qPCR, and then NGS. Along with these lab assumptions, the time taken to deconvolute the data and identify fish species has been calculated based on the experience during the project. The length of time needed to complete these steps is likely to decrease as experience is gained.

5.2 Results

The analysis of the three scenarios of surveying identified the varying limitations as well as benefits for each method. Some of the limitations highlighted were; fyke netting, the number of work hours needed to get an accurate survey result; single source surveying, not a complete analysis of the waterbody; and molecular multiplex surveying, the costs involved with setting up facilities for analysis of eDNA. This highlighted the need for robust design of each survey so that the best method could be utilised.

For Fyke net surveying, one of the restrictions identified was the number of nets that could be set each week. As the nets have to be set and then collected, weekly numbers set by a team of two is restricted to ten waterbodies (Table 5.1). From the assumptions outlined, day two will be the longest workday with 11.3 hours designated. Day 4 is the shortest day of the week as there are no new waterbodies to be surveyed so only the nets from the previous day need to be collected and scored. The fifth day has no fieldwork as the number of field hours has already reached 34.3 hours, and is close to the maximum of a 38-hour working week. The limitation of needing to retrieve nets means that the fifth day is not utilised in the field and this contributes to the lower number of waterbodies that can be surveyed each week.
Table 5.1 Fyke netting maximum number of waterbodies that can be surveyed each week by a team of two working together in the field. Each person works the same number of hours (h).

<table>
<thead>
<tr>
<th></th>
<th>Work Days</th>
<th></th>
<th></th>
<th></th>
<th>Work week</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Number of waterbodies</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Travel time (h)</td>
<td>1.5</td>
<td>4.5</td>
<td>4.5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Setting Fyke nets (h)</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Scoring Fyke nets (h)</td>
<td>0</td>
<td>4.5</td>
<td>3</td>
<td>4.5</td>
<td>0</td>
</tr>
<tr>
<td>Net Maintenance (h)</td>
<td>0</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>Workload (h)</td>
<td>4.5</td>
<td>11.3</td>
<td>10.7</td>
<td>7.8</td>
<td>0</td>
</tr>
</tbody>
</table>

The model for traditional surveying shows that over the course of 40 weeks (1,520 hours) a team of two can comfortably sample up to 340 waterbodies a year, using the assumptions made here (Figure 5.1). Molecular surveying does not have this restriction. Molecular multiplex surveying, with separate field and lab teams, can comfortably survey 560 waterbodies a year when the lab processing is added to the weekly hours (Table 5.2). If a single species is the target of the molecular surveying, and Sanger sequencing is utilised to process the eDNA extracted from the water samples, up to 780 waterbodies including lab processing can be surveyed each year.

A larger number of waterbodies can be surveyed using molecular methods allowing for a greater depth of coverage. There is opportunity using this method either to target more waterbodies, or to monitor waterbodies more regularly. Regular monitoring may lead to earlier detection of problem species or, by targeting more waterbodies; it means there is a greater coverage of an area that may have a suspected invasion.
Table 5.2 Total number of waterbodies multiplexed surveyed each week, combining lab and field team hours.

<table>
<thead>
<tr>
<th>Work Days</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of waterbodies</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Travel time (h)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2.25</td>
</tr>
<tr>
<td>Molecular Surveying (h)</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>1.5</td>
</tr>
<tr>
<td>Workload (h)</td>
<td>5.25</td>
<td>5.25</td>
<td>5.25</td>
<td>5.25</td>
<td>3.75</td>
</tr>
<tr>
<td>Lab time 14 waterbodies (h)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.10</td>
</tr>
<tr>
<td>Total weekly hours</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>39.35</td>
</tr>
</tbody>
</table>

Table 5.3 Total number of waterbodies single target surveyed each week, combining lab and field team hours.

<table>
<thead>
<tr>
<th>Work Days</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of waterbodies</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Travel time (h)</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3</td>
</tr>
<tr>
<td>Molecular Surveying (h)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Workload (h)</td>
<td>6.75</td>
<td>6.75</td>
<td>6.75</td>
<td>6.75</td>
<td>6.00</td>
</tr>
<tr>
<td>Lab time 19 waterbodies (h)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.55</td>
</tr>
<tr>
<td>Total weekly hours</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38.75</td>
</tr>
</tbody>
</table>
Figure 5.1 Cumulative time to sample waterbodies for presence of fish species not including lab, or Fyke netting set up costs. A comparison of three methods of sampling showing the number of waterbodies sampled weekly for 40 weeks (~38-hour week). Molecular (multiplex) and molecular (single) both include field and laboratory. Horizontal lines shows the number of waterbodies that can be surveyed by each method in a year. Number of waterbodies rounded to the nearest 10.

One concern that may be raised for molecular surveying is the cost of setting up a lab for processing of eDNA samples. Specialised equipment can be very expensive and for this chapter start-up costs have been estimated at AUD66,300 for multiplexed molecular surveying, AUD51,300 for single target analysis and AUD16,300 for traditional surveying. The model (Figure 5.2) shows that even with the high start-up costs involved with molecular surveying it takes less than a year to break even. When surveying for a single target, the break-even point is 95 waterbodies, and for multispecies diversity 145 waterbodies.

When plotted against each other (Figure 5.3) it is clear that single source species targeting has the greatest depth of surveying coverage, though with the high numbers of lakes that can be surveyed it means that the overall costs are higher. Fyke net surveying is the second most cost-effective method, but has the limitation of being highly labour intensive. Multiplex surveying which has the added benefit of observing a ‘snapshot’ of the lake biodiversity, also this method has found to be the cheapest of the methods analysed (Table 5.4).
Figure 5.2 Comparison between the three methods including start-up costs. Showing break-even intercepts of both molecular (multiplex) surveying and molecular (single target) surveying with traditional Fyke net surveying.
Figure 5.3 Cumulative costs for each method of detection over one year of surveying. Showing intercepts for each three methods including yearly cost and the number of waterbodies surveyed. Dollar values rounded to the nearest 1000 (AUD), number of waterbodies rounded to the nearest 10.
Table 5.4 Weekly and yearly break down of cumulative hours and costs needed to survey waterbodies using all three methods, including weekly lab time. Based on a 38-hour week and a 1520-hour year.

<table>
<thead>
<tr>
<th>Waterbodies</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>340</th>
<th>560</th>
<th>780</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traditional</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lake costs (AUD)</td>
<td>1116</td>
<td>1694</td>
<td>2273</td>
<td>2958</td>
<td>3539</td>
<td>4222</td>
<td>4803</td>
<td>5381</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>4.10</td>
<td>8.20</td>
<td>14.25</td>
<td>17.15</td>
<td>21.25</td>
<td>26.10</td>
<td>30.20</td>
<td>34.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Km travelled</td>
<td>120</td>
<td>180</td>
<td>240</td>
<td>330</td>
<td>390</td>
<td>450</td>
<td>540</td>
<td>600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Multiplex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lake costs (AUD)</td>
<td>676</td>
<td>935</td>
<td>1194</td>
<td>1453</td>
<td>1780</td>
<td>2144</td>
<td>2403</td>
<td>2662</td>
<td>3011</td>
<td>3270</td>
<td>3633</td>
<td>3892</td>
<td>4220</td>
<td>4479</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>182,029</td>
</tr>
<tr>
<td>Time</td>
<td>8.75</td>
<td>10.8</td>
<td>12.85</td>
<td>14.9</td>
<td>17.7</td>
<td>21.25</td>
<td>21.8</td>
<td>23.85</td>
<td>27.25</td>
<td>29.3</td>
<td>31.35</td>
<td>33.4</td>
<td>36.2</td>
<td>38.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,527.4</td>
</tr>
<tr>
<td>Km travelled</td>
<td>60</td>
<td>90</td>
<td>120</td>
<td>150</td>
<td>180</td>
<td>210</td>
<td>270</td>
<td>300</td>
<td>330</td>
<td>360</td>
<td>420</td>
<td>450</td>
<td>480</td>
<td>510</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20,160</td>
</tr>
<tr>
<td><strong>Single target</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lake costs (AUD)</td>
<td>507</td>
<td>741</td>
<td>975</td>
<td>1209</td>
<td>1617</td>
<td>1746</td>
<td>1980</td>
<td>2319</td>
<td>2779</td>
<td>3013</td>
<td>3247</td>
<td>3481</td>
<td>3889</td>
<td>4123</td>
<td>4358</td>
<td>4592</td>
<td>5086</td>
<td>5313</td>
<td>5633</td>
<td>5867</td>
<td></td>
<td>228,817</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>4.75</td>
<td>6.25</td>
<td>7.75</td>
<td>9.25</td>
<td>10.75</td>
<td>13</td>
<td>14.5</td>
<td>16</td>
<td>19</td>
<td>20.5</td>
<td>22.75</td>
<td>24.25</td>
<td>25.75</td>
<td>27.25</td>
<td>28.75</td>
<td>31</td>
<td>34.25</td>
<td>35.75</td>
<td>37.25</td>
<td>38.75</td>
<td></td>
<td>1,511.25</td>
<td></td>
</tr>
<tr>
<td>Km travelled</td>
<td>60</td>
<td>90</td>
<td>120</td>
<td>150</td>
<td>180</td>
<td>210</td>
<td>270</td>
<td>300</td>
<td>330</td>
<td>360</td>
<td>420</td>
<td>450</td>
<td>480</td>
<td>510</td>
<td>540</td>
<td>600</td>
<td>630</td>
<td>660</td>
<td>690</td>
<td>720</td>
<td></td>
<td>23,550</td>
<td></td>
</tr>
</tbody>
</table>
### 5.3 Discussion

Of the three scenarios, single target molecular surveying is the most cost-effective method long term. It is cheap, and allows for the largest number of waterbodies to be sampled each year. This high number of surveyed waterbodies each year means that a great depth of detection can be achieved, allowing for a higher likelihood that target species will be detected early. One of the limitations associated with single species detection is that if there are other feral species in the water body, they will not be detected. This approach will have the greatest benefits when targeting a possible invasion front, or if trying to detect a single species either endemic or feral.

Multiplexing eDNA allows for a complete survey of the lake, and has the potential to detect multiple species, so while this method is moderately more expensive than single target sequencing, it does enable a more complete snapshot of the diversity of the waterbody. With the multiplexing method, multiple invasive species can be detected, along with any native and endemic species. This means that conservation policies and management plans can be put into place for the entire water body, not just for dealing with the feral species.

When processing eDNA weekly alongside field collection, 560 waterbodies a year can be surveyed when multiplexing, and 800 waterbodies can be surveyed when targeting a single species (Figure 5.3). Along with a greater coverage than traditional methods, multiplex surveying is ~AUD31,000 a year cheaper than traditional Fyke net surveying, and single target surveying is ~AUD24,000 more expensive than traditional Fyke net surveying but double the number of waterbodies can be surveyed over the course of a 40-week year. These costs include the lab processing costs that may even work out to be cheaper if the next-generation sequencing can be done in-house. It is clear to see that molecular surveying is the most cost-effective and efficient method compared with traditional netting (Figure 5.3).

A comparison of the number of waterbodies that can be surveyed using each method on a weekly (38 hour) and yearly (1520-hour) basis including lab time, indicates that molecular single target surveying is the most efficient of the methods (Table 5.4). Even though it costs more each year if the maximum number of waterbodies is surveyed, the benefits mean that invasion fronts are likely to be detected earlier than using traditional methods. Molecular multiplex surveying is not as efficient as single target surveying, but there are benefits associated with this method.
For example, this method allows for a greater understanding of how feral species interact with endemic species over both short term and long term monitoring. The increased number of waterbodies sampled using this method is still less expensive than using traditional methods, so it is both more cost effective and efficient than traditional Fyke net surveying. When looking at this method only from a sampling point of view, the efficiency of this method is clearly shown (Figure 5.2). Without including lab hours, 900 waterbodies can be surveyed each year. It is unlikely that the team collecting the water is the same team processing the samples therefore it is feasible that this many waterbodies could be surveyed each year. Paying a separate lab team to multiplex the samples would mean that 900 waterbodies would cost AUD294,300. This is AUD68,000 more expensive than the yearly netting costs, but there is almost a three-fold increase in the waterbodies that could be surveyed. Alternatively, money and time can be saved using molecular methods if the number of waterbodies surveyed is reduced.
6. Conclusion

Invasive species invasions are understood to be a "wicked" problem: complex with no total solutions, engendering value conflicts, and requiring multiple actions now to prevent damage that might not manifest until far in the future (Balint et al., 2006). Intervention is needed at multiple scales and on various issues, firstly to stop overseas invaders from crossing borders, secondly to prevent those that already have populations from spreading, thirdly to eliminate any newly detected invaders, where possible, and finally to control any new species that threaten biodiversity (Figure 6.1).

![Figure 6.1 Invasion processes and management options (Invasive Species Council, 2010)](image)

Environmental DNA surveying is developing, and the results of this study have shown that it has the potential to be used with confidence in aquatic biodiversity surveying. Three outcomes have emerged: molecular multiplex surveying can be as accurate as traditional sampling methods, is a time efficient method of sampling, and is cost effective.

Biodiversity surveying, along with invasive species monitoring, is a difficult undertaking, and it is important to be as accurate as possible. Detection of finfish species using the molecular approach were comparable with Fyke results (despite the results of the Fyke surveying being outdated). This accuracy, along with the ability to detect low-density organisms suggests that the molecular method has the potential to be used with confidence. As long as the molecular environmental DNA surveying has a robust design, it is unlikely that any species present will not be detected, and an awareness of habitat and fish ecology, along with stringent field and lab protocols will reduce the possibility that false positives will be recorded.
It is simpler to collect water from various locations around a waterbody than it is to set nets and return to score and identify fish. The development of methods such as next generation sequencing of these water samples is also becoming accessible meaning more waterbodies can be surveyed and therefore anomalies i.e. invasion fronts, are likely to be detected earlier. Early detection along with detection of low-density (rare) species means that management plans can be put in place saving both money and the environment in the long term.

Third, using the model outlined in this study eDNA surveying, both multiplexed and single target, has been shown to be cost-effective compared to the more traditional method of Fyke net surveying. Single target molecular surveying breaks even after 95 waterbodies, while multiplex surveying breaks even after surveying 145 waterbodies. Thus, even taking into consideration the cost of setting up labs for molecular analysis, within the first year eDNA surveying is cheaper than the traditional methods.

The cost-benefit analysis demonstrates that molecular surveying is the way of the future and it may contribute to solving the wicked problem of invasive species. The molecular approach has a relatively low impact on the environment (Valentini et al., 2016) and this study has shown that there is a strong potential for this approach to be a more effective and efficient method of monitoring invasion fronts and species present in low numbers i.e. rare or critically endangered species than traditional Fyke net surveying.

There are other benefits to using eDNA for biodiversity monitoring including the need for only one molecular biologist to identify all species present, providing sequences are in a database such as NCBI or FISH-BOL, along with the reduced chance of misidentification of a species. With further optimisation, this method has potential to complement current trapping surveys along with providing a low-cost, logistically simple method of obtaining basic genetic data such as species presence.

This method has shown to be highly sensitive, with the detection of potentially a single organism in a small waterbody. Early detection of such small numbers of organisms could mean that government bodies such as the Department of Fisheries could save money in the long term; it is cheaper to eradicate a few invaders in a water body than to try to eradicate thousands (Jerde et al., 2013; Mantyka-Pringle et al., 2016).
Attempting to put a dollar value on native ecosystems is extraordinarily complex, and when making up-to-date decisions about land-use and land management it is vital that there is absolute cost accounting of economic values of a flourishing, native ecosystem (ANZECC & Committee, 2001). Advantages and disadvantages along with trade-offs about land-use and land management must be as knowledgeable as possible, but the monetary value that is placed on biodiversity conservation can only be calculated after a comprehensive identification of both the environmental and social value of ecosystem services, along with any commercial activities that may be dependent on that ecosystem (ANZECC & Committee, 2001).

When ecosystem processes are examined and identification of ecosystem processes are concluded, the environmental value of biodiversity can be ascertained (ANZECC & Committee, 2001; Atkins et al., 2015). For example, in swamplands, vegetation captures water-carried sediment and soil organisms break down a range of nutrients and pollutants washed into the area (ANZECC & Committee, 2001; Camp & Heath-Camp, 2015). These processes provide the ecosystem a service of purifying water. These areas also behave as spawning and nursery grounds for some fish and provide a refuge for animals during dry weather patterns (ANZECC & Committee, 2001). Removal of deep-rooted vegetation in many areas may lead to a rise in the water table and in turn increase salinity in waterbodies (ANZECC & Committee, 2001). This allows invasive species such as *Cy. carpio*, which can endure low oxygen levels, pollutants, and turbid water bodies better than many native fish to dominate in degraded habitats (Harris, 1996; Lougheed et al., 1998). Variations in water flow, deteriorating water quality and fluctuations to freshwater habitats over the past decades have had a negative effect on many native fish while favouring carp (Nelson, 2015; Strayer, 2010). The ability to detect carp invasion fronts early has potential to save both the ecosystem they are found in and native species from being forced from the system.

There have been a few non-native species introduced to help boost the economy, whether sport, aquaculture or as a food source. This means that in the short term these invasions may be seen as beneficial (McNeely, 2001). Other introduced species such as plants can produce food, timber and energy while introduced insects and toads can provide biological control (McNeely, 2001). A decision cannot automatically be determined once the costs and benefits have been identified because value judgements and distributional questions are commonly also involved.
Sometime the costs involved can be so high that the action is rendered politically unacceptable, even if the benefits are likely to be even greater. The benefits of the control of the invasive species may not be seen for many years meaning that the public may not be supportive of control measures, especially as the costs for control may need to be covered quickly and the money may be raised by increase in taxes. However, on the other hand, not identifying invasion fronts can cause their own problems.

Missing invasion fronts can have serious ongoing economic, environmental, and biological cost associated with it. The estimated economic impact of carp is around AUD4 million a year, a figure made up of ~AUD2 million a year spent by the public sector on carp control and the remaining money on researching how to contain or eliminate the species from waterways around Australia (Dawson, 2005). Other economic impacts have yet to be quantified, though carp are known to have impacts on commercial fishing, water quality, tourism and lead to a decline in native fish species, they have also been known to impact agriculture through damage to irrigation channels (Dawson, 2005). In 2002 the carp industry (the fish is used in fertilisers and fish meal) had a total gross value of ~AUD1.7 million (McLeod, 2004). The effects of carp in the Gippsland Lakes in Victoria, a rough estimate of the costs to the community over 5 years was AUD175 million (Dawson, 2005). These costs included the loss to native commercial fishing, along with losses to recreational fishing, tourism and commerce (Bomford & Hart, 2002). The discrepancy between the two figures is likely due to the first estimate only including public sector costs and research and the second included other factors such as tourism and fishing, both recreational and commercial.

Along with Cy. carpio, other introduced freshwater species that have been mentioned in previous chapters that have become invasive and are having a negative impact on native fish and other aquatic activities include G. holbrooki, P. fluviatilis, and Oreochromis spp., are suspected to have a potentially negative economic impact. This impact could be greatly reduced if the species are detected early. Early detection means that stocks of native, endemic fish available for recreational fishing, and general irrigation and reduced water quality impacts would not be adversely affected by the non-native species. So far, there have been no estimates of the economic impacts of species other than carp.
Advances in sample preparation, and sequencing technology mean that an environmental sample can supply eDNA from many species and therefore information on species detection, biomonitoring, presence/absence studies, and population genetics (Bohmann et al., 2014). The use of eDNA as a tool of analysis is accelerating the rate of new species discovery (Bohmann et al., 2014; Galimberti et al., 2015; Stutz, 2009) and there is no need for a priori information about probable species found in a specific environment. Environmental DNA techniques have allowed researchers working with non-native species, population, and ecosystem processes along with biodiversity and functional diversity to benefit greatly, there is less need to sample target organisms as environmental sampling means less impact on rare and endangered species.

Environmental DNA allows researchers to detect previously undetected invasive species along with detection of species that may have been unrecorded due to difficulty in locating species, i.e. hard to reach ecosystems, along with species that have had the ability to evade conventional sampling methods such as netting (Bohmann et al., 2014; Cooke et al., 2013; Cote et al., 2010; Mehta et al., 2007). So far, projects such as this and other eDNA studies have looked at species identification along with detection of pathogenic, endangered, invasive, or genetically modified organisms, but there are many questions that have arisen from these projects. Some of these include questioning the accuracy of results, along with how unbiased these results are, especially as there are questions over how detailed the eDNA record is.

Not every species has been barcoded so are new species truly new species, sequencing error, or simply artefacts due to the process? It is important to remember that DNA degrades after being exposed to environmental variables, such as oxygen, UV, DNases and even water (Barnes et al., 2014; Eichmiller et al., 2016; Lindahl, 1993) which is why rigorous standards and controls are required to ensure that results collected are not misleading (Bohmann et al., 2014). Environmental DNA analysis allows for a more holistic, ecosystem based approach (Clarke & Jupiter, 2010) as it means that trophic, energetic, and terrestrial-aquatic interactions can be detected and tracked rather than just single species studies (Bohmann et al., 2014).
eDNA is a powerful tool that has many possible applications from functional traits and DNA metabarcoding studies (Barberán et al., 2012) to a recent study where eDNA was used for community analysis in an ecotoxicology setting (Chariton et al., 2014). Complementary multidisciplinary approaches, i.e. eDNA and aDNA (ancient DNA) combined with morphological analyses of both micro and macrofossils show promise for clarifying the impacts of climate change on species and communities over time (Anderson-Carpenter et al., 2011; Jorgensen et al., 2012; Sarkissian et al., 2014; Willerslev et al., 2014).

One of the most important things to consider with the use of eDNA is the relative ease with which eDNA samples can be collected. This has opened a new way of analysing community diversity and dynamics through time (Bohmann et al., 2014). Instead of looking at a community that is reliant on observation of individual organisms captured using traditional methods, researchers can now sample the system as often as permitted by terrain, weather and need. The ability to simply and quickly sample locations may allow researchers to identify niche based and stochastic processes that shape both species distribution and abundance, along with roles played by the species in the system (Haegeman & Loreau, 2011).

Also important to consider is the sensitivity required when targeting species that are likely present in low numbers. For low copy numbers a larger number of PCR cycles are needed to ensure that the target DNA is amplified. The problem with this increased sensitivity is that with universal primers there is a greater likelihood of contamination. This means that a balance must be found with this technique. A larger number of PCR cycles is also needed if the species has a high C\textsubscript{T} value in single source laboratory testing. Sensitivity is one of the key advantages of this method; the goal is to be able to detect those animals that are present in low numbers.

Surveying organisms that are present in large numbers is relatively easy; all that needs to be accounted for is the organism size and habitat preference, so that the correct surveying gear can be selected. To overcome problems of sensitivity it may be better to use primers that identify single species, and not universal primers. When multiplexing, it is possible to use multiple primer sets as up to 96 samples can be run each time, this means that DNA from eight different waterbody samples could be sequenced by up to twelve different primer sets, allowing for up to twelve species to be detected from each sample.
This project, and others like it, have benefits for biosecurity, whether aquatic, terrestrial, or airborne as environmental DNA surveying has shown that it is useful when targeting organisms of interest, such as invasive or pest species. This coupled with the potential cost effectiveness of the method; support the need for further studies into biodiversity analysis using environmental DNA.
7. References


Ashburner, L. (1976). Australian fish face threat from foreign diseases [with specific reference to salmonids]. *Australian Fisheries (Australia).*


Balasingham, K. D. (2016). Environmental DNA Detection as an Effective Tool For Delimiting Spatial Distribution of Rare and Invasive Species, and Assessing Community in Flowing, Freshwater Ecosystems.


Boothroyd, M., Mandrak, N. E., Fox, M., & Wilson, C. C. (2016). Environmental DNA (eDNA) detection and habitat occupancy of threatened spotted gar (*Lepisosteus oculatus*). *Aquatic Conservation: Marine and Freshwater Ecosystems*.


Fletcher, W., & Santoro, K. (2014). Status reports of the fisheries and aquatic resources of Western Australia 2013/14: the state of the fisheries. WA Department of Fisheries, Perth.


91


### 8. Appendices

**Appendix 1** Fish species common names, scientific names, classification, and abbreviated name used in this thesis.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Scientific name</th>
<th>Common name</th>
<th>Abbreviated name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyprinidae/Cyprinidae</td>
<td>Cyprinus carpio (Linnaeus, 1758) / Carassius auratus (Linnaeus, 1758)</td>
<td>Common Carp/Goldfish</td>
<td>Cy. Carpio/Ca. auratus</td>
</tr>
<tr>
<td>Poecilidae</td>
<td>Gambusia holbrooki (Girard, 1859)</td>
<td>Eastern Mosquitofish</td>
<td>G. holbrooki</td>
</tr>
<tr>
<td>Terapontidae</td>
<td>Leiopotherapon unicolor (Günther, 1859)</td>
<td>Spangled Perch</td>
<td>L. unicolor</td>
</tr>
<tr>
<td>Percichthyidae</td>
<td>Maccullochella peelii (T. L. Mitchell, 1838)</td>
<td>Murray Cod</td>
<td>M. peelii</td>
</tr>
<tr>
<td>Percichthyidae</td>
<td>Nannoperca vittata (Castelnau, 1873) Synonyms Edelia vittata (Castelnau, 1873)</td>
<td>Western Pygmy Perch</td>
<td>N. vittata</td>
</tr>
<tr>
<td>Percidae</td>
<td>Perca fluviatilis (Linnaeus, 1758)</td>
<td>Redfin Perch</td>
<td>Pe. fluviatilis</td>
</tr>
<tr>
<td>Poecilidae</td>
<td>Poecilia sphenops (Valenciennes, 1846)</td>
<td>Molly</td>
<td>Po. sphenops</td>
</tr>
<tr>
<td>Gobiidae</td>
<td>Pseudogobius olorum (Sauvage, 1880)</td>
<td>Blue-spot Goby/Swan River Goby</td>
<td>Ps. olorum</td>
</tr>
</tbody>
</table>
## Appendix 2 Accession numbers for target COI and 16S sequences

<table>
<thead>
<tr>
<th>Common Name</th>
<th>Scientific name</th>
<th>COI Accession Number</th>
<th>16S Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balston's Pygmy Perch</td>
<td><em>Nannatherina balstoni</em></td>
<td>KJ669542.1</td>
<td>AY254564.1</td>
</tr>
<tr>
<td>Black-striped Minnow</td>
<td><em>Galaxiella nigrostriata</em></td>
<td>KJ669462.1</td>
<td>AF112324.1</td>
</tr>
<tr>
<td>Common Jollytail</td>
<td><em>Galaxias maculatus</em></td>
<td>KJ669433.1</td>
<td>AF007034.1</td>
</tr>
<tr>
<td>Freshwater Cobbler</td>
<td><em>Tandanus bostocki</em></td>
<td>KJ669639.1</td>
<td>Tandanus sp. EU307875.1</td>
</tr>
<tr>
<td>Little Pygmy Perch</td>
<td><em>Nannonperca sp.</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mud Minnow</td>
<td><em>Galaxiella munda</em></td>
<td>KJ669460.1</td>
<td>AF112329.1</td>
</tr>
<tr>
<td>Nightfish</td>
<td><em>Bostockia porosa</em></td>
<td>KJ669398.1</td>
<td>DQ532844.1</td>
</tr>
<tr>
<td>Pouched Lamprey</td>
<td><em>Geotria australis</em></td>
<td>KT185629.1 Complete genome</td>
<td></td>
</tr>
<tr>
<td>Salamanderfish</td>
<td><em>Lepidogalaxias salamandroides</em></td>
<td>KJ669501.1</td>
<td>AF112340.1</td>
</tr>
<tr>
<td>Trout Minnow</td>
<td><em>Galaxias truttaceus</em></td>
<td>KJ669457.1</td>
<td>AF022100.1</td>
</tr>
<tr>
<td>Western Minnow</td>
<td><em>Galaxias occidentalis</em></td>
<td>KJ669437.1</td>
<td>JN232552.1</td>
</tr>
<tr>
<td>Western Pygmy Perch</td>
<td><em>Edelia vittata</em></td>
<td>KJ669556.1</td>
<td>AY254562.1</td>
</tr>
<tr>
<td>Brown Trout</td>
<td><em>Salmo trutta</em></td>
<td>KM287119.1</td>
<td>KC984250.1</td>
</tr>
<tr>
<td>Common Carp</td>
<td><em>Cyprinus carpio</em></td>
<td>JX983284.1</td>
<td>DQ868866.1</td>
</tr>
<tr>
<td>Convict Cichlid</td>
<td><em>Amatitlania nigrofasciata</em></td>
<td>JN024725.1</td>
<td>GU737208.1</td>
</tr>
<tr>
<td>Freshwater Catfish; Eeltailed Catfish</td>
<td><em>Tandanus tandanus</em></td>
<td>KJ669643.1</td>
<td></td>
</tr>
<tr>
<td>Common Name</td>
<td>Scientific name</td>
<td>COI Accession Number</td>
<td>16S Accession Number</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------------------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Golden Perch</td>
<td><em>Macquaria ambigua</em></td>
<td>KJ669512.1</td>
<td>AY254553.1</td>
</tr>
<tr>
<td>Goldfish</td>
<td><em>Carassius auratus</em></td>
<td>KP112186.1</td>
<td>KC984245.1</td>
</tr>
<tr>
<td>Guppy</td>
<td><em>Poecilia reticulata</em></td>
<td>JX968696.1</td>
<td>U80051.1</td>
</tr>
<tr>
<td>Mosquitofish</td>
<td><em>Gambusia holbrooki</em></td>
<td>HQ937054.1</td>
<td>U80050.1</td>
</tr>
<tr>
<td>Murray Cod</td>
<td><em>Maccullochella peeli</em></td>
<td>DQ107941.1</td>
<td>AY254559.1</td>
</tr>
<tr>
<td>Pearl Cichlid</td>
<td><em>Geophagus brasiliensis</em></td>
<td>KM897636.1</td>
<td>GU737148.1</td>
</tr>
<tr>
<td>Rainbow Trout</td>
<td><em>Oncorhynchus mykiss</em></td>
<td>KM373668.1</td>
<td>HQ592250.1</td>
</tr>
<tr>
<td>Redfin Perch</td>
<td><em>Perca fluviatilis</em></td>
<td>KM373677.1</td>
<td>KR476954.1</td>
</tr>
<tr>
<td>Rosy Barb</td>
<td><em>Puntius conchonius</em></td>
<td>JN965201.1</td>
<td>JX416153.1</td>
</tr>
<tr>
<td>Silver Perch</td>
<td><em>Bidyanus bidyanus</em></td>
<td>KF999850.1</td>
<td>NC_024854.1 (Complete genome)</td>
</tr>
<tr>
<td>Southern platyfish</td>
<td><em>Xiphophorus maculatus</em></td>
<td>JQ667593.1</td>
<td>EF017600.1</td>
</tr>
<tr>
<td>Spangled Perch</td>
<td><em>Leiopotherapon unicolor</em></td>
<td>KJ669500.1</td>
<td>AY935337.1</td>
</tr>
<tr>
<td>Speckled Mosquitofish; One-spot Livebearer</td>
<td><em>Phalloceros caudimaculatus</em></td>
<td>KJ669580.1</td>
<td>U80053.1</td>
</tr>
<tr>
<td>Swordtail</td>
<td><em>Xiphophorus helleri</em></td>
<td>KJ669651.1</td>
<td>U80047.1</td>
</tr>
<tr>
<td>Tilapia</td>
<td><em>Oreochromis mossambicus</em></td>
<td>AY597335.1</td>
<td>DQ426661.1</td>
</tr>
<tr>
<td>Black Bream</td>
<td><em>Acanthopagrus butcheri</em></td>
<td>KJ767814.1</td>
<td></td>
</tr>
<tr>
<td>Swan River/ Blue-Spot Goby</td>
<td><em>Pseudogobius olorum</em></td>
<td>Supplied by Mr. Jason Ledger</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3 DNA of common species expected to be found in environmental samples

<table>
<thead>
<tr>
<th>Common Name</th>
<th>Scientific name</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dog</td>
<td><em>Canis lupis familiaris</em></td>
<td>KC985188.1</td>
</tr>
<tr>
<td>Human</td>
<td><em>Homo sapiens</em></td>
<td>KC750830.1</td>
</tr>
<tr>
<td>Pacific Black Duck</td>
<td><em>Anas superciliosa</em></td>
<td>JN801369.1</td>
</tr>
<tr>
<td>Western Swamp Turtle</td>
<td><em>Psuedemydura umbrina</em></td>
<td>HQ329635.1</td>
</tr>
</tbody>
</table>

Appendix 4 Assumptions made for the analysis of the three scenarios.

Assumptions

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fyke nets</td>
<td>Based on 10, the number cited as practical for a week’s surveying in Metro Perth</td>
</tr>
<tr>
<td>Average number of km travelled per week</td>
<td>Based on an assumption that each water body is 30 km apart</td>
</tr>
<tr>
<td>Fuel (c/L)</td>
<td>National 12 month average</td>
</tr>
<tr>
<td>Nissan Qashqai</td>
<td>Through Hertz unlimited free Km included</td>
</tr>
<tr>
<td>Number of L of fuel bought each week</td>
<td>6.9L/100km economy</td>
</tr>
<tr>
<td>Fuel costs each week</td>
<td>Fuel costs calculated each week based on the above assumptions</td>
</tr>
<tr>
<td>Next generation sequencing</td>
<td>Based on 96 samples per sequencing run</td>
</tr>
<tr>
<td>One person in the lab</td>
<td>Salary for lab staff (AUD1330)</td>
</tr>
<tr>
<td>Hourly rate</td>
<td>Salary divided by 38 working hours a week</td>
</tr>
<tr>
<td>2 people in the field</td>
<td>Salary for field staff (AUD2660)</td>
</tr>
<tr>
<td>Hourly rate</td>
<td>Salary divided by 38 working hours a week</td>
</tr>
<tr>
<td>Travel to waterbody</td>
<td>Based on 45 min to drive 30 km</td>
</tr>
<tr>
<td>Travel between lakes</td>
<td>Based on 45 min to drive 30 km</td>
</tr>
<tr>
<td>Travel from Waterbody</td>
<td>Based on 45 min to drive 30 km</td>
</tr>
<tr>
<td>Molecular surveying</td>
<td>Collection of water samples at each waterbody (45 min)</td>
</tr>
<tr>
<td>Set nets</td>
<td>Time needed to set the Fyke nets at each location (1 hour)</td>
</tr>
<tr>
<td>Clear nets</td>
<td>Time needed to score the fish caught in each net at each waterbody (1.5 hours)</td>
</tr>
<tr>
<td>Net Maintenance</td>
<td>This is approximately 1 day per year for each net</td>
</tr>
<tr>
<td>Filtering</td>
<td>Based on filtering 5 samples simultaneously (30 min)</td>
</tr>
<tr>
<td>Extraction</td>
<td>Lab time needed to extract the DNA from the filters (30 min)</td>
</tr>
<tr>
<td>qPCR</td>
<td>Prep for qPCR, includes loading samples and making master mix (1 hour)</td>
</tr>
<tr>
<td>Next Gen Prep</td>
<td>Prep, includes pooling samples, cleaning samples and loading samples on to the NGS machine (2 hours)</td>
</tr>
<tr>
<td>Data processing</td>
<td>Deconvolution of samples and bioinformatics (30 min/sample)</td>
</tr>
</tbody>
</table>