School of Molecular and Life Sciences

Exploring the uses and value of invertebrate DNA metabarcoding

Kristen Marie Rose Fernandes

0000-0003-3388-5354

This thesis is presented to obtain the double-degree of
Doctor of Philosophy
at
Curtin University
and at the
University of Copenhagen

June 2022

Name of department:	School of Molecular and Life Science (Curtin University), School of Science (University of Copenhagen)
Author:	Kristen Marie Rose Fernandes
Title:	Exploring the uses and value of invertebrate DNA metabarcoding
Topic description:	DNA metabarcoding is used to examine how invertebrates can detect changes in the environment.
Supervisor:	Paul Nevill and Kristine Bohmann
Co-supervisor:	Michael Bunce, Benjamin Saunders, and Mark Gibberd
Submitted on:	30 June 2022

DECLARATION

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signed: Kristen Fernandes

Date: 30 June 2022

This page has been left intentionally blank

ENGLISH SUMMARY

Arthropods play important roles in ecosystem function and are widely used as indicator species in terrestrial biomonitoring surveys. Traditionally, arthropod surveys have relied on morphological taxonomic identification, which is typically labour intensive and requires expertise in identification across multiple taxonomic groups. This makes undertaking in-depth arthropod surveys costly. In recent years, DNA metabarcoding has emerged as a complementary tool that can help combat some of the disadvantages of morphological surveys due to its ability to assess biodiversity from hundreds of samples in parallel rapidly. Another advantage of DNA metabarcoding is that it allows for the detection of organisms that the arthropods have in one way, or another interacted with, for instance, plants or vertebrates. This thesis explores how DNA metabarcoding of invertebrates can monitor ecosystem responses to change by looking at invertebrates as a proxy for broader ecosystem functioning and health. I begin with a review of the various uses that metabarcoding has for monitoring terrestrial ecosystems **(Chapter 2).** I then follow with four experimental chapters covering some of the main ways that metabarcoding can be used to monitor environments with the common themes of a) the scale of the impact, b) measuring species interactions, and c) how this information can be used for conservation and land management. I look at the scale at which invertebrates can be used to detect changes in the ecosystem through the lens of overall arthropod communities at mine sites (**Chapter 3**). I then delve into how the interactions between plants and invertebrates change in response to urbanisation **(Chapter 4)**. My final chapters look at the interactions of arthropods with vertebrates, and I aim to uncover how we can use invertebrate derived DNA (iDNA) from carrion flies in Australian ecosystems to track vertebrate populations **(Chapters 5 and 6).** The results from this body of work add depth and clarity to the tremendous potential of invertebrate metabarcoding to monitor terrestrial ecosystems and therefore opens the door to many new and exciting avenues for future research that will eventually make invertebrate metabarcoding more widespread in terrestrial ecosystem monitoring.

DANSK RESUMÉ

ler en vigtige rolle i økosystemets funktioner og er meget udbredt som indikatorarter i jordbaserede biomonitoreringsundersøgelser. Traditionelt har leddyrundersøgelser været afhængige af morfologisk taksonomisk identifikation, som typisk er arbejdskrævende, samt kræver ekspertise til identifikation af arter på tværs af flere taksonomiske grupper, hvilket gør det dyrt at foretage dybdegående undersøgelser af leddyr. I de senere år er DNA-metabarcoding dukket op som et komplementært værktøj, der kan hjælpe med at bekæmpe nogle af ulemperne ved morfologiske undersøgelser, grundet dets evne til hurtigt at vurdere biodiversitet ud fra hundredvis af prøver parallelt. En anden fordel ved DNA-metabarcoding er, at det giver mulighed for påvisning af organismer der på den ene eller anden måde har interageret med leddyr, for eksempel planter eller hvirveldyr. Denne afhandling undersøger hvordan DNA-metabarcoding a hvirvelløse dyr kan bruges til monitorering af økosystemers respons til ændringer i miljøet, ved at se på hvirvelløse dyr som en proxy for bredere økosystemfunktion og sundhed. Jeg begynder med en gennemgang af de forskellige anvendelser metabarcoding har til overvågning af terrestriske økosystemer (**kapitel 2**). Jeg følger derefter op med fire eksperimentelle kapitler, der dækker nogle af de vigtigste måder metabarcoding kan bruges til at overvåge miljøer med de fælles temaer som a) størrelse af miljøpåvirkningen, b) mål for artsinteraktioner og c) hvordan denne information kan blive brugt til natur forvaltning og land management. Jeg undersøger under hvilket omfang hvirvelløse dyr kan bruges til at detektere ændringer i økosystemet, ved brug a leddyr communities omkring miner **(kapitel 3).** Jeg dykker derefter ned i, hvordan interaktionerne mellem planter og hvirvelløse dyr ændrer sig som reaktion på urbanisering **(kapitel 4).** Mine sidste kapitler ser på leddyrs interaktioner med hvirveldyr, hvor jeg hvordan vi kan bruge DNA fra hvirvelløse dyr (iDNA), specifikt ådselfluer i australske økosystemer, til at spore hvirveldyrpopulationer **(kapitel 5 og 6).** Resultaterne fra dette arbejde demonstrere det enorme potentiale metabarcoding af hvirvelløse dyr har til at overvåge terrestriske økosystemer og åbner derfor døren til mange nye og spændende muligheder for fremtidig forskning, der i sidste ende vil gøre metabarcoding af hvirvelløse dyr mere udbredt i terrestrisk økosystemovervågning.

ABSTRACT

Arthropods play various roles in ecosystem function and are widely used as indicator species in terrestrial biomonitoring surveys. Traditionally, arthropod surveys have relied on morphological taxonomic identification, which is typically labour intensive and requires expertise in identification across multiple taxonomic groups. This makes undertaking in-depth arthropod surveys costly. In recent years, DNA metabarcoding has emerged as a complementary tool that can help combat some of the disadvantages of morphological surveys due to its ability to assess biodiversity from hundreds of samples in parallel rapidly. Another advantage of DNA metabarcoding is that it allows for the detection of organisms that the arthropods have in one way, or another interacted with, for instance, plants or vertebrates. This thesis explores how DNA metabarcoding of invertebrates can monitor ecosystem responses to change by looking at invertebrates as a proxy for broader ecosystem functioning and health.

The research described within this thesis encompasses a literature review and four different case studies, with the common theme of using DNA metabarcoding on invertebrates. I sought to answer whether invertebrate DNA metabarcoding can be used to monitor changes in terrestrial ecosystems through the lens of three key areas; the ability and sensitivity to detect change, uncover species interactions, and finally, its utility for conservation and land management.

Chapter 2 presents a comprehensive literature review describing the history and applications of bulk arthropod sample metabarcoding. Here, I review previous literature to explore the use of this technique in monitoring arthropod diversity, habitat quality assessment, the study of plant-pollinator interactions, vertebrate monitoring, and the transmission of vector-borne pathogens. I also discuss current limitations and perspectives on future research areas.

In Chapter 3, I aimed to examine the sensitivity of the DNA metabarcoding method to detect change through a case study at three mine sites in the Midwest region of Western

Australia. Here I aimed to uncover if exploration infrastructure impacted ground-dwelling arthropods and the scale of impacts if they occurred. I found that the communities of arthropods were strongly affected by the type of exploration disturbance, with community composition shifting within 100 metres from more extreme disturbances. This chapter showed that DNA metabarcoding is sensitive to detect change at a very small scale, with implications for ecosystem management and remediation of these impacted landscapes.

In Chapter 4, I examine how invertebrate DNA metabarcoding can be used to look at species interactions through the case study of native bees in the urban Perth Metropolitan Region of Western Australia. In this study, I examined eight species of cavity-nesting bees to determine how their foraging sources changed in urban and natural environments by looking at the contents of nesting tubes. I discovered that DNA metabarcoding could identify patterns of change in foraging behaviour, with specialised bee species identified to access a much broader range of forage sources than previously thought when found in urban environments. Furthermore, for two bee species in this study that were difficult to survey visually this was the first time that forage sources and potential host plants had been recorded. These results demonstrate that DNA metabarcoding can be used to reveal species interactions and uncover knowledge on species ecology that traditional methods had been unable to show.

In Chapters 5 and 6, I combine the three key areas of my thesis to examine how sensitive DNA metabarcoding is to identify species interactions and how uncovering this data can be used in conservation and land management. Here I look at invertebrate derived DNA (iDNA) from carrion flies to monitor vertebrates. First, I aimed to examine the scale and preference of carrion flies by looking at different classes of animals and the distance that these animals can be detected using a case study of Perth Zoo (Chapter 5). I uncovered that iDNA from carrion flies is predominantly mammalian and can be detected at an approximate distance of 1.25 km from a source point. This technique was then applied in a naturalised environment to detect the distribution of several key mammalian species within a 742 km^2 area of fragmented reserves and pastoral land (Chapter 6). Here I found that these iDNA detections were representative of known distributions of the key taxa and was able to detect several threatened mammal species within reserves.

In a broader context, the data and analyses presented in this thesis emphasise how essential invertebrate biomonitoring is for understanding ecosystem health. Through the case studies, I demonstrate that invertebrate DNA metabarcoding technology can be used beyond simple arthropod biodiversity surveys to provide valuable insight into various areas of ecological interest. This thesis creates new knowledge and tests novel methods, opening new ways to interpret and design DNA metabarcoding studies - ultimately leading to better detection of biodiversity changes in the environment. The results from this body of work add depth and clarity to the tremendous potential of metabarcoding as a tool to monitor terrestrial ecosystems and therefore opens the door to many new and exciting avenues for future research that will eventually make invertebrate DNA metabarcoding more widespread in terrestrial ecosystem monitoring.

This page has been left intentionally blank

ACKNOWLEDGEMENTS

It takes a village to raise a child, and in the same way finishing a PhD requires the tireless efforts of a team. So, I would like to acknowledge the support from team of people over the last few years that made finishing this project possible.

I am grateful to my supervisory team, Paul Nevill, Ben Saunders, Mark Gibberd, Michael Bunce, and Kristine Bohmann, and to my chairperson, Euan Harvey. Thank you for your patience and assistance throughout this PhD process. It certainly wasn't easy supervising and managing a project through a pandemic, and I appreciate all your hard work and patience throughout these last few years. A special thank you to Mike and Kristine for giving me an opportunity of a lifetime, allowing me to travel the world and opening doors to opportunities I never imagined I would have. Without your encouragement, guidance, and enthusiasm for my project, I don't think I would have ever made it to the end.

My gratitude then naturally follows to the team of academics and co-authors who supported me on this journey. I would like to give special thanks to Kit Prendergast and to Sophia Clark-Ioannou, for letting me use the samples that you worked so hard to collect. Thank you to Bill Bateman for your helpful advice, pep talks, and for reading all those terrible drafts. I would also like to extend a thank you to my academic mentor, Julia Easton, for your support and for teaching me how to believe in myself. Further, a big thank you to Franca Jones, the DGR for all the coffee and for taking time to listen to me.

To the TrEnD Lab team (both past and present): Arne Adam, Katrina West, Matt Heydenreych, Mahsa Mousavi-Derazmahalleh, Miwa Takahashi, Mieke van der Heyde, Nicole White, Josh Kestel, Jackie Jones, Alishum Ali, David Juszkiewicz, Matt Campbell, Jason Alexander, Tina Berry, Frederik Seersholm, Georgia Nester, Sophie Preston, and Matt Power. You have been like family to me. I feel like it has been a rite of passage in the office to see me cry, so thank you for seeing that and still wanting to be my friends. Thank you for listening to my whinging, celebrating my wins, facilitating my caffeine addiction, and for all the help troubleshooting over the past four years.

To the team in Copenhagen (and Greenland), thank you to Christina Lynggaard, Physilia Chua, Luke Holman, Martin Nielsen, and Emilia Rolander Langkjaer for making me feel welcome as part of the team (even though I was located so far away) but also for being there to support my research and give me pep talks when I needed.

A big thank you to the Food Agility students, especially the Curtin group; Cassandra Chong, Sylvia Soltyk, and Patrick Duong for the support and the very relevant memes. It has been so lovely to know that no matter what field you are in, the PhD journey is the same! I'm glad to have found such talented and kind peers.

Thank you to Perth Zoo and Peter Mawson and Caroline Lawrence for their help to get my carrion flies (and for all your enthusiasm for early morning sample collections).

Thank you to the researchers at ECU, Anna Hopkins and the wonderful MEEG lab for taking me in and making me feel so welcome, and the team at DBCA Ecosystem Services branch, particularly Josie Hyde and Katinka Ruthrof. You showed me that I could be a scientist outside of the PhD and gave me the push I needed to finish so I could look forward to the next steps.

Thank you to my family, my parents (Suzette and Derrick) and my sister (Xian). All my gratitude goes to you for your undying encouragement and enthusiasm; although I am sure you have no idea what I spent all these years doing, this thesis would not have been born without your love. Naturally, thanks also follow to my dear friends who kept me going and ensured that I always had plenty of reasons to smile, Michelle Talbot, Declan Stevens, Nikki Regnard, Chris Miller, Stacey Sutijono, Maddie Foo, Sarah English, Jo Newman, Simon Martin, Caroline The, Carina Mancinone, Matt Smith, and Jarryd Reeves. I'm sure you now know just as much about metabarcoding and bugs as I do.

Thank you especially to Andrew Sugden, my partner and best friend. The wins and losses that we have shared through these last few years have been enormous, and I couldn't have picked a better person to come with me along the way. Thank you for your unwavering love, kindness, and optimism. Without you there to cheer me on I wouldn't have had the courage to finish.

Pets are not our whole life, but they make our lives whole. I would like to say a very special thank you to the pets that have been part of this journey, some who sadly didn't see its conclusion, but whose presence brightened the path along the way. To the paw-rents of Maisy, Maca, Amber, Neo, Millie, and Banjo, thank you for facilitating the many, many, many hours of pet therapy I needed to get me through this journey. Special thanks to my three little rats, Penelope, Mango, and Snona who filled even the darkest of days with whiskers and love.

Finally, I would like to recognise financial support for my PhD projects from the Australian Research Council, Food Agility CRC, and the Holsworth Wildlife Research Endowment. I was personally supported by an Australian Government Research Training Program (RTP) Scholarship over the duration of my PhD research.

This page has been left intentionally blank

ACKNOWLEDGEMENT OF COUNTRY

We acknowledge that Curtin University works across hundreds of traditional lands and custodial groups in Australia, and with First Nations people around the globe. We wish to pay our deepest respects to their ancestors and members of their communities, past, present, and to their emerging leaders. Our passion and commitment to work with all Australians and peoples from across the world, including our First Nations peoples are at the core of the work we do, reflective of our institutions' values and commitment to our role as leaders in the Reconciliation space in Australia.

This page has been left intentionally blank

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page has been left intentionally blank

CHAPTER 1

GENERAL INTRODUCTION

1.1 Biomonitoring

Environmental change has profoundly altered biodiversity and the provision of ecosystem services across the globe (Bohan et al., 2017). Since the Industrial Revolution, biomonitoring programs have been used to measure these environmental changes and their effects on the natural world (Gray et al., 2014). Biomonitoring refers to the monitoring of biological indicators of stressors in the environment; this can refer to monitoring at a molecular level of biochemical markers of stress or monitoring large-scale impacts, such as the changes in structure and composition of whole communities of organisms (Derocles et al., 2018). Hereafter I refer to biomonitoring at this community scale. At a community level, biomonitoring refers to recording species diversity and abundances across locations using various ecosystem census techniques and taxonomic identification (Derocles et al., 2018). This census information is then assessed against a baseline level relative to a reference system or idealised target (for example, a target based on an acceptable level of a response variable or a restoration goal) (Gray et al., 2014). Biomonitoring studies can measure the response and recovery of communities from disturbances, protect biodiversity, evaluate compliance, and build on the understanding of the relationship between the abiotic and biotic components of ecosystems (Hershey, Lamberti, Chaloner, & Northington, 2010). Hence, biomonitoring underpins many areas of environmental policy and land management (Niemeijer, 2002). Many different organism groups can be used in biomonitoring; however, invertebrates are most commonly used in both terrestrial and aquatic ecosystems (Gerlach, Samways, & Pryke, 2013).

1.1.1 Biomonitoring using invertebrates

Invertebrates are a group of highly diverse organisms that constitute a large proportion of the species and biomass on Earth (Wagner, 2020). They exhibit a diversity of size, life histories, and habitat specialisations essential to various critical ecosystem services, such as pollination (Obrist & Duelli, 2010). These factors make invertebrates ideal ecosystem indicators (Kim, 1993). Furthermore, the diversity and ubiquity in many environments, ease of sampling, and sensitivity to disturbance have made invertebrates standard measures of ecosystem stress and health (Birk et al., 2012; Gerlach et al., 2013).

Invertebrate biomonitoring programs usually involve the mass collection of invertebrates in traps, followed by morphological identification of the individual specimens within traps, calculating biodiversity and abundance (Gerlach et al., 2013). However, despite the value of invertebrates as indicators in the environment, their usage has been limited due to the high taxonomic diversity of invertebrate groups, which includes numerous cryptic and undescribed species with poorly known ecology (McGeoch, 1998). Therefore, to carry out biomonitoring using invertebrates requires a deal of expertise in morphological identification across invertebrate groups, which can be resource and time-intensive depending on the availability of identification keys, local experts, or the volume of individuals within a trap (Gerlach et al., 2013; McGeoch, 1998; Yu et al., 2012).

To navigate this limitation, some studies will identify taxa with coarser taxonomic resolution (de Oliveira, Ortega, Ribas, Lopes, & Bini, 2020; Souza et al., 2016) or use well studied focal groups to act as a proxy (Pik, Oliver, & Beattie, 1999) instead of studying whole arthropod composition within a system. These species groups are used as indicators of system responses to stress because their ecology and species are well known, they respond quickly to disturbance, and their responses can reflect overall ecosystem responses (Pik et al., 1999). This comes with its own sets of caveats. Some studies have found that coarse taxonomic resolution is unable to identify diverse responses to ecosystem stress (Macher et al., 2016). Additionally, some invertebrate species are so spatially distinct from other groups (for example, ants on the forest floor, compared to flying insects within the canopy), the disturbance impacts to one species group may not transcend into all groups (Silva, Karunarathna, & Karunaratne, 2017). With global biodiversity in decline, especially in arthropods and other invertebrate groups (Wagner, 2020), there is a motivation to document ecosystem biodiversity changes and develop and communicate methods to mitigate the threats to global ecosystems (Wilson & Fox, 2021). Technological advancements will be an important part of increasing the accuracy, speed, and breadth of biodiversity monitoring. One such advancement has been using DNA-based methods to monitor arthropod communities in ecosystems.

1.2 DNA metabarcoding as a tool for biomonitoring

DNA-based strategies for identifying species have been proposed since the 1980s using various techniques (Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012). However, the development of polymer chain reaction (PCR) based amplification of DNA (Mullis & Faloona, 1987; Saiki et al., 1988) and the discovery of universal primers (T. D. Kocher et al., 1989; Taberlet, Gielly, Pautou, & Bouvet, 1991) have given way to a standardised approach for 'DNA barcoding' for species identification. In the last two decades, DNA barcoding has been applied globally to catalogue life on the planet. This impressive effort has facilitated the creation of global taxonomic reference libraries, such as the Barcode of Life database (http://www.barcodeoflife.org/), to document and curate the world's taxonomic biodiversity. However, DNA barcoding was designed for the identification of species from DNA isolated from intact single specimens using Sanger sequencing (Taberlet et al., 2012), with a focus on taxonomic delimitation (Hubert & Hanner, 2015). With the high number of invertebrate specimens collected from ecological studies, DNA barcoding by itself may not be practical for use as it can be very time-consuming or near impossible to select appropriate specimens for such analysis (Taberlet et al., 2012). The development of high-throughput sequencing (HTS) (also known as next-generation sequencing (NGS)) has made it possible to generate millions of sequencing reads in parallel – paving the way for metagenomic applications that could be used to identify genetic material from communities of mixed organisms. One such metagenomic technique is metabarcoding.

The term metabarcoding was first used in 2011 (Pompanon, Coissac, & Taberlet, 2011; Riaz et al., 2011), describing the simultaneous PCR amplification and sequencing of DNA from various species in one sample using 'universal' mini-barcode gene regions (Taberlet et al., 2012). These regions are selected because they are shared amongst a wider group of organisms (universal) but have enough variation within the region to distinguish taxa apart (Riaz et al., 2011). The metabarcoding methodology uses PCR primers tagged with molecular identification tags (MIDS) or indexes. MID tags are generally small nucleotide sequences (between 5-8 bp) that can identify one sample from another. Amplifying each sample using primers with a unique MID tag allows for the pooling of multiple samples (multiplexing) to be sequenced together (Taberlet et al., 2012). Sequences can then be associated back to the sample they originated from, and the taxa within the sample can be identified. This means that biodiversity audits can now be conducted across hundreds of samples simultaneously, significantly decreasing the time it would take to perform the same studies using morphological identification (Ji et al., 2013).

1.3 DNA metabarcoding of invertebrates

The development of metabarcoding techniques has dramatically increased the scope of biodiversity assessments, especially for invertebrates with large numbers of cryptic or undescribed species. The use of metabarcoding to identify arthropods was first attempted by Yu et al. (2012) to characterise the biodiversity of arthropods from bulk samples of canopy sweepings and malaise traps in the tropics. This study showed that it was possible to use metabarcoding to assess biodiversity rapidly and accurately across a range of taxa. Since then, DNA metabarcoding based analysis of invertebrate groups has been applied in a wide variety of ecological applications across the globe, from ecosystem restoration monitoring (Fernandes et al., 2019) to tracking pathogens (Miller, Hopkins, Inward, & Vogler, 2016). Metabarcoding of whole invertebrate communities has been shown to be faster and more cost effective than traditional morphological identification across samples (Ji et al., 2013). Further, because metabarcoding assays can be used to target specific groups of organisms, collecting an invertebrate sample for metabarcoding is not restricted to identifying just the invertebrates within it. Invertebrates are used for biomonitoring because of their importance to ecosystems through their interactions with the many different trophic groups (McGeoch, 1998). Using invertebrate metabarcoding makes it possible to look at the interactions of multiple groups within an environment, from plants to vertebrates (Schnell et al., 2015; van der Heyde et al., 2020). Measuring species interactions and networks in ecosystems has been recommended as a critical next step in improving biomonitoring science (Gray et al., 2014).

1.3.1 Plant-insect interactions

One of the most pivotal interactions in ecological systems is between insects and plants (Potts et al., 2010). These relationships are crucial for many different ecosystem functions, from pollination to disease spread, and hence have direct links to human food production systems and overall ecosystem health. The traditional methodologies associated with measuring and studying plant-insect interactions require observing the plant and insect species and assessing the nature of their interaction (Burkle & Alarcón, 2011). However, it is difficult to conduct these studies at large scales because of the time and skill required to undertake observation studies (Alarcón, 2009). Hence, the development of DNA based techniques in recent years has been explored as a beneficial complement to traditional methods (Evans & Kitson, 2020). Here, metabarcoding can be used on the gut contents of insects for dietary studies (Masonick, Hernandez, & Weirauch, 2019), or metabarcoding can even be applied to flowers to uncover plant-pollinator networks by detecting the invertebrates that have visited (Thomsen & Sigsgaard, 2019). Furthermore, it is possible to study insect behaviour and foraging patterns through metabarcoding insect pollen loads (Lucas et al., 2018; Macgregor et al., 2019; Pornon et al., 2016) or insect brood cells (Gresty et al., 2018). For species with little knowledge on ecology or host-plant preference, the use of metabarcoding can be beneficial to bridge these knowledge gaps and aid in the conservation and management of ecosystem functioning and diversity.

1.3.2 Vertebrate-invertebrate interactions

Many invertebrate groups interact with vertebrates by feeding on their blood, faeces, carcasses, or flesh. Metabarcoding can be used on the gut contents of these invertebrates to detect these food sources and hence monitor vertebrate species in the environment (Calvignac-Spencer, Leendertz, Gilbert, & Schubert, 2013). This is known as invertebrate derived DNA (iDNA) and has been applied to a variety of invertebrate groups such as leeches (Ji et al., 2022; Schnell et al., 2015), carrion flies (Calvignac-Spencer et al., 2013; Rodgers et al., 2017; Schubert & Stockhausen, 2015), mosquitoes (Massey et al., 2021; Reeves, Holderman, Gillett-Kaufman, Kawahara, & Kaufman, 2016), sandflies (A. Kocher et al., 2017; Massey et al., 2021), and ticks (Gariepy, Lindsay, Ogden, & Gregory, 2012). Although, it is also possible to use arthropods sourced from bulk arthropod samples such as malaise traps to identify vertebrates (Lynggaard et al., 2019). Traditional methods for monitoring vertebrates require the direct observation of an animal or its traces, which is usually carried out by trapping animals or conducting observational surveys (Lee, Gan, Clements, & Wilson, 2016). In recent years, non-invasive methods have developed as technology has advanced, such as acoustic or camera trapping. However, these methods require considerable person-hours to process audio or visual recordings, and often the identification of smaller-bodied or non-mammalian species can be challenging (Gogarten et al., 2019). In comparison studies, iDNA has shown to be a valuable tool for detecting smaller-bodied and arboreal mammals and can detect groups such as amphibians and birds that can be difficult to detect using camera traps alone (Gogarten et al., 2019). Hence, iDNA metabarcoding has immense utility in conservation and land management.

1.4 The DNA metabarcoding workflow

The applications of invertebrate DNA metabarcoding for terrestrial monitoring are seemingly endless. However, several key considerations need to be made across the workflow to use the metabarcoding method effectively. The following section will briefly describe the DNA metabarcoding workflow to explain the key terms and methodology used throughout this body of work (simplified in Fig 1-1).

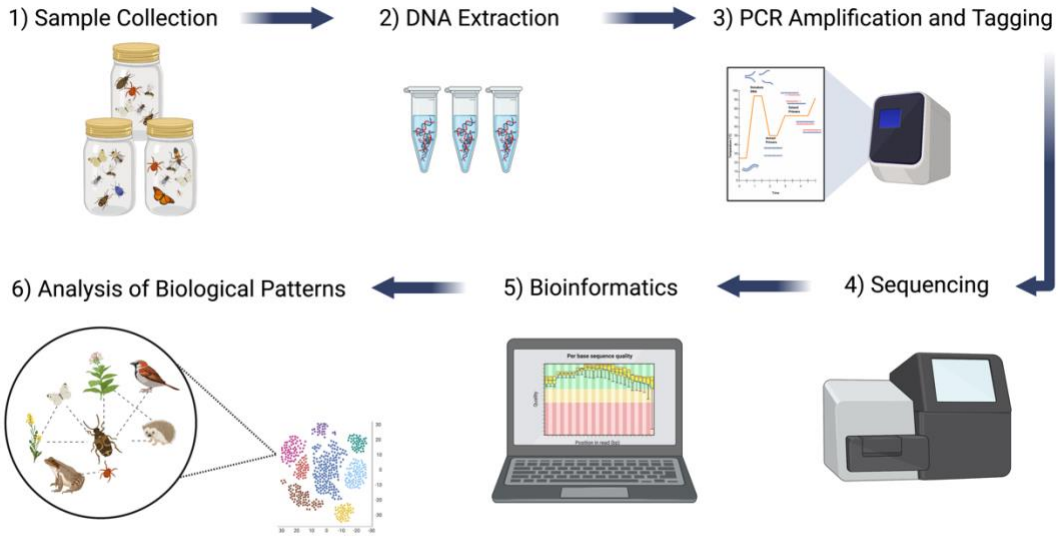


Figure 1-1: Basic Invertebrate DNA metabarcoding workflow. Once samples are collected, DNA is extracted. Next, PCR amplification is used to add MID tags. Sequencing libraries are then prepared and pooled for sequencing. After sequencing, samples are quality-filtered, demultiplexed, and dereplicated using various bioinformatics protocols. The sequences in the samples can then be taxonomically identified and used in biodiversity analysis. Created with BioRender.com.

1.4.1 Sources of DNA for metabarcoding

Invertebrate DNA metabarcoding has a unique versatility in a biomonitoring context in that it is a methodology that can be applied in a variety of ways. There are a few main approaches to metabarcoding that studies have used for invertebrates. The dominant method has been through the collection of bulk arthropod samples from traps which are commonly used for standard biodiversity monitoring (for example, pitfall traps, malaise traps, and leaf litter traps) in both aquatic and terrestrial ecosystems (Beng et al., 2016; Kuntke, de Jonge, Hesselsøe, & Lund Nielsen, 2020; Liu, Clarke, Baker, Jordan, & Burridge, 2019). These are often called 'bulk tissue samples' or 'bulk samples.' It is also possible to extract invertebrate DNA from environmental substrates such as soil (Oliverio, Gan, Wickings, & Fierer, 2018), water (Mychek-Londer, Balasingham, & Heath, 2019), and even air (Roger et al., 2022). This is known as environmental DNA or eDNA. Further, studies have branched out encompassing more than just biodiversity audits of invertebrates, but into the interactions within ecosystems. As discussed previously, this has considerable benefit in dietary studies (Masonick et al., 2019), pollination studies (Pornon et al., 2016), and detecting vertebrates (Schnell et al., 2015).

Processing and extracting DNA from these samples is entirely dependent on the type of sample collected. Typically, with bulk sample invertebrates, it is best practice to size select specimens and then homogenise whole samples together to help overcome bias associated with uneven biomass in a mixed sample (Creedy, Ng, & Vogler, 2019; Elbrecht, Peinert, & Leese, 2017). Recently, new techniques have emerged in non-destructive DNA extractions from arthropods. For example, a digestion buffer can be used that still retains the integrity of the exoskeletons of arthropods (Nielsen, Gilbert, Pape, & Bohmann, 2019). It is also possible to extract DNA for metabarcoding invertebrate communities from filtered preservative ethanol used in insect traps (Erdozain et al., 2019). The benefit of these non-destructive techniques is the retention of diagnostic features for taxonomic identification if reference databases are lacking, building reference databases, or undertaking abundance quantification if required. Although, the success of extracting DNA from these specimens the second time around may be limited (Carew, Coleman, & Hoffmann, 2018).

1.4.2 Amplicon sequencing

Metabarcoding is also sometimes referred to as amplicon sequencing (Taberlet, Bonin, Zinger, & Coissac, 2018), and this is because it works through the sequencing of PCR amplicons. Once DNA has been extracted from a sample, PCR can be carried out to amplify targeted barcoding gene regions. Primer pairs, or assays, are short fragments of synthetic oligonucleotides that anneal and amplify target DNA; this target is an amplicon. There is no universal barcode region with taxonomic resolution across the entire tree of life. Therefore, assays are designed and chosen for a target species or a broader taxonomic group. The choices made in assay selection impact the biodiversity you can detect from a sample (Alberdi, Aizpurua, Gilbert, & Bohmann, 2018), and trade-offs need to be made when considering size, breadth, resolution of the assay, and the availability of reference barcodes to identify the species in a sample. For example, for arthropods, the cytochrome c oxidase subunit I (COI) mitochondrial DNA region is generally the most used in invertebrate metabarcoding studies. There has been considerable effort into designing and testing invertebrate metabarcoding assays from this gene region (Elbrecht et al., 2019). Although the COI region is not perfect (Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014; Porter & Hajibabaei, 2020), there is enough representation across reference sequence databases for major invertebrate taxonomic groups to make it a popular choice, as opposed to other regions such as the mitochondrial 12S ribosomal RNA (rRNA) gene used for vertebrates (Riaz et al., 2011). The taxa targeted by PCR assays, however, are not limited to just one group of organisms, and it is possible to multiplex multiple metabarcoding assays together on one sample to uncover a range of biodiversity simultaneously. Invertebrate DNA metabarcoding has excellent potential for application in broad biodiversity audits and in uncovering ecosystem interactions between different trophic levels.

PCR involves three main stages, denaturing, annealing, and extending. At high temperatures, double-stranded DNA is denatured and separates into two single strands; the temperature is then lowered to allow the primers to anneal to target regions on the single strands, and the temperature is then raised slightly to allow for DNA polymerases to extend the target DNA using synthetic deoxynucleotide triphosphates (dNTPs). These stages are repeated as a cycle (denaturing, annealing, extending), doubling the number of DNA copies every single time. The number of cycles is variable but can range from 20-50 cycles depending on the efficiency of the reaction and the initial starting quantity of target DNA (Taberlet et al., 2018). Using PCR, it is possible to generate thousands to millions of copies of target DNA from a very small initial amount. Therefore, metabarcoding by essence is very sensitive to contamination. Furthermore, if matching the target of the metabarcoding assays, any small amount of DNA will also amplify alongside the 'real' sample. Therefore, it is good practice to include negative controls in PCR and sequence these alongside real samples to assess potential contamination and background contamination levels. PCR is also recommended to be conducted with the sterile practice in mind, with physical separation between the areas where DNA extractions occur and PCR amplification is performed (Goldberg, Strickler, & Fremier, 2018). The final number of amplicon copies can be quantified through visualisation on gel electrophoresis or quantitative PCR (qPCR), which measures the incorporation of a fluorophore during the extension phase to quantify DNA copy numbers in real-time.

The efficiency of PCR is known to be variable depending on several different factors. Whilst, in some cases, biomolecules extracted alongside DNA, such as tannins, can inhibit PCR amplification, other factors such as the quality of the polymerases, the GC content of the primers chosen, or reaction conditions can impact reaction efficiency (Best, Oakes, Heather, Shawe-Taylor, & Chain, 2015). Therefore, amplicon sequencing workflows require quality and inhibition control checking to minimise bias (Murray, Coghlan, & Bunce, 2015). This will usually involve the selection of an optimal dilution of a sample to maximise the efficiency of PCR (Best et al., 2015; Murray et al., 2015). In this case, positive controls (of known quantities of uninhibited DNA) can also be included alongside real samples to assess inhibition levels of the reaction.

The basis of metabarcoding is the addition of MID tags to samples so that sequences can be identified from the sample in which they originated. These tags are added in the PCR reaction, whereby the metabarcoding primers also include a MID tag. The uniquely tagged amplicons can then be pooled and built into sequencing libraries, and the contents of many samples can be sequenced in parallel (Bohmann et al., 2011). The most common platform used today for amplicon sequencing is Illumina technology (Carøe & Bohmann, 2020). Illumina platforms rely on the addition of flanking adapters to amplicons which are complementary to the surface of the sequencing flow cell along with an index that acts as a unique tag for each sample. There are a few main approaches to adding these adapters to samples. 1) a one-step PCR approach, where PCR is conducted using regular gene-specific assays, the adapters and sample indexes are then added through ligation in a PCR free manner; 2) a two-step PCR approach, where samples are first amplified with the metabarcoding primers that carry sequence overhangs, and the second round of PCR incorporates the sequencing adapters and indexes; and 3) a one-step PCR approach with fusion tagged primers, that contains the forward and reverse adapters and indexes already as part of the primers (Taberlet et al., 2018). Amplicons are combined in equimolar quantities in a sequencing library to ensure equal sequencing depth across all samples. The fusion tagging approach allows for individual replicates to be incorporated into ready-to-sequence libraries in one go, this minimises the risks associated with PCR contamination. Further steps to library preparation include fragment analysis and size selection of amplicons to remove primer dimer and purification and quantification of the library for optimal sequencing efficiency (Stat et al., 2017).

1.4.3 Bioinformatics and sequence analysis

The most important part of working with metabarcoding data lies in the processing of the huge amounts of data that comes from next-generation sequencing. Whilst computational and storage requirements for metabarcoding data are high, another crucial aspect is dealing with the noise associated with amplicon datasets (Taberlet et al., 2018). Improvements in sequencing technology have allowed for more detailed descriptions of biological diversity, on both an intra and interspecific level, and for the detection of molecular artifacts from the amplicon sequencing methodologies. Bioinformatic processing, therefore, is an essential part of metabarcoding analyses to filter out these artifacts and reveal true ecological patterns. There are a range of tools available that can be used for quality filtering, including OBITools (Boyer et al., 2016), USEARCH (Edgar and Flyvbjerg, 2015), and R packages such as DADA2 (Callahan, McMurdie, Rosen, & Han, 2015). The steps and parameters that are used in bioinformatic processing vary in accordance with the multiplexing and sequencing strategies used. However, the main steps revolve around quality filtering and curating sequences based on quality scores (termed 'Phred quality scores') associated with the base call reading from the sequencing run. Further steps include merging reads (if using paired-end sequencing); sequence demultiplexing, where the MIDs used to identify samples are used to place each sequence to the sample it originated from; dereplicating sequences, where identical sequences are clustered together to reduce the size of the data; and the removal of chimeras and only partially sequenced metabarcodes. These steps are used to minimise and remove obvious artifacts that could stem from PCR or the sequencing methodology. However, PCR and sequencing errors in metabarcoding and numerous and unique sequences are more numerous than the actual expressions of true taxonomic diversity within any taxonomic group (Schloss, Gevers, & Westcott, 2011).

Whilst singleton unique sequences are often considered spurious and removed from datasets (Clare, Chain, Littlefair, & Cristescu, 2016), constructing groups of similar sequences, known as clustering, is an important part of seeing past the artifactual inflation in metabarcoding datasets to find the true biological patterns. Sequences can then be clustered into operational taxonomic units (OTUs) where a certain similarity threshold between sequences is used to group sequences together (usually between 97-99%) or remain as unique amplicon sequence variants (ASVs). ASVs can also be known as zero-distance OTUs or zOTUs. ASVs or OTUs can then be matched against sequence reference databases to taxonomically identify the sequences clusters. While publicly available databases such as BOLD and GenBank are popular options, custom reference databases of sequences generated from single source specimens can also be used for querying against. Taxonomic assignments are made in accordance with percentage identity matching and query coverage of a sequence to those available in a reference database. The quality of the match to the database can leave taxonomic assignments of sequences ranging from a phylum to species level assignment. For some geographic regions, there is a lack of available reference sequence databases and available biodiversity information (Ashfaq & Hebert, 2016). In these cases, it can be possible to identify patterns in a taxonomy-independent space using ASVs or OTUs instead. The use of these sequence clusters (taxonomically identified or otherwise) provides community composition data that can then be used to make alpha or beta diversity assessments, however, with the caveat that nuanced ecological or functional inferences can be lost in taxonomy-independent space. Nevertheless, the ability to move beyond a taxonomy-dependent space means that metabarcoding analysis can be a valuable tool across multiple scales, and allows for rapid, but meaningful analysis even in the absence of complete databases.

1.5 Research gaps

Although there has been considerable development in metabarcoding technologies, the predominant effort has been on assessing the benefits and limitations of its application in aquatic environments (Thomsen & Willerslev, 2015). Furthermore, this extends to invertebrate metabarcoding, where there has been considerable emphasis on using metabarcoding to monitor aquatic macroinvertebrate communities (Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017). The lack of terrestrial studies has been attributed mainly to the scarcity of knowledge available for terrestrial invertebrate groups. In aquatic systems, macroinvertebrate communities have been used routinely in habitat quality assessment protocols. This is because there is available knowledge on how habitat quality indicators (for example, pH level or nutrient load) impact the abundance and diversity of species in these invertebrate communities (Macher et al., 2016). However, whilst terrestrial systems utilise invertebrates in habitat quality assessment as well, there remain some significant gaps in our knowledge of species diversity and ecology (Paknia, Rajaei, & Koch, 2015). Therefore, developing terrestrial invertebrate metabarcoding studies represents a way to bridge both gaps in knowledge in the applications of terrestrial metabarcoding and in our understanding of terrestrial invertebrate diversity and ecology.

This thesis discusses case studies that aim to bridge three main research gaps. Firstly, understanding the sensitivity of this technique to detect change. It is known that using arthropod metabarcoding is just as sensitive as biodiversity surrogacy measures in determining changes in the environment (Barsoum, Bruce, Forster, Ji, & Yu, 2019). When it comes to environments impacted by subtle impacts, little is known about the sensitivity of metabarcoding to detect these changes at varying spatial scales. Further, this leads to the second research gap; knowledge of terrestrial invertebrate ecology. Understanding the scale and extent of environmental impacts is only possible when related to understanding how these disturbances impact invertebrate ecology. For many species in biodiversity hotspots, there is a significant gap in the knowledge of ecology and taxonomy (Ashfaq & Hebert, 2016; Gálvez-Reyes et al., 2020). This make interpreting the results from any biodiversity audit, including metabarcoding, difficult if there is insufficient data to draw meaningful ecological assessments. Therefore, building knowledge on species ecology will be an important part of developing metabarcoding studies for terrestrial biomonitoring. Finally, the third area requiring knowledge is the calibration of iDNA studies. To draw meaningful conclusions, it is necessary to understand any monitoring tool's full strengths and limitations. Although iDNA studies have shown to be effective in detecting vertebrate biodiversity (Schnell et al., 2015), studies are just beginning to explore the parameters for successful vertebrate detection using this method. This includes studies on the length of time that DNA persists (Owings et al., 2019), the effects of using different invertebrate groups (Massey et al., 2021), or host selection preferences (Hanya et al., 2019). However, this calibration effort is not evenly distributed across all the groups of invertebrates used in iDNA studies, and more research is required for specific groups, such as carrion flies.

1.6 Research question and thesis overview

Understanding the scope and potential usages of metabarcoding methodologies will be essential for developing terrestrial ecosystem monitoring tools in the future. My thesis therefore addresses the overarching question of: 'how can invertebrate DNA metabarcoding be used to monitor the response of terrestrial ecosystems to change?' (Fig 1-2). This central research question will be answered across studies based in Western Australia. This thesis contains one literature review chapter (Chapter 2) and four data chapters (Chapters 3 to 6). Chapters 2 to 6 compromise papers in review for publication or prepared for publication in academic journals and follow the referencing style and formatting of the respective journals that they are prepared for. For continuity between chapters, each chapter will begin with a preface of the chapter within the context of the overall thesis. The final discussion chapter (Chapter 7) will synthesise the main findings of the thesis and make comment on the scope of this field of research as a technique for monitoring terrestrial ecosystems.

In Chapter 2, I discuss the current and future applications of bulk arthropod metabarcoding in terrestrial ecological monitoring. The aim of this chapter was to give a current examination of the field and to look at the technique's capabilities for terrestrial biomonitoring. Here gaps and future directions in the research are identified through the lens of bulk-sample metabarcoding.

In Chapter 3, I investigate the impacts of exploration infrastructure on ground-dwelling arthropod communities using bulk sample metabarcoding at three mine sites in Australia's Midwest region. Here I look at the impacts of three different types of infrastructure and examine the impacts on the communities at each site. This chapter represents a common use of metabarcoding to examine biodiversity of arthropods, showing the sensitivity of the technique to detect fine scale differences in invertebrate communities across broader landscape level disturbances.

In Chapter 4, I investigate the impacts of urbanisation on the foraging habits of solitary cavity nesting bees in the Perth metropolitan region. This chapter represents the use of the ecology of an invertebrate species to reveal subtle impacts of urbanising effects on pollinator behaviour. This chapter further builds on how metabarcoding of invertebrates can be used, looking beyond biodiversity audits of invertebrates and into species interactions.

In Chapter 5 and 6, I use invertebrate metabarcoding to investigate how invertebrate interactions in the environment can be used for conservation and land management. In Chapter 5 I explore iDNA monitoring through carrion flies by looking at three main areas; distance from a source point, biomass, taxonomic group, and how these will impact the detectability of vertebrates from carrion flies. The results from this study are used to inform the interpretations from Chapter 6.

In Chapter 6, I use the methodology established in Chapter 5 in a field-based study. Here I use the iDNA method to explore the distribution of mammals within the highly fragmented Wheatbelt region of Western Australia focusing on conservation reserves and road edges. This chapter explores how iDNA metabarcoding can be used in a conservation and land management setting.

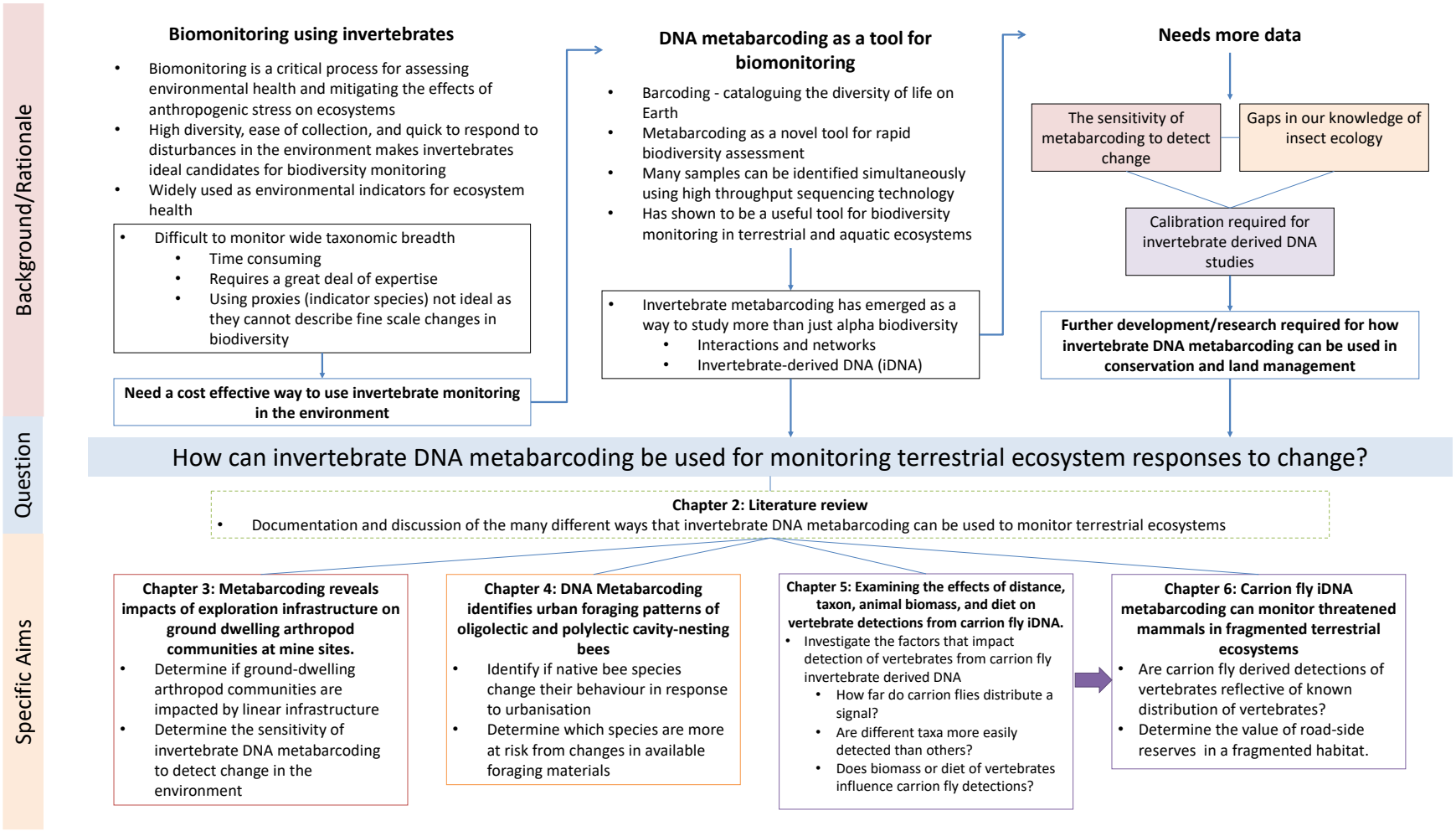


Figure 1-2 Conceptual framework relating the background, key research question and specific aims of thesis

1.7 References

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged

Alarcón, R. (2009). Congruence between visitation and pollen-transport networks in a California plant-pollinator community. *Oikos* , *119*(1), 35–44. doi:10.1111/j.1600-0706.2009.17694.x

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*(1), 134–147. doi:10.1111/2041-210X.12849

Ashfaq, M., & Hebert, P. D. N. (2016). DNA barcodes for bio-surveillance: regulated and economically important arthropod plant pests. *Genome*, *59*(11), 933–945. doi:10.1139/gen-2016-0024

Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, *21*(8), 2039–2044. doi:10.1111/j.1365-294x.2012.05519.x

Barsoum, N., Bruce, C., Forster, J., Ji, Y.-Q., & Yu, D. W. (2019). The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity. *Ecological Indicators*, *101*, 313–323. doi:10.1016/j.ecolind.2019.01.023

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., & Slik, J. W. F. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific Reports*, *6*(1), 24965. doi:10.1038/srep24965

Best, K., Oakes, T., Heather, J. M., Shawe-Taylor, J., & Chain, B. (2015). Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding. *Scientific Reports*, *5*, 14629. doi:10.1038/srep14629

Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., … Hering, D. (2012). Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. *Ecological Indicators*, *18*, 31–41. doi:10.1016/j.ecolind.2011.10.009

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., & Woodward, G. (2017). Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks. *Trends in Ecology & Evolution*, *32*(7), 477–487. doi:10.1016/j.tree.2017.03.001

Bohmann, K., Monadjem, A., Noer, C. L., Rasmussen, M., Zeale, M. R. K., Clare, E., … Gilbert, M. T. P. (2011). Molecular diet analysis of two African free-tailed bats (Molossidae) using high throughput sequencing. *PloS One*, *6*. doi:10.1371/journal.pone.0021441

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*. doi:10.1111/1755-0998.12428

Burkle, L. A., & Alarcón, R. (2011). The future of plant-pollinator diversity: understanding interaction networks across time, space, and global change. *American Journal of Botany*, *98*(3), 528–538. doi:10.3732/ajb.1000391

Callahan, B. J., McMurdie, P. J., Rosen, M. J., & Han, A. W. (2015). DADA2: High resolution sample inference from amplicon data. *BioRxiv*. https://www.biorxiv.org/content/10.1101/024034v1.abstract

Calvignac-Spencer, S., Leendertz, F. H., Gilbert, M. T. P., & Schubert, G. (2013). An invertebrate stomach's view on vertebrate ecology: certain invertebrates could be used as "vertebrate samplers" and deliver DNA-based information on many aspects of vertebrate ecology. *BioEssays*, *35*(11), 1004–1013. doi:10.1002/bies.201300060

Calvignac-Spencer, S., Merkel, K., Kutzner, N., Kühl, H., Boesch, C., Kappeler, P. M., … Leendertz, F. H. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Molecular Ecology*, *22*(4), 915–924. doi:10.1111/mec.12183

Carew, M. E., Coleman, R. A., & Hoffmann, A. A. (2018). Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ*, *6*, e4980. doi:10.7717/peerj.4980

Carøe, C., & Bohmann, K. (2020). Tagsteady: A metabarcoding library preparation protocol to avoid false assignment of sequences to samples. *Molecular Ecology Resources*, *20*(6), 1620–1631. doi:10.1111/1755-0998.13227

Clare, E. L., Chain, F. J. J., Littlefair, J. E., & Cristescu, M. E. (2016). The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. *Genome*, *59*(11), 981–990. doi:10.1139/gen-2015-0184

Creedy, T. J., Ng, W. S., & Vogler, A. P. (2019). Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy. *Ecology and Evolution*, *9*(6), 3105–3116. doi:10.1002/ece3.4839

de Oliveira, S. S., Ortega, J. C. G., Ribas, L. G. dos S., Lopes, V. G., & Bini, L. M. (2020). Higher taxa are sufficient to represent biodiversity patterns. *Ecological Indicators*, *111*, 105994. doi:10.1016/j.ecolind.2019.105994

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, *10*(9), 20140562. doi:10.1098/rsbl.2014.0562

Derocles, S. A. P., Bohan, D. A., Dumbrell, A. J., Kitson, J. J. N., Massol, F., Pauvert, C., … Evans, D. M. (2018). Biomonitoring for the 21st Century: Integrating Next-Generation Sequencing Into Ecological Network Analysis. In D. A. Bohan, A. J. Dumbrell, G. Woodward, & M. Jackson (Eds.), *Advances in Ecological Research* (Vol. 58, pp. 1–62). doi:10.1016/bs.aecr.2017.12.001

Edgar, R. C., Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476–3482.

Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., … Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, *7*, e7745. doi:10.7717/peerj.7745

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*(17), 6918–6926. doi:10.1002/ece3.3192

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution.* doi:10.1111/2041-210X.12789

Erdozain, M., Thompson, D. G., Porter, T. M., Kidd, K. A., Kreutzweiser, D. P., Sibley, P. K., … Hajibabaei, M. (2019). Metabarcoding of storage ethanol vs. conventional morphometric identification in relation to the use of stream macroinvertebrates as ecological indicators in forest management. *Ecological Indicators*, *101*, 173–184. doi:10.1016/j.ecolind.2019.01.014

Evans, D. M., & Kitson, J. J. (2020). Molecular ecology as a tool for understanding pollination and other plant-insect interactions. *Current Opinion in Insect Science*, *38*, 26–33. doi:10.1016/j.cois.2020.01.005

Fernandes, K., Heyde, M., Coghlan, M., Wardell-Johnson, G., Bunce, M., Harris, R., & Nevill, P. (2019). Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences. *Restoration Ecology*, *27*(5), 1177–1186. doi:10.1111/rec.12976

Gálvez-Reyes, N., Arribas, P., Andujar, C., Emerson, B., Piñero, D., & Mastretta-Yanes, A. (2020). Local-scale dispersal constraints promote spatial structure and arthropod diversity within a tropical sky-island. *Authorea Preprints*. Retrieved from https://www.authorea.com/doi/full/10.22541/au.160193334.45224582

Gariepy, T. D., Lindsay, R., Ogden, N., & Gregory, T. R. (2012). Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks. *Molecular Ecology Resources*, *12*(4), 646–652. doi:10.1111/j.1755-0998.2012.03140.x

Gerlach, J., Samways, M., & Pryke, J. (2013). Terrestrial invertebrates as bioindicators: an overview of available taxonomic groups. *Journal of Insect Conservation*, *17*(4), 831–850. doi:10.1007/s10841-013-9565-9

Gogarten, J. F., Hoffmann, C., Arandjelovic, M., Sachse, A., Merkel, K., Dieguez, P., … Calvignac-Spencer, S. (2019). Fly-derived DNA and camera traps are complementary tools for assessing mammalian biodiversity. *Environmental DNA*, *9*, 79. doi:10.1002/edn3.46

Goldberg, C. S., Strickler, K. M., & Fremier, A. K. (2018). Degradation and dispersion limit environmental DNA detection of rare amphibians in wetlands: Increasing efficacy of sampling designs. *The Science of the Total Environment*, *633*, 695–703. doi:10.1016/j.scitotenv.2018.02.295

Gray, C., Baird, D. J., Baumgartner, S., Jacob, U., Jenkins, G. B., O'Gorman, E. J., … Woodward, G. (2014). FORUM: Ecological networks: the missing links in biomonitoring science. *The Journal of Applied Ecology*, *51*(5), 1444–1449. doi:10.1111/1365-2664.12300

Gresty, C. E. A., Clare, E., Devey, D. S., Cowan, R. S., Csiba, L., Malakasi, P., … Willis, K. J. (2018). Flower preferences and pollen transport networks for cavity-nesting solitary bees: Implications for the design of agri-environment schemes. *Ecology and Evolution*, *8*(15), 7574–7587. doi:10.1002/ece3.4234

Hanya, G., Morishima, K., Koide, T., Otani, Y., Hongo, S., Honda, T., … Nakano, T. (2019). Host selection of hematophagous leeches (Haemadipsa japonica): Implications for iDNA studies. *Ecological Research*, *34*(6), 842–855. doi:10.1111/1440-1703.12059

Hershey, A. E., Lamberti, G. A., Chaloner, D. T., & Northington, R. M. (2010). Chapter 17 - Aquatic Insect Ecology. In J. H. Thorp & A. P. Covich (Eds.), *Ecology and Classification of North American Freshwater Invertebrates (Third Edition)* (pp. 659–694). doi:10.1016/B978-0-12-374855-3.00017-0

Hubert, N., & Hanner, R. (2015). DNA Barcoding, species delineation and taxonomy: a historical perspective. *DNA Barcodes* , *3*(1). doi:10.1515/dna-2015-0006

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. doi:10.1111/ele.12162

Ji, Y., Baker, C. C. M., Popescu, V. D., Wang, J., Wu, C., Wang, Z., … Yu, D. W. (2022). Measuring protected-area effectiveness using vertebrate distributions from leech iDNA. *Nature Communications*, *13*(1), 1555. doi:10.1038/s41467-022-28778-8

Kim, K. C. (1993). Biodiversity, conservation and inventory: why insects matter. *Biodiversity and Conservation*, *2*(3), 191–214. doi:10.1007/bf00056668

Kocher, A., de Thoisy, B., Catzeflis, F., Valière, S., Bañuls, A.-L., & Murienne, J. (2017). iDNA screening: Disease vectors as vertebrate samplers. *Molecular Ecology*, *26*(22), 6478–6486. doi:10.1111/mec.14362

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(16), 6196–6200. doi:10.1073/pnas.86.16.6196

Kuntke, F., de Jonge, N., Hesselsøe, M., & Lund Nielsen, J. (2020). Stream water quality assessment by metabarcoding of invertebrates. *Ecological Indicators*, *111*, 105982. doi:10.1016/j.ecolind.2019.105982

Lee, P.-S., Gan, H. M., Clements, G. R., & Wilson, J.-J. (2016). Field calibration of blowfly-derived DNA against traditional methods for assessing mammal diversity in tropical forests. *Genome*, *59*(11), 1008–1022. doi:10.1139/gen-2015-0193

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2019). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, *6*, e27295v2. doi:10.1111/een.12831

Lucas, A., Bodger, O., Brosi, B. J., Ford, C. R., Forman, D. W., Greig, C., … de Vere, N. (2018). Floral resource partitioning by individuals within generalised hoverfly pollination networks revealed by DNA metabarcoding. *Scientific Reports*, *8*(1), 5133. doi:10.1038/s41598-018-23103-0

Lynggaard, C., Nielsen, M., Santos-Bay, L., Gastauer, M., Oliveira, G., & Bohmann, K. (2019). Vertebrate diversity revealed by metabarcoding of bulk arthropod samples from tropical forests. *Environmental DNA*, *1*(4), 329–341. doi:10.1002/edn3.34

Macgregor, C. J., Kitson, J. J. N., Fox, R., Hahn, C., Lunt, D. H., Pocock, M. J. O., & Evans, D. M. (2019). Construction, validation, and application of nocturnal pollen transport networks in an agro-ecosystem: a comparison using light microscopy and DNA metabarcoding. *Ecological Entomology*, *44*(1), 17–29. doi:10.1111/een.12674

Macher, J. N., Salis, R. K., Blakemore, K. S., Tollrian, R., Matthaei, C. D., & Leese, F. (2016). Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators*, *61*, 159–169. doi:10.1016/J.ECOLIND.2015.08.024

Masonick, P., Hernandez, M., & Weirauch, C. (2019). No guts, no glory: Gut content metabarcoding unveils the diet of a flower-associated coastal sage scrub predator. *Ecosphere* , *10*(5), e02712. doi:10.1002/ecs2.2712

Massey, A. L., Bronzoni, R. V. de M., da Silva, D. J. F., Allen, J., de Lazari, P. R., Dos Santos Filho, M., … Levi, T. (2021). Invertebrates for vertebrate biodiversity monitoring: comparisons using three insect taxa as iDNA samplers. *Molecular Ecology Resources*. doi:10.1111/1755-0998.13525

McGeoch, M. A. (1998). The selection, testing and application of terrestrial insects as bioindicators. *Biological Reviews of the Cambridge Philosophical Society*, *73*(2), 181–201. doi:10.1017/s000632319700515x

Miller, K. E., Hopkins, K., Inward, D. J. G., & Vogler, A. P. (2016). Metabarcoding of fungal communities associated with bark beetles. *Ecology and Evolution*, *6*(6), 1590–1600. doi:10.1002/ece3.1925

Mullis, K. B., & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology*, *155*, 335-350. doi:10.1016/0076-6879(87)55023-6

Murray, D. C., Coghlan, M. L., & Bunce, M. (2015). From benchtop to desktop: important considerations when designing amplicon sequencing workflows. *PloS One*, *10*(4), e0124671. doi:10.1371/journal.pone.0124671

Mychek-Londer, J. G., Balasingham, K. D., & Heath, D. D. (2019). Using environmental DNA metabarcoding to map invasive and native invertebrates in two Great Lakes tributaries. *Environmental DNA*, *95*, 323. doi:10.1002/edn3.56

Nielsen, M., Gilbert, M. T. P., Pape, T., & Bohmann, K. (2019). A simplified DNA extraction protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity. *Environmental DNA*, edn3.16. doi:10.1002/edn3.16

Niemeijer, D. (2002). Developing indicators for environmental policy: data-driven and theory-driven approaches examined by example. *Environmental Science & Policy*, *5*(2), 91–103. doi:10.1016/S1462-9011(02)00026-6

Obrist, M. K., & Duelli, P. (2010). Rapid biodiversity assessment of arthropods for monitoring average local species richness and related ecosystem services. *Biodiversity and Conservation*, *19*(8), 2201–2220. doi:10.1007/s10531-010-9832-y

Oliverio, A. M., Gan, H., Wickings, K., & Fierer, N. (2018). A DNA metabarcoding approach to characterize soil arthropod communities. *Soil Biology & Biochemistry*, *125*, 37–43. doi:10.1016/J.SOILBIO.2018.06.026

Owings, C. G., Banerjee, A., Asher, T. M. D., Gilhooly, W. P., 3rd, Tuceryan, A., Huffine, M., … Picard, C. J. (2019). Female Blow Flies As Vertebrate Resource Indicators. *Scientific Reports*, *9*(1), 10594. doi:10.1038/s41598-019-46758-9

Paknia, O., Rajaei, H. S., & Koch, A. (2015). Lack of well-maintained natural history collections and taxonomists in megadiverse developing countries hampers global biodiversity exploration. *Organisms, Diversity & Evolution*, *15*(3), 619–629. doi:10.1007/s13127-015-0202-1

Pik, A. J., Oliver, I., & Beattie, A. J. (1999). Taxonomic sufficiency in ecological studies of terrestrial invertebrates. *Australian Journal of Ecology*, *24*(5), 555–562. doi:10.1046/j.1442-9993.1999.01003.x

Pompanon, F., Coissac, E., & Taberlet, P. (2011). Metabarcoding a new way to analyze biodiversity. *Biofutur, 319,* 30- 32

Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., … Andalo, C. (2016). Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific Reports*, *6*, 27282. doi:10.1038/srep27282

Porter, T. M., & Hajibabaei, M. (2020). Putting COI Metabarcoding in Context: The Utility of Exact Sequence Variants (ESVs) in Biodiversity Analysis. *Frontiers in Ecology and Evolution*, *8*, 248. doi:10.3389/fevo.2020.00248

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution*, *25*(6), 345–353. doi:10.1016/J.TREE.2010.01.007

Reeves, L. E., Holderman, C. J., Gillett-Kaufman, J. L., Kawahara, A. Y., & Kaufman, P. E. (2016). Maintenance of host DNA integrity in field-preserved mosquito (Diptera: Culicidae) blood meals for identification by DNA barcoding. *Parasites & Vectors*, *9*(1), 503. doi:10.1186/s13071-016-1791-z

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, *39*(21), e145. doi:10.1093/nar/gkr732

Rodgers, T. W., Xu, C. C. Y., Giacalone, J., Kapheim, K. M., Saltonstall, K., Vargas, M., … Jansen, P. A. (2017). Carrion fly-derived DNA metabarcoding is an effective tool for mammal surveys: Evidence from a known tropical mammal community. *Molecular Ecology Resources*, *17*(6), e133–e145. doi:10.1111/1755-0998.12701

Roger, F., Ghanavi, H., Danielsson, N., Wahlberg, N., Löndahl, J., Pettersson, L. B., … Clough, Y. (2022). Airborne environmental DNA metabarcoding for the monitoring of terrestrial insects - a proof of concept from the field. *Environmental DNA.* doi: 10.1002/edn3.290

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., … Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, *239*(4839), 487–491. doi:10.1126/science.2448875

Schloss, P. D., Gevers, D., & Westcott, S. L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PloS One*, *6*(12), e27310. doi:10.1371/journal.pone.0027310

Schnell, I. B., Sollmann, R., Calvignac-Spencer, S., Siddall, M. E., Yu, D. W., Wilting, A., & Gilbert, M. T. P. (2015). iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool - prospects, pitfalls and avenues to be developed. *Frontiers in Zoology*, *12*, 24. doi:10.1186/s12983-015-0115-z

Schubert, G., & Stockhausen, M. (2015). Targeted detection of mammalian species using carrion fly-derived DNA. *Molecular Ecology,* 15(2), 285-294. doi:10.1111/1755-0998.12306

Silva, T. H. S. E., Karunarathna, N. B., & Karunaratne, W. A. I. P. (2017). Competence of Litter Ants for Rapid Biodiversity Assessments. *International Journal of Biodiversity*, *2017*, 1–10. doi:10.1155/2017/6582191

Souza, J. L. P., Baccaro, F. B., Landeiro, V. L., Franklin, E., Magnusson, W. E., Pequeno, P. A. C. L., & Fernandes, I. O. (2016). Taxonomic sufficiency and indicator taxa reduce sampling costs and increase monitoring effectiveness for ants. *Diversity & Distributions*, *22*(1), 111–122. doi:10.1111/ddi.12371

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., … Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, *7*(1), 12240. doi:10.1038/s41598-017-12501-5

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For Biodiversity Research and Monitoring*. doi:10.1093/oso/9780198767220.001.0001

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050. doi:10.1111/j.1365-294X.2012.05470.x

Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, *17*(5), 1105–1109. doi:10.1007/BF00037152

Thomsen, P. F., & Sigsgaard, E. E. (2019). Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. *Ecology and Evolution*, *9*(4), 1665–1679. doi:10.1002/ece3.4809

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. doi:10.1016/J.BIOCON.2014.11.019

van der Heyde, M., Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., & Nevill, P. (2020). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Molecular Ecology Resources*, *20*(3). doi:10.1111/1755-0998.13148

Wagner, D. L. (2020). Insect Declines in the Anthropocene. *Annual Review of Entomology*, *65*, 457–480. doi:10.1146/annurev-ento-011019-025151

Wilson, R. J., & Fox, R. (2021). Insect responses to global change offer signposts for biodiversity and conservation. *Ecological Entomology*, *46*(4), 699–717. doi:10.1111/een.12970

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, *3*(4), 613–623. doi:10.1111/j.2041-210X.2012.00198.x

CHAPTER 2

CURRENT AND FUTURE APPLICATIONS OF BULK ARTHROPOD SAMPLE METABARCODING IN TERRESTRIAL ECOLOGICAL MONITORING

2.1 Preface

This chapter reviews the current and past literature on the use of metabarcoding bulk arthropods for terrestrial ecological assessment and makes comment on the potential future of this technique's application in ecological monitoring.

In recent years there has been a rapid increase in articles published using metabarcoding of arthropods as valuable ecological sensors towards changes or disturbance to the environment. The growing number of studies sheds light on the importance of arthropods to the health of many ecological systems, not only through their diversity but through the many trophic scales that arthropods are able to access and interact with. However, there have not been any reviews to our knowledge that sum up the extent to which bulk arthropod sample metabarcoding has been applied, or the potential future applications of this methodology for use in monitoring terrestrial ecological systems. Here we explore the current and potential future use of metabarcoding bulk arthropod samples in arthropod biodiversity surveying (and its various applications in habitat quality assessment), plant-pollinator interactions, vertebrate monitoring, and the study of pathogen-transmission. We then discuss considerations for bulk arthropod sample metabarcoding, particularly the future developments that need to be made to advance the field. We feel that metabarcoding and other molecular techniques will become a common place addition to monitoring terrestrial ecosystems.

This chapter consists of a manuscript that is currently being prepared for submission to the peer reviewed journal, *Environmental DNA*. It has been reformatted for inclusion in the thesis, but some elements of the formatting such as the structure and citation style remain.

2.1.1 Acknowledgements

KF was partially supported by funding from Food Agility CRC Ltd, funded under the Commonwealth Government CRC Program. CL was supported by the National Council of Science and Technology in Mexico (CONACyT) Grant CVU 582667 and the Danish Council for Independent Research, DFF grant 7079-00029B. PC is part of the H2020 MSCA-ITN-ETN Plant.ID network and has received funding from the European Union's Horizon 2020 research innovation programme under grant agreement No 765000. PN was supported by the Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration (ICI150100041).

2.1.2 Author Contributions

KF, CL, MN, KB conceived the manuscript; KB and CL created the figure. KF coordinated and edited the final manuscript. All authors contributed to writing and editing the manuscript and gave final approval for publication.

Current and future applications of bulk arthropod sample metabarcoding in terrestrial ecological monitoring

Kristen Fernandes[1,2,3], Christina Lynggaard[2], Martin Nielsen[2,4], Physilia Chua[2], Paul Nevill[1,5], Michael Bunce[1,6], Kristine Bohmann[2]

1 Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia
2 Section for Evolutionary Genomics, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen K, Denmark

3 Food Agility CRC Ltd, 81 Broadway, Ultimo, NSW, Australia 2007

4Arctic Station, Greenland, University of Copenhagen

5ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

6 Environmental Protection Authority, 215 Lambton Quay, Wellington 6011, New Zealand

2.2 Abstract

Arthropods play various roles in ecosystem function and are widely used as indicator species in terrestrial biomonitoring surveys. Traditionally, arthropod surveys have relied on morphological taxonomic identification, which is typically labour intensive and requires expertise in identification across multiple taxonomic groups. This makes undertaking in-depth arthropod surveys costly. In recent years, DNA metabarcoding has emerged as a complementary tool that can help combat some of the disadvantages of morphological surveys due to its ability to rapidly assess biodiversity from hundreds of samples in parallel. Another advantage of DNA metabarcoding is that it allows for the detection of organisms that the arthropods have in one way or another interacted with, for instance, vertebrates, plants, fungi, bacteria, and pathogens. Therefore, for terrestrial ecological monitoring, metabarcoding is uniquely versatile and beneficial. A growing number of studies have recently used bulk arthropod sample metabarcoding to assess ecosystem biodiversity in various contexts. Accordingly, a review is timely to sum up how this technique has been used in terrestrial biomonitoring studies and to examine the potential future directions of the field. Here, we review studies using metabarcoding of terrestrial bulk arthropod samples to explore the multiple uses for this technique in monitoring arthropod diversity, habitat quality assessment, the study of plant-pollinator interactions, vertebrate monitoring, and the transmission of vector-borne pathogens. We also discuss current limitations and perspectives on future research areas that need to be conducted for bulk arthropod metabarcoding to become more commonplace in the biomonitoring toolkit.

2.3 Introduction

Arthropods are the most successful taxonomic group to roam the Earth in terms of the number of taxa, habitat adaptation, biomass, and ecological impact (Wagner, 2020). With millions of species, presence in a wide diversity of habitats, variety of ecosystem functions, sensitivity to disturbance, and relative ease of collection, arthropods have been widely used in environmental assessments of both aquatic and terrestrial ecosystems across the globe (da Rocha, De Almeida, Lins, and Durval, 2010; Tylianakis et al., 2004). For example, arthropod monitoring has been routinely used to assess ecosystem quality and health (Peck, Mcquaid, and Campbell, 1998; Segat et al., 2017; Tiede et al., 2017) and to monitor pollution levels (Celli and Maccagnani, 2003; Ghannem, Touaylia, and Boumaiza, 2018; Salkova and Panayotova-Pencheva, 2016). However, their sensitivity to ecosystem disturbance has led to notable declines of arthropod species across the globe (Wagner, 2020), and many more will go unnoticed due to the large gap in our knowledge of their diversity (Sánchez-Bayo and Wyckhuys, 2019; van Klink et al., 2020; Wagner, 2020). As arthropods are essential to functionality in ecosystems, arthropod communities must be documented to establish biodiversity baselines and inform conservation practices where these ecosystems are threatened.

Traditionally, arthropod diversity assessment involves the mass collection of individuals (hereafter referred to as a 'bulk sample' or a 'bulk arthropod sample') using various trapping techniques, followed by morphological identification of each specimen to compile community assemblages. However, undertaking such assessment across multiple taxonomic groups is resource and time-intensive and requires specialised taxonomic expertise to identify species in each arthropod group (Basset et al., 2012; Ji et al., 2013; Montgomery et al., 2021). Therefore, studies are often restricted to selected taxonomic groups such as ants, beetles, or butterflies, which are used as indicators of ecosystem change (Oliver and Beattie, 1996; Pik, Oliver, and Beattie, 1999; Terlizzi et al., 2003). Unfortunately, focusing on only select taxonomic groups has been shown to inaccurately depict true biodiversity, as many of these arthropod groups do not interact or overlap with each other spatially or in taxonomic richness (Khosravi and Hemami, 2019; Prendergast et al., 1993; van Jaarsveld et al., 1998). Put simply, sometimes, the arthropods that might be easy to identify (or count) may not be the best indicators for a given study.

Recent developments in DNA sequencing technologies have overcome some of the limitations of morphology-based studies, enabling the survey of multiple taxonomic groups with accuracy (Yu et al., 2012). For this, the most common method is metabarcoding. Metabarcoding is a method that can be used to simultaneously identify the taxa from multiple mixed-species samples in parallel, reducing the time taken to undertake biodiversity surveying (Taberlet et al. 2012). Metabarcoding uses a targeted approach in which only taxonomically informative markers (typically mitochondrial DNA, mtDNA) from specific taxonomic groups are sequenced (Taberlet et al., 2012). Polymerase chain reaction (PCR) amplification is carried out with primers that broadly target a taxonomically informative 'mini-barcode marker' in a selected taxonomic group. The foundation of metabarcoding is incorporating sample-specific nucleotide identifiers (tags and/or indexes) into the PCR amplicons (Bohmann et al., 2021). These identifiers allow the parallel sequencing of multiple samples by assigning the metabarcoding sequence reads back to the samples from which they originated. Before taxonomic assignment, sequence clusters are generated. Sequence clusters can be referred to as molecular operational taxonomic units (MOTUs) or amplicon sequence variants (ASVs) based on the algorithm used to cluster the sequences together. MOTUs are sequence clusters algorithmically determined by the degree of sequence similarity (typically from 97-99% similarity between sequences) (Clare et al., 2016; Floyd et al., 2002), whilst ASVs (also known as zero-distance OTUs (zOTUs)) are unique DNA sequences, distinguished by single nucleotide differences, generated from high-throughput sequencing methods (Callahan et al., 2016; Edgar, 2013). Here, we will use the terms MOTUs and ASVs interchangeably. MOTUs can then be taxonomically assigned by matching them against a reference database or used independent of taxonomy, with MOTU richness used as a proxy to estimate species diversity when reference databases are unavailable or poorly populated (Ashfaq and Hebert, 2016; Gálvez-Reyes et al., 2020). Thus, the application of metabarcoding to terrestrial biodiversity surveys is valuable, especially in regions with poorly studied biodiversity or limited resources available for conducting surveys.

Since the first study was published on metabarcoding of bulk arthropod samples in 2012 (Yu et al., 2012), bulk arthropod metabarcoding have been used to document biodiversity in a range of contexts (Fig 2-1). Some examples include habitat quality assessment (Beng et al., 2016; M. Liu et al., 2020; Yu et al., 2012), assessing restoration trajectories (Fernandes et al., 2019; Gervan et al., 2020; Lynggaard et al., 2020; van der Heyde et al., 2022), description of the phylogeographic structure and intraspecific variation within arthropod groups (Arribas et al., 2020; Elbrecht et al., 2018; Turon et al., 2019) and identification of vertebrates (Lynggaard et al., 2019). Here, we review the use of bulk sample metabarcoding for ecological monitoring in arthropod diversity studies, habitat quality assessments, vertebrate surveys, studies of plant-insect interactions, and vector-borne pathogens. We also discuss current limitations of the technique and perspectives on future development required for bulk arthropod metabarcoding studies to become commonplace in the biomonitoring toolkit.

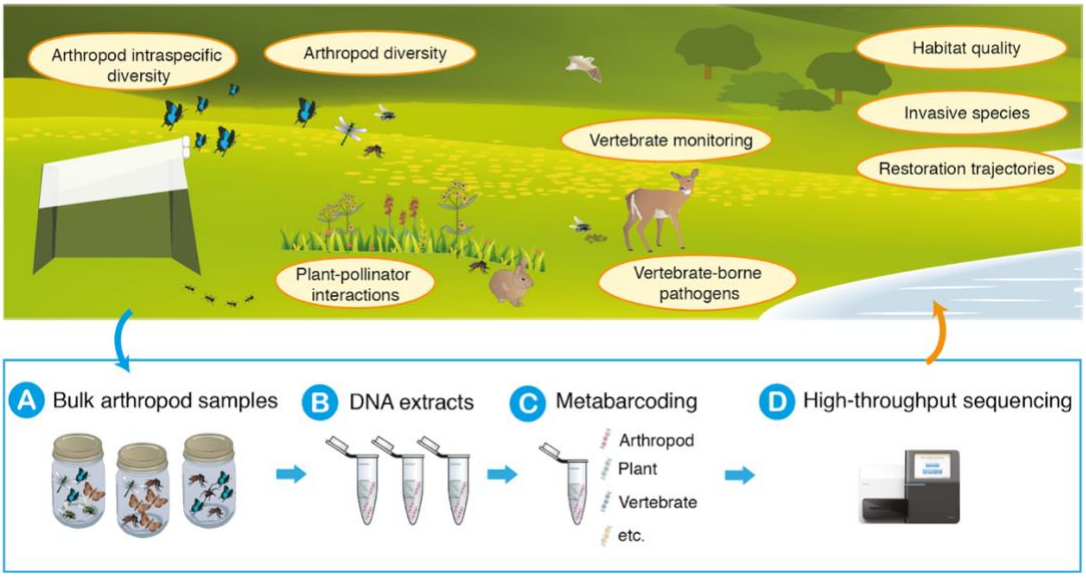


Figure 2-1 Illustrative example of methods and potential outputs of metabarcoding bulk arthropod samples for biomonitoring in terrestrial ecology.

2.4 Assessment of arthropod diversity

Assessment of arthropod taxonomic composition is generally the primary purpose when carrying out metabarcoding of bulk arthropod samples. Some of the first studies that demonstrated the ability of metabarcoding to assess arthropod taxonomic composition did so on 'mock' arthropod community samples. Mock community samples are produced by combining DNA extracted from multiple identified arthropod species in known quantities. Metabarcoding of mock samples has been used to validate the taxa detected from metabarcoding of field-collected bulk arthropod samples (S. Liu et al., 2013; Yu et al., 2012). Later studies have used mock samples as positive controls sequenced alongside field samples to validate laboratory and sequencing protocols (Braukmann et al., 2019; Carew, Coleman, and Hoffmann, 2018; Elbrecht et al., 2019; Elbrecht, Peinert, and Leese, 2017; S. Liu et al., 2013; Morinière et al., 2016; Nielsen et al., 2019).

Following these mock studies, the efficacy of metabarcoding has been tested in comparison to traditional morphological surveys, showing that metabarcoding can assess taxonomic composition as well, if not better, than morphological assessments of the same samples (Beng et al., 2016; Ji et al., 2013; M. Liu et al., 2020). Furthermore, bulk arthropod sample metabarcoding can be used to survey arthropods in areas where the effort and the associated cost for morphology-based surveys would be prohibitive. For example, metabarcoding of bulk arthropod samples have been used to characterise arthropod communities at scale in tropical forests and plantations (Beng et al., 2016; Zhang et al., 2016), mine sites (Fernandes et al., 2019; Lynggaard et al., 2020), soil (Oliverio et al., 2018), grasslands (Elbrecht et al., 2019; Foster et al., 2020), national parks (Geiger et al., 2016; Hardulak et al., 2020), farms (Agerbo Rasmussen et al., 2021; Hausmann et al., 2020), and forest canopies (Creedy, Ng, and Vogler, 2019). Metabarcoding has been used in these studies to provide powerful insights into ecosystem recovery, and the impacts of land use change and/or management on arthropod biodiversity through species composition and richness measures. In the following section we discuss how using arthropod diversity from metabarcoding studies can be applied to various terrestrial ecological monitoring applications, such as habitat quality assessment and invasive species detection. We also dive further into the application of arthropod diversity data gathered from metabarcoding and explore its usage for the assessment of intraspecific variation within arthropod communities.

2.4.1 Habitat quality assessment

Assessing habitat quality is an important part of ecosystem monitoring applications, and it is particularly useful in environmental impact studies, long term monitoring (Johnson, 2007), and monitoring ecosystem restoration trajectories (Choi, 2004). Most habitat quality assessments are conducted by measuring various degrees of ecosystem functioning or condition through documenting ecological indicators (McGeoch et al. 2011). These can be abiotic, for example, water or soil quality, or they can be biotic, for example, floral or faunal biodiversity. The application of effective habitat quality monitoring allows for appropriate management strategies to be implemented to mitigate damage. Whilst vegetation community monitoring has been a large focus for many habitat monitoring projects, vegetation alone cannot be used to assess ecosystem functioning (Cross, Bateman, and Cross, 2020). Arthropod diversity is increasingly recognised as an important biotic indicator in habitat quality assessments (Hackman, Gong, and Venevsky, 2017). This is because arthropods display a sensitive and quantifiable response to disturbances (McGeoch, 1998), are often essential to many ecological functions (Rosenberg, Danks, and Lehmkuhl, 1986), and are relatively easy to capture (McGeoch et al., 2011). Studies using arthropods for habitat quality assessment will usually measure the changes in arthropod community dynamics over time or before and after disturbance. Such changes can be measured in arthropod density, abundance, or taxonomic composition (Benson, Dinsmore, and Hohman, 2007; Happe et al., 2019; Kremen et al., 1993; Wettstein and Schmid, 1999).

In long term habitat quality assessment, continual monitoring of arthropod communities is vital for accuracy in detecting impacts of disturbances (Herrick, Schuman, and Rango, 2006). However, continual monitoring is rarely implemented effectively in terrestrial ecosystems because of the difficulties and costs associated with arthropod monitoring (Bang and Faeth, 2011; Gillespie et al., 2020). Thus, metabarcoding for arthropod community monitoring has emerged as an essential tool in habitat quality assessment. For example, bulk arthropod sample metabarcoding has been used to evaluate the impacts of historical logging in forests on beetle community recovery (M. Liu et al., 2020), monitor the effects of dieback and climate change on invertebrate communities in forests (Sire et al., 2017), and assess the status of ecosystem restoration trajectories and outcomes in the context of mining (Lynggaard et al., 2020; Gervan et al., 2020; Fernandes et al., 2019; van der Heyde et al., 2022). However, continual habitat monitoring using bulk sample metabarcoding is yet to be formally explored in terrestrial environments to the same degree that it has been in aquatic environments (Aylagas et al., 2018; Beermann et al., 2021; Li et al., 2018).

Macroinvertebrate species have long been a part of aquatic habitat quality assessment. Unlike terrestrial systems, there has been considerable efforts undertaken to assign these species to functional aspects of habitat quality and health. For example, characteristics such as altered pH, temperature, or pollution have been found to impact aggregation, activity, or reproduction of certain aquatic macroinvertebrate groups (Moretti et al., 2017). Subsequently, this information can be related to the overall levels of nutrient cycling, reproduction, or other key functional aspects of habitat quality. In European freshwater systems, macroinvertebrates have become ideal candidates to use in biomonitoring for habitat quality and function (Hering et al., 2006; Leese et al., 2018; Macher et al., 2016), and DNA metabarcoding of macroinvertebrate communities is being explored as a way to monitor and inform the management of these systems to protect and restore ecosystem health (i.e., DNAqua-net; https://dnaqua.net/). In terrestrial systems, it is much harder to undertake such monitoring regimes for several reasons, including the lack of taxonomic knowledge for terrestrial arthropod species (Paknia, Rajaei, and Koch, 2015). Further, this extends into lack of species ecology and life history traits, and therefore identification of functional traits to assess habitat quality is made even more difficult.

In this case, metabarcoding can be used effectively in a taxonomy independent space. Whilst taxonomic identification in biotic surveys will always be essential for biomonitoring purposes, MOTUs from metabarcoding datasets can be used as a proxy for some diversity metrics. By combining arthropod MOTU diversity matrices and environmental parameters, such as soil quality indicators or vegetation indexes, it is possible to gain insightful correlations towards habitat quality through the identification of 'indicator MOTUs' of ecosystem health and functioning (Beentjes et al., 2019; Lanzén et al., 2016). Further, if MOTUs repeatedly occur with specific environmental conditions, targeted efforts can be made to build upon reference databases to taxonomically identify them at finer taxonomic resolutions.

2.4.2 Detection of invasive arthropod species

Bulk arthropod sample metabarcoding can also be an effective technique to detect invasive arthropod species. Invasive arthropod species have spread globally through the increase in trade (Liebhold et al., 2016) and can have devastating economic, ecosystem, and human health impacts (Dahlsten and Garcia, 1989; Liebhold et al., 2016). For example, a recent review on soybean production in Brazil found that the invasions of several invertebrate species such as the silverleaf whitefly (*Bemisia tabaci*), red spider mite (*Tetranychus urticae*), and the cotton bollworm (*Helicoverpa armigera*) contributed to the loss of approximately US$1.51 billion annually (Pozebon et al., 2020). Effective monitoring programs across broad spatial and temporal scales, and diverse taxonomic groups are thus essential to reducing the spread of invasive arthropod species (Cuthbert et al., 2020; Davidovitch et al., 2009; Piper et al., 2019; Whittle et al., 2013). However, morphological identifications from arthropod surveillance traps (such as pitfall traps and baited traps) can become costly and time-consuming to obtain because they capture large numbers of specimens including bycatch (Piper et al., 2019). Therefore, metabarcoding can be implemented as a method for rapid arthropod identification. Whilst this method has been trialled on a small-scale basis (Hardulak et al., 2020), the approach is yet to be implemented across wider geographic and temporal scales.

There are also opportunities to use metabarcoding to build knowledge on mechanisms of biological control for pest and invasive arthropod species by looking at host-parasitoid relationships. Parasitoids can function as biological control agents if used in a targeted manner. For example, the parasitoid wasp *Encarsia formosa* is used to control greenhouse whitefly populations to improve crop production (Bale, van Lenteren, and Bigler, 2008). These interactions have traditionally been identified through rearing of hosts and parasitoids together (Pocock, Evans, and Memmott, 2012). But this is labour intensive, and species can be challenging to identify morphologically due to the lack of diagnostic features in eggs or larval stages (Evans and Kitson, 2020) and the presence of cryptic species (Kaartinen et al., 2010; Smith et al., 2008; Smith et al., 2007). There is also evidence that rearing approaches can underestimate the rate of parasitism (Day, 1994). Applying metabarcoding to bulk samples to identify associations with parasitoids would help resolve unknown or cryptic host-parasitoid interactions with less time invested than lab-based rearing methods.

Whilst using metabarcoding in these contexts seems promising, regulatory bodies require the standardisation and validation of new techniques before implementation. Therefore, a consensus on the methodologies, processing, and analysis of metabarcoding terrestrial bulk arthropod samples will be paramount to its application in invasive species monitoring in the future (Piper et al. 2019; Darling et al. 2020). This can be challenging because standardisation in a developing field is difficult to achieve. Nevertheless, the application of bulk sample metabarcoding shows promise in invasive arthropod monitoring because rapid species identification is essential for enacting management strategies that prevent damages from invasive arthropod species to crop yield, human, or environmental health.

2.4.3 Metabarcoding of bulk samples to assess intraspecific diversity of arthropods

In addition to assessing the taxonomic constituents of bulk arthropod samples, metabarcoding can provide insights into the arthropod's intraspecific genetic diversity by looking at DNA sequence variations between and within the taxa detected (Elbrecht et al., 2018; Turon et al., 2019; Wares and Pappalardo, 2015). These intraspecific variations are known as haplotypes. Haplotype analysis provides information on the diversity within a population as well as population size and demographics (Adams et al., 2019; Sigsgaard et al., 2019). This genetic information can be used to track population genetic variability, which could be applied in dispersal studies to understand connectivity and biogeography (Turon et al., 2019). For example, Pedro et al. (2017) were able to differentiate arthropod communities between separate farms and assess the dispersal capabilities of *Xyleborus* beetles at these locations by examining the haplotypes generated from metabarcoding.

Whilst haplotype studies initially focused on using 18s ribosomal RNA sequences for MOTU delimitation, there is limited within-MOTU variation within this gene region (Turon et al., 2019). Thus, studies on arthropod population dynamics have focused on the much more variable COI (cytochrome c oxidase subunit I) gene (Elbrecht et al., 2019). The COI gene has previously been proven as a useful marker in population genetics and phylogeographical studies across terrestrial and aquatic systems (Doorenweerd et al., 2020; Emerson et al., 2011). Further, increasing efforts globally to sequence the COI region across many taxonomic groups (i.e., the Barcode of Life Data Systems) gives a wealth of information on the intraspecies variation that will be continuously increasing in breadth and coverage (Ratnasingham and Hebert, 2007). This means that there will be data available on the intraspecific variation that spans countless studies across species and geographic regions that could be useful in documenting variation among and within species. With this, the new field of 'metaphylogeography' has been proposed where metabarcoding datasets using COI can be used to study intraspecific diversity and phylogeographic structure of many species simultaneously, instead of focusing on a single species at a time (Turon et al., 2019).

Some of the primary issues identified in obtaining haplotype data through metabarcoding of bulk arthropod samples have been; primer biases, loss of sequences with low abundances, PCR and sequencing error, and variation in the size of individual arthropods in the samples, which can inflate and confound haplotype diversity assessment through unequal sequencing depth (Edgar and Flyvbjerg, 2015; Elbrecht et al., 2018; Turon et al., 2019). Although it is possible to eliminate these errors through the use of MOTU clusters, clustering can cause the loss of some of the important intraspecific variation of populations from unique sequences (Elbrecht et al., 2018). Therefore, denoising protocols and improvements to bioinformatic workflow have been crucial in separating false MOTUs and true MOTUs (Frøslev et al., 2017; Turon et al., 2019). Subsequently, developing these protocols and further analysis using bulk arthropod sample metabarcoding will be important in closing the knowledge gaps surrounding arthropod ecology, moving beyond cataloguing just alpha and beta diversity.

2.5 Understanding plant-insect interactions

Plant-insect interactions play a major role in many important ecosystem functions crucial for food production and ecosystem health (Potts et al., 2010). There are two main types of interactions between insects and plants: mutualistic or antagonistic (Bernays, 1992). In the mutualistic relationship, both plants and insects benefit from these interactions (e.g., pollination) while antagonistic relationships are when only one party benefits while the other is harmed (e.g., phytophagy by insect pests). The main methodology used to study plant-insect interactions is through observation and identification of both plant and insect species and the nature of their interaction (Burkle and Alarcón, 2011). However, this approach is inefficient at a large scale because of the time and skills required for observation and identification (Alarcón, 2009). Instead, it is possible to use bulk arthropod metabarcoding to detect the plant species that an insect has been in contact with using plant-specific primers (van der Heyde et al., 2020). This bypasses the need for the morphological identification of both plants and insects. However, there are a few limitations to this approach. First, it is not possible to tease apart species-specific interactions due to the nature of the bulk samples where extracted DNA from many individuals are mixed. Second, it is not possible to quantify these relationships, and only associations between groups can be measured (Evans and Kitson, 2020). Lastly, it is difficult to determine if these associations represent mutualistic or antagonistic relationships. Nonetheless, metabarcoding of bulk arthropod samples can be valuable in identifying key associations between insect and plant groups, and monitoring changes in associations can help determine the impacts of threats to ecosystems.

Quantification of the interactions from bulk samples represents an important area of future research if metabarcoding is to be used as an informative ecological monitoring tool. Frequency and types of interactions are needed for undertaking ecosystem network analysis. Whilst it is not possible to calculate frequencies of associations within bulk arthropod samples at this stage, a few approaches can be used to identify species-specific associations. Firstly, metabarcoding individual insects from the bulk samples to identify both arthropods and plants can determine the frequency of interaction (Evans and Kitson, 2020). However, arthropods must first be separated by species or morphotype before DNA extraction. Another approach could be to apply metabarcoding to plant material, such as flowers to detect interactions with pollinators (Thomsen and Sigsgaard, 2019), or pest species from herbivory marks on plants (Kudoh, Minamoto, and Yamamoto, 2020). Lastly, metabarcoding of pollen loads from individual insects (Lucas et al., 2018; Macgregor et al., 2019; Pornon et al., 2016), or brood cells of cavity nesting bees (Gresty et al., 2018) can be used to calculate interaction strength. However, these methods require the processing and sequencing of a large number of samples to reach adequate replication and sampling depth, which can increase costs prohibitively (Evans and Kitson, 2020). Therefore, the reduction of costs associated with processing and sequencing will also be an important part of the future development of these techniques.

2.6 Vertebrate monitoring

Many arthropods feed on vertebrate remains such as blood, faeces, flesh and carcasses, and the use of metabarcoding of the vertebrate DNA in their stomach contents is a valuable monitoring tool to complement traditional vertebrate monitoring (Calvignac-Spencer et al., 2013). This is known as invertebrate-derived DNA (iDNA) and the technique has been used to detect vertebrate taxa in the guts of arthropods such as flies (Calvignac-Spencer et al., 2013; Hoffmann et al., 2018; Lee et al., 2016; Lee, Sing, and Wilson, 2015; Rodgers et al., 2017; Schubert et al., 2014), mosquitoes (Reeves et al., 2016), sand flies (Kocher, de Thoisy, Huguin, et al., 2017) and ticks (Gariepy et al., 2012). While iDNA studies are based on the targeted collection of arthropods known to feed on vertebrates specifically, bulk arthropod samples contain a mix of individuals across many taxonomic groups of arthropods. This means that by applying vertebrate-specific primer sets to bulk sample metabarcoding, there is an opportunity to understand both vertebrate and invertebrate diversity simultaneously. For example, the application of vertebrate metabarcoding primers to bulk arthropod samples collected in tropical forests was used to detect 32 vertebrate taxa across birds, mammalian, and amphibian groups (Lynggaard et al., 2019). This represents a promising methodology for non-invasive sampling of vertebrate diversity.

There are some technical considerations that need to be made when using metabarcoding to detect vertebrate taxa in DNA extracted from bulk arthropod samples. Firstly, the collection season of bulk arthropods plays an essential role in the success of vertebrate detection, as some vertebrate-feeding arthropods may be more abundant at different seasons (Lynggaard et al., 2019). The taxon-specificity of metabarcoding primers is another important consideration for detecting vertebrate species. Whilst metabarcoding primers can be selected to target only specific taxonomic groups (such as birds, fish, or mammals), some 'universal' primer sets may have biases through preferential amplification of one taxonomic group over another. For example, using a mammal primer set in addition to a general vertebrate primer set has been shown to reveal more vertebrate taxa in bulk arthropod or soil samples (Leempoel, Hebert, and Hadly, 2020; Lynggaard et al., 2019). Additionally, because of the high number of arthropods in bulk samples, amplifying low levels of vertebrate DNA may be challenging. This is a common concern in environmental DNA (eDNA) studies that rely on detecting trace amounts of DNA in substrates (Murray, Coghlan, and Bunce, 2015) and some eDNA methods can be used to optimise the amplification of trace vertebrate DNA in bulk arthropod samples. For instance, this can include the inclusion of multiple PCR replicates per sample to ensure PCR stochasticity does not impact the detection of vertebrate diversity and the sequencing of positive and negative PCR controls to validate bioinformatic filtering thresholds and avoid the detection of false positives or negatives (Alberdi et al., 2018; Lynggaard et al., 2019).

2.7 Studying vector-borne disease

By using metabarcoding to detect vertebrates from bulk arthropod samples, it is also possible to study the transmission of vector-borne diseases (Kocher, de Thoisy, Catzeflis, et al., 2017; Reeves et al., 2016). Many emerging infectious diseases have a zoonotic origin and can be traced back to wildlife (Jones et al., 2008; Kent, 2009; Onwugamba et al., 2018). Arthropods that feed on vertebrates (such as mosquitoes, ticks and flies) can transmit pathogens when they feed on infected hosts and then on uninfected hosts (Maldonado and Centeno, 2003). Therefore, the study of these arthropods can be a valuable source of wildlife infectious disease information (Gogarten et al., 2019; Hoffmann et al., 2016), including host-use patterns (Brugman et al., 2017; West et al., 2020). Whilst high-throughput sequencing has been used to study diet and pathogen information from either individual or small pools of arthropods (Gogarten et al., 2019; Hoffmann et al., 2016; Junqueira et al., 2017; Onwugamba et al., 2018), the application of metabarcoding to detect pathogenic nucleic acids has, to our knowledge, not been applied to bulk arthropod samples.

The capability of metabarcoding to monitor the arthropods, vertebrates and pathogens from one bulk arthropod sample provides a valuable addition to understanding wildlife infectious disease spread as the microorganism diversity can be detected along with all plausible hosts and vectors (Hoffmann et al., 2016). As detection rates for pathogenic nucleic acids from arthropods are considerably lower than other sources (such as faeces) (Hoffmann et al., 2016), metabarcoding methodologies in disease transmission studies will need to be carefully undertaken to avoid possible false positive or negative results. Some studies have used shotgun sequencing to reveal pathogen loads in arthropods (Junqueira et al., 2017; Ng et al., 2011; Shi et al., 2015), which can similarly be employed to bulk arthropod samples to find viral or bacterial pathogens. As technology is becoming more portable, for example, the MinION sequencer (Oxford Nanopore Technologies, Oxford, United Kingdom), the capability is being developed to diagnose pathogens in the field (Donoso and Valenzuela, 2018). In the same way, bulk arthropod samples could be used in the future as a tool to diagnose and study vector-borne infectious diseases, possibly in real-time.

2.8 Biodiversity across the tree of life – a multimarker approach

As arthropods interact with many different groups of organisms in the environment, it is possible to use metabarcoding of bulk arthropods for more than just arthropod diversity assessment. Whilst most studies have approached the use of bulk arthropod sample metabarcoding with a single goal in mind (e.g., to detect arthropod or vertebrate diversity from bulk samples separately), the power of metabarcoding is that using a multi-marker approach, targeting several taxonomic groups, can be used to study biodiversity across the tree of life. This is a significant benefit over non-molecular surveying techniques as it reduces the effort that is needed for biodiversity surveying across multiple taxonomic groups.

Whilst sequencing costs associated with metabarcoding using multiple marker sets can be relatively high, several factors can make this approach more cost-effective than non-molecular biodiversity surveys. First, collecting one sample type (bulk arthropod samples) reduces the time spent in the field instead of different trapping and monitoring methods that need to be undertaken to survey across multiple taxonomic groups (e.g., camera trapping, bird surveys, pitfall trapping, etc.). Second, there is higher availability of researchers trained in molecular laboratory skills than those with specialised taxonomic expertise, meaning that processing large numbers of samples is more feasible through metabarcoding (Ji et al., 2013). Third, the lowered cost in person-hours compared to field surveying multiple groups is preferable in studies with limited funding available for biodiversity research. Lastly, molecular data is more accessible to archive and share (Jarman, Berry, and Bunce, 2018), which is beneficial to assess changes in biodiversity over time. The benefits of molecular biomonitoring techniques over non-molecular techniques are more apparent when looking at specific examples of how a multi-marker approach can be directly applied, such as when examining insect-dispersed fungal pathogens.

Certain arthropod groups are known to play a mutually beneficial role in the dispersal of pathogenic fungi through ecosystems (e.g., phytopathogenic fungi weakening host tree, thereby increasing chances of successful attack by bark beetles) (Paine, Raffa, and Harrington, 1997). There are two predominant ways to study fungi-insect interactions. With trees, for example, mature trees need to be felled for the collection of wood samples which then must be examined for insect marks. These insect marks then need to be used to determine insect species. Additionally, insects can be collected directly from the felled trees and can be sampled for fungal isolates, either morphologically or molecularly to determine insect species (Müller et al., 2002; Persson, Ihrmark, and Stenlid, 2011; Strid et al., 2014). This process can be sped up through bulk arthropod metabarcoding with primers targeting both arthropods and fungal DNA, allowing for the identification of the arthropod communities and their associated fungi (K. E. Miller et al., 2016), decreasing the costs of biodiversity assessments as compared to morphological identification and single-species barcoding.

A consideration with a multi-marker approach though, is the secondary or even tertiary transport of eDNA between species which cannot always be accounted for. Much like with the issues surrounding assessing animal-plant interactions with bulk arthropod sampling, it is difficult to directly link DNA sequences between organisms. For example, an individual arthropod may bring with it eDNA from a multitude of sources, such as from other arthropods, plants, or even airborne eDNA (Johnson, Cox, and Barnes, 2019; Lynggaard et al. 2022). This can mean that it can be difficult to distinguish ‘real’ interactions from ‘accidental’ interactions. Therefore, caution needs to be applied when interpreting the broader networks from bulk samples, and potentially the application of a ‘site-level’ controls for background eDNA signal may be beneficial in accounting for the secondary transport of eDNA.

2.9 Current barriers for bulk arthropod sample metabarcoding studies

As new applications and technological advances are adapted for metabarcoding studies, there are still barriers to the widespread implementation of metabarcoding of bulk arthropod samples in biomonitoring. Here we identify four main barriers that need to be taken into consideration when using bulk arthropod metabarcoding and represent areas for further development: validation of detections from iDNA data, non-destructive DNA extraction, abundance quantification, the validation of detections from iDNA data, and data accessibility (including reference databases).

One of the methods that have been used to extract DNA from bulk samples is the physical homogenisation of samples, which is destructive and leads to the loss of valuable morphological information (Buchner, Haase, and Leese, 2021; Gibson et al., 2014; Gillett et al., 2014; Morinière et al., 2016; Yu et al., 2012). Alternative non-destructive DNA extraction approaches have been developed where whole bulk arthropod samples are digested in a digestion buffer that is used for extractions (Batovska et al., 2021; Carew et al., 2018; Nielsen et al., 2019), or where the preservative ethanol from bulk arthropod samples is filtered and DNA then extracted from the filter (Martins et al., 2019; Zizka et al., 2019). Both these methods retain the structural integrity of the arthropod specimens in the bulk sample which has several advantages. First, they allow for individual species to be re-sequenced for creating reference data. This is particularly useful when insects with underrepresented reference barcode databases are detected in bulk arthropod samples, as it allows for reference databases to be populated on an *ad hoc* basis. Although, the success of extracting from these specimens the second time around may be limited (Carew et al., 2018). Second, it allows for more detailed identification of arthropods, as specimens can be morphologically identified after such non-destructive DNA extraction. This is especially useful if metabarcoding does not yield detailed enough taxonomic identifications and species need to be verified morphologically. Last, abundance quantification is possible when samples are intact and specimens can be physically counted. However, this is dependent on the taxon, and it is likely that softer bodied arthropods (such as Dipterans or Hemipterans) may be more fragile and morphological diagnostic features may be damaged when digestions of whole bulk samples are carried out (Nielsen et al., 2019). There is also an increased likelihood that the more soft-bodied arthropods are detected from the non-destructive extraction as their tissues would be easier to digest which can skew diversity estimates (Carew et al., 2018). As such, when using a non-destructive extraction protocol, it is essential that results are interpreted with this in mind.

Estimating species abundance is an integral part of conducting biomonitoring studies (van Klink et al., 2020), but it is difficult to collect true species abundance data from metabarcoding results. This is because sequence abundance can be skewed by many factors. Firstly, detection skews can be influenced by the effect of unequal biomass in mixed samples, as higher biomass can result in more sequences of a certain species. Nevertheless, this does not necessarily translate to a higher abundance of that species in a sample (Elbrecht et al., 2017). Secondly, PCR introduces biases that can skew abundance estimates and misrepresent true composition in samples (Fonseca, 2018). This is because PCR conditions (for example, number of PCR cycles, number of PCR replicates, or choice of polymerase) can have a large influence on the abundance of sequences generated based on polymerase preferential binding or PCR stochasticity (Alberdi et al., 2018; Nichols et al., 2018). Therefore, rank abundance (or relative species abundance) is often used as a surrogate to true abundance in environmental DNA metabarcoding studies (Hänfling et al., 2016; Klymus, Marshall, and Stepien, 2017; Richardson et al., 2015). Rank abundance measures are based on giving an estimate of abundance of species based on the repeated occurrence of taxa in site replicates. There are some recent developments that may improve species abundance estimation from metabarcoding bulk samples in the future. The development of SPIKEPIPE, a pipeline used to account for stochasticity through DNA spike-ins and filtering thresholds, can be used to quantify abundance of a single species across bulk arthropod samples based on DNA mass (Ji et al., 2019). In a similar way, the use of digital-droplet PCR techniques to detect gene copy numbers and hence number of cells in a sample (Yarimizu et al., 2021) can be used to estimate abundance of an arthropod species in a bulk sample if there is adequate calibration. Although, at this stage the utility of digital-droplet PCR for estimating species abundance in bulk samples is speculative and further research is required.

The use of iDNA has become a valuable tool in complementing biomonitoring, allowing for more than just arthropod biodiversity to be documented from bulk arthropod traps. However, there is little known about the confounding factors that can affect iDNA detections from bulk arthropod samples, and how this can influence the interpretation of results. Most studies investigating iDNA recovery are limited to individual arthropods from targeted collections (Calvignac-Spencer et al., 2013; Schubert et al., 2014). However, future studies investigating how complex, mixed arthropod samples interfere with the DNA quality of iDNA will be needed to interpret results from such datasets (Krehenwinkel et al., 2017). This can be achieved by looking at how sample storage or collection impacts iDNA detections. Additionally, there is scope to delve further into how various extraction methods, PCR amplification techniques and sequencing chemistry will impact iDNA biodiversity detected through bulk samples.

Another barrier to the uptake of metabarcoding in terrestrial arthropod monitoring applications has been the availability and accessibility of data. Accessible historical information of biodiversity will be essential to establish biodiversity baselines and allow the predictions of future environmental health (Jarman et al., 2018). Furthermore, there is a need from a conservation management perspective for a global repository of biodiversity information to be made available freely (Stephenson and Stengel, 2020). As metabarcoding data is easy to archive and share, there will be considerable benefits to the maintenance and collation of global metabarcoding survey information for future uses. However, the value of a biodiversity database with metabarcoding data is only as useful as the quality and depth of detail of this data. Although MOTU information is informative, taxonomic information is still important for biodiversity assessment, and assigning taxonomic identities to MOTUs is only possible with well populated barcode reference databases. While DNA barcode reference databases are continually growing, coverage within particular families and genera are typically low (Weigand et al., 2019), particularly for arthropods (Pompanon and Samadi, 2015; Yu et al., 2012), and for some geographic regions such as the tropics (J. A. Miller et al., 2014). Part of this absence of barcode reference data is due to lack of taxonomic information for these groups and regions (Cardoso et al., 2011). It becomes imperative that the status of both flora and fauna in various regions worldwide are known for future and current conservation of species and habitats, especially in the face of global biodiversity decline (Wagner, 2020). Currently, there has been considerable effort dedicated to populating databases across Europe (Weigand et al., 2019), but this needs to extend across the globe – especially in regions where biodiversity is under threat. Techniques like genome skimming can be used to generate large amounts of DNA references and may be a solution to overcoming these database gaps, especially because reliable reference barcodes can be obtained using historical and preserved specimens (Alsos et al., 2020; Margaryan et al., 2021; Nevill et al., 2020; Wu et al., 2020). Further, it is not only important that these reference databases grow, but they also need to be maintained indefinitely for long-term baseline monitoring.

2.10 Perspectives

The power and precision of metabarcoding makes it a valuable tool in terrestrial ecological applications. Applying metabarcoding to bulk arthropod collections allows for more than simple biodiversity assessments, instead, the sensitivity and depth of biodiversity information retrieved through metabarcoding has opened the door to many ways to monitor ecosystems and the network of trophic interactions that underpin them. We consider that metabarcoding will become commonplace amongst ecological studies to comprehensively undertake biodiversity surveying. Here we will discuss some perspectives for future directions of the field.

One of the major benefits that have been discussed for metabarcoding in ecological studies has been in long-term monitoring (Bálint et al., 2018); particularly because data can be generated consistently across time across many taxa simultaneously. Time-series using morphological identification of arthropods has previously been used to assess long-term changes in insect fauna in response to environmental changes (Borges et al., 2020; Matthews et al., 2019; Seibold et al., 2019; Thomsen et al., 2016). The application of metabarcoding to such studies would advance the analysis of drivers and impacts of environmental changes - allowing for the assessment of how variation between and within species changes over time.

Additionally, the improvements in sequencing technologies and capabilities of modern computers have allowed for in-depth analysis of large volumes of DNA sequences relatively quickly. With the sheer volume of data generated from metabarcoding datasets, there is potential for integration with machine learning to elevate analysis (Makiola et al., 2020). For example, in the study by Cordier, Forster and Dufresne (2018), supervised machine learning could predict the status of aquatic ecosystems using metabarcoding data from soil sediment across a bacterial and eukaryotic marker set. This same technology used with bulk arthropod metabarcoding datasets could be trained for evaluating terrestrial biodiversity and ecosystem networks (Bohan et al., 2017). Incorporating machine learning and predictive algorithms can lead to a deeper understanding of the relationships between arthropods and the functionality of ecosystems (Makiola et al., 2020). Further, incorporating time-series data, a multi-marker approach, and machine learning would elevate the value of biomonitoring using DNA based methods. Investing into the development of such techniques would give the unprecedented ability to predict trends in ecosystem health from a whole ecosystem perspective (Bohan et al., 2017).

In such a short period, the field of bulk arthropod sample metabarcoding has come a long way. As the field continues to develop, coupled with the reduction in costs and technology advancements, the potential of using bulk arthropod metabarcoding as a much needed ‘game-changer’ will be more fully realised. However, the uptake of metabarcoding whole ecosystem diversity through bulk arthropod collections is subject to practitioners embracing (and integrating) the metabarcoding technology and tools, which presents its own set of challenges. The future of arthropod monitoring will always lie in a combination of approaches, both molecular and morphological. However, the large increase in studies confirming the benefits of metabarcoding suggests that these DNA-based tools will play an increasingly important role in global biological monitoring of terrestrial ecosystems.

2.11 References

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged

Adams, C. I. M., Knapp, M., Gemmell, N. J., Jeunen, G.-J., Bunce, M., Lamare, M. D., Taylor, H. R. (2019). Beyond Biodiversity: Can Environmental DNA (eDNA) Cut It as a Population Genetics Tool? *Genes*, *10*(3). doi:10.3390/genes10030192

Agerbo Rasmussen, J., Nielsen, M., Mak, S. S. T., Döring, J., Klincke, F., Gopalakrishnan, S., Dunn, R. R, Kauer, R., Gilbert, M. T. P. (2021). eDNA-based biomonitoring at an experimental German vineyard to characterize how management regimes shape ecosystem diversity. *Environmental DNA*, *3*(1), 70–82.

Alarcón, R. (2009). Congruence between visitation and pollen-transport networks in a California plant-pollinator community. *Oikos*, *119*(1), 35–44.

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., Bohmann, K. (2018). Scrutinizing key steps for reliable metabaroding of environmental samples. *Methods in Ecology and Evolution*, *9*(1), 134–147.

Alsos, I. G., Lavergne, S., Merkel, M. K. F., Boleda, M., Lammers, Y., Alberti, A., … Coissac, E. (2020). The Treasure Vault Can be Opened: Large-Scale Genome Skimming Works Well Using Herbarium and Silica Gel Dried Material. *Plants*, *9*(4). doi:10.3390/plants9040432

Arribas, P., Andújar, C., Salces-Castellano, A., Emerson, B. C., Vogler, A. P. (2020). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Molecular Ecology*. doi:10.1111/mec.15591

Ashfaq, M., Hebert, P. D. N. (2016). DNA barcodes for bio-surveillance: regulated and economically important arthropod plant pests. *Genome*, *59*(11), 933–945.

Aylagas, E., Borja, Á., Muxika, I., Rodríguez-Ezpeleta, N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecological Indicators*, *95*, 194–202.

Bale, J. S., van Lenteren, J. C., Bigler, F. (2008). Biological control and sustainable food production. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1492), 761–776.

Bálint, M., Pfenninger, M., Grossart, H.-P., Taberlet, P., Vellend, M., Leibold, M. A., … Bowler, D. (2018). Environmental DNA Time Series in Ecology. *Trends in Ecology Evolution*, *33*(12), 945–957.

Bang, C., Faeth, S. H. (2011). Variation in arthropod communities in response to urbanization: Seven years of arthropod monitoring in a desert city. *Landscape and Urban Planning*, *103*(3), 383–399.

Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Guilhaumon, F., Missa, O., … Leponce, M. (2012). Arthropod diversity in a tropical forest. *Science*, *338*(6113), 1481–1484.

Batovska, J., Piper, A. M., Valenzuela, I., Cunningham, J. P., Blacket, M. J. (2021). Developing a non-destructive metabarcoding protocol for detection of pest insects in bulk trap catches. *Scientific Reports*, *11*(1), 7946.

Beentjes, K. K., Speksnijder, A. G. C. L., Schilthuizen, M., Hoogeveen, M., van der Hoorn, B. B. (2019). The effects of spatial and temporal replicate sampling on eDNA metabarcoding. *PeerJ*, *7*, e7335.

Beermann, A. J., Werner, M.-T., Elbrecht, V., Zizka, V. M. A., Leese, F. (2021). DNA metabarcoding improves the detection of multiple stressor responses of stream invertebrates to increased salinity, fine sediment deposition and reduced flow velocity. *The Science of the Total Environment*, *750*, 141969.

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., Slik, J. W. F. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific Reports*, *6*(1), 24965.

Benson, T. J., Dinsmore, J. J., Hohman, W. L. (2007). Responses of Plants and Arthropods to Burning and Disking of Riparian Habitats. *The Journal of Wildlife Management*, *71*(6), 1949–1957.

Bernays, E. A. (1992). Interaction of insects and plants. *Science Progress*, *76*(2), 247–271.

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., Woodward, G. (2017). Next-Generation Global Biomonitoring: Large-scale, Automated Reconstructin of Ecological Networks. *Trends in Ecology and Evolution*, *32*(7), 477–487.

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D.W., Seymour, M., Dumbrell, A. J., Creer, S. (2021). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Authorea Preprints*. doi:10.22541/au.162141261.10649593/v1

Borges, P. A. V., Rigal, F., Ros-Prieto, A., Cardoso, P. (2020). Increase of insular exotic arthropod diversity is a fundamental dimension of the current biodiversity crisis. *Insect Conservation and Diversity*, *13*(5), 508–518.

Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Elbrecht, V., Steinke, D., Ratnasingham, S., … Hebert, P. D. N. (2019). Metabarcoding a diverse arthropod mock community. *Molecular Ecology Resources*, *19*(3), 711–727.

Brugman, V. A., Hernández-Triana, L. M., England, M. E., Medlock, J. M., Mertens, P. P. C., Logan, J. G., … Carpenter, S. (2017). Blood-feeding patterns of native mosquitoes and insights into their potential role as pathogen vectors in the Thames estuary region of the United Kingdom. *Parasites and Vectors*, *10*(1), 163.

Buchner, D., Haase, P., Leese, F. (2021). Wet grinding of invertebrate bulk samples – a scalable and cost-efficient protocol for metabarcoding and metagenomics. *Metabarcoding and Metagenomics*, *5*, e67533.

Burkle, L. A., Alarcón, R. (2011). The future of plant-pollinator diversity: understanding interaction networks across time, space, and global change. *American Journal of Botany*, *98*(3), 528–538.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583.

Calvignac-Spencer, S., Merkel, K., Kutzner, N., Kühl, H., Boesch, C., Kappeler, P. M., … Leendertz, F. H. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Molecular Ecology*, *22*(4), 915–924.

Cardoso, P., Erwin, T. L., Borges, P. A. V., New, Tim R. (2011). The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, *144*(11), 2647–2655.
Carew, M. E., Coleman, R. A., Hoffmann, A. A. (2018). Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ*, *6*, e4980.
Celli, G., Maccagnani, B. (2003). Honey bees as bioindicators of environmental pollution. *Bulletin of Insectology*, *56*(1), 137–139.
Choi, Y. D. (2004). Theories for ecological restoration in changing environment: Toward "futuristic" restoration. *Ecological Research*, *19*(1), 75–81.
Clare, E. L., Chain, F. J. J., Littlefair, J. E., Cristescu, M. E. (2016). The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. *Genome*, *59*(11), 981–990.
Cordier, T., Forster, D., Dufresne, Y. (2018). Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. *Molecular Ecology*. doi:10.1111/1755-0998.12926
Creedy, T. J., Ng, W. S., Vogler, A. P. (2019). Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy. *Ecology and Evolution*, *9*(6), 3105–3116.
Cross, S. L., Bateman, P. W., Cross, A. T. (2020). Restoration goals: Why are fauna still overlooked in the process of recovering functioning ecosystems and what can be done about it? *Ecological Management Restoration*, *21*(1), 4–8.
Cuthbert, R. N., Cunningham, E. M., Crane, K., Dick, J. T. A., Callaghan, A., Coughlan, N. E. (2020). In for the kill: novel biosecurity approaches for invasive and medically important mosquito species. *Management of Biological Invasions, 11*(1): 9—25.
da Rocha, J. R. M., De Almeida, J. R., Lins, G. A., Durval, A. (2010). Insects as indicators of environmental changing and pollution: a review of appropriate species and their monitoring. *Holos Environment*, *10*(2), 250–262.
Dahlsten, D. L., Garcia, R. (1989). *Eradication of exotic pests: analysis with case histories*. Yale University Press.
Davidovitch, L., Stoklosa, R., Majer, J. D., Nietrzeba, A., Whittle, P., Mengersen, K., Ben-Haim, Y. (2009). Info-gap theory and robust design of surveillance for invasive species: the case study of Barrow Island. *Journal of Environmental Management*, *90*(8), 2785–2793.
Day, W. H. (1994). Estimating mortality caused by parasites and diseases of insects: comparisons of the dissection and rearing methods. *Environmental Entomology*. doi:10.1093/ee/23.3.543
Donoso, A., Valenzuela, S. (2018). In-field molecular diagnosis of plant pathogens: recent trends and future perspectives. *Plant Pathology*, *67*(7), 1451–1461.
Doorenweerd, C., San Jose, M., Barr, N., Leblanc, L., Rubinoff, D. (2020). Highly variable COI haplotype diversity between three species of invasive pest fruit fly reflects remarkably incongruent demographic histories. *Scientific Reports*, *10*(1), 6887.
Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*(10), 996–998.
Edgar, R. C., Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476–3482.

Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., … Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, *7*, e7745.

Elbrecht, V., Peinert, B., Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*(17), 6918–6926.

Elbrecht, V., Vamos, E. E., Steinke, D., Leese, F. (2018). Estimating intraspecific genetic diversity from community DNA metabarcoding data. *PeerJ*, *6*, e4644.

Emerson, B. C., Cicconardi, F., Fanciulli, P. P., Shaw, P. J. A. (2011). Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1576), 2391–2402.

Evans, D. M., Kitson, J. J. N. (2020). Molecular ecology as a tool for understanding pollination and other plant–insect interactions. *Current Opinion in Insect Science*, *38*, 26–33.

Fernandes, K., van der Heyde, M., Coghlan, M., Wardell-Johnson, G., Bunce, M., Harris, R., Nevill, P. (2019). Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences. *Restoration Ecology*, *215*, 403.

Floyd, R., Abebe, E., Papert, A., Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology*, *11*(4), 839–850.

Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. *Molecular Ecology Resources*, *18*(5), 923–926.

Foster, J. G., Ploughe, L. W., Akin-Fajiye, M., Singh, J. P., Bottos, E., Van Hamme, J., Fraser, L. H. (2020). Exploring trophic effects of spotted knapweed (*Centaurea stoebe* L.) on arthropod diversity using DNA metabarcoding. *Food Webs*, *24*, e00157.

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications*, *8*(1), 1188.

Gálvez-Reyes, N., Arribas, P., Andujar, C., Emerson, B., Piñero, D., Mastretta-Yanes, A. (2020). Local-scale dispersal constraints promote spatial structure and arthropod diversity within a tropical sky-island. *Authorea Preprints*. doi:10.22541/au.160193334.45224582

Gariepy, T. D., Lindsay, R., Ogden, N., Gregory, T. R. (2012). Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks. *Molecular Ecology Resources*, *12*(4), 646–652.

Geiger, M., Moriniere, J., Hausmann, A., Haszprunar, G., Wägele, W., Hebert, P., Rulik, B. (2016). Testing the Global Malaise Trap Program – How well does the current barcode reference library identify flying insects in Germany? *Biodiversity Data Journal*. doi:10.3897/bdj.4.e10671

Gervan, C. A., Gardner, W. C., Van Hamme, J. D., Higgins, R. J., Bottos, E. M. (2020). *Invertebrate Response to Mine Reclamation (South-Central British Columbia): The Effects of Reclamation Age on Arthropod Assemblages*. Retrieved from http://www.geosciencebc.com/i/pdf/SummaryofActivities2019/Minerals/Sch_Gervan_Minerals_SOA2019.pdf

Ghannem, S., Touaylia, S., Boumaiza, M. (2018). Beetles (Insecta: Coleoptera) as bioindicators of the assessment of environmental pollution. *Human and Ecological Risk Assessment: An International Journal*, *24*(2), 456–464.

Gibson, J., Shokralla, S., Porter, T. M., King, I., van Konynenburg, S., Janzen, D. H., … Hajibabaei, M. (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(22), 8007–8012.

Gillespie, M. A. K., Alfredsson, M., Barrio, I. C., Bowden, J., Convey, P., Coulson, S. J., … Høye, T. T. (2020). Circumpolar terrestrial arthropod monitoring: A review of ongoing activities, opportunities and challenges, with a focus on spiders. *Ambio*, *49*(3), 704–717.

Gillett, C. P. D. T., Crampton-Platt, A., Timmermans, M. J. T. N., Jordal, B. H., Emerson, B. C., Vogler, A. P. (2014). Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). *Molecular Biology and Evolution*, *31*(8), 2223–2237.

Gogarten, J. F., Hoffmann, C., Arandjelovic, M., Sachse, A., Merkel, K., Dieguez, P., … Calvignac-Spencer, S. (2019). Fly-derived DNA and camera traps are complementary tools for assessing mammalian biodiversity. *Environmental DNA*, *9*, 79.

Gresty, C. E. A., Clare, E., Devey, D. S., Cowan, R. S., Csiba, L., Malakasi, P., … Willis, K. J. (2018). Flower preferences and pollen transport networks for cavity-nesting solitary bees: Implications for the design of agri-environment schemes. *Ecology and Evolution*, *8*(15), 7574–7587.

Hackman, K. O., Gong, P., Venevsky, S. (2017). A rapid assessment of landscape biodiversity using diversity profiles of arthropod morphospecies. *Landscape Ecology*, *32*(1), 209–223.

Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., … Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, *25*(13), 3101–3119.

Happe, A.-K., Alins, G., Blüthgen, N., Boreux, V., Bosch, J., García, D., … Mody, K. (2019). Predatory arthropods in apple orchards across Europe: Responses to agricultural management, adjacent habitat, landscape composition and country. *Agriculture, Ecosystems Environment*, *273*, 141–150.

Hardulak, L. A., Morinière, J., Hausmann, A., Hendrich, L., Schmidt, S., Doczkal, D., … Haszprunar, G. (2020). DNA metabarcoding for biodiversity monitoring in a national park: Screening for invasive and pest species. *Molecular Ecology Resources*. doi:10.1111/1755-0998.13212

Hausmann, A., Segerer, A. H., Greifenstein, T., Knubben, J., Morinière, J., Bozicevic, V., … Habel, J. C. (2020). Toward a standardized quantitative and qualitative insect monitoring scheme. *Ecology and Evolution*, *10*(9), 4009–4020.

Hering, D., Feld, C. K., Moog, O., Ofenböck, T. (2006). Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. In M. T. Furse, D. Hering, K. Brabec, A. Buffagni, L. Sandin, P. F. M. Verdonschot (Eds.), *The Ecological Status of European Rivers: Evaluation and Intercalibration of Assessment Methods* (pp. 311–324). Dordrecht: Springer Netherlands.

Herrick, J. E., Schuman, G. E., Rango, A. (2006). Monitoring ecological processes for restoration projects. *Journal for Nature Conservation*, *14*(3–4), 161–171.
Hoffmann, C., Merkel, K., Sachse, A., Rodríguez, P., Leendertz, F. H., Calvignac-Spencer, S. (2018). Blow flies as urban wildlife sensors. *Molecular Ecology Resources*, *18*(3), 502–510.
Hoffmann, C., Stockhausen, M., Merkel, K., Calvignac-Spencer, S., Leendertz, F. H. (2016). Assessing the feasibility of fly based surveillance of wildlife infectious diseases. *Scientific Reports*, *6*, 37952.
Jarman, S. N., Berry, O., Bunce, M. (2018). The value of environmental DNA biobanking for long-term biomonitoring. *Nature Ecology Evolution*, *2*(8), 1192–1193.
Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257.
Ji, Y., Huotari, T., Roslin, T., Schmidt, N. M., Wang, J., Yu, D. W., Ovaskainen, O. (2019). SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes. *Molecular Ecology Resources.* doi:10.1101/533737
Johnson, M. D. (2007). Measuring Habitat Quality: A Review. *The Condor*, *109*(3), 489–504.
Johnson, M. D., Cox, R. D., & Barnes, M. A. (2019). Analyzing airborne environmental DNA: A comparison of extraction methods, primer type, and trap type on the ability to detect airborne eDNA from terrestrial plant communities. *Environmental DNA*, *1*(2), 176-185.
Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990–993.
Junqueira, A. C. M., Ratan, A., Acerbi, E., Drautz-Moses, D. I., Premkrishnan, B. N. V., Costea, P. I., … Schuster, S. C. (2017). The microbiomes of blowflies and houseflies as bacterial transmission reservoirs. *Scientific Reports*, *7*(1), 16324.
Kaartinen, R., Stone, G. N., Hearn, J., Lohse, K., Roslin, T. (2010). Revealing secret liaisons: DNA barcoding changes our understanding of food webs. *Ecological Entomology*, *35*(5), 623–638.
Kent, R. J. (2009). Molecular methods for arthropod bloodmeal identification and applications to ecological and vector-borne disease studies. *Molecular Ecology Resources*, *9*(1), 4–18.
Khosravi, R., Hemami, M.-R. (2019). Identifying landscape species for ecological planning. *Ecological Indicators*, *99*, 140–148.
Klymus, K. E., Marshall, N. T., Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PloS One*, *12*(5), e0177643.
Kocher, A., de Thoisy, B., Catzeflis, F., Valière, S., Bañuls, A., Murienne, J. (2017). iDNA screening: Disease vectors as vertebrate samplers. *Molecular Ecology*, *26*(22), 6478–6486.
Kocher, A., de Thoisy, B., Huguin, M., Valière, S, Zinger, L., Bañuls, A., Murienne, J. (2017). Evaluation of short mitochondrial metabarcodes for the identification of Amazonian mammals. *Methods in Ecology and Evolution*, *8*(10), 1276–1283.

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*. doi:10.1038/s41598-017-17333-x

Kremen, C., Colwell, R. K., Erwin, T. L., Murphy, D. D., Noss, R. F., Sanjayan, M. A. (1993). Terrestrial Arthropod Assemblages: Their Use in Conservation Planning. *Conservation Biology*, *7*(4), 796–808.

Kudoh, A., Minamoto, T., Yamamoto, S. (2020). Detection of herbivory: eDNA detection from feeding marks on leaves. *Environmental DNA*, 81.

Lanzén, A., Lekang, K., Jonassen, I., Thompson, E. M., Troedsson, C. (2016). High-throughput metabarcoding of eukaryotic diversity for environmental monitoring of offshore oil-drilling activities. *Molecular Ecology*, *25*(17), 4392–4406.

Lee, P.-S., Gan, H. M., Clements, G. R., Wilson, J.-J. (2016). Field calibration of blowfly-derived DNA against traditional methods for assessing mammal diversity in tropical forests. *Genome, 59*(11), 1008–1022.

Lee, P.-S., Sing, K.-W., Wilson, J.-J. (2015). Reading Mammal Diversity from Flies: The Persistence Period of Amplifiable Mammal mtDNA in Blowfly Guts (Chrysomya megacephala) and a New DNA Mini-Barcode Target. *PloS One*, *10*(4), e0123871.

Leempoel, K., Hebert, T., Hadly, E. A. (2020). A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. *Proceedings of the Royal Society B, Biological Sciences*. doi: 10.1098/rspb.2019.2353

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, Á., Bruce, K., … Weigand, A. M. (2018). Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived From the DNAqua-Net COST Action. *Advances in Ecological Research*, *58*, 63–99.

Li, F., Peng, Y., Fang, W., Altermatt, F., Xie, Y., Yang, J., Zhang, X. (2018). Application of Environmental DNA Metabarcoding for Predicting Anthropogenic Pollution in Rivers. *Environmental Science Technology*, *52*(20), 11708–11719.

Liebhold, A. M., Berec, L., Brockerhoff, E. G., Epanchin-Niell, R. S., Hastings, A., Herms, D. A., … Yamanaka, T. (2016). Eradication of Invading Insect Populations: From Concepts to Applications. *Annual Review of Entomology*, *61*, 335–352.

Liu, M., Baker, S. C., Burridge, C. P., Jordan, G. J., Clarke, L. J. (2020). DNA metabarcoding captures subtle differences in forest beetle communities following disturbance. *Restoration Ecology*. doi:10.1111/rec.13236

Liu, S., Li, Y., Lu, J., Su, X., Tang, M., Zhang, R., … Zhou, X. (2013). SOAPBarcode: revealing arthropod biodiversity through assembly of Illumina shotgun sequences of PCR amplicons. *Methods in Ecology and Evolution*, *4*(12), 1142–1150.

Lucas, A., Bodger, O., Brosi, B. J., Ford, C. R., Forman, D. W., Greig, C., … de Vere, N. (2018). Floral resource partitioning by individuals within generalised hoverfly pollination networks revealed by DNA metabarcoding. *Scientific Reports*, *8*(1), 5133.

Lynggaard, C., Nielsen, M., Santos-Bay, L., Gastauer, M., Oliveira, G., Bohmann, K. (2019). Vertebrate diversity revealed by metabarcoding of bulk arthropod samples from tropical forests. *Environmental DNA*, *1*(4), 329–341.

Lynggaard, C., Yu, D. W., Oliveira, G., Caldeira, C. F., Ramos, S. J., Ellegaard, M. R., … Bohmann, K. (2020). DNA-Based Arthropod Diversity Assessment in Amazonian Iron Mine Lands Show Ecological Succession Towards Undisturbed Reference Sites. *Frontiers in Ecology and Evolution*, *8*, 426.

Macgregor, C. J., Kitson, J. J. N., Fox, R., Hahn, C., Lunt, D. H., Pocock, M. J. O., Evans, D. M. (2019). Construction, validation, and application of nocturnal pollen transport networks in an agro-ecosystem: a comparison using light microscopy and DNA metabarcoding. *Ecological Entomology*, *44*(1), 17–29.

Macher, J. N., Salis, R. K., Blakemore, K. S., Tollrian, R., Matthaei, C. D., Leese, F. (2016). Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators*, *61*, 159–169.

Makiola, A., Compson, Z. G., Baird, D. J., Barnes, M. A., Boerlijst, S. P., Bouchez, A., … Bohan, D. A. (2020). Key Questions for Next-Generation Biomonitoring. *Frontiers of Environmental Science*. doi:10.3389/fenvs.2019.00197

Maldonado, M. A., Centeno, N. (2003). Quantifying the potential pathogens transmission of the blowflies (Diptera: Calliphoridae). *Memorias Do Instituto Oswaldo Cruz*, *98*(2), 213–216.

Margaryan, A., Noer, C. L., Richter, S. R., Restrup, M. E., Bülow-Hansen, J. L., Leerhøi, F., … Bohmann, K. (2021). Mitochondrial genomes of Danish vertebrate species generated for the national DNA reference database, DNAmark. *Environmental DNA*, *3*(2), 472–480.

Martins, F. M. S., Galhardo, M., Filipe, A. F., Teixeira, A., Pinheiro, P., Paupério, J., … Beja, P. (2019). Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. *Molecular Ecology Resources*, *19*(4), 863–876.

Matthews, T. J., Sadler, J., Carvalho, R., Nunes, R. (2019). Differential temporal beta-diversity patterns of native and non-native arthropod species in a fragmented native forest landscape. *Ecography*. doi:10.1111/ecog.03812

McGeoch, M. A. (1998). The selection, testing and application of terrestrial insects as bioindicators. *Biological Reviews of the Cambridge Philosophical Society*, *73*(2), 181–201.

McGeoch, M. A., Sithole, H., Samways, M. J., Simaika, J. P., Pryke, J. S., Picker, M., … Others. (2011). Conservation and monitoring of invertebrates in terrestrial protected areas. *Koedoe*, *53*(2), 131–143.

Miller, J. A., Miller, J. H., Pham, D.-S., Beentjes, K. K. (2014). Cyberdiversity: improving the informatic value of diverse tropical arthropod inventories. *PloS One*, *9*(12), e115750.

Miller, K. E., Hopkins, K., Inward, D. J. G., Vogler, A. P. (2016). Metabarcoding of fungal communities associated with bark beetles. *Ecology and Evolution*, *6*(6), 1590–1600.

Montgomery, G. A., Belitz, M. W., Guralnick, R. P., Tingley, M. W. (2021). Standards and Best Practices for Monitoring and Benchmarking Insects. *Frontiers in Ecology and Evolution*, *8*, 513.

Moretti, M., Dias, A. T. C., Bello, F., Altermatt, F., Chown, S. L., Azcárate, F. M., … Berg, M. P. (2017). Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology*, *31*(3), 558–567.

Morinière, J., Cancian de Araujo, B., Lam, A. W., Hausmann, A., Balke, M., Schmidt, S., … Haszprunar, G. (2016). Species Identification in Malaise Trap Samples by DNA Barcoding Based on NGS Technologies and a Scoring Matrix. *PloS One*, *11*(5), e0155497.

Müller, M. M., Varama, M., Heinonen, J., Hallaksela, A.-M. (2002). Influence of insects on the diversity of fungi in decaying spruce wood in managed and natural forests. *Forest Ecology and Management*, *166*(1), 165–181.

Murray, D. C., Coghlan, M. L., Bunce, M. (2015). From benchtop to desktop: important considerations when designing amplicon sequencing workflows. *PloS One*, *10*(4), e0124671.

Nevill, P. G., Zhong, X., Tonti-Filippini, J., Byrne, M., Hislop, M., Thiele, K., … Small, I. (2020). Large scale genome skimming from herbarium material for accurate plant identification and phylogenomics. *Plant Methods*, *16*, 1.

Ng, T. F. F., Willner, D. L., Lim, Y. W., Schmieder, R., Chau, B., Nilsson, C., … Breitbart, M. (2011). Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PloS One*, *6*(6), e20579.

Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., … Shapiro, B. (2018). Minimizing polymerase biases in metabarcoding. *Molecular Ecology Resources*. doi:10.1111/1755-0998.12895

Nielsen, M., Gilbert, M. T. P., Pape, T., Bohmann, K. (2019). A simplified DNA extraction protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity. *Environmental DNA*. doi:10.1002/edn3.16

Oliver, I., Beattie, A. J. (1996). Designing a Cost-Effective Invertebrate Survey: A Test of Methods for Rapid Assessment of Biodiversity. *Ecological Applications*, *6*(2), 594–607.

Oliverio, A. M., Gan, H., Wickings, K., Fierer, N. (2018). A DNA metabarcoding approach to characterize soil arthropod communities. *Soil Biology Biochemistry*, *125*, 37–43.

Onwugamba, F. C., Fitzgerald, J. R., Rochon, K., Guardabassi, L., Alabi, A., Kühne, S., … Schaumburg, F. (2018). The role of "filth flies" in the spread of antimicrobial resistance. *Travel Medicine and Infectious Disease*, *22*, 8–17.

Paine, T. D., Raffa, K. F., Harrington, T. C. (1997). Interactions among Scolytid bark beetles, their associated fungi, and live host conifers. *Annual Review of Entomology*, *42*, 179–206.

Paknia, O., Rajaei, H. S., Koch, A. (2015). Lack of well-maintained natural history collections and taxonomists in megadiverse developing countries hampers global biodiversity exploration. *Organisms Diversity and Evolution*, *15*(3), 619–629.

Peck, S. L., Mcquaid, B., Campbell, C. L. (1998). Using ant species (Hymenoptera: Formicidae) as a biological indicator of agroecosystem condition. *Environmental Entomology*. doi:10.1093/ee/27.5.1102

Pedro, P. M., Piper, R., Bazilli Neto, P., Cullen, L., Jr, Dropa, M., Lorencao, R., … Turati, D. T. (2017). Metabarcoding Analyses Enable Differentiation of Both Interspecific Assemblages and Intraspecific Divergence in Habitats With Differing Management Practices. *Environmental Entomology*, *46*(6), 1381–1389.

Persson, Y., Ihrmark, K., Stenlid, J. (2011). Do bark beetles facilitate the establishment of rot fungi in Norway spruce? *Fungal Ecology*, *4*(4), 262–269.

Pik, A. J., Oliver, I., Beattie, A. J. (1999). Taxonomic sufficiency in ecological studies of terrestrial invertebrates. *Australian Journal of Ecology*, *24*(5), 555–562.

Piper, A. M., Batovska, J., Cogan, N. O. I., Weiss, J., Cunningham, J. P., Rodoni, B. C., Blacket, M. J. (2019). Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *GigaScience*, *8*(8). doi:10.1093/gigascience/giz092

Pocock, M. J. O., Evans, D. M., Memmott, J. (2012). The robustness and restoration of a network of ecological networks. *Science*, *335*(6071), 973–977.

Pompanon, F., Samadi, S. (2015). Next generation sequencing for characterizing biodiversity: promises and challenges. *Genetica*, *143*(2), 133–138.

Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., … Andalo, C. (2016). Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific Reports*, *6*, 27282.

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., Kunin, W. E. (2010). Global pollinator declines: trends, impacts and drivers. *Trends in Ecology Evolution*, *25*(6), 345–353.

Pozebon, H., Marques, R. P., Padilha, G., O Neal, M., Valmorbida, I., Bevilaqua, J. G., … Arnemann, J. A. (2020). Arthropod Invasions Versus Soybean Production in Brazil: A Review. *Journal of Economic Entomology*, *113*(4), 1591–1608.

Prendergast, J. R., Quinn, R. M., Lawton, J. H., Eversham, B. C., Gibbons, D. W. (1993). Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature*, *365*(6444), 335–337.

Ratnasingham, S., Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes*. doi:10.1111/j.1471-8286.2006.01678.x

Reeves, L. E., Holderman, C. J., Gillett-Kaufman, J. L., Kawahara, A. Y., Kaufman, P. E. (2016). Maintenance of host DNA integrity in field-preserved mosquito (Diptera: Culicidae) blood meals for identification by DNA barcoding. *Parasites Vectors*, *9*(1), 503.

Richardson, R. T., Lin, C., Quijia, J. O., Riusech, N. S., Goodell, K., Johnson, R. M. (2015). *Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach* [Data set]. doi:10.3732/apps.1500043

Rodgers, T. W., Xu, C. C. Y., Giacalone, J., Kapheim, K. M., Saltonstall, K., Vargas, M., … Jansen, P. A. (2017). Carrion fly-derived DNA metabarcoding is an effective tool for mammal surveys: Evidence from a known tropical mammal community. *Molecular Ecology Resources*, *17*(6), e133–e145.

Rosenberg, D. M., Danks, H. V., Lehmkuhl, D. M. (1986). Importance of insects in environmental impact assessment. *Environmental Management*, *10*(6), 773–783.

Salkova, D., Panayotova-Pencheva, M. (2016). Honey bees and their products as indicators of environmental pollution: a review. *Agricultural Science and Technology*, *8*(3), 175-182.

Sánchez-Bayo, F., Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, *232*, 8–27.

Schubert, G., Stockhausen, M., Hoffmann, C., Merkel, K., Vigilant, L., Leendertz, F. H., Calvignac-Spencer, S. (2014). Targeted detection of mammalian species using carrion fly-derived DNA. *Molecular Ecology Resources.* doi:10.1111/1755-0998.12306

Segat, J. C., Figueiredo Vasconcellos, R. L., Paiva Silva, D., Baretta, D., Cardoso, E. J. B. N. (2017). Ants as indicators of soil quality in an on-going recovery of riparian forests. *Forest Ecology and Management*, *404*, 338–343.

Seibold, S., Gossner, M. M., Simons, N. K., Blüthgen, N., Müller, J., Ambarlı, D., … Weisser, W. W. (2019). Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature*, *574*(7780), 671–674.

Shi, C., Liu, Y., Hu, X., Xiong, J., Zhang, B., Yuan, Z. (2015). A metagenomic survey of viral abundance and diversity in mosquitoes from Hubei province. *PloS One*, *10*(6), e0129845.

Sigsgaard, E. E., Jensen, M. R., Winkelmann, I. E., Møller, P. R., Hansen, M. M., Thomsen, P. F. (2019). Population-level inferences from environmental DNA-Current status and future perspectives. *Evolutionary Applications*, *13*(2), 245–262.

Sire, L., Rougerie, R., Larrieu, L., Bouget, C., Courtial, B., Bézier, A., … Vaamonde, C. L. (2017). Quantifying the impact of tree-diebacks and salvage logging on mountain forest biodiversity using metabarcoding. *4. Colloque de Génomique Environnementale*, Marseille, France.

Smith, M. A., Rodriguez, J. J., Whitfield, J. B., Deans, A. R., Janzen, D. H., Hallwachs, W., Hebert, P. D. N. (2008). Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(34), 12359–12364.

Smith, M. A., Wood, D. M., Janzen, D. H., Hallwachs, W., Hebert, P. D. N. (2007). DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 4967–4972.

Stephenson, P. J., Stengel, C. (2020). An inventory of biodiversity data sources for conservation monitoring. *PloS One*, *15*(12), e0242923.

Strid, Y., Schroeder, M., Lindahl, B., Ihrmark, K., Stenlid, J. (2014). Bark beetles have a decisive impact on fungal communities in Norway spruce stem sections. *Fungal Ecology*, *7*, 47–58.

Taberlet, P., Coissac, E., Hajibabaei, M., Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793.

Terlizzi, A., Bevilacqua, S., Fraschetti, S., Boero, F. (2003). Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. *Marine Pollution Bulletin*, *46*(5), 556–561.

Thomsen, P. F., Jørgensen, P. S., Bruun, H. H., Pedersen, J., Riis-Nielsen, T., Jonko, K., … Karsholt, O. (2016). Resource specialists lead local insect community turnover associated with temperature - analysis of an 18-year full-seasonal record of moths and beetles. *The Journal of Animal Ecology*, *85*(1), 251–261.

Thomsen, P. F., Sigsgaard, E. E. (2019). Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. *Ecology and Evolution*, *9*(4), 1665–1679.

Tiede, Y., Schlautmann, J., Donoso, D. A., Wallis, C. I. B., Bendix, J., Brandl, R., Farwig, N. (2017). Ants as indicators of environmental change and ecosystem processes. *Ecological Indicators*, *83*, 527–537.

Turon, X., Antich, A., Palacín, C., Praebel, K., Wangensteen, O. S. (2019). From metabarcoding to metaphylogeography: separating the wheat from the chaff. *Ecological Applications*, *30*(2), e02036.

Tylianakis, J., Veddeler, D., Lozada, T., Lopez, R. M. (2004). Biodiversity of land-use systems in coastal Ecuador and bioindication using trap-nesting bees, wasps, and their natural enemies. *Lyonia*, *6*(2), 7-15.

van der Heyde, M., Bunce, M., Dixon, K. W., Fernandes, K., Majer, J., Wardell-Johnson, G., White, N. E., & Nevill, P. (2022). Evaluating restoration trajectories using DNA metabarcoding of ground-dwelling and airborne invertebrates and associated plant communities. *Molecular Ecology, 31*, 2172– 2188. doi:10.1111/mec.16375

van der Heyde, M., Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., Nevill, P. (2020). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Molecular Ecology Resources*, *20*(3). doi:10.1111/1755-0998.13148

van Jaarsveld, A. S., Freitag, S., Chown, S. L., Muller, C., Koch, S., Hull, H., … Scholtz, C. H. (1998). Biodiversity assessment and conservation strategies. *Science*, *279*(5359), 2106–2108.

van Klink, R., Bowler, D. E., Gongalsky, K. B., Swengel, A. B., Gentile, A., Chase, J. M. (2020). Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science*, *368*(6489), 417–420.

Wagner, D. L. (2020). Insect Declines in the Anthropocene. *Annual Review of Entomology*, *65*, 457–480.

Wares, J. P., Pappalardo, P. (2015). Can Theory Improve the Scope of Quantitative Metazoan Metabarcoding? *Diversity*, *8*(1), 1.

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., … Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *The Science of the Total Environment*, *678*, 499–524.

West, R. G., Mathias, D. K., Day, J. F., Acevedo, C., Unnasch, T. R., Burkett-Cadena, N. D. (2020). Seasonal Changes of Host Use by Culiseta melanura (Diptera: Culicidae) in Central Florida. *Journal of Medical Entomology*, *57*(5), 1627–1634.

Wettstein, W., Schmid, B. (1999). Conservation of arthropod diversity in montane wetlands: effect of altitude, habitat quality and habitat fragmentation on butterflies and grasshoppers. *The Journal of Applied Ecology*, *36*(3), 363–373.

Whittle, P. J. L., Stoklosa, R., Barrett, S., Jarrad, F. C., Majer, J. D., Martin, P. A. J., Mengersen, K. (2013). A method for designing complex biosecurity surveillance systems: detecting non-indigenous species of invertebrates on Barrow Island. *Diversity Distributions*, *19*(5–6), 629–639.

Wu, C.-S., Sudianto, E., Hung, Y.-M., Wang, B.-C., Huang, C.-J., Chen, C.-T., Chaw, S.-M. (2020). Genome skimming and exploration of DNA barcodes for Taiwan endemic cypresses. *Scientific Reports*, *10*(1), 20650.

Yarimizu, K., Sildever, S., Hamamoto, Y., Tazawa, S., Oikawa, H., Yamaguchi, H., … Nagai, S. (2021). Development of an absolute quantification method for ribosomal RNA gene copy numbers per eukaryotic single cell by digital PCR. *Harmful Algae*, *103*, 102008.

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., Ding, Z. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, *3*(4), 613–623.

Zhang, K., Lin, S., Ji, Y., Yang, C., Wang, X., Yang, C., … Yu, D. W. (2016). Plant diversity accurately predicts insect diversity in two tropical landscapes. *Molecular Ecology*, *25*(17), 4407–4419.

Zizka, V. M. A., Leese, F., Peinert, B., Geiger, M. F. (2019). DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method 1. *Genome*, *62*(3), 122–136.

CHAPTER 3

METABARCODING REVEALS THE IMPACTS OF EXPLORATION INFRASTRUCTURE ON ARTHROPOD COMMUNITIES AT MINE SITES

3.1 Preface

This chapter presents a case study for one of the most common applications of invertebrate DNA metabarcoding, biomonitoring invertebrates. Here, metabarcoding was applied to study ground-dwelling arthropods in a study system where they had yet to be studied. Particularly of interest was the sensitivity of the invertebrates to detect changes in the system and how they would respond to different mining exploration infrastructure.

Mining exploration infrastructure is rarely considered as part of the disturbance footprint of mining operations. Typically, this is because these disturbance types require minimal clearing of land, whilst the surrounding vegetation remains untouched. However, clearing even small areas cause edge effects in the landscapes and create microclimates between the border of cleared land and intact vegetation; increasing the light, heat, and access to these areas by pests or weed species. Within the three sites investigated in this chapter, the arthropod communities had yet to be studied and therefore this study presented an opportunity to be the first to document the impacts of exploration infrastructure on these communities.

This chapter consists of a manuscript that is currently being prepared for submission to a peer reviewed journal.

3.1.1 Acknowledgements

This project was supported by the Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration. Computational resources were provided by the Pawsey Supercomputing Centre, funding from the Australian Government, and the Government of Western Australia. KF was partially supported by funding from the Food Agility CRC. We would like to thank Marti J Anderson and Adam H Smith for their helpful advice on statistical analysis for this study. We also thank the mining companies, Sinosteel Midwest Corporation Ltd, Karara Mining Ltd, and Mount Gibson Iron Ltd for access to sites for sampling.

3.1.2 Data Accessibility

Sample information, zOTU Tables, and de-convoluted unique sequences are available at: 10.5281/zenodo.6395458

3.1.3 Author Contributions

SC-I, and PN conceptualised the study. SC-I collected the samples and processed them. KF conducted the experimentation and analysis. BJS assisted with statistical analysis. KF wrote the original draft of the manuscript. All authors contributed to review and editing of the manuscript.

Metabarcoding reveals the impacts of exploration infrastructure on arthropod communities at mine sites.

Kristen Fernandes[1,2,3], Sophia Clark-Ioannou[4], Benjamin J Saunders[4], Philip W Bateman[5], Michael Bunce[1,6], Paul Nevill[1,7]

1 Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

2 Section for Molecular Ecology and Evolution, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen K, Denmark

3 Food Agility CRC Ltd, 175 Pitt St Sydney, NSW 2000, Australia

4 School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

5 Behavioural Ecology Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

6 Environmental Science and Research (ESR), Kenepuru, Porirua 5022, New Zealand.

7 ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

3.2 Abstract

For most mining projects, regular monitoring of impacts on the environment are typically measured solely for the main operation footprint and not the supporting network of exploration infrastructure. Exploration infrastructure includes roads or drill pads, where small areas are cleared of native vegetation and the surrounding habitat is left undisturbed. However, this exploration infrastructure contributes to fragmentation and edge effects. Here we use DNA metabarcoding to study the impacts of mining exploration infrastructure on the ground-dwelling arthropod communities found on Banded Ironstone Formations (BIFs) in the Midwest region of Western Australia. Further, we examined whether the two indicator species groups, Formicidae and Collembola, reflect the impacts on overall arthropod communities. The study was conducted at three mine sites, each with three infrastructure types: maintained tracks, unmaintained tracks, and drill pads. We collected ground-dwelling arthropods and recorded habitat characteristics (e.g., coarse woody debris, soil depth) along transects extending 100 m away from the disturbance into remnant vegetation. We found that all three infrastructure types impacted the habitat characteristics in the same way, with the changes in the habitat extending out to 10 m into the remnant vegetation. This trend was similarly observed in the overall arthropod communities and Formicidae composition, but not Collembola, and also not for zOTU richness. Among all three groups of arthropods (overall arthropods, Formicidae and Collembola), we observed a clear distinction between the exploration infrastructure types at each site, indicating the homogenisation of these arthropod communities around these disturbances. This effect potentially extends more than 100 m into remnant vegetation. This suggests that Formicidae can detect fine-scale differences in habitat alterations to a larger extent than Collembola. Here, we demonstrate the utility of DNA metabarcoding to document fine-scale impacts on arthropod communities. Our results highlight the need to consider different types of disturbances associated with mining activity in rehabilitation and build research into managing the impacts associated with the extensive legacy of unrestored mining exploration infrastructure in Australian ecosystems.

3.3 Introduction

The main drivers of species extinction across the globe are human activities that cause degradation and fragmentation of landscapes (Ellis et al., 2010) of which some are more impactful than others (Redondo-Vega et al., 2017). Although mining operations only occupy 0.3% of land globally (Hooke et al., 2012), they can have more impacts on natural habitats than other visible and wide-ranging activities such as agriculture and urbanisation (Redondo-Vega et al., 2017). This is because mining activities will often target geologically unique formations in the landscape that are surrounded by pristine and endemic ecosystems and can affect these ecosystems in various ways (Guijón et al., 2011; Redondo-Vega et al., 2017; Tibbett, 2015; Wycisk et al., 2009).

In addition to direct impacts on natural habitats, mining poses significant consequences and long-term effects that extend beyond the initial footprint through habitat fragmentation, which can result from mining operations themselves or from the associated infrastructure linked with mining exploration. The ecological impacts of habitat fragmentation are complex and pervasive, affecting population dynamics and community structure for both flora and fauna (Bolger et al., 2000; Fahrig, 2003). For example, fragmentation increases the number of edges around habitats. These edges naturally occur between two habitats where there is a varying level of light, humidity, temperature, and hence vegetation community structure and species (Ries & Sisk, 2004; Santos et al., 2008). These biotic and abiotic changes between the contrasting ecotones are known as ‘edge effects’ (Ewers & Didham, 2006; Murcia, 1995). With the creation of more edges in anthropogenically-altered landscapes, the increased presence of microclimates with variable light availability, temperature and humidity also allows the invasions of weeds and other alien species and can change the overall community structure, and potentially, the functionality of ecosystems (Goosem, 2012; Pohlman et al., 2009). The extent to which the impacts of the edge effect can penetrate adjoining habitats can extend from a few metres to several hundreds of metres into habitat patches (Goosem, 2007; Laurance et al., 2009).

In the context of mining, habitat fragmentation and the resulting edge effects are further complicated by the associated necessary mining exploration and linear infrastructure that is constructed. Mining exploration and linear infrastructure

encompass haul roads, unpaved tracks, exploration (or seismic) lines, drill pads, power lines, pipelines, railways, and fence lines (Raiter et al., 2014). While the impacts of individual roads and tracks may only require a small area to be cleared, extensive exploration infrastructure to support mining operations means that the overall impacts can be quite significant (Raiter et al., 2014). These impacts are known as 'cumulative' or 'enigmatic impacts' (Canter & Ross, 2010; Raiter et al., 2017; Therivel & Ross, 2007). As these impacts are rarely accounted for in impact assessment protocols and regulatory frameworks (McKenney & Kiesecker, 2010), it is important to understand the extent of their impact on species and communities and the biophysical drivers behind them to improve how mining operations are managed (Cross et al. 2021).

One way of assessing edge effects is to monitor changes in species richness and community composition, and how far these changes extend into a habitat from an edge (Harper et al., 2005). Invertebrates are one of the most diverse taxa on Earth (Wilson, 1987), and their high abundance, importance in ecosystem function, sensitivity to disturbance and ease of sampling makes them ideal to study as environmental indicators (Kremen et al., 1993). In both aquatic and terrestrial systems, invertebrates are used in biomonitoring programs to assess disturbance and ecosystem health (Hodkinson & Jackson, 2005; Kwok et al., 2016; McGeoch, 1998; Moir et al., 2005; Roy et al., 2003). Invertebrates are also used widely in a mining context to monitor environmental impacts and restoration trajectories (Casimiro et al., 2019; Fernandes et al., 2019; Hoffmann & Andersen, 2003; Majer & Nichols, 1998; Moir et al., 2005; van der Heyde et al., 2022). Traditionally, biomonitoring using invertebrates is undertaken by trapping followed by morphological identification of every individual to determine the taxonomic constituents and their abundance in the sample. However, this methodology is resource-intensive, relying on a considerable amount of time and taxonomic expertise (over multiple taxonomic groups) to identify large volumes of invertebrate taxa. This can be costly, especially in biodiverse areas or over large spatial scales.

DNA metabarcoding is a useful molecular tool that can address some of the problems associated with traditional morphology-based monitoring techniques (Ji et al., 2013; Yu et al., 2012). With metabarcoding, it is possible to identify species across multiple taxonomic groups simultaneously and with a fraction of the person-

hours of morphology-based techniques (Ji et al. 2013). Thus, DNA metabarcoding of invertebrates could be a beneficial addition to mine site monitoring programs and provide valuable insights into environmental impacts across the whole community of arthropods inhabiting these areas. In recent years, DNA metabarcoding has been used in various systems, including a mine site monitoring context, to help understand and document the impact of environmental changes on invertebrates (Andújar et al., 2018; Beng et al., 2016; Fernandes et al., 2019; Kuntke et al., 2020; Lynggaard et al., 2020; van der Heyde et al. 2022).

In traditional biomonitoring studies, indicator taxa are often used instead of whole arthropod communities, to measure the impacts of disturbance (Majer & Nichols, 1998; Heino, 2014; Kerr et al., 2000). Indicator taxa are focal groups of organisms with well-known taxonomy and ecology that can be used as indicators for overall changes in the ecosystem. Some of the most common arthropod groups used as indicator taxa in terrestrial ecosystems include beetles, ants, spiders, and butterflies (Oliver and Beattie, 1996; Pik, Oliver, and Beattie, 1999; Terlizzi et al., 2003). Ants (Formicidae) are favoured as an indicator taxon due to the consistency of their response to disturbance (Hoffmann & Andersen, 2003). Springtails (Collembola) have also been recommended as indicator species in Australian systems as they are highly sensitive to disturbance in habitat structure (Greenslade, 2007). However, unlike more well-studied focal groups, such as ants, there is limited knowledge available on springtail taxonomy and many species are yet to be described (Jordana & Greenslade, 2020). Further, focusing only on certain groups of taxa may not be reflective of whole ecosystem dynamics as some of these arthropod groups may not interact with each other spatially, or have comparable taxonomic richness (Khosravi and Hemami, 2019; Prendergast et al., 1993; van Jaarsveld et al., 1998) and thus they may not fully encapsulate the changes and impacts on the diversity of multiple taxonomic groups and total arthropod assemblages (Siddig et al., 2016; Silva et al., 2017). Therefore, a direct comparison of patterns detected by all arthropods, and of particular groups like ants and springtails, is valuable as a way of validating the use of indicator groups.

In this study, we used invertebrate DNA metabarcoding to explore the impacts of mining exploration infrastructure on ground-dwelling arthropod communities associated with banded iron formations (BIFs) in the Midwest region of Western

Australia. Here we aimed to identify the response of arthropod community diversity and composition to mining exploration disturbances and to identify if ants (Formicidae) - an established indicator taxon - and springtails (Collembola) - a recommended indicator taxon but less well-known in Western Australia - could function as indicator groups that reflect patterns in the overall arthropod community.

3.4 Materials and Methods

3.4.1 Experimental Design

Three mine sites were selected for this study. At each site, three exploration infrastructure types were chosen: maintained tracks (MT), unmaintained tracks (UT), and drill pads (Pad). A maintained track was identified as a frequently used and disturbed track, which comprised no evidence of vegetation regrowth. By contrast, an unmaintained track would display vegetation growth, typically along the centre of the track or in old tyre tracks, and it may have vehicular access blocked off. Drill pads were areas where land had been cleared of all vegetation to drill into the BIF from which iron-ore is extracted. These were small, cleared areas that, once disturbed, had been left relatively untouched, with many having some degree of vegetation regrowth on them.

Sampling was carried out along a transect that ran from the interior of each chosen exploration disturbance to around 100 m into the interior of the adjacent BIF shrublands along the natural contour of the BIF. For each infrastructure type at each site, four transects were used. Sampling points were located along a transect at six different quadrat locations: (i) disturbance interior, (ii) disturbance edge, (iii) edge, (iv) natural edge, (v) 10 m from natural edge into natural vegetation, (vi) natural interior (Fig 3-1B). In total, there were 72 sampling points per site and 216 sampling points in total (3 sites x 12 transects per site x 6 sample points per transect). Transects varied slightly in length. This is appropriate because exploration disturbance features vary in both shape and size, so the interior of a disturbance and its edges are unique distances apart. Moreover, the edge of exploration land uses is not always clear-cut so soil disturbance and movement were used as a subjective indication of edge locations. Lastly, due to the fragmented landscape mosaic across these BIF sites, the natural interior (or control) sampling point was also at varying distances from the sampled edge. For comparative purposes the criteria for this sampling point were as follows: at least 100 m away along the same contour from its paired edge sampling point and if possible, at least 100 m away from any other disturbance edges (Swart et al., 2018).

3.4.2 Study sites

This study was conducted across three iron-ore mines: Koolanooka Hills (KO), Karara (KA), and Mount Gibson (MG) (Fig 3-1A) within the Midwest region of Western Australia. The Midwest region is characterised by a dry-warm Mediterranean climate, with mild wet winters and hot dry summers, receiving an average rainfall of 300-400 mm, mostly occurring during the winter months (Meissner & Caruso, 2008a). Disturbance age varied across all three sites. The age of the drill pads is estimated at each site; at KA, drill pads were constructed after 2012, at KO drill pads were constructed between 2005-2012 and at MG drill pads were constructed before 2000 (Stephens, 2018). The age of the unmaintained and maintained tracks is unknown, although maintained tracks were tracks that were still in use at the time of the sampling.

Vegetation communities differed between the three sites, but all were *Acacia*-dominated shrublands. Vegetation communities at the KO site had considerable disturbance, mainly from feral goats grazing on the natural vegetation.

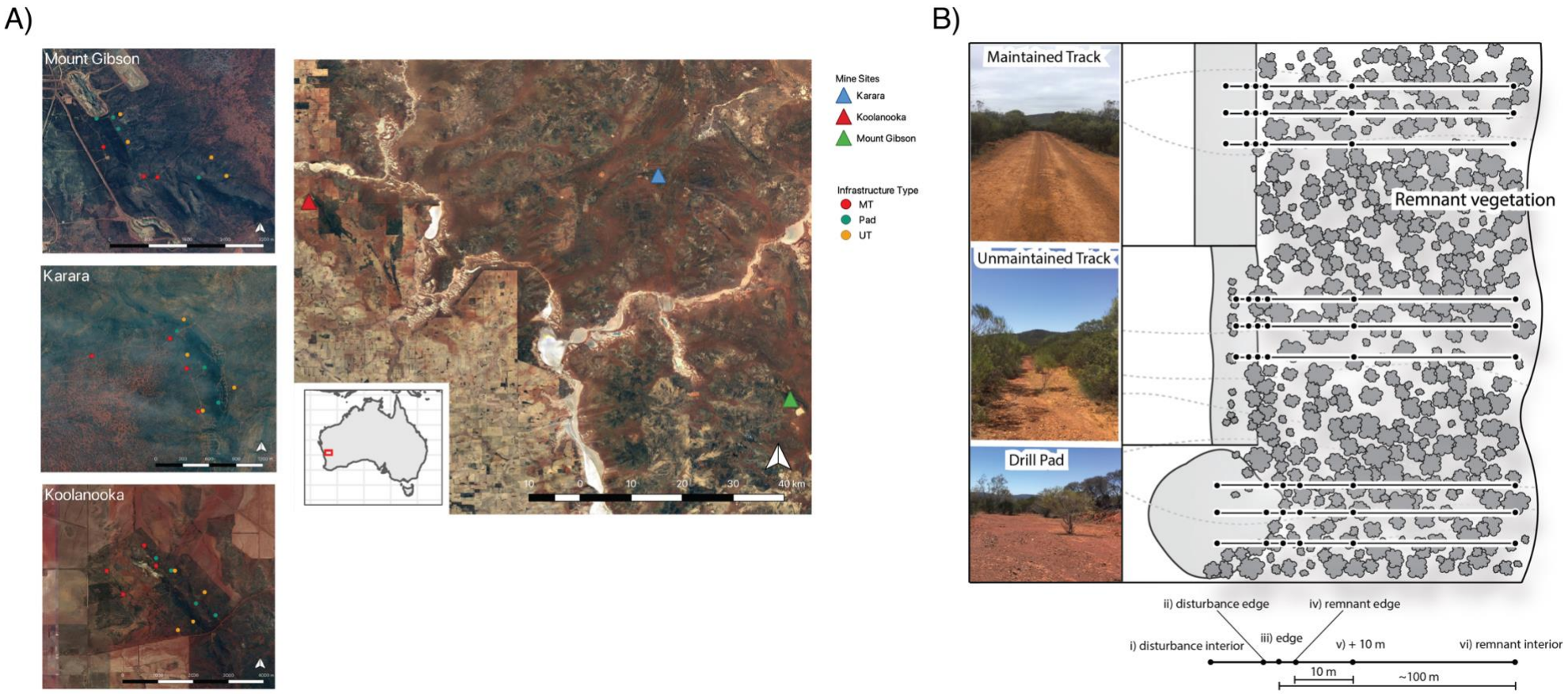


Figure 3-1 A) Study site showing location of three mine sites and the locations of the different infrastructure types at each site, where MT is Maintained Track, Pad is Drill Pad, and UT is Unmaintained track B) Stylised image of experimental design and pitfall trap placement along each transect across the ecotone between the disturbance (maintained track, unmaintained track, or drill pad) and remnant vegetation. The grey dashed lines indicate prospective topographic lines.

3.4.3 Sample Collection

Each sampling point along the transect comprised a 1 m2 plot. Within each plot, three pitfall traps with a diameter of 4 cm and a capacity of 120 mL, containing ~ 40 mL of ethylene glycol (undiluted commercial engine coolant, "Nulon Premium Long Life Coolant 100% Concentrate") were dug into the ground, with the rim flush with the soil surface. The pitfall traps were left open for approximately nine days and once collected the three pitfall traps were pooled into one sample.

3.4.4 Habitat Characteristics

A selection of habitat characteristics was measured and recorded within each 1 m2 plot. Slope aspect was measured using a compass and slope angle (degrees) measured using a Suunto PM-5 clinometer. A modified Braun-Blanquet scale (1, < 1 %; 2, 1-5 %; 3, 5-10 %; 4, 10-25 %; 5, 25-50 %; 6, 50-75 %; 7, >75 %) was used to approximate total vegetation cover, bare ground cover, rock cover and litter cover within the quadrat. Soil depth (soil compaction) was measured at the centre of each quadrat using a soil probe. Rugosity was measured with a chain along the 1 m horizontal and 1 m vertical transect which crossed the centre of the quadrat. All fine woody debris (FWD; 0 < diameter < 0.5 cm) touching the 1 m horizontal transect within the quadrat was counted. All medium woody debris (MWD; 0.5 ≤ diameter < 2.5 cm) was counted along both 1 m horizontal and vertical transects and all coarse woody debris (CWD; 2.5 ≤ diameter) present in the quadrat had its diameter measured and recorded.

3.4.5 Sample Processing and DNA extraction

Pitfall traps were rinsed with deionised water using 20-µm sieves to remove any ethylene glycol (coolant). Sieves were sterilised with bleach and UV light between each sample. Arthropod samples were roughly size sorted, where any organisms bigger than the size of a honey-bee had only two legs added to the overall sample (as per Ji et al., 2013; Meier et al., 2016). Samples were stored at -18˚C until extraction.

Samples were homogenised using a Qiagen TissueLyzer II using 5 mm metal beads for 2 minutes. Approximately 750 µL of homogenate was used for DNA extraction.

DNA extraction was carried out using a modified Qiagen QIAamp DNA Mini Kit with an overnight digest and a 400 µL digest volume as well as a final 100 µL elution in EB buffer. All extractions were carried out on an automated Qiacube (Qiagen). DNA extraction controls were carried out every 24 samples.

3.4.6 DNA Amplification and Sequencing

DNA extracts were diluted to 1/10 and 1/100 dilutions to test the amplification efficiency and detect the presence of any PCR inhibitors on the diluted and undiluted extracts using quantitative PCR (qPCR). The qPCR assay was run using an invertebrate mitochondrial COI primer set (fwhF2: 5'-GGDACWGGWTGAACWGTWTAYCCHCC-3' and fwhR2n: 5'-GTRATWGCHCCDGCTARWACWGG-3'; (Vamos et al., 2017)). This assay was determined as the ideal candidate for achieving optimal taxonomic resolution whilst still having a small fragment size able to be used with metabarcoding workflows (Elbrecht et al., 2019). PCR was carried out in 25 µL containing 2.5 mm MgCl2 (Applied Biosystems), 1x PCR Gold Buffer (Applied Biosystems), 0.25 mm dNTPs (Astral Scientific), 0.4 mg/mL bovine serum albumin (Fisher Biotec), 0.4 µmol/L forward and reverse primer, 1U AmpliTaq Gold DNA polymerase (Applied Biosystems), and 0.6 µL of a 1:10000 SYBR Green dye solution (Life Technologies). They were run on a StepOnePlus (Applied Biosystems) real-time qPCR instrument with the following conditions: 5 min at 95˚C, 45 cycles of 95˚C at 30s, 50˚C at 30s and 72˚C for 2 mins, followed by a 10 min 72˚C elongation. Extraction negative, PCR negative and PCR positive controls were also included in this initial screening. PCR positive controls were a mixed arthropod sample from a previous study that showed optimal amplification using laboratory workflows were used as a way of determining PCR efficiency. Contamination was minimised using a dedicated clean-lab space for preparing PCR reactions and a PCR-free laboratory for adding DNA extracts.

Samples with sufficient amplifiable DNA (as determined by the qPCR screening) were assigned a unique combination of fusion tag primers incorporated with multiplex identifier (MID) tags and Illumina sequencing adaptors (P5 and P7). Fusion PCR reactions were carried out in duplicate on DNA extracts using the same thermocycling conditions and reagents as the screening qPCR reactions. PCR and

extraction negative controls, but not positive controls, were also fusion-tagged during this. Samples were then pooled into pools of 10 samples consisting of approximately equimolar concentrations based on the qPCR amplification curves. These sample pools were quantified using a QIAxcel Advanced System (Qiagen). Sample pools were combined in approximate equimolar ratios as determined by this quantification to create a sequencing library. The library was size-selected using a Pippin Prep 2% agarose Marker B cassette (Sage Science) for fragments between 100-400 bp and then purified using a QIAquick PCR purification kit (Qiagen) modified with a 5-minute bench top incubation prior to final elution volume of 40 µL. The final library was quantified on a QuBit (Invitrogen) using double-stranded DNA high sensitivity reagents. The amplicon library was sequenced in a single direction on an Illumina Miseq instrument (Illumina) using a 300 cycle V2 reagent kit with a standard V2 flow cell as per manufacturer's protocol.

3.4.7 Sequence Analysis

Bioinformatic processing was conducted using the 'eDNAFlow' pipeline (Mousavi-Derazmahalleh et al., 2021). Briefly, raw sequences were demultiplexed to their sample using MID-tag combinations and length filtered using 'OBITools' (Boyer et al., 2016) with a minimum length of 50. Sequences were then quality filtered with errors and chimeras removed using USEARCH (Edgar, 2010) with a minimum Phred quality score of 20 and a minimum abundance of 8 sequences for zero-distance OTU (zOTU) creation. zOTUs were curated using 'LULU' at default parameters (Frøslev et al., 2017) on R v 3.6.1 (R Core Team, 2019). Appropriate sample depth per sample was assessed using rarefaction curves with the 'rarecurve' function in vegan (Oksanen et al., 2019) and samples with less than 2,000 sequences were removed. A 0.05% minimum abundance filtering threshold was used to combat false, low abundance zOTUS (Prodan et al., 2020). Low occurrence zOTUs with less than five sequences and occurring in only one sample were also removed. zOTUs were matched to the GenBank Reference database (www.ncbi.nlm.nih.gov/genbank/) and the Barcode of Life Database (Ratnasingham & Hebert, 2007) using the Basic Local Alignment Search Tool (BLAST) for taxonomic assignment. BLASTn results returned the top ten hits with a minimum query coverage of 95% and a minimum percentage identity of 85% to account for low representation of Australian arthropod

species in reference databases. zOTUs present in the extraction and PCR controls (Fungal zOTUs) were then removed from the dataset. The dataset was then divided into three categories; firstly, all zOTUs that were able to be identified to the phylum Arthropoda (overall arthropods), secondly all zOTUs belonging to the family Formicidae (ants), and lastly all zOTUs belonging to the class Collembola (springtails).

3.4.8 Statistical analysis

Statistical analysis was performed using PRIMER v.7 (Clarke & Gorley, 2015) with the PERMANOVA+ add on (Anderson et al., 2008). Firstly, to illustrate the scale of environmental disturbance the environmental variables from all samples were normalised and a principal component analysis (PCA) was performed with Euclidean distance on the centroids of quadrat number, located at each treatment and each site. A permutational ANOVA (PERMANOVA) was then conducted on the Euclidian distance matrix with 9999 permutations. The factors were as follows: Site (Fixed; KA, KO, MG), Infrastructure type (Fixed; MT, Pad, UT), Transect (Random; Nested in Site and Infrastructure type), and Quadrat location (Fixed; i – vi). The terms of interest to assess the impacts of linear infrastructure on arthropod communities were: the interaction between site, infrastructure type, and quadrat location; the interaction between infrastructure type and quadrat location; and the interaction between site and infrastructure type. For the overall arthropod community data, Formicidae community, and Collembola community, number of zOTUs (zOTU richness) was calculated and tested using univariate PERMANOVA with Euclidean distance and 9999 permutations using the same design as above. Arthropod community, Formicidae community, and Collembola communities were then transformed to presence-absence, and community composition was illustrated using nonmetric multidimensional scaling (nMDS) on Jaccard similarity on centroids of quadrat number located at each treatment and each site. A PERMANOVA was then conducted with the design as above using Jaccard similarity and 9999 permutations.

3.5 Results

3.5.1 Habitat characteristics

A significant interaction was detected between the infrastructure type and quadrat (Table 3-1). There was no significant interaction among site, infrastructure type, and quadrat, nor between site and infrastructure type. The different patterns of change with quadrat position among infrastructure types (interaction) were reflected in the PCA plots (Fig 3-2). However, there was a common pattern observed between all three sites and infrastructure types. There was a trajectory of change as the quadrats moved further from the disturbance (Quadrat i) into the remnant vegetation (Quadrat vi). This directional change was characterised by increased woody debris and soil depth moving away from the infrastructure type. This was a pattern observed across all the infrastructure types and sites.

Table 3-1 PERMANOVA results from the habitat characteristics compared at each site, infrastructure type, quadrat, transect and their interactions. Factors of interest are highlighted in bold. Significance codes are as follows: '*'0.001, '**'0.01, '*'0.05, ' '1.**

Factor	df	Pseudo-F	P value
Site	2	3.957	<0.001***
Infrastructure Type	2	4.159	<0.001***
Quadrat	5	14.409	<0.001***
Site x Infrastructure Type	**4**	**0.712**	**0.823**
Site x Quadrat	10	1.118	0.239
Infrastructure Type x Quadrat	**10**	**1.537**	**0.004****
Transect(Site x Infrastructure Type)	27	3.704	<0.001***
Site x Infrastructure Type x Quadrat	**20**	**0.974**	**0.572**

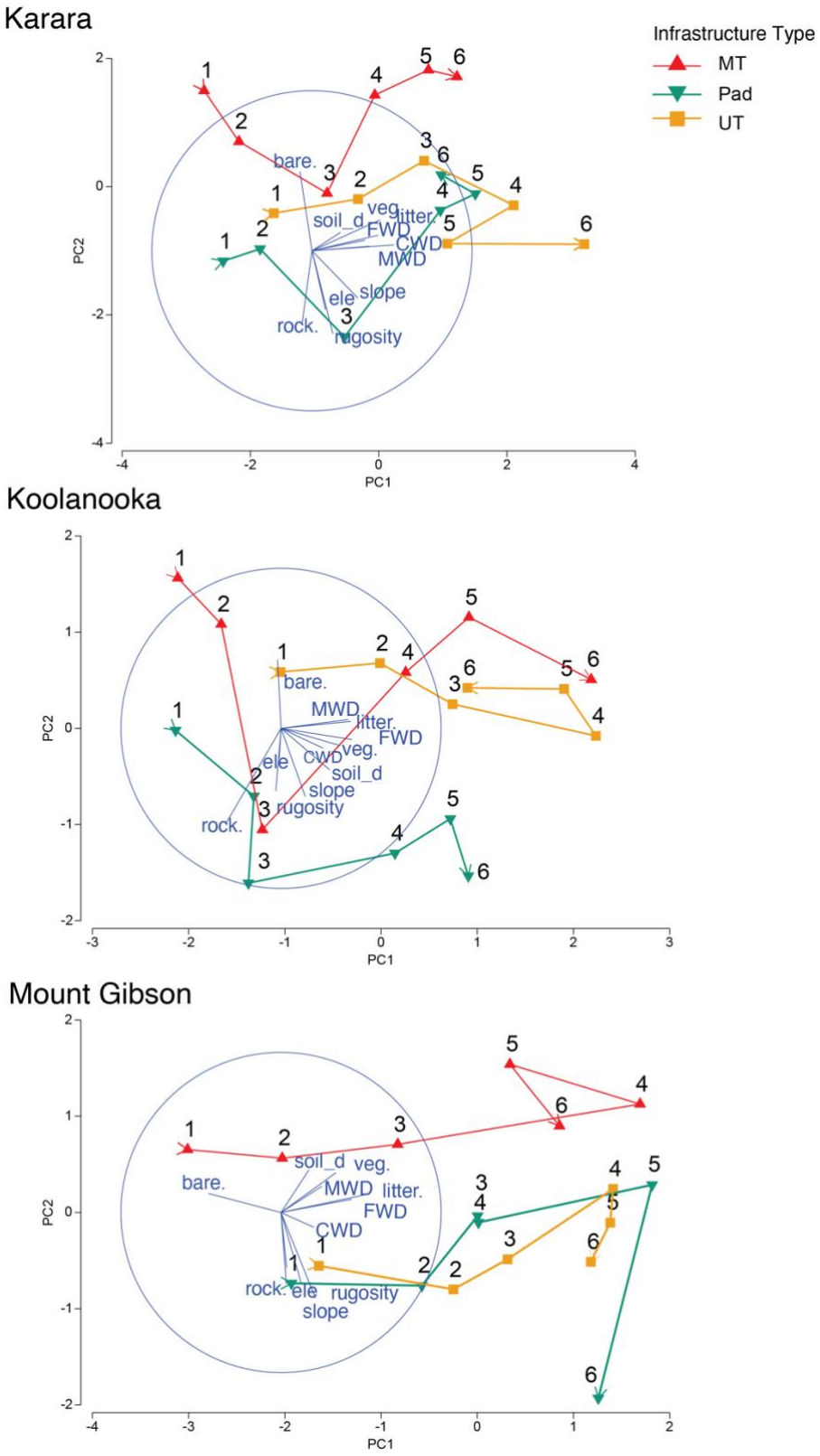


Figure 3-2 Principal component analysis (PCA) plot of the habitat characteristics at each site showing the three different infrastructure types where MT is Maintained Track, Pad is Drill Pad, and UT is Unmaintained track. Trajectories are overlaid starting from Quadrat i (1) on the disturbance to Quadrat vi (6) approximately 100 m into the remnant vegetation. Vectors overlaid of the correlation between each habitat characteristic and the principal component axes.

3.5.2 DNA Metabarcoding Results

A total of 15,507,193 quality-filtered sequences were generated across the 217 samples. A total of 212 samples had the minimum sequencing depth required of 2,000 sequences per sample and 208 samples passed the minimum required quality filtering thresholds. Only 5 zOTUs were detected in the control samples; however, none of these zOTUs were from the phylum Arthropoda and thus were removed from the data set. In total 2,250 zOTUs were generated, however, there were only 1,101 zOTUs that met filtering criteria and were used in this study. These zOTUs represented 4 classes, 19 orders and 84 families of arthropod (Table S 3-1). Formicidae made up 21% of the total zOTUS from this data set (238 zOTUs) and Collembola made up 14% (170 zOTUs).

3.5.2.1 zOTU richness

There were significant differences in the zOTU richness of the overall arthropod community, Formicidae and Collembola communities between sites. The overall arthropod richness differed significantly with infrastructure type, but Formicidae and Collembola richness did not. For the terms of interest that might indicate a disturbance effect, there was a significant interaction between site, infrastructure and quadrat detected for the Formicidae community (Table 3-2), but not for the overall arthropod community or the Collembola community. There was a significant interaction at the site and infrastructure level for all three arthropod communities (Fig 3-3).

Table 3-2 PERMANOVA results from the arthropod zOTU richness and zOTU composition compared at each site, infrastructure type, quadrat, transect and their interactions. Factors of interest are highlighted in bold. Significance codes are as follows: '*'0.001, '**'0.01, '*'0.05, ' '1.**

zOTU richness							
		Overall Arthropods		*Formicidae*		*Collembola*	
Factor	df	Pseudo-F	P value	Pseudo-F	P value	Pseudo-F	P value
Site	2	30.512	<0.001***	15.184	<0.001***	21.064	<0.001***
Infrastructure Type	2	6.366	0.006**	2.922	0.066	1.930	0.164
Quadrat	5	1.055	0.393	1.531	0.189	2.297	0.053
Site x Infrastructure Type	**4**	**7.844**	**<0.001*** **	**5.804**	**0.002****	**4.634**	**0.006****
Site x Quadrat	10	1.919	0.046*	1.355	0.214	0.753	0.668
Infrastructure Type x Quadrat	**10**	**1.208**	**0.292**	**1.112**	**0.357**	**1.425**	**0.177**
Transect (Site x Infrastructure Type)	27	2.563	<0.001***	4.861	<0.001***	3.445	<0.001***
Site x Infrastructure Type x Quadrat	**20**	**1.529**	**0.085**	**1.716**	**0.042***	**1.210**	**0.254**
zOTU Composition							
		Overall Arthropods		*Formicidae*		*Collembola*	
Factor	df	Pseudo-F	P value	Pseudo-F	P value	Pseudo-F	P value
Site	2	6.837	<0.001***	4.434	<0.001***	10.571	<0.001***
Infrastructure Type	2	1.917	<0.001***	2.005	<0.001***	2.516	<0.001***
Quadrat	5	1.231	<0.001***	1.245	0.012*	1.052	0.283
Site x Infrastructure Type	**4**	**1.596**	**<0.001*** **	**1.793**	**<0.001*** **	**1.690**	**<0.001*** **
Site x Quadrat	10	1.108	0.006**	1.155	0.017*	1.071	0.149
Infrastructure Type x Quadrat	**10**	**1.026**	**0.254**	**1.019**	**0.374**	**1.015**	**0.404**
Transect (Site x Infrastructure Type)	27	2.476	<0.001***	2.412	<0.001***	4.247	<0.001***
Site x Infrastructure Type x Quadrat	**20**	**1.067**	**0.012***	**1.140**	**0.004****	**1.063**	**0.107**

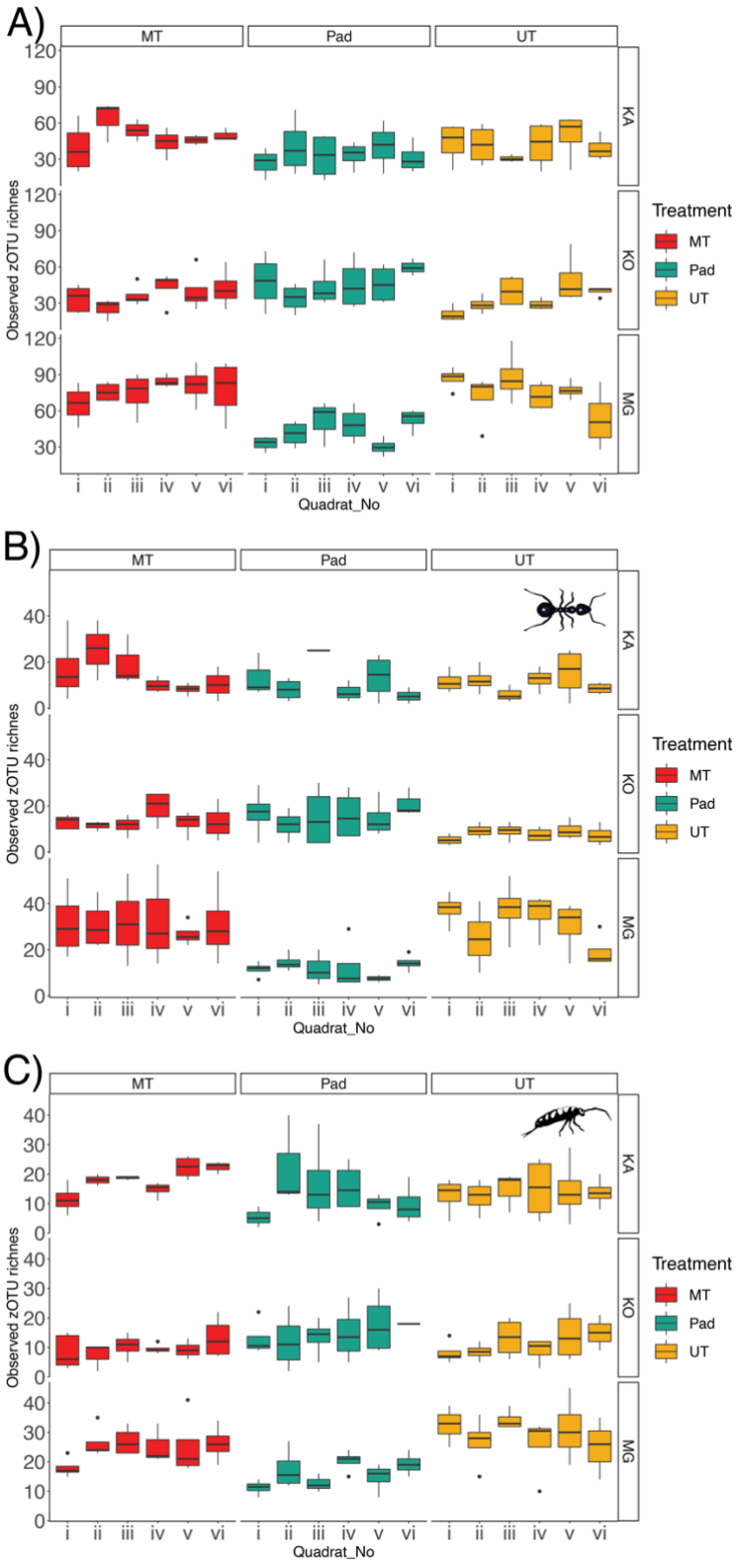


Figure 3-3 zOTU richness) at the different quadrat locations (Quadrat Number i-vi) at the different sites and infrastructure types for the A) overall arthropod communities, B) Formicidae communities, and C) Collembola communities. Sites are notated as KA for Karara, KO for Koolanooka, and MG for Mount Gibson. Different infrastructure types and sites are highlighted where MT is Maintained Track, Pad is Drill Pad, and UT is Unmaintained track.

3.5.2.2 Community composition

For both the overall arthropod communities and the Formicidae community there was a significant difference detected at the site, infrastructure type and quadrat interaction, together with an interaction of site and quadrat, and an interaction of site and infrastructure type (Table 3-2). However, unlike the habitat characteristics, there was no significant interaction detected between the infrastructure type and the quadrat location (Table 3-2). For the Collembola communities, only the site and infrastructure type interaction were found to be significantly different from our terms of interest (Table 3-2). For the overall arthropod communities, the different sites showed the most distinct separation (Fig 3-4A), and this pattern was reflected partially for the Formicidae communities (Fig 3-4B) and strongly for the Collembola communities (Fig 3-4C).

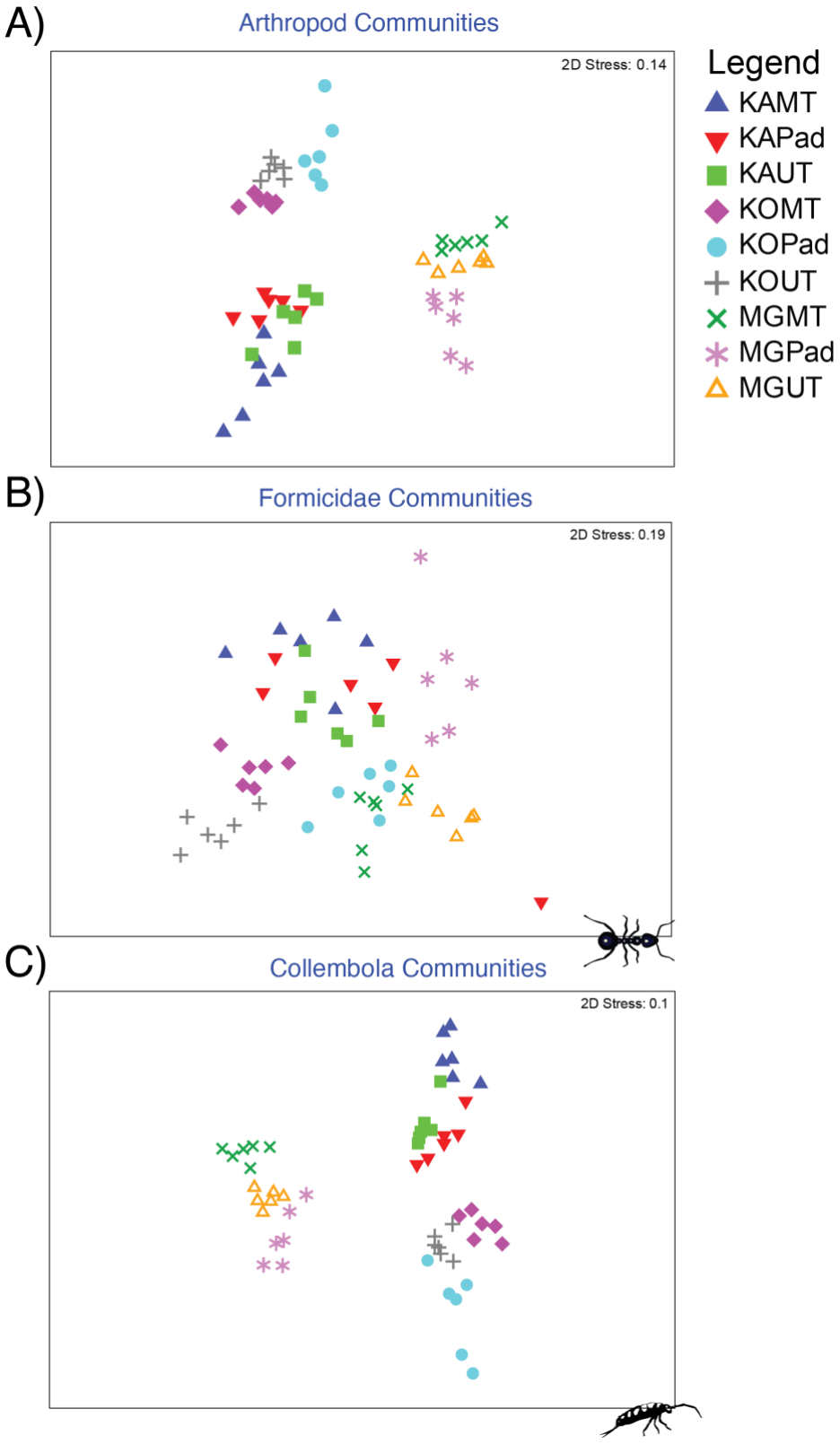


Figure 3-4 Non-metric multidimensional scaling (NMDS) plots showing A) overall arthropod communities, B) Formicidae communities, and C) Collembola communities at the three sites. Sites are notated as KA for Karara, KO for Koolanooka, and MG for Mount Gibson. Different infrastructure types and sites are highlighted where MT is Maintained Track, Pad is Drill Pad, and UT is Unmaintained track.

The site and infrastructure type interaction was also influential in differentiating the communities of overall arthropods, Formicidae, and Collembola (Fig 3-4). When examining the communities at a site-by-site level, the different infrastructure types were distinctly separated (Fig 3-5). For the overall arthropod communities, at the Karara site there was a distinct trajectory observed from the quadrat on the disturbance (1) to the quadrat in the remnant vegetation (6) for both the maintained track and drill pad infrastructure type (Fig 3-5A). However, at the other sites, this pattern was less clear, and, particularly at Mount Gibson, indicated a homogenization of the overall arthropod communities across the transect from which quadrats were sampled (100 m).

For the Formicidae communities at the Karara site, the distinction between the infrastructure types was not as clear as for the overall arthropod community (Fig 3-5B). However, at the other sites (Koolanooka and Mount Gibson), there was a separation of the Formicidae communities between the infrastructure type (Fig 3-5B). This distinction between the quadrats did not show the same directionality as the overall arthropod community. Similarly, with the Collembola communities, although there was a distinction between the different infrastructure types (Fig 3-5C), there was no clear distinction between the different quadrat locations (1-6) (Table 3-2).

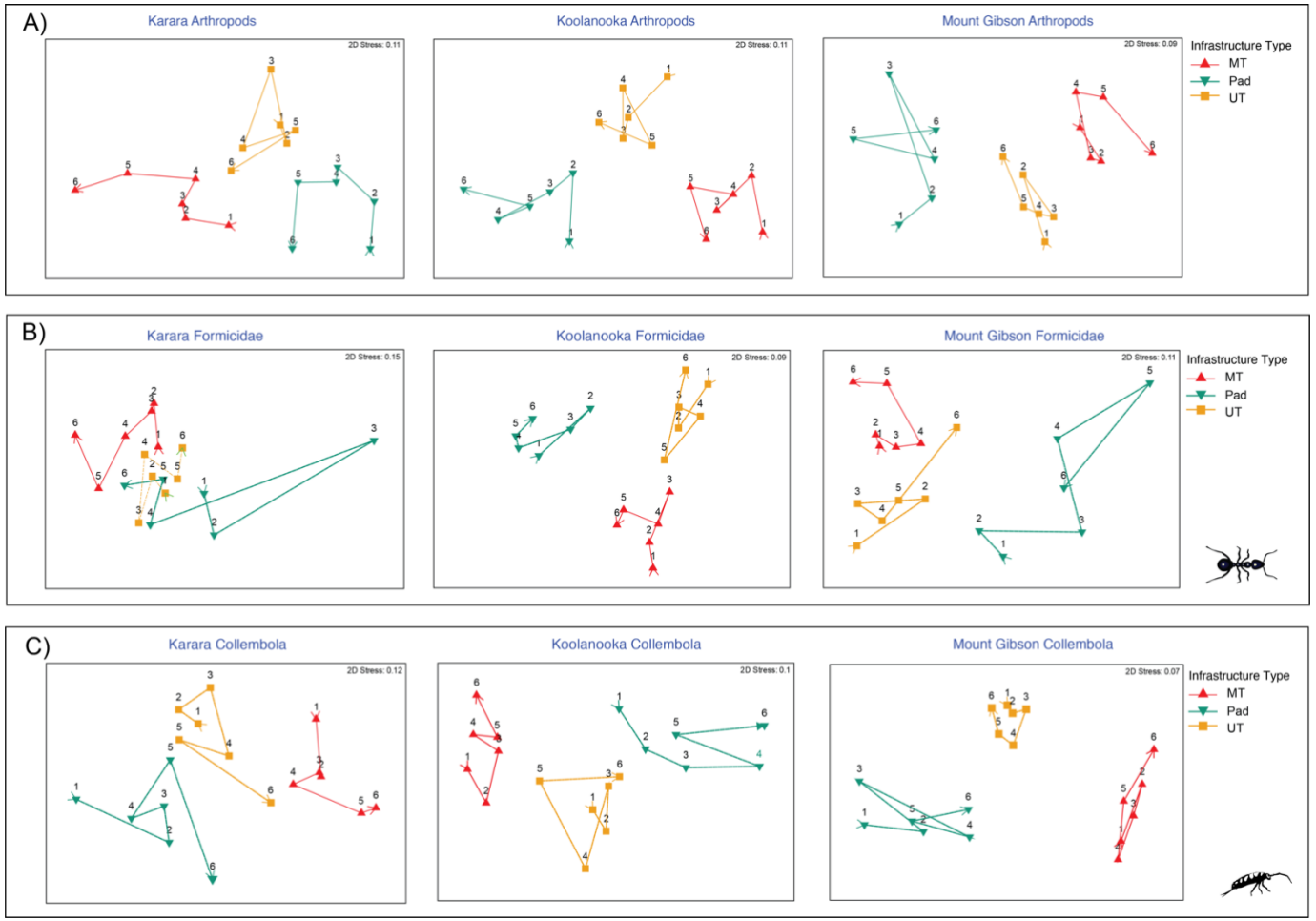


Figure 3-5 Non-metric multidimensional scaling (NMDS) plots showing the A) arthropod communities, B) Formicidae communities, and C) Collembola communities at the three sites. Different infrastructure types are highlighted where MT is Maintained Track, Pad is Drill Pad, and UT is Unmaintained track. Trajectories are overlaid starting from Quadrat i (1) to Quadrat vi (6).

3.6 Discussion

Using DNA metabarcoding we were able to show the impacts of mining exploration disturbances on invertebrate communities. Although the three infrastructure types impacted the habitat conditions similarly, the arthropod communities' changes were more complex. In general, we found similar patterns expressed in the compositional changes for arthropods overall and between our chosen indicator species groups (Formicidae and Collembola), but they were not the same across all sites and infrastructure types.

3.6.1 Habitat disturbance around linear infrastructure.

We found that habitat characteristics across the different infrastructure types were altered in very similar ways by anthropogenic disturbance. This change was broadly characterised by decreased woody debris and shallower soil depth (i.e., increased compaction) closer to the infrastructure type. This pattern was observed across all the infrastructure types and sites and was consistent with the findings of previous studies into the impacts of linear infrastructure on natural vegetation (Dániel-Ferreira et al., 2020; Goosem, 2012). For some sites, this change in habitat characteristics extended out to 10 m from the disturbance, suggesting edge effects at least at this spatial scale.

The overall habitat characteristics of quadrats on the drill pads and unmaintained tracks remained the same as for the regularly maintained tracks. Even years after the initial disturbance, the infrastructure areas we studied were relatively clear, and restoration to the surrounding vegetation community structure had not occurred. This has been attributed to specialisation of vegetation communities to particular soil chemical and physical characteristics, which remain altered following disturbance, and limit the return of vegetation (Gibson et al., 2012; Yates et al., 2011). Rehabilitation of such sites around exploration infrastructure need to be considered and planned for, just as other aspects of mining impacts are in the development of mine rehabilitation plans (Manero et al., 2020). At our study sites, exploration disturbances were either currently still maintained (maintained tracks) or left to revegetate passively (unmaintained tracks and drill pads), without any intervention methods such as recontouring or seeding (Meli et al., 2017). Tracks and roads are

considered as part of mine infrastructure, and mine closure requirements now consider rehabilitation of tracks as part of the overall environmental rehabilitation and closure obligations (DMIRS, 2021; Young et al., 2019). However, there is an extensive legacy of unrestored exploration infrastructure in Australia (Mills, 2022)

3.6.2 Impacts on ground-dwelling arthropod communities.

There was an influence of both site and infrastructure type on overall zOTU richness; however, not at the finer scale among quadrats, which suggests that we cannot measure edge effects via zOTU richness. A potential explanation for this finding is that richness (number of zOTUs) does not respond to disturbance in the same way as other metrics such as community composition (Fernandes et al., 2019). In general, a habitat can only support a certain number of species, and as the habitat changes over time, or because of disturbances, the composition of species changes rather than species richness (Lengyel et al., 2016). Other traditional univariate diversity measures that consider abundance alongside number of species (e.g., Shannon's, Maurer & McGill, 2011) may be more useful to measure these differences; however, it is not possible to obtain robust species abundance data using metabarcoding techniques that are needed to calculate these metrics. Therefore, looking at both zOTU richness and composition together is more effective at determining the responses of communities to change (Fernandes et al., 2019; van der Heyde et al., 2022).

In contrast to zOTU richness, using zOTU composition we could detect finer scale effects between quadrats that indicated disturbance. For some site-infrastructure combinations, such as the maintained tracks at Karara, there was a pattern of change in zOTU composition among quadrats from communities on the infrastructure out to those in remnant bushland, that indicated an edge effect extending 10 meters into the remnant bushland. However, this pattern was not consistent across all the maintained tracks at the three sites. This is not unexpected, as a previous study found that physical changes in the landscape caused by seismic lines were very visible, but they did not have a great impact on ant communities beyond one year post-disturbance (Watts et al., 2002). This was attributed to changes in vegetation within 10-100 metres of the infrastructure resulting in only minor reorganisation of foraging patterns for ants, but not the overall community structure (Watts et al. 2002).

One of our study aims was to examine whether two indicator groups of arthropods, Formicidae and Collembola, reflected overall patterns in arthropods at our sites in response to exploration disturbances. Whilst metabarcoding can be used to detect whole assemblages of arthropods, applying metabarcoding to indicator species can be a valuable way to understand functional changes in the environment, especially for well-studied groups where this information is available (Hajibabaei et al. 2019; Liu et al., 2020). Further, it is easier to build reference libraries, improve the taxonomic resolution of assays, and improve the understanding of specific species ecology when focusing only on select indicator taxon groups compared to whole species assemblages (Stat et al., 2017). Therefore, choosing an indicator taxon group can be an important consideration for designing metabarcoding studies. Formicidae and Collembola are valuable indicator taxa in their own right (Greenslade, 2007; Tiede et al. 2017). For example, Formicidae are reported to have consistent responses to disturbance (Hoffmann & Andersen, 2003) and hence make good indicators. Collembolans, unlike social invertebrates such as Formicidae, can be highly spatially aggregated in the soil (Greenslade, 2007), and are very sensitive to subtle changes in soil conditions (e.g., pH) (Sławski and Sławska, 2000) and theoretically could be useful for understanding the fine-scale impacts of disturbance. However, collembolan species composition is impacted by weather events or seasonal conditions (Yin et al., 2019) and patterns of disturbance may be confounded by these variables. Therefore, comparing the different groups (whole arthropods, Formicidae, and Collembola) is important to determine the utility of these groups as indicators.

While we observed similar patterns between the three sites and infrastructure types across the Formicidae and Collembola communities, the patterns were not identical across both groups. Further, the effect of site, infrastructure type, and quadrat was not significant for Collembola community composition, but significant for the whole arthropod communities and Formicidae. Overall, the Formicidae communities were more reflective of the differences in the total arthropod communities than Collembola, presumably reflecting their consistent responses to disturbance (Hoffmann & Andersen, 2003). This implies that while Collembola communities may be useful in the broader context of examining disturbances in the environment

(Greenslade and Majer, 1993), to examine the edge effects at a fine scale, Formicidae communities are more informative.

3.6.3 DNA metabarcoding as a tool to assess arthropod communities.

Here we demonstrated that DNA metabarcoding can be a useful tool for measuring arthropod community responses to changes in the environment. We were able to detect over 84 different families of arthropods and find differences between sites, infrastructure types, quadrats, and different indicator groups. Like any survey method, the sensitivity and efficacy of the DNA metabarcoding methodology is influenced by the inherent nature of the technology, study design and execution (van der Heyde et al., 2020). One of the most pressing issues in DNA metabarcoding studies is the reliance on reference databases, which are lacking in systems like ours that have poorly studied arthropod fauna (van der Heyde et al., 2020). We detected 2,250 zOTUs but only 1,101 could be identified as the phylum Arthropoda (~49%). Greater representation of species in databases is essential to increase the resolution of zOTU identifications and when combined with better understanding of species ecology, enables the assignment of ecological function to taxonomic groups (Pawlowski et al., 2018).

Further, improving the representation of understudied taxonomic groups (such as Collembola) would also reduce the number of sequences that are discarded because their identity could not be assigned (van der Heyde, et al. 2022). As we could only use less than 50% of the overall data that we generated, we could potentially only be looking at half the picture of disturbance impacts at these sites. The beauty of metabarcoding is that global barcode databases are constantly being populated. In the future, it may be possible to re-analyse these results and build a more complete picture of the impact of linear disturbance on arthropod communities. However, even with an incomplete arthropod barcode database, we detected similar responses to disturbance to those found in previous studies using morphological identification (Majer et al. 2020; Yu et al. 2006). Therefore, we have confidence that our observed trends reflect real patterns in arthropod communities at our study sites.

Visual observations of our samples strongly suggested that there were differences in the abundance of some taxa in samples collected from the disturbances and within

the remnant vegetation. However, because reliable abundance is difficult to obtain from metabarcoding data (Elbrecht & Leese, 2015), we were only able to collect presence/absence of our arthropod zOTUs and did not examine abundance-related differences. However, abundance data is critical for calculating different diversity metrics and may allow for a more nuanced interpretation of community responses to disturbance at our sites. There are methods to assess species abundance in bulk invertebrate samples (Elbrecht & Leese, 2015; Serrana et al., 2019); but these methods have either only been tested in aquatic systems or account for the relative abundance of a single taxon within a sample, not multiple groups simultaneously.

Seasonal variation is known to impact the composition of both arthropods (Santorufo et al., 2014) and plant communities, and their interactions (CaraDonna et al., 2017). Therefore, the resolution of our findings may have been improved by surveying arthropod communities over multiple time points. Thus, the use of repeated surveys and other monitoring methodologies (e.g., indicator species studies) may be required to determine if the observed patterns are one-off occurrences or a result of broader ecological trends.

Furthermore, metabarcoding of arthropod communities in this way is subject to the same considerations as traditional methods. Study design may have also played a role in the patterns observed. Here, we used multiple 100 m transects across infrastructure types at one site and we detected differences in arthropod communities at this scale, an interesting comparison would be to expand these transects out further distances into remnant bushland to determine broader landscape scale differences because of habitat modification within the same site. Furthermore, increasing the replicates within a quadrat may also be beneficial in adding support to these results.

3.7 Conclusion

Mining exploration infrastructure had significant impacts on both physical habitat and arthropod communities. These disturbances altered physical habitat characteristics in similar ways across infrastructure types at all sites. There were detectable edge effects, but these impacts did not extend far into remnant bushland (<10 m). For arthropod communities as a whole and Formicidae communities, the effect of disturbance extended into the remnant bushland in a pattern similar to the habitat. Our study highlighted the utility of DNA metabarcoding as a tool to identify changes in arthropod communities impacted by mining. We recommend continued monitoring of arthropod communities in our study system, especially over different spatial scales, seasons, and time points, to develop a more holistic understanding of their responses to disturbance. Our data emphasise the need to consider different types of disturbances associated with mining activity when assessing the impact of new mining projects, as they may all have the potential to impact surrounding faunal communities to a significant degree.

3.8 References

Andersen, A. N. (2019). Responses of ant communities to disturbance: Five principles for understanding the disturbance dynamics of a globally dominant faunal group. *The Journal of Animal Ecology*, *88*(3), 350–362.

Anderson, M., Gorley, R., & Clarke, K. P. (2008). *PERMANOVA+ for PRIMER: guide to software and statistical methods*. PRIMER-E Ltd.

Andújar, C., Arribas, P., Gray, C., Bruce, C., Woodward, G., Yu, D. W., & Vogler, A. P. (2018). Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, *27*(1), 146–166.

Arribas, P., Andújar, C., Salces-Castellano, A., Emerson, B. C., & Vogler, A. P. (2020). The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Molecular Ecology*. https://doi.org/10.1111/mec.15591

Barsoum, N., Bruce, C., Forster, J., Ji, Y.-Q., & Yu, D. W. (2019). The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity. *Ecological Indicators*, *101*, 313–323.

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., & Slik, J. W. F. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific Reports*, *6*(1), 24965.

Bolger, D. T., Suarez, A. V., Crooks, K. R., Morrison, S. A., & Case, T. J. (2000). Arthropods in urban habitat fragments in Southern California: Area, age, and edge effects. *Ecological Applications*, *10*(4), 1230–1248.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12428

Canter, L., & Ross, B. (2010). State of practice of cumulative effects assessment and management: the good, the bad and the ugly. *Impact Assessment and Project Appraisal*, *28*(4), 261–268.

CaraDonna, P. J., Petry, W. K., Brennan, R. M., Cunningham, J. L., Bronstein, J. L., Waser, N. M., & Sanders, N. J. (2017). Interaction rewiring and the rapid turnover of plant–pollinator networks. *Ecology Letters*, *20*(3), 385–394. https://doi.org/10.1111/ele.12740

Casimiro, M. S., Sansevero, J. B. B., & Queiroz, J. M. (2019). What can ants tell us about ecological restoration? A global meta-analysis. *Ecological Indicators*, *102*, 593–598.

Clarke, K. R., & Gorley, R. (2015). *Getting started with PRIMER v7*. Marine Laboratory, Plymouth.

Cross, S. L., Cross, A. T., Tomlinson, S., Clark-Ioannou, S. M., Nevill, P. G., & Bateman, P. W. (2021). Mitigation and management plans should consider all anthropogenic disturbances to fauna. *Global Ecology and Conservation, 26,* e01500.

Dániel-Ferreira, J., Bommarco, R., Wissman, J., & Öckinger, E. (2020). Linear infrastructure habitats increase landscape-scale diversity of plants but not of flower-visiting insects. *Scientific Reports, 10*(1), 1-11.

Department of Mines, Industry Regulation, and Safety (DMIRS). (2021). *Mine Closure Completion Guideline – For demonstrating completion of mine closure in accordance with an approved Mine Closure Plan*. Department of Mines, Industry Regulation and Safety, Government of Western Australia, .

Dormontt, E. E., Van Dijk, K. J., Bell, K. L., Biffin, E., Breed, M. F., Byrne, M., ... & Lowe, A. J. (2018). Advancing DNA barcoding and metabarcoding applications for plants requires systematic analysis of herbarium collections—an Australian perspective. *Frontiers in Ecology and Evolution*, *6*, 134.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461.

Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., Zakharov, E. V., Hebert, P. D. N., & Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, *7*, e7745.

Elbrecht, V., & Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. *PloS One*, *10*(7), e0130324.

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, *8*(10), 1265–1275.

Ellis, E. C., Klein Goldewijk, K., Siebert, S., Lightman, D., & Ramankutty, N. (2010). Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography: A Journal of Macroecology*, *19*(5), 589–606.

Environmental Protection Authority. (2009). *Karara Iron Ore Project* (No. 1321). Environmental Protection Authority Perth, Western Australia . http://www.epa.wa.gov.au/sites/default/files/EPA_Report/2930_Rep1321KararaPER28409.pdf

Ewers, R. M., & Didham, R. K. (2006). Continuous response functions for quantifying the strength of edge effects. *The Journal of Applied Ecology*, *43*(3), 527–536.

Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 487–515.

Fernandes, K., van der Heyde, M., Coghlan, M., Wardell-Johnson, G., Bunce, M., Harris, R., & Nevill, P. (2019). Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences. *Restoration Ecology*, *215*, 403.

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications*, *8*(1), 1188.

Gardiner, T., Pye, M., Field, R., & Hill, J. (2002). The influence of sward height and vegetation composition in determining the habitat preferences of three Chorthippus species (Orthoptera: Acrididae) in Chelmsford, Essex, UK. *Journal of Orthoptera Research*, *11*(2), 207–213.

Gibson, N., Meissner, R., Markey, A. S., & Thompson, W. A. (2012). Patterns of plant diversity in ironstone ranges in arid south western Australia. *Journal of Arid Environments*, *77*, 25-31.

Gibson, N., Yates, C. J., & Dillon, R. (2010). Plant communities of the ironstone ranges of South Western Australia: hotspots for plant diversity and mineral deposits. *Biodiversity and Conservation*, *19*(14), 3951-3962.

Gleason, J. E., Elbrecht, V., Braukmann, T. W. A., Hanner, R. H., & Cottenie, K. (2020). Assessment of stream macroinvertebrate communities with eDNA is not congruent with tissue-based metabarcoding. *Molecular Ecology*. https://doi.org/10.1111/mec.15597

Goosem, M. (2007). Fragmentation impacts caused by roads through rainforests. *Current Science*, *93*(11), 1587–1595.

Goosem, M. (2012). Mitigating the impacts of rainforest roads in Queensland's Wet Tropics: Effective or are further evaluations and new mitigation strategies required? *Ecological Management & Restoration*, *13*(3), 254–258.

Grant, C., Loch, R., McCaffrey, N., Anstee, S., & Doley, D. (2016). *Mine Rehabilitation: Leading Practice Sustainable Development Program for the Mining Industry*. Commonwealth of Australia.

Greenslade, P. (2007). The potential of Collembola to act as indicators of landscape stress in Australia. *Australian Journal of Experimental Agriculture*, *47*(4), 424–434.

Greenslade, P., & Majer, J. D. (1993). Recolonization by Collembola of rehabilitated bauxite mines in Western Australia. *Australian Journal of Ecology*, *18*(4), 385-394.

Guariguata, M. R., & Dupuy, J. M. (1997). Forest Regeneration in Abandoned Logging Roads in Lowland Costa Rica. *Biotropica*, *29*(1), 15–28.

Guijón, R., Henríquez, F., & Naranjo, J. A. (2011). Geological, Geographical and Legal Considerations for the Conservation of Unique Iron Oxide and Sulphur Flows at El Laco and Lastarria Volcanic Complexes, Central Andes, Northern Chile. *Geoheritage*, *3*(4), 299–315.

Hajibabaei, M., Porter, T. M., Wright, M., & Rudar, J. (2019). COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems. *PloS One*, 14(9), e0220953

Harper, K. A., Macdonald, S. E., Burton, P. J., Chen, J., Brosofske, K. D., Saunders, S. C., Euskirchen, E. S., Roberts, D., Jaiteh, M. S., & Esseen, P.-A. (2005). Edge Influence on Forest Structure and Composition in Fragmented Landscapes. *Conservation Biology*, *19*(3), 768–782.

Haycock, R., & Cox, C. (2015). *Sinosteel Midwest Corporation Limited: Koolanooka vegetation monitoring program, Spring 2015*. Maia Environmental Consultancy Pty Ltd. https://www.smcl.com.au/bluehills-perdocs/Appendix-C/Flora-and-vegetation/Maia-2015b.pdf

Heino, J. (2014). Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: Are the inferences transferable among regions? *Ecological Indicators*, *36*, 186–194.

Hodkinson, I. D., & Jackson, J. K. (2005). Terrestrial and Aquatic Invertebrates as Bioindicators for Environmental Monitoring, with Particular Reference to Mountain Ecosystems. *Environmental Management*, *35*(5), 649–666.

Hoffmann, B. D., & Andersen, A. N. (2003). Responses of ants to disturbance in Australia, with particular reference to functional groups. *Austral Ecology*, *28*(4), 444–464.

Hooke, R. L., Martín-Duque, J. F., & Pedraza, J. (2012). Land transformation by humans: a review. *GSA Today: A Publication of the Geological Society of America*, *22*(12), 4–10.

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257.

Ji, Y., Huotari, T., Roslin, T., Schmidt, N. M., Wang, J., Yu, D. W., & Ovaskainen, O. (2020). SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes. *Molecular Ecology Resources*, *20*(1), 256–267.

Johnson, C. N., Balmford, A., Brook, B. W., Buettel, J. C., Galetti, M., Guangchun, L., & Wilmshurst, J. M. (2017). Biodiversity losses and conservation responses in the Anthropocene. *Science*, *356*(6335), 270–275.

Jones, I. L., Bull, J. W., Milner-Gulland, E. J., Esipov, A. V., & Suttle, K. B. (2014). Quantifying habitat impacts of natural gas infrastructure to facilitate biodiversity offsetting. *Ecology and Evolution*, *4*(1), 79–90.

Jordana, R., & Greenslade, P. (2020). Biogeographical and ecological insights from Australasian faunas: the megadiverse collembolan genus, Entomobrya (Entomobryidae). *Zootaxa*, *4770*(1), 1–104.

Kerr, J. T., Sugar, A., & Packer, L. (2000). Indicator Taxa, Rapid Biodiversity Assessment, and Nestedness in an Endangered Ecosystem. *Conservation Biology*, *14*(6), 1726–1734.

Khosravi, R., Hemami, M.-R. (2019). Identifying landscape species for ecological planning. *Ecological Indicators*, *99*, 140–148.

Kremen, C., Colwell, R. K., Erwin, T. L., Murphy, D. D., Noss, R. F., & Sanjayan, M. A. (1993). Terrestrial Arthropod Assemblages: Their Use in Conservation Planning. *Conservation Biology*, *7*(4), 796–808.

Kuntke, F., de Jonge, N., Hesselsøe, M., & Lund Nielsen, J. (2020). Stream water quality assessment by metabarcoding of invertebrates. *Ecological Indicators*, *111*, 105982.

Kwok, A. B. C., Wardle, G. M., Greenville, A. C., & Dickman, C. R. (2016). Long-term patterns of invertebrate abundance and relationships to environmental factors in arid Australia. *Austral Ecology*, *41*(5), 480–491.

Laurance, W. F., Goosem, M., & Laurance, S. G. W. (2009). Impacts of roads and linear clearings on tropical forests. *Trends in Ecology & Evolution*, *24*(12), 659–669.

Lengyel, S., Déri, E., & Magura, T. (2016). Species Richness Responses to Structural or Compositional Habitat Diversity between and within Grassland Patches: A Multi-Taxon Approach. *PloS One*, *11*(2), e0149662.

Liu, L., Guo, Z., Zhong, C., & Shi, S. (2018). DNA barcoding reveals insect diversity in the mangrove ecosystems of Hainan Island, China. *Genome*, *61*(11), 797–806.

Liu, M., Baker, S. C., Burridge, C. P., Jordan, G. J., Clarke, L. J. (2020). DNA metabarcoding captures subtle differences in forest beetle communities following disturbance. *Restoration Ecology*. doi:10.1111/rec.13236

Lynggaard, C., Yu, D. W., Oliveira, G., Caldeira, C. F., Ramos, S. J., Ellegaard, M. R., Gilbert, M. T. P., Gastauer, M., & Bohmann, K. (2020). DNA-Based Arthropod Diversity Assessment in Amazonian Iron Mine Lands Show Ecological Succession Towards Undisturbed Reference Sites. *Frontiers in Ecology and Evolution*, *8*, 426.

Majer, J. D., de Sousa-Majer, M. J., & Heterick, B. E. (2020). Partial clearing of a road corridor leads to homogenisation of the invertebrate fauna. *Pacific Conservation Biology*. https://doi.org/10.1071/PC20039

Majer, J. D., & Nichols, O. O. (1998). Long-Term Recolonization Patterns of Ants in Western Australian Rehabilitated Bauxite Mines with Reference to Their Use as Indicators of Restoration Success. *The Journal of Applied Ecology*, *35*(1), 161–182.

Manero, A., Kragt, M., Standish, R., Miller, B., Jasper, D., Boggs, G., & Young, R. (2020). A framework for developing completion criteria for mine closure and rehabilitation. *Journal of Environmental Management*, *273*, 111078.

Maurer, B. A., & McGill, B. J. (2011). Measurement of species diversity. *Biological Diversity: Frontiers in Measurement and Assessment*, 55-65.

McGeoch, M. A. (1998). The selection, testing and application of terrestrial insects as bioindicators. *Biological Reviews of the Cambridge Philosophical Society*, *73*(2), 181–201.

McKenney, B. A., & Kiesecker, J. M. (2010). Policy development for biodiversity offsets: a review of offset frameworks. *Environmental Management*, *45*(1), 165–176.

Meier, R., Wong, W., Srivathsan, A., & Foo, M. (2016). $1 DNA barcodes for reconstructing complex phenomes and finding rare species in specimen-rich samples. *Cladistics*, *32*(1), 100–110.

Meissner, R., & Caruso, Y. (2008a). Flora and vegetation of banded iron formations of the Yilgarn Craton: Koolanooka and Perenjori Hills. *Conservation Science Western Australia*, *7*(1), 73–88.

Meissner, R., & Caruso, Y. (2008b). Flora and vegetation of banded iron formations of the Yilgarn Craton: Mount Gibson and surrounding area. *Conservation Science Western Australia*, *7*(1), 105–120.

Meli, P., Holl, K. D., Rey Benayas, J. M., Jones, H. P., Jones, P. C., Montoya, D., & Moreno Mateos, D. (2017). A global review of past land use, climate, and active vs. passive restoration effects on forest recovery. *PloS One*, *12*(2), e0171368.

Mills, L. N. (2022). Getting closure? Mining rehabilitation reform in Queensland and Western Australia. *The Extractive Industries and Society*, 101097.

Moir, M. L., Brennan, K. E. C., Koch, J. M., Majer, J. D., & Fletcher, M. J. (2005). Restoration of a forest ecosystem: The effects of vegetation and dispersal capabilities on the reassembly of plant-dwelling arthropods. *Forest Ecology and Management*, *217*(2), 294–306.

Mousavi-Derazmahalleh, M., Stott, A., Lines, R., Peverley, G., Nester, G., Simpson, T., Zawierta, M., De La Pierre, M., Bunce, M., & Christophersen, C. T. (2021). eDNAFlow, an automated, reproducible and scalable workflow for analysis of environmental DNA (eDNA) sequences exploiting Nextflow and Singularity. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13356

Murcia, C. (1995). Edge effects in fragmented forests: implications for conservation. *Trends in Ecology & Evolution*, *10*(2), 58–62.

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (Version 2.5-6) [Computer software]. https://cran.ism.ac.jp/web/packages/vegan/vegan.pdf

Oliver, I., Beattie, A. J. (1996). Designing a Cost-Effective Invertebrate Survey: A Test of Methods for Rapid Assessment of Biodiversity. *Ecological Applications*, *6*(2), 594–607.

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., ... & Kahlert, M. (2018). The future of biotic indices in the ecogenomic era: Integrating (e) DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment*, *637*, 1295-1310.

Pik, A. J., Oliver, I., Beattie, A. J. (1999). Taxonomic sufficiency in ecological studies of terrestrial invertebrates. *Australian Journal of Ecology*, *24*(5), 555–562.

Pohlman, C. L., Turton, S. M., & Goosem, M. (2009). Temporal variation in microclimatic edge effects near powerlines, highways and streams in Australian tropical rainforest. *Agricultural and Forest Meteorology*, *149*(1), 84–95.

Prendergast, J. R., Quinn, R. M., Lawton, J. H., Eversham, B. C., Gibbons, D. W. (1993). Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature*, *365*(6444), 335–337.

Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A. H., Nieuwdorp, M., & Levin, E. (2020). Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PloS One*, *15*(1), e0227434.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing* (Version 3.6.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Raiter, K. G., Possingham, H. P., Prober, S. M., & Hobbs, R. J. (2014). Under the radar: mitigating enigmatic ecological impacts. *Trends in Ecology & Evolution*, *29*(11), 635–644.

Raiter, K. G., Prober, S. M., Hobbs, R. J., & Possingham, H. P. (2017). Lines in the sand: quantifying the cumulative development footprint in the world's largest remaining temperate woodland. *Landscape Ecology*, *32*(10), 1969–1986.

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes*. https://doi.org/10.1111/j.1471-8286.2006.01678.x

Redondo-Vega, J. M., Gómez-Villar, A., Santos-González, J., González-Gutiérrez, R. B., & Álvarez-Martínez, J. (2017). Changes in land use due to mining in the north-western mountains of Spain during the previous 50 years. *Catena*, *149*, 844–856.

Ries, L., & Sisk, T. D. (2004). A predictive model of edge effects. *Ecology*, *85*(11), 2917-2926.

Roy, A. H., Rosemond, A. D., Leigh, D. S., Paul, M. J., & Wallace, J. B. (2003). Habitat-specific responses of stream insects to land cover disturbance: biological consequences and monitoring implications. *Journal of the North American Benthological Society*, *22*(2), 292–307.

Santorufo, L., Van Gestel, C. A., & Maisto, G. (2014). Sampling season affects conclusions on soil arthropod community structure responses to metal pollution in Mediterranean urban soils. *Geoderma*, *226*, 47-53.

Santos, B. A., Peres, C. A., Oliveira, M. A., Grillo, A., Alves-Costa, C. P., & Tabarelli, M. (2008). Drastic erosion in functional attributes of tree assemblages in Atlantic forest fragments of northeastern Brazil. *Biological Conservation*, *141*(1), 249–260.

Serrana, J. M., Miyake, Y., Gamboa, M., & Watanabe, K. (2019). Comparison of DNA metabarcoding and morphological identification for stream macroinvertebrate biodiversity assessment and monitoring. *Ecological Indicators*, *101*, 963–972.

Siddig, A. A. H., Ellison, A. M., Ochs, A., Villar-Leeman, C., & Lau, M. K. (2016). How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in Ecological Indicators. *Ecological Indicators*, *60*, 223–230.

Silva, T. H. S. E., Karunarathna, N. B., & Karunaratne, W. A. I. P. (2017). Competence of Litter Ants for Rapid Biodiversity Assessments. *Lundiana: International Journal of Biodiversity*, *2017*, 1–10.

Sławski, M., & Sławska, M. (2000). The forest edge as a border between forest and meadow. Vegetation and Collembola communities. *Pedobiologia*, *44*(3-4), 442-450.

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., Harvey, E.S, & Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, 7(1), 1-11.

Stephens, C. (2018). *Exploration infrastructure impacts on Banded Iron Formation (BIF) ranges in the Midwest, Western Australia* [Honours Dissertation, Curtin University]. Curtin eSpace.

Swart, R. C., Pryke, J. S., & Roets, F. (2018). Arthropod assemblages deep in natural forests show different responses to surrounding land use. *Biodiversity and Conservation*, *27*(3), 583–606.

Sylvain, Z. A., Espeland, E. K., Rand, T. A., West, N. M., & Branson, D. H. (2019). Oilfield Reclamation Recovers Productivity but not Composition of Arthropod Herbivores and Predators. *Environmental Entomology*, *48*(2), 299–308.

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050.

Terlizzi, A., Bevilacqua, S., Fraschetti, S., Boero, F. (2003). Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. *Marine Pollution Bulletin*, *46*(5), 556–561.

Therivel, R., & Ross, B. (2007). Cumulative effects assessment: Does scale matter? *Environmental Impact Assessment Review*, *27*(5), 365–385.

Tibbett, M. (2015). *Mining in Ecologically Sensitive Landscapes*. CSIRO Publishing.

Tiede, Y., Schlautmann, J., Donoso, D. A., Wallis, C. I. B., Bendix, J., Brandl, R., & Farwig, N. (2017). Ants as indicators of environmental change and ecosystem processes. *Ecological Indicators*, *83*, 527–537.

Tousignant, M.-Ê., Pellerin, S., & Brisson, J. (2010). The Relative Impact of Human Disturbances on the Vegetation of a Large Wetland Complex. *Wetlands*, *30*(2), 333–344.

Vamos, E. E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics, 1*, e14625.

van der Heyde, M., Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., & Nevill, P. (2020). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Molecular Ecology Resources, 20*(3). https://doi.org/10.1111/1755-0998.13148

van der Heyde, M., Bunce, M., Dixon, K. W., Fernandes, K., Majer, J., Wardell-Johnson, G., White, N. E., & Nevill, P. (2022). Evaluating restoration trajectories using DNA metabarcoding of ground-dwelling and airborne invertebrates and associated plant communities. *Molecular Ecology, 31*, 2172– 2188. doi:10.1111/mec.16375

van Jaarsveld, A. S., Freitag, S., Chown, S. L., Muller, C., Koch, S., Hull, H., … Scholtz, C. H. (1998). Biodiversity assessment and conservation strategies. *Science*, *279*(5359), 2106–2108.

Watts, C., McArthur, A., Oakey, H., & Verbyla, A. (2002). *The impact of seismic lines on ant communities in the Cooper Basin and potential use of ants as bioindicators of ecological recovery rates*. South Australian Museum and BiometricsSA. https://energymining.sa.gov.au/__data/assets/pdf_file/0005/256235/ants_report_mar2002.pdf

Wilson, E. O. (1987). The Little Things That Run the World (The Importance and Conservation of Invertebrates). *Conservation Biology: The Journal of the Society for Conservation Biology*, *1*(4), 344–346.

Wycisk, P., Hubert, T., Gossel, W., & Neumann, C. (2009). High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites. *Computers & Geosciences*, *35*(1), 165–182.

Yates, C. J., Gibson, N., Pettit, N. E., Dillon, R., & Palmer, R. (2011). The ecological relationships and demography of restricted ironstone endemic plant species: implications for conservation. *Australian Journal of Botany*, *59*(7), 692-700.

Young, R.E., Manero, A., Miller, B.P., Kragt, M.E., Standish, R.J., Jasper, D.A., & Boggs, G.S. (2019). *A framework for developing mine-site completion criteria in Western Australia.* The Western Australian Biodiversity Science Institute, Perth, Western Australia.

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, *3*(4), 613–623

Yu, X. D., Luo, T. H., Yang, J., & Zhou, H. Z. (2006). Distribution of ground-dwelling beetles (Coleoptera) across a forest-clearcut ecotone in Wolong Natural Reserve, southwestern China. *Insect Science*, *13*(2), 127-135.

3.9 Supplementary Tables

Table S 3-1 Full list of families identified from zOTUs across all mine sites.

Class	Order	Family
Arachnida	Araneae	Clubionidae
Arachnida	Araneae	Corinnidae
Arachnida	Araneae	Ctenidae
Arachnida	Araneae	Desidae
Arachnida	Araneae	Gnaphosidae
Arachnida	Araneae	Hahniidae
Arachnida	Araneae	Linyphiidae
Arachnida	Araneae	Lycosidae
Arachnida	Araneae	Nemesiidae
Arachnida	Araneae	Oxyopidae
Arachnida	Araneae	Salticidae
Arachnida	Araneae	Sparassidae
Arachnida	Araneae	Tetragnathidae
Arachnida	Araneae	Zodariidae
Arachnida	Scorpiones	Buthidae
Arachnida	Trombidiformes	Erythraeidae
Arachnida	Trombidiformes	Eupodidae
Collembola	Entomobryomorpha	Entomobryidae
Collembola	Entomobryomorpha	Isotomidae
Collembola	Poduromorpha	Hypogastruridae
Collembola	Poduromorpha	Onychiuridae
Collembola	Symphypleona	Bourletiellidae
Insecta	Blattodea	Blattidae
Insecta	Blattodea	Polyzosteriinae
Insecta	Blattodea	Rhinotermitidae
Insecta	Blattodea	Termitidae
Insecta	Blattodea	Termitoidae
Insecta	Coleoptera	Carabidae
Insecta	Coleoptera	Cleridae
Insecta	Coleoptera	Coccinellidae
Insecta	Coleoptera	Curculionidae
Insecta	Coleoptera	Dytiscidae
Insecta	Coleoptera	Elateridae

Insecta	Coleoptera	Geotrupidae
Insecta	Coleoptera	Tenebrionidae
Insecta	Diptera	Asilidae
Insecta	Diptera	Cecidomyiidae
Insecta	Diptera	Chironomidae
Insecta	Diptera	Chloropidae
Insecta	Diptera	Muscidae
Insecta	Diptera	Phoridae
Insecta	Diptera	Psychodidae
Insecta	Diptera	Sciaridae
Insecta	Diptera	Sphaeroceridae
Insecta	Diptera	Tachinidae
Insecta	Diptera	Tephritidae
Insecta	Embioptera	Oligotomidae
Insecta	Hemiptera	Aphididae
Insecta	Hemiptera	Cicadellidae
Insecta	Hemiptera	Nogodinidae
Insecta	Hemiptera	Pachygronthidae
Insecta	Hemiptera	Rhopalidae
Insecta	Hemiptera	Triozidae
Insecta	Hymenoptera	Aphelinidae
Insecta	Hymenoptera	Bethylidae
Insecta	Hymenoptera	Braconidae
Insecta	Hymenoptera	Crabronidae
Insecta	Hymenoptera	Diapriidae
Insecta	Hymenoptera	Dryinidae
Insecta	Hymenoptera	Encyrtidae
Insecta	Hymenoptera	Eulophidae
Insecta	Hymenoptera	Eurytomidae
Insecta	Hymenoptera	Formicidae
Insecta	Hymenoptera	Mymaridae
Insecta	Hymenoptera	Platygastridae
Insecta	Hymenoptera	Pompilidae
Insecta	Hymenoptera	Scelionidae
Insecta	Hymenoptera	Sphecidae
Insecta	Lepidoptera	Cosmopterigidae
Insecta	Lepidoptera	Cossidae
Insecta	Lepidoptera	Erebidae
Insecta	Lepidoptera	Gelechiidae
Insecta	Lepidoptera	Lycaenidae

Insecta	Lepidoptera	Noctuidae
Insecta	Lepidoptera	Oecophoridae
Insecta	Neuroptera	Mantispidae
Insecta	Neuropterida	Berothidae
Insecta	Neuropterida	Mantispidae
Insecta	Orthoptera	Acrididae
Insecta	Orthoptera	Gryllidae
Insecta	Orthoptera	Morabidae
Insecta	Orthoptera	Phalangopsidae
Insecta	Thysanoptera	Phlaeothripidae
Insecta	Thysanoptera	Thripidae
Myriapoda	Chilopoda	Scolopendridae

This page has been left intentionally blank

CHAPTER 4

DNA METABARCODING IDENTIFIES URBAN FORAGING PATTERNS OF OLIGOLECTIC AND POLYLECTIC CAVITY-NESTING BEES

1.1 Preface

The previous chapter introduces one, and possibly the most common, use of DNA metabarcoding for invertebrate communities. Whilst broad invertebrate community biodiversity assessment is a useful tool, metabarcoding can be used to examine biodiversity at different scales. In this chapter I delve into the use of metabarcoding on a species scale, looking at the way that foraging behaviour of eight species of cavity-nesting native bees change in response to urbanisation.

There is not a lot known about Australian native bee species, with much of the literature and focus being centred around European honeybees. However, native bee biodiversity and populations are just as much, if not more threatened, by anthropogenic factors. As some native bees can be very specialised in their foraging habits, they are at considerable risk of population decline and extinction if nothing is done to preserve their habitat and food sources. Therefore, this chapter is a timely and important inclusion into the current research being conducted on native bee populations to build knowledge on foraging sources for these bees. This study represents the first use of metabarcoding to identify the contents of nesting tubes in urban environments and builds knowledge of foraging for two species of native bees where there remains no published information on their preferred host plants.

This chapter, in a manuscript format, has been accepted with major revisions to the journal *Oecologia*. The following chapter is the revised manuscript acting on two annonymous reviewer's comments that has been formatted for inclusion in the thesis.

The data from this paper has been presented in a poster format at the John Glover Symposium for the Royal Society of Western Australia in September 2021 and was awarded a poster prize, and as a presentation for the Ecological Society of Australia 2021 Conference in November 2021, where the presentation was highly commended.

1.1.1 Chapter acknowledgements

This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. KF was partially supported by funding from Food Agility CRC Ltd, funded under the Commonwealth Government CRC Program. KP was supported by a Forrest Research Foundation Scholarship. We would like to acknowledge Kristine Bohmann for their helpful comments and advice in drafting the manuscript.

1.1.2 Data accessibility

Sequencing data, Bioinformatic scripts, and sample information can be found at: 10.5281/zenodo.5564128

1.1.3 Author contributions

KF, KP, MG, MB, and PN conceptualised the study. KP supplied the samples. KF conducted the experimentation and analysis. BJS assisted with statistical analysis. KF wrote the original draft of the manuscript. All authors contributed to review and editing of the manuscript.

DNA Metabarcoding identifies urban foraging patterns of oligolectic and polylectic cavity-nesting bees

Kristen Fernandes[1,2,3*], Kit Prendergast[4], Philip W Bateman[5], Benjamin J Saunders[4], Mark Gibberd[6], Michael Bunce[1,7], Paul Nevill[1,8]

1 Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

2 Section for Molecular Ecology and Evolution, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen K, Denmark

3 Food Agility CRC Ltd, 175 Pitt St Sydney, NSW 2000, Australia

4 School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

5 Behavioural Ecology Lab, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

6 Centre for Crop and Disease Management, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

7 Environmental Science and Research (ESR), Kenepuru, Porirua 5022, New Zealand.

8 ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

* Corresponding Author: Kristen.fernandes@postgrad.curtin.edu.au

4.1 Abstract

Urbanisation modifies natural landscapes resulting in built-up space that is covered by buildings or hard surfaces and managed green spaces that often substitute native plant species with exotics. Some native bee species have been able to adapt to urban environments, foraging and reproducing in these highly modified areas. However, little is known on how the foraging ecology of native bees is affected by urbanised environments, and whether impacts vary among species with different degrees of specialisation for pollen collection. Here, we aim to investigate the responses of native bee foraging behaviour to urbanisation, using DNA metabarcoding to identify the resources within nesting tubes. We targeted oligolectic (specialist) and polylectic (generalist) cavity-nesting bee species in residential gardens and remnant bushland habitats. We were able to identify 40 families, 50 genera, and 23 species of plants, including exotic species, from the contents of nesting tubes. Oligolectic bee species had higher diversity of plant pollen in their nesting tubes in residential gardens compared to bushland habitats, along with significantly different forage composition between the two habitats. This result implies a greater degree of forage flexibility for oligolectic bee species than previously thought. In contrast, the diversity and composition of plant forage in polylectic bee nesting tubes did not vary between the two habitat types. Our results suggest a complex response of cavity-nesting bees to urbanisation and support the need for additional research to understand how the shifts in foraging resources impact overall bee health.

4.2 Introduction

Within the next century, urban areas around the world will grow rapidly, with some models predicting that by 2100 the global area of urban land will increase to 5.9 times the area that it was in 2000, to cover over 3.6 million km^2 of land (Gao and O'Neill 2020). Higher human population density and associated urbanisation can cause the loss of biodiversity and endemic species (McDonald et al. 2018), not only by clearing native vegetation, but by permanently modifying the natural landscape through the creation of built-up space (buildings, roads, and other structures) and managed green spaces (Harrison and Winfree 2015). Even in regions with high biodiversity, such as the southwest of Australia, urban green spaces have a generally higher level of plant diversity than remnant bushland, stemming from the increased planting of exotic species (Prendergast, 2020). This pattern has also been observed in other areas of the globe, such as the United Kingdom (Davies et al. 2009). However, native species relying on the pollen and nectar resources from native plants, may not always be able to access exotic floral resources in urbanised environments. This is because in regions with high endemism and species richness, ecosystem dependencies are common between groups of flora and fauna species (Johnson 2010). Additionally, there are many ornamental varieties of plants in residential gardens that offer minimal nectar or pollen rewards for insects (Corbet et al. 2001). Therefore, the clearing of native habitat in these ecosystems due to urbanisation can cause the destabilisation of dependant ecosystem networks, resulting in local extinctions and ecosystem functional collapse (Sánchez-Bayo and Wyckhuys 2019).

Overall, the impacts of urbanisation on organisms are highly varied, and how a species will respond depends on its ecological requirements, functional and life history traits, the spatial scale of investigation, geographic region, and the intensity of urbanisation (Theodorou, 2022). Species more at risk from urbanisation include specialist cavity-nesting birds, short distance migrants, and narrowly distributed species (Luck and Smallbone, 2010). Whilst increased degrees of urbanisation generally result in a decline in species diversity, paradoxically some urban areas become a refuge for native biodiversity (Goddard et al. 2010). For example, urban parks in San Francisco, USA, supported higher abundances of generalist native bumblebee (*Bombus spp.*) than parks outside of the urban area (McFrederick and LeBuhn, 2006). Additionally, populations of the European common brown frog (*Rana temporaria*) have shown increases in urban gardens and parks, whilst declining in rural areas (Carrier and Beebee, 2003). For insects, urban areas have been found to benefit cavity-nesting, small-bodied, generalist, and exotic species (Buccholz and Egerer, 2020; Fitch et al. 2019). Partially, this is due to the value of certain traits of urban gardens that can enhance the retention of biodiversity. The value of a particular urban garden for insects will depend on the built form, vegetation cover, vegetation composition, management procedures, interconnectivity with other green spaces, and human population density (Persson et al. 2020).

Native bees play a key role in functional ecosystems and maintaining their populations in urban areas is crucial. They perform pollinator services across the globe, for both crop and native plant species (Winfree et al. 2008). The survival of native bees in urban areas is dependent on species' ecology and foraging preferences: in some regions, there is evidence of co-evolution of bee species with specific native flowering plants (Phillips et al. 2010; Menz et al. 2011), implying that the loss of certain plant groups can have profound impacts on resource availability for their associated visitors. The level of forage flexibility of individual bee species will determine whether: a) the species can access a variety of floral resources (i.e., a generalist) or b) whether they are restricted to a certain group of plants (i.e., a specialist). The loss of native flora can restrict the resource availability for specialist bee species to a narrower range of available flora (Prendergast and Ollerton 2021). For many native bee species, there can also be a preferential avoidance of exotic plant species (Buchholz and Kowarik 2019). Lecty refers to the degree of trophic specialisation for pollen collection (Cane and Sipes 2006). Bees that exhibit specialisation in their diets for pollen from a particular taxon are known as "oligolectic" bees; these bees are believed to be constrained to a narrow resource breadth by physiological, temporal, or environmental factors (Fox and Morrow 1981; Devictor et al. 2010). "Polylectic" bee species, however, can feed on a wide variety of pollen sources from different families of plants.

There is a current lack of available knowledge on floral specialisation for many bee species (Bogusch et al. 2020). To capture the full spectrum of floral resources used by bees requires a combination of foraging observations and pollen analysis from netted bees (Cane and Sipes 2006). Quantification of floral resource usage by many native bees has been largely based on observation data, rather than on pollen collection (Roulston and Cane 2000; Bosch et al. 2009). Additionally, there is evidence to suggest that lecty is a spectrum, rather than binary, and that resource usage can be varied based on sex or blooming phase of preferred flowering plants (Ritchie et al. 2016). For oligolectic bees, some species have been documented to access nutrition from nectar, floral oils, or pollen from less preferred plants where preferred host plants may be rare or have limited blooming periods (Wcislo and Cane 1996). However, there is still limited understanding of how less preferred forage resources can impact reproduction or overall bee health (Filipiak and Filipiak 2020). Therefore, if conservation actions are needed to protect native bee populations, it is important to understand the preferred foraging resources and the range of forage flexibility of native bees in an area under threat.

Artificial nesting blocks - ‘trap nests’ or ‘bee hotels’ - can be beneficial in understanding foraging behaviour of solitary cavity-nesting bees (MacIvor 2017; Staab et al. 2018) and pollen-bee and host-parasitoid interactions between cavity-nesting bee taxa and the surrounding environment (Krombein 1967). Within the cavities, female bees construct brood cells, which they provision with pollen and nectar and then lay an egg. Although not all bee species use trap nests, appropriately designed trap nests can allow for the detection of a wide diversity of bee species, including both males and females of the same species that may not otherwise be observed in field surveys (Prendergast et al. 2020). Additionally, as cavity-nesting bees are central place foragers, the species that use trap nests forage in an area around the nest that is limited by their flight range (Zurbuchen et al. 2010b). This means that cavity-nesting bees can be considered indicators that help understand changes in the local environment (Tscharntke et al. 1998). Studying the larval provisions (nectar and pollen) within trap nests can be a valuable tool to understand foraging resource availability within a season. Forage resource availability for solitary bees partly determines the number, size, and sex-ratio of offspring (Pitts-Singer, 2015). This is because female bees can control the offspring sex and body size, and a shortage of resources can result in reduced maternal investment favouring the production of fewer young that require less resources, often males (Seidelmann et al. 2010). Therefore, studying the resources within trap nests can provide valuable information into the future health and functionality of changed ecosystems.

As morphological identification of plant materials requires expertise in taxonomic identification across multiple families of plants, genetic tools are being increasingly implemented to aid in pollen identification, primarily DNA metabarcoding (Pornon et al. 2016; Bell et al. 2017). The value of DNA metabarcoding is its ability to identify species accurately and rapidly, which in turn can reveal fine-scale interactions that may not be detected from the observation of pollinator-plant interactions alone (Pornon et al. 2016, 2017). This is especially useful in understanding the impacts of urbanisation on native bees, especially in regions where these species may be understudied. DNA metabarcoding works by (i) extracting DNA from environmental or bulk specimen samples, (ii) amplifying the DNA using nucleotide-labelled primers (Bohmann et al. 2021), (iii) sequencing on high-throughput sequencing platforms and identifying the resulting sequences using reference sequence databases (Taberlet et al. 2012) or via taxon-independent approaches (e.g., OTUS). For taxonomic assessment of pollen, metabarcoding has allowed simultaneous identification of plant taxa across multiple species and samples (Taberlet et al. 2012). DNA metabarcoding has been used to identify the taxonomic constituents of pollen loads from pollinators (Pornon et al. 2016, 2017; Bell et al. 2017), brood cells within trap nests (Gresty et al. 2018; Voulgari-Kokota and Ankenbrand 2019), honey (De Vere et al. 2017) and pollen traps at the entrances of beehives (Keller et al. 2015). To our knowledge, DNA metabarcoding has yet to be used to document foraging behaviour and preferences from cavity-nesting bee species in urban environments or in Australian ecosystems.

We used DNA metabarcoding of the biological material from trap nests to investigate how eight species of Australian oligolectic or polylectic cavity-nesting bees utilise forage resources in urban bushland remnants compared to residential gardens. Our hypothesis was that polylectic bees will gather a greater diversity of plant material in their trap nests compared to oligolectic bee species in both habitat types (residential gardens and bushland remnants). Further, we predict that because of the greater floral diversity in residential gardens (Prendergast and Ollerton, 2021), there will be a higher diversity and varied composition of plants collected by polylectic bee species in residential gardens compared to bushland remnants. We anticipate that in residential gardens, the forage composition within trap nests of oligolectic bees will not change due to their specialisation or be reduced because only a subset of plant species will be present.

4.3 Materials and Methods

4.3.1 Experimental design

To investigate the impacts of urbanisation on oligolectic and polylectic native bee foraging behaviour, we collected nesting tubes from trap nests from 14 sites across the Perth metropolitan region, in southwest Australia (Fig 4-1A and B). This region is a known biodiversity hotspot with high levels of species endemism and diversity but it is under threat from various anthropogenic factors, such as urbanisation (Phillips et al. 2010). The 14 sites represented two habitat types: native bushland remnants and residential gardens, seven in each (Fig 4-1A). Trap nests and the recorded habitat characteristics used in this study were sourced from a previous study investigating sampling methods for Western Australian native bees (Prendergast et al. 2020). From these trap nests, we examined eight species of native bee that had varying levels of diet specialisation (lecty). Oligolectic bees collect pollen from one plant family (specialists) and polylectic bees collect pollen from a greater diversity of plant families (generalists) (Cane & Sipes, 2006). Our study included three specialist and five generalist bee species (Table S 4-1). The oligolectic bee species were: *Megachile* (*Hackeriapis*) *canifrons* (Smith, 1853), *Megachile* (*Mitchellapis*) *fabricator* (Smith, 1868), and *Rozenapis ignita* (Smith, 1853). The polylectic bee species were: *Hylaeus* (*Euprosopis*) *violaceus* (Smith, 1853), *Megachile aurifrons* (Smith, 1853), *Megachile erythropyga* (Smith., 1853), *Megachile* (*Hackeriapis*) *oblonga* (Smith, 1879), and *Megachile* (*Hackeriapis*) *tosticauda* (Cockerell, 1912). Lecty for each species were designated based on observations of bees foraging on flowers across southwest WA from 2016 – 2021 by K. S. Prendergast (unpub), observations from Prendergast and Ollerton (2021), and, if present, records in Houston (2000). Where possible, equal numbers of nesting tubes were selected from each habitat type, this ranged from a minimum of 6 to a maximum of 10 nesting tubes for each bee species (Table S 4-1)

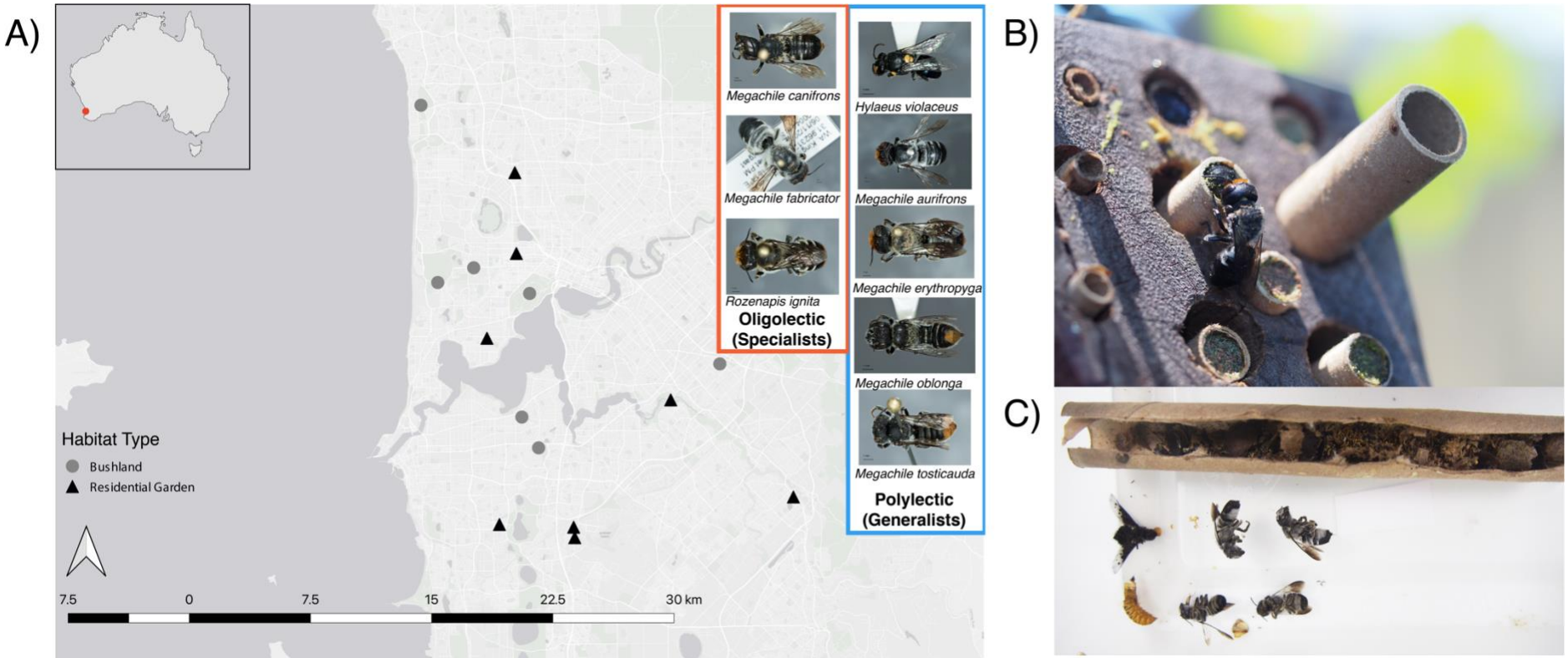


Figure 4-1 **A) Map of the study sites in Perth, Western Australia showing locations of bushland remnant (grey circle) and residential garden (black triangle) habitat types with images of bee species included in this study alongside. Species in the blue box are polylectic (generalists) whilst those in the orange box are oligolectic (specialists). Images of bees were taken by K.S. Prendergast using the WA Museum's imaging microscope and stacking software. B) Bee visitor to a trap nest. Image by K.S Prendergast. C) Inside a *M. fabricator* nesting tube showing four mature adults, larvae, a parasitic bombyliid *Anthrax* sp. fly, and the remaining pollen and plant material debris. Images by K.S Prendergast.**

4.4 Habitat Characteristics and Native Bee Observations

We used the following habitat information gathered from Prendergast et al. (2020) and Prendergast and Ollerton (2021) at each of the 14 sites to distinguish between remnant bushland and residential gardens: bare ground cover (a proxy for nesting space for ground-nesting bee species); the number of woody plants (a proxy for nesting material for cavity-nesting bee species); the total area of the site; percentage of built space; native floral species richness; the number of native flowers; and the proportion of native flowers to horticultural species both in richness and in number (for descriptions see Table S 4-2).

Floral hosts for each species were designated based on observations of bees foraging on flowers across southwest WA from 2016 – 2021 by K. S. Prendergast (unpub) and, if present, records in Houston (2000). Intertegular span (Cane 1987) was measured from dorsal stacked photos of a female of each species (Canon DLSR, 100mm lens, 1:1 magnification, f-stop 8). The images were imported into Adobe Photoshop and measured using the set measurement scale and ruler features. Intertegular span is the distance between the points where the wings attach to the thorax. It has been used as an estimate of bee size and flight abilities (Cane 1987). Greater intertegular span is a proxy for greater potential foraging distance (Wright et al. 2015). The largest bee species in our study was the oligolectic *Megachile (Mitchellapis) fabricator*, and the smallest species was the polylectic *Hylaeus violaceus* (Table S 4-1).

4.4.1 Sample Processing

Once young bees had emerged from nesting tubes, each tube was separated by site and species, constituting a sample. In total, we sampled 148 nesting tubes. Where possible, equal numbers of nesting tubes were selected for each species from each habitat type (ranging from five to ten tubes per habitat type) (Table S 4-1). Sterilised forceps were used for each sample to scrape the insides of nesting tubes of frass (larvae faecal matter), pollen and, for some species, resin debris (Fig 4-1C). Scrapings were then homogenised using a PreCellLys 24 2.8 mm Ceramic Bead Kit and a Minilys Personal Homogeniser for 3 minutes at 5000 rpm (Bertin Instruments, France).

4.4.2 DNA extraction, PCR Amplification, and Sequencing

DNA extraction was conducted using a DNeasy Plant Mini Kit on an automated Qiacube (Qiagen, Netherlands) modified with a 450 µL starting volume of digest and a 100 µL elution volume. Negative extraction controls were included for every 48 samples (n = 4).

Two plant metabarcoding assays were used to analyse the bee nesting tube contents across two gene regions of varying lengths. A shorter assay of ~30-143 bp targeting the P6 loop of the chloroplast trnL (UAA) intron (primers g and h; Taberlet et al. 2007) and a longer ~563 bp ITS2 assay (ITS2_S2F/S3R; Chen et al. 2010). Quantitative PCR (qPCR) was carried out on all samples to assess the amplification efficiency and presence of PCR inhibitors using serial dilutions of undiluted, 1:10 and 1:100. qPCR reactions were carried out in 25 µl reactions containing: 1 U of AmpliTaq gold, 1 x PCR Gold Buffer and 2 mM MgCl2 (all from Applied Biosystems, USA), 0.4 mg/mL bovine serum albumin (Fisher Biotec, Australia), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4 µM of each forward and reverse primer, 0.6 µL of 1/1000 SYBR Green (Invitrogen, USA), and 2 µL of template DNA. The qPCR conditions for trnL were as follows: 95˚C for 5 minutes, followed by 45 cycles of 95˚C for 30s, 52˚C for 30s, and 72˚C for 45s, with a final elongation at 72˚C for 10 minutes. For ITS2, the qPCR conditions were as follows: 94˚C for 5 minutes, followed by 45 cycles of 94˚C for 30s, 56˚C for 30s, and 72˚C for 45s, with a final elongation at 72˚C for 10 minutes. Negative extraction, qPCR and positive (*Brassica oleracea*, cauliflower DNA) controls were also included in the reactions. The positive control was chosen as a species that displayed optimal amplification in laboratory workflows to provide a baseline comparison for other samples. Further, this species was not anticipated to occur in any of our study sites, and therefore any sources of cross-contamination from this positive would be easily recognised in the resulting sequences (Bohmann et al. 2022).

Following qPCR, dilutions that showed the optimal level of amplification (template amount relative to any inhibition), were amplified with ‘fusion primers’, which are gene-specific primers labelled on both the forward and reverse with 6-8bp molecular identification (MID) tags coupled to Illumina sequencing adaptors. Each sample was tagged with a unique combination of forward and reverse MID tags not previously used within the laboratory, and qPCR reactions were prepared in an ultra-clean laboratory free from extracted or amplified DNA to minimise the possibility of contamination. Samples were amplified in duplicate using the qPCR conditions mentioned above to reduce the effects of PCR stochasticity (Murray et al. 2015). This included extraction and qPCR negative controls, but not qPCR positive controls. Using the qPCR results, PCR products were pooled in approximate equimolar concentration pools based on amplification curves, including negative controls. Pools were then quantified using a QIAxcel Advanced System (Qiagen) with the QIAxcel DNA High-Resolution Kit. As per the results of the quantification, sample pools were then combined in approximate equimolar ratios to create a sequencing library for each assay (trnL and ITS2). The trnL library was size-selected using a Pippin Prep 2% agarose Marker B cassette (Sage Science, USA) for fragments between 160-450bp long, and the ITS2 library was size-selected for 200-650bp on a Pippin Prep 1.5% Marker K cassette (Sage Science). Library pools were then purified using a QIAquick PCR purification kit (Qiagen) as per the manufacturer's instructions with the addition of a 5-minute incubation at room temperature before elution. The purified library was eluted in 40 µl and quantified with a QuBit (Invitrogen, USA) using double-stranded DNA high-sensitivity reagents to determine the optimal volume of the library required for sequencing. Both libraries were sequenced on an Illumina MiSeq (Illumina, USA). The trnL libraries were sequenced on a single-end 300 cycle V2 kit, and the ITS2 libraries were sequenced on a paired-end 600 cycle V3 kit as per the manufacturer's directions.

4.4.3 Bioinformatics and Sequence Processing

Unidirectional and unmerged paired-end sequencing reads were demultiplexed (assigned to their appropriate sample using the MID-tag combos) using 'Obitools' (Boyer et al. 2016) for the *trn*L dataset. To retain the paired-end data in the ITS2 dataset as unmerged reads for analysis using the 'DADA2' package (Callahan et al. 2016), demultiplexing was carried out using the default parameters in the 'insect' package (Wilkinson et al. 2018) in R v 3.6.1 (R Core Team 2019). Sequencing data was then quality filtered (*trn*L: minimum length= 50, maximum expected error = 2, no ambiguous nucleotides; ITS2: minimum length= 100, maximum expected error = 2, no ambiguous nucleotides), denoised, with paired-end reads (ITS2) merged with a minimum overlap length of 12, sequences identified as chimeras removed, and then dereplicated using the 'DADA2' package (Callahan et al. 2016) to produce Amplicon Sequence Variants (ASVs). ASVs were then curated using the 'LULU' package at default parameters (Frøslev et al. 2017). ASVs were matched to the NCBI GenBank reference database (www.ncbi.nlm.nih.gov/genbank/) using the Basic Local Alignment Search Tool (BLAST) for taxonomic assignment on a high-performance cluster computer (Pawsey Supercomputing Centre). BLAST results returned the top 10 hits with a minimum query coverage of 95% and a minimum percentage identity of 85%. These values were set based on of the poor availability of reference sequences in GenBank, and therefore improve likelihood of detection (Ryan et al. 2022; van der Heyde et al. 2020). Taxonomic assignments were made to the lowest common ancestor (LCA) using MEGAN (METAGenome Analyser v 6.13.5 (Huson 2018)) with a minimum score of 50 for *trn*L and 150 for ITS2. Plant taxa were cross-referenced to the Atlas of Living Australia (www.ala.org.au) and plant surveys of the sites (Prendergast and Ollerton 2021).

To determine the plant communities associated with the bee nesting tubes, only ASVs identified as plants (Phylum: Streptophyta) were retained in the analysis. ASV tables from both markers were then combined, retaining their ASV identity from each assay independent of taxonomy. Further filtering was then carried out on the entire data set. Any ASVs present in the negative control samples were removed using the 'phyloseq' package (McMurdie and Holmes 2013). Using the combined ASV table, a 0.05% minimum abundance filtering threshold was set within each sample to combat false, low abundance ASVs from each sample across with R 3.6.1 (R Core Team 2019). Minimum abundance filtering is equivalent to conducting rarefaction on the dataset without the need to remove low abundance samples (Prodan et al. 2020). Using 'phyloseq' (McMurdie and Holmes 2013), low occurrence ASVs with less than 5 sequences and occurring in only one sample were also removed. We removed any samples with less than 12,000 reads as this was where most samples had reached asymptote on a rarefaction curve (Fig S 4-1).

4.4.4 Statistical Analysis

To establish the differences between the two different habitat types, a one-way PERMANOVA (fixed factor of 'habitat' with two levels: 'residential garden' and 'bushland remnant') was conducted on the normalised habitat characteristic values outlined above with Euclidian distance and 9999 permutations using the PERMANOVA+ add on (Anderson et al. 2008) for PRIMER 7 (Clarke and Gorley 2015). A Principal Coordinate Analysis (PCO) based on Euclidian distance was performed on the normalised data from the measured habitat characteristics. The number of flowers and the number of native flowers were found to be co-linear variables (r=0.96), however, because of the importance of these two characters for describing the habitat, they were retained for the analysis despite collinearity. The relative contribution of each habitat characteristic to the differences between habitat types was evaluated using the strength of the correlation coefficient to the PCO axes. Vectors were plotted to illustrate the strength and direction of the association.

Statistical analysis on sequencing data was performed using R 3.6.1 (R Core Team 2019) and PRIMER 7 (Clarke and Gorley 2015). The ASV abundance matrix was converted to presence-absence data, and all plant community statistics were calculated from this matrix. As a measure of alpha diversity, ASV richness was calculated using the DIVERSE function in PRIMER 7. A Euclidean distance resemblance matrix was made. ASV richness was tested using a univariate Permutational Analysis of Variance (PERMANOVA) with three factors, Habitat (fixed, two levels), Lecty (fixed, 2 levels) and Species, (random, nested within Lecty, varying levels). A Pearson correlation test was also conducted on the observed plant ASV richness and the number of nesting tubes using R 3.6.1 (R Core Team 2019).

The effect of Habitat, Lecty and Species on the plant community composition was tested in the same way with a Permutational Multivariate Analysis of Variance (PERMANOVA) using Jaccard similarity with 9999 permutations using the PERMANOVA+ add on (Anderson et al. 2008) for PRIMER 7 (Clarke and Gorley 2015). The multivariate dispersions around the centroid for habitat was tested for each of the bee species using the PERMDISP function in the PERMANOVA+ add on (Anderson et al. 2008) for PRIMER 7 (Clarke and Gorley 2015). The plant community composition was illustrated with Non-metric Multidimensional Scaling (NMDS) using Jaccard Similarity with the 'vegan' package (Oksanen et al. 2019) and 'ggplot2' (Wickham 2016). Similarity percentage analysis (SIMPER) was used to identify plant families responsible for the differences between habitat type using PRIMER 7 (Clarke and Gorley 2015) based on the ASVs that could be identified to a family level. A distance-based linear model (DistLM) was used to characterise the relationship between the measured habitat characteristics and plant ASVs found in nesting tube contents. This model also included the factors Habitat, Lecty and Species. The DistLM was done using the BEST selection procedure and the Akaike Information Criterion with correction (AICc) selection criterion using PERMANOVA+ add on (Anderson et al. 2008) for PRIMER 7 (Clarke and Gorley 2015).

4.5 Results

4.5.1 Residential Gardens and Remnant Bushland Habitats

There was a significant difference between the habitat types (residential garden and bushland remnant) based on the measured habitat characteristics (PERMANOVA, $F_{(1,131)}$= 89.1, p=0.001). PCO showed that 69.2% of the variation among the two habitat types was explained by axes 1 and 2 (Fig 4-2). Residential gardens were associated with a greater percentage of built space and floral species richness, whilst remnant bushland was associated with the greater richness and abundance of native plant flowers and bee species, woody plants, and bare ground (Fig 4-2, see also Prendergast et al. 2021)

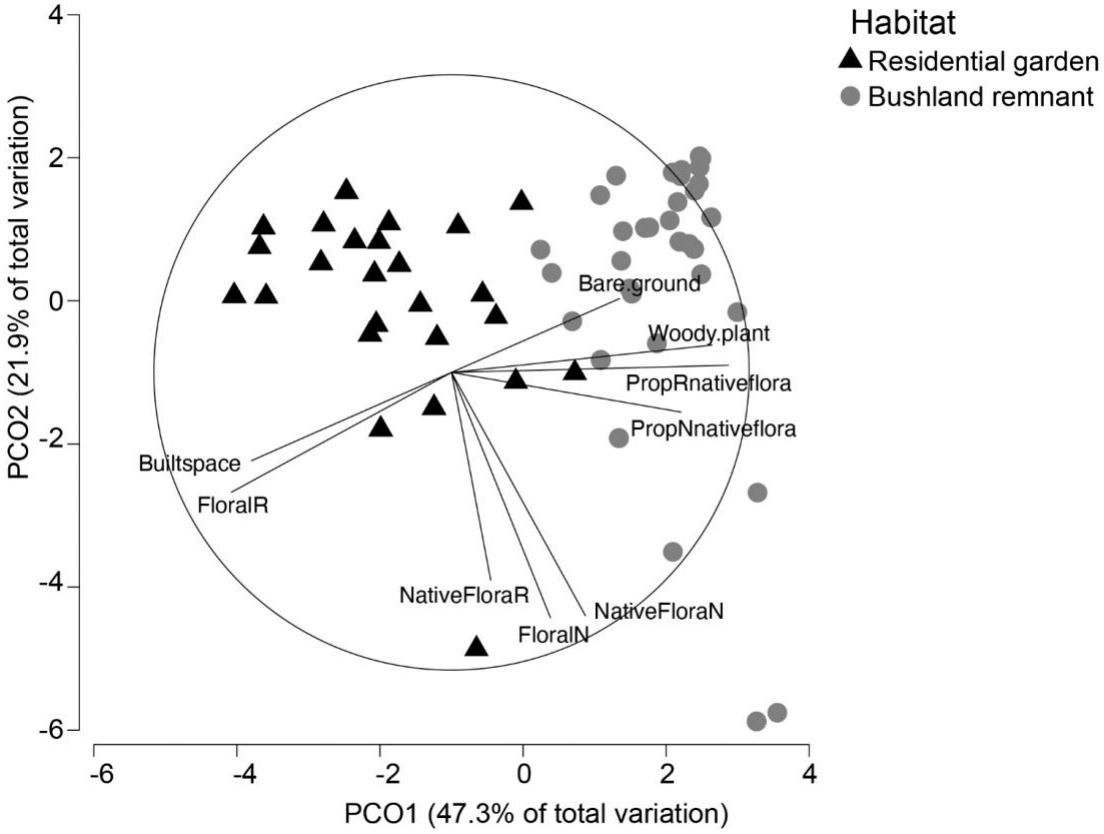


Figure 4-2 Principal Coordinates Analysis (PCO) plot of the measured habitat characteristics between bushland remnant (grey circle) and residential garden (black triangle) habitat types. The vectors plotted illustrate the strength and direction of the correlations of habitat characteristics to the PCO axes. For descriptions of abbreviations of habitat characteristics see Table S 4-3.

4.5.2 Sequencing Results

The *trn*L and the ITS2 assays generated 8,949,032 (mean = 113,114 ± 812 SE sequences per sample) and 17,419,536 (mean = 72,646 ± 501 SE sequences per sample) quality filtered and 'LULU'-curated sequences, respectively. Only six ASVs were detected within the negative control samples, two from the ITS2 assay and four from the *trn*L assay. As per Bell et al. (2017), these signified low levels of contamination either in the reagents or from sampling/laboratory workflows as the ASVs were from a subset of some of the most common taxa detected (Myrtaceae spp. and Fabaceae spp.). These ASVs were removed from further analysis. Analysis was then conducted on 14,521,974 sequences from 213 ASVs and 115 samples.

In total there were 40 families, 50 genera, and 23 species of terrestrial vascular plants detected through metabarcoding of the bee nesting tubes. The majority of the metabarcoding detections belonged to the family Myrtaceae (103 ASVs), followed by Fabaceae (23 ASVs), Poaceae (10 ASVs) and Asteraceae (10 ASVs). There were several plant families detected through metabarcoding of the nesting tubes that were not observed as floral-hosts in plant-pollinator surveys within the same geographic region (Table S 4-1). These families include both native and exotic plant species (Fig S 4-2, Table S 4-3) that are either native to the area or can be found in residential gardens or road-side verges. Here we define exotic plants as those that are exotic to Australia.

Both assays performed similarly at higher taxonomic levels, at least 99.1% of ITS2 ASVs and 88.7% of *trn*L ASVs were able to be identified to family level. This was markedly reduced at finer taxonomic levels for the *trn*L assay where only 43.9% of ASVs could be identified to genus level while 97.4% of ITS2 ASVs could be identified to genus level; however, this was predominantly *Eucalyptus* ASV detections. At a species level, both assays performed similarly with 29.6% of the ITS2 ASVs identified to species and 23.5% of *trn*L ASVs identified to species. Even though the *trn*L assay had a limited taxonomic resolution, it detected a broader range of plant families (36) than the ITS2 assay (15), with 10 plant families shared between the two (Fig S 4-2). For both ITS2 and *trn*L there was a higher relative sequence abundance (from presence-absence data) from the Myrtaceae family than any other plant family within the dataset (Fig S 4-3). However, whilst the ITS2 data was dominated by Myrtaceae sequences, the trnL dataset showed higher proportions of other families, such as Fabaceae (Fig S 4-3).

4.5.3 Native bee nest provision in residential gardens and urban bushland fragment habitats

A univariate PERMANOVA on the observed ASV richness showed a significant interaction between habitat type and lecty, as well as a main effect of both habitat and lecty (Table 4-1). There was not a significant effect of bee species, or a significant interaction between habitat and species. Post-hoc tests on the interaction of habitat and lecty identified that oligolectic (specialist) bees had greater ASV richness within their nest tubes in residential gardens than in Bushland (Table 4-1, Fig 4-3A). However, there was no significant difference in the ASV richness in gardens or bushland for polylectic (generalist) bees (Table 4-1, Fig 4-3A). Between the habitat types, residential gardens had a higher observed ASV richness (mean 35.3 ± 2.3 SE) than bushland habitat types (mean = 29.6 ± 1.4 SE, Table 4-1).

Table 4-4-1 Analysis of variance results for plant ASVs detected through metabarcoding. PERMANOVA main tests of ASV Richness and Plant ASV composition are presented. These are followed by post-hoc pairwise comparisons for the significant interaction terms, Habitat x Lecty for ASV richness and Habitat x Species for Plant ASV composition. Habitat types are Residential Gardens or Remnant Bushland. P-Adj values are adjusted with a post-hoc Holm-Bonferroni correction due to multiple comparisons. * Indicates significance at α = 0.05.

		ASV Richness				Plant ASV composition			
Term	**DF**	**MS**	**Pseudo-F**	**P(perm)**	**Unique perms**	**MS**	**Pseudo-F**	**P(perm)**	**Unique perms**
Habitat	1	1678	22.792	0.001*	9862	13853	3.19	0.018*	9959
Lecty	1	1467	4.809	0.052	8538	7169	0.92	0.495	8815
Species (Lecty)	6	306	1.319	0.257	9938	7791	2.95	<0.001*	9786
Habitat x Lecty	1	1508	20.482	0.001*	9864	4691	1.08	0.390	9956
Habitat x Species(Lecty)	6	73	0.315	0.929	9957	4347	1.64	<0.001*	9748
Residual	117	232				2645			

Post-hoc pairwise comparisons

		Observed ASV Richness Lecty x Habitat			Plant ASV Composition Species x Habitat		
Bee species	**Lecty (oligolectic (specialist) vs polylectic (generalist))**	**t value**	**P value**	**P-Adj**	**t value**	**P value**	**P-Adj**
Megachile canifrons	Oligolectic	4.843	0.036*	0.072	1.324	0.039*	0.156
Megachile fabricator					1.904	<0.001*	0.008*
Rozenapis ignita					1.769	<0.001*	0.008*
Hylaeus violaceus	Polylectic	0.229	0.829	0.829	1.247	0.089	0.267
Megachile aurifrons					1.562	0.001*	0.008*
Megachile erythropyga					1.311	0.012*	0.060
Megachile oblonga					1.189	0.096	0.267
Megachile tosticauda					1.249	0.098	0.267

The PERMANOVA on the composition of plant ASVs in nesting tubes found that there was a significant interaction of bee species and habitat type (Table 4-1, Fig 3-3B), and also differences between habitat type and bee species (Table 4-1). Further assessment of the interaction difference in the composition of plant taxa ASVs detected from residential garden and bushland nesting tubes for all the oligolectic species (*M. canifrons*, *M. fabricator*, and *R. ignita*) but only two of the polylectic species (*M. erythropyga* and *M. aurifrons*) (Fig 3-3B, Table 4-1).

Analysis of multivariate dispersions (PERMDISP) indicated no significant difference in the diversity of forage between residential gardens and bushland remnants for most bee species ($p>0.05$). The exception was the polylectic species *M. oblonga* ($F_{(1,15)}=6.562$, $p=0.021$), with forage ASVs in the Residential Gardens being less variable (mean dispersion 46.76 ± 1.68 SE) than in the Bushland Remnants (52.74 ± 1.61 SE).

SIMPER analysis indicated that Myrtaceae and Fabaceae were the most common detections in residential gardens and bushland remnant habitat types (Table S 4-4). This was to be expected as Myrtaceae and Fabaceae were the most common plant families detected across the dataset (Fig S 4-3). Results from the SIMPER analysis based on the plant families of detected ASVs indicated that there was an observed decrease in the frequency of Fabaceae ASVs contributing to the similarity within residential gardens (Table S 4-4).

We found that although there were some habitat characteristics that had statistically significant relationships with the observed plant ASV composition, these variables could only explain a very low percentage of the variation in composition (Table S 4-5). The overall BEST solution indicated that the three factors Habitat, Species and Lecty, together explained 17% of the variation in plant ASV composition. Additional variables that were included in alternative models within 2 AICc of the BEST model were related to the number of plant species or the number of native plant species (floral richness, native floral richness, proportion of richness which is native flora) or distance to bushland.

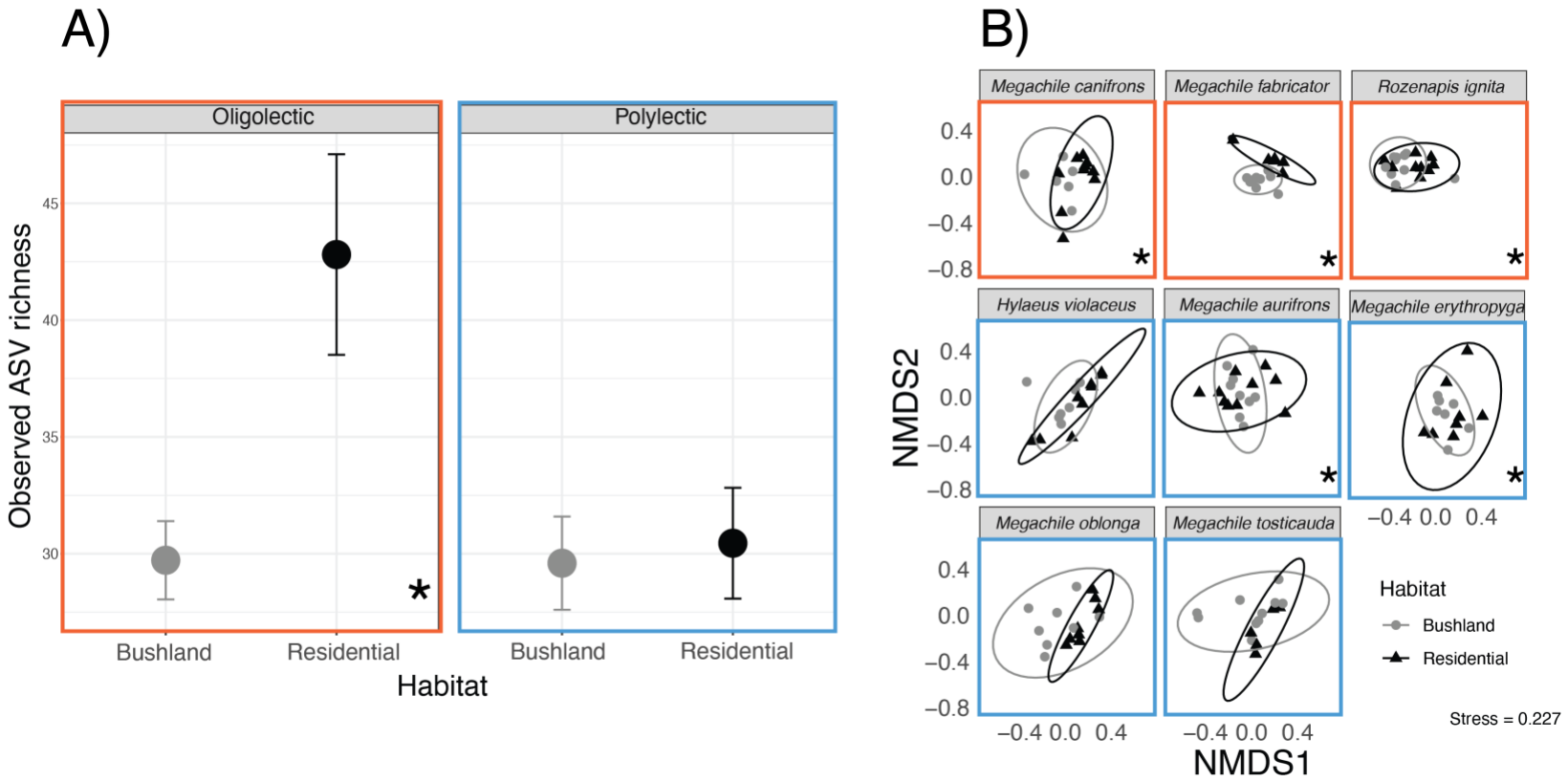


Figure 4-3 **A) Mean observed ASV richness (± S.E.) of plant taxa detected within nesting tubes of oligolectic (specialists, orange box) and polylectic (generalists, blue box) bees for bushland remnant (grey circle) and residential garden (black circle) habitat type. Asterisk (*) indicates pairwise significant difference (α = 0.05). B) Non-metric Multi-dimensional Scaling (NMDS) plot with Jaccard Similarity showing species of bees and the forage composition of nesting tubes between those in bushland remnant (grey circle) and residential garden (black triangle) habitat type. 95% confidence intervals illustrated with circles corresponding to colour of bushland remnants (grey) and residential gardens (black). Species in the blue box are polylectic whilst those in the orange box are oligolectic.**

4.6 Discussion

Our study showed that eDNA metabarcoding can reveal the contents of nesting tubes, using eight native, cavity-nesting bee species in bushland remnants and residential gardens in Western Australia. Contrary to our hypothesis, oligolectic (specialist) bee species identified in our study (designations defined by Houston, 2000 and Prendergast and Ollerton, 2021) showed significantly higher species richness of plant hosts in residential gardens than in bushland remnant sites, and all our oligolectic species showed significantly different forage composition between habitat types. In comparison, for the majority of polylectic bee species there was no significant difference between habitat types in richness or forage composition derived from the nesting tubes. This suggests a much more complex response of native bee species to urbanisation than previously thought.

4.6.1 Urban survival and forage flexibility for oligolectic bee species

Contrary to our hypothesis, the oligolectic bee species had different ASV richness between habitat types, with higher richness observed in the residential gardens. There are several potential explanations for this. Higher diversity could indicate greater availability of forage in these habitat types for native bees, although, considering the co-evolution of native bees to their native host plants (Houston 2000; Phillips et al. 2010) and that these species are oligolectic, this seems unlikely. Instead, we suggest that this is an indicator of lower availability of *preferred* resources. This was supported by the increase in similarity of forage composition within residential gardens for these bees. As a result, we suggest that even these specialist native bees can expand their diet breadth to meet their resource requirements in suboptimal habitats. It should be noted that these oligolectic bee species were chosen as they were commonly found in our residential garden and remnant bushland study sites and therefore allowed us to achieve an adequate sampling size. As such, these species could be considered 'urban adapters' (McKinney 2002) in these spaces, as they have broad ecological adaptations that have positively translated in urban environments to allow them to forage and reproduce efficiently enough to allow populations to be maintained. One generalised adaptation is that the oligolectic bee species in our study could diversify their forage sources, indicating phenotypic plasticity (also known as behavioural flexibility). Behavioural flexibility is an important characteristic required for animals to be successful in urban environments (Lowry et al. 2012). This finding is supported by previous observations in these same residential garden sites, where native bees would visit native plants even if they were not native to the local area (Prendergast and Ollerton 2021). Phenotypic plasticity, generalisation, and dispersal ability have been identified as important characteristics required for survival in urban environments (Santini et al. 2019).

The oligolectic bee species in our study also had a generally larger intertegular span than did the polylectic species in this study, indicating that they can theoretically fly farther (Wright et al. 2015) to meet resource requirements. They may therefore be able to increase the diversity of their forage, as reflected by the contents of the nesting tubes. With residential gardens in our study characterised by higher floral richness and increased built space than bushland remnants, this could mean that these larger oligolectic bees were able to navigate through these spaces to find adequate forage resources. The habitats within our study were only surveyed within a 100 x 100 m area and bees have been documented to forage from 300 m – 1km depending on their body size (Greenleaf et al. 2007). This is supported by previous research where bee communities in urbanised, fragmented vegetation were dominated by bee species with a greater flight range than in nature reserve areas (Hung et al. 2019). However, longer flight distance to forage for resources may reduce fitness of solitary bees by reducing their offspring production (Zurbuchen et al. 2010a) and lifespan due to the wear and stress posed on the exoskeleton and flight muscles (Torchio and Tepedino 1980). While these oligolectic bees may survive in urban areas, there may be unknown physiological and reproductive consequences to living in urbanised areas that could impact the overall health of these bee populations.

In contrast, only two of our polylectic bee species, showed no significant differences in the forage composition between habitat types. Although there are more exotic plant species in residential gardens than in bushland fragments, residential gardens were not devoid of native flowering plant species (Prendergast 2020). This suggests that the generalist bee species can access the same range of forage in residential gardens that they would in bushland remnants. These results reflect those of Buchholz et al. (2020) who found that urbanisation leads to an increase in the number of polylectic bee species. However, even though the polylectic bee species *M. oblonga* showed no significant difference in the ASV composition of its forage between habitat types, there was significantly smaller dispersion observed for the ASVs in the residential gardens than in the bushland habitats. This indicates reduced diversity of forage availability for this species in urban areas. Similarly, oligolectic bee species with significant differences in forage richness between habitat types also demonstrated a significant difference in the composition of forage in nesting tubes. A significant difference in composition could indicate that these bee species are able to access the varying resources - exotic or ornamental native plant species - available in urban environments, even if these foraging sources may not be preferred. As lecty is also considered through family level specialisation, this might mean that oligolectic bee species are feeding from multiple species within a plant family. Further, the distinction between urban and bushland environments in forage resources, especially for oligolectic bees, can suggest that these species are having to change their foraging behaviour to a higher degree than the polylectic species that showed no effect.

For both the oligolectic *M. canifrons* and *M. fabricator*, composition of forage resources in nesting tubes was characterised by *Eucalyptus* ASVs (family Myrtaceae), which is a common native genus and frequent in horticultural plantings (Prendergast and Ollerton 2021). Myrtaceae ASVs also contributed to a significant percentage of the similarity in residential gardens, potentially in the absence of preferred Fabaceae forage. Whilst *M. canifrons* and *M. fabricator* are Fabaceae specialists, lecty specialisation refers to pollen specialisation and not nectar (Cane and Sipes 2006); it may be that these additional plant taxa recorded in the specialist bees' tubes represent DNA from nectar sources. One of the limitations of the current methods in that they cannot accurately quantify the relative proportions of plant species present, nor determine whether the sources were derived from nectar or pollen foraging. Further, these detections could also represent resin gathered from *Eucalyptus* trees to create partitioning between brood cells in nesting tubes (Houston 2000). Additional research is required to determine the fitness consequences, if any, of how these differences in pollen diversity and composition in nesting tubes affect the native bee progeny (Filipiak and Filipiak 2020).

4.6.2 DNA metabarcoding for taxonomic identification of plants

Prior studies have shown that DNA metabarcoding of pollen samples is simpler and provides greater taxonomic resolution than does traditional palynological approaches (Galimberti et al. 2014; Bell et al. 2017). However, this approach is not without limitations. Assays targeting shorter DNA fragments have been recommended for metabarcoding studies because this DNA can be heavily degraded (Taberlet et al. 2007, 2012), but short fragments may lack the resolving power to discriminate at finer taxonomic levels (Pornon et al. 2016). The ITS2 region has been previously suggested as a useful region for molecular identification of eukaryotes because it has fairly conserved regions across many taxonomic groups and contains a great deal of variability to distinguish closely related species (Chen et al. 2010; Yao et al. 2010). Nevertheless, both the assays used in our study showed limited species-level identification. This might also be explained by inadequate taxonomic representation in reference databases (Gous et al. 2021), which are limited for many floral taxonomic groups in Australia (Dormontt et al. 2018). Therefore, to compare the richness and composition of forage between bee species, we left ASVs independent of their taxonomy. This approach has been found to be an accurate proxy to estimate species diversity in the absence of adequate reference sequence databases (Ashfaq and Hebert, 2016; Gálvez-Reyes et al. 2020). However, taxonomic identification is still crucial for conservation because species-level identification is important for effective conservation and management. These findings support the need for increased coverage of reference databases across a variety of Australian plant taxonomic groups to aid molecular taxonomic assignment (Bell et al. 2016).

The shorter *trn*L assay (~30-143 bp) detected a much wider range of plant families than did the ITS2 assay (~563 bp), which could suggest that larval digestion or other environmental factors may have degraded the eDNA and thus favours short amplicons. Dietary analysis using ITS2 plant assays is somewhat problematic as the amplicon length is too large to be reliably detected in degraded dietary items (Moorhouse-Gann et al. 2018). Further, interpretation of ITS2 data presents a challenge because of paralogous gene copies (Hollingsworth et al. 2011; Moorhouse-Gann et al. 2018), which may be a particular issue for eucalypts (Bayly et al. 2008). The ITS2 assay had very high numbers of *Eucalyptus* ASVs, which appeared to amplify in our samples preferentially. Thus, there needs to be a balance between taxonomic resolution and taxonomic breadth when choosing assays for metabarcoding studies. Therefore, we suggest a multi-assay approach, such as ours, to better distinguish plant communities from insect-gathered pollen (Pornon et al. 2016). Additionally, the ongoing development of group-specific (e.g. family) assays may help complement the use of assays such as ITS2 and *trn*L whose role is to provide a high-level assessment.

While there are no known visual observations of any of the cavity-nesting bee species in our study foraging on members of Poaceae (Houston 2000; Prendergast 2020; Prendergast and Ollerton 2021), Poaceae ASVs were detected from 48 out of the 114 samples, equally among bee species and habitat types. In addition, a recent study that used metabarcoding of pollen from Australian native beehives similarly found unexpected detections of Poaceae (Wilson et al. 2021). Although Wilson et al. (2021) propose that these detections represent actual foraging activity, the results from the previous pollinator surveys at the sites in our study (Prendergast 2020; Prendergast and Ollerton 2021) do not support this hypothesis. Additionally, Poaceae constitutes a large proportion of total airborne pollen (Brennan et al. 2019), which suggests that Poaceae detections in our samples may instead have been airborne. Likewise, we cannot discern whether the detection of exotic plant ASVs from the nesting tubes represents actual foraging activity or background environmental accumulation. Previous observations from pollinator surveys showed no interactions between native bee species and exotic plants (Prendergast 2020; Prendergast and Ollerton 2021). Still, pollinator surveys undertaken through visual observation can be affected by bias based on observer, method, and context (O'Connor et al. 2019). Therefore, these exotic plant detections from nesting tubes represent directions for future research to explore the value of exotic plant species as a foraging resource for endemic native bees. For example, collection of data across different seasons to explore the persistence of the signal, and/or dissection of gastro-intestinal tracts directly from bees to avoid environmental background from the nesting tubes.

Our finding of a differential response of oligolectic and polylectic bee species to urbanisation adds to growing recognition that not all bees respond uniformly to ecosystem changes (Banaszak-Cibicka and Żmihorski 2012; Rader et al. 2014; De Palma et al. 2015). The bee species in our study were chosen as polylectic and oligolectic species that readily use urban environments, and these designations were defined through observation of their foraging behaviour in these environments (Prendergast and Ollerton, 2021; Prendergast et al. 2021). The oligolectic bee species in our study demonstrated a shift from their preferred forage in bushland remnants to forage that was available in residential gardens. This same shift was not observed for polylectic species. Therefore, these species represent urban adapters in this region, with a degree of plasticity in their foraging preferences and resources. The shift in these resources currently has unknown future impacts for the health of bee species in urbanised areas. Previous work on cavity-nesting bees advocates for increasing the diversity of native forage available, especially in anthropogenically impacted landscapes (Gresty et al. 2018). However, there is little knowledge available on species-specific ecology and preferred host plants in urban environments for many Australian bee species, as most studies have been conducted in Europe or the Americas (Staab et al. 2018; Wenzel et al. 2020). Although studies on Australian native bees are increasing (Threlfall et al. 2015; Prendergast and Ollerton 2021; Prendergast et al. 2021), it is still imperative to continue research into the natural history of bee species here and around the globe.

The use of DNA metabarcoding can provide a valuable complement to dietary and other species-interaction studies through the ability to rapidly identify the composition of forage resources collected by bees and other organisms. As areas become more urbanised, future research on the impacts that changes in forage availability and composition have on the health and reproduction of fauna will be invaluable in conserving native fauna populations. For example, where metabarcoding was applied to the faecal material of songbirds, differences in diet in urban areas could be linked to decreases in offspring growth (Jarrett, et al. 2020). Such research builds upon our knowledge of which species will be able to survive in urban environments, and which will completely avoid these areas or disappear from them. It is important that a more nuanced approach is taken to studying foraging preferences. For native bees, our results support the idea that lecty is a spectrum (Ritchie et al. 2016) and an individual species' behavioural flexibility will have an influence on their survival in urban areas. Understanding the complexities of foraging behaviour in different organisms will be an important part of designing interventions to mitigate threats and build healthier urban ecosystems that can support high biodiversity.

4.7 References

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged

Anderson M, Gorley R, Clarke KP (2008) PERMANOVA+ for PRIMER: guide to software and statistical methods. PRIMER-E Ltd.

Ashfaq, M, Hebert, PDN (2016) DNA barcodes for bio-surveillance: regulated and economically important arthropod plant pests. Genome, 59:933–945.

Banaszak-Cibicka W, Żmihorski M (2012) Wild bees along an urban gradient: winners and losers. J Insect Conserv 16:331–343. https://doi.org/10.1007/s10841-011-9419-2

Bayly MJ, Udovicic F, Gibbs AK, Parra-O C, Ladiges PY (2008) Ribosomal DNA pseudogenes are widespread in the eucalypt group (Myrtaceae): implications for phylogenetic analysis. Cladistics 24:131–146. https://doi.org/10.1111/j.1096-0031.2007.00175.x

Bell KL, de Vere N, Keller A, Richardson RT, Gous A, Burgess KS, Brosi BJ (2016) Pollen DNA barcoding: current applications and future prospects. Genome 59:629–640. https://doi.org/10.1139/gen-2015-0200

Bell KL, Fowler J, Burgess KS, Dobbs EK, Gruenewald D, Lawley B, Morozumi C, Brosi BJ (2017) Applying Pollen DNA Metabarcoding to the Study of Plant–Pollinator Interactions. Appl Plant Sci 5:1600124. https://doi.org/10.3732/apps.1600124

Bogusch P, Bláhová E, Horák J (2020) Pollen specialists are more endangered than non-specialised bees even though they collect pollen on flowers of non-endangered plants. Arthropod Plant Interact 14:759–769. https://doi.org/10.1007/s11829-020-09789-y

Bohmann, K, Elbrecht, V, Carøe, C, Bista, I, Leese, F, Bunce, M, Yu, DW, Seymour, M, Dumbrell, AJ, Creer, S (2022) Strategies for sample labelling and library preparation in DNA metabarcoding studies. Mol Ecol Resour 22:1231– 1246. https://doi.org/10.1111/1755-0998.13512

Bosch J, Martín González AM, Rodrigo A, Navarro D (2009) Plant-pollinator networks: adding the pollinator's perspective. Ecol Lett 12:409–419. https://doi.org/10.1111/j.1461-0248.2009.01296.x

Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) obitools: A unix-inspired software package for DNA metabarcoding. Mol Ecol Resour https://doi.org/10.1111/1755-0998.12428

Brennan GL, Potter C, de Vere N, Griffith GW, Skjøth CA, Osborne NJ, Wheeler BW, McInnes RN, Clewlow Y, Barber A, Hanlon HM, Hegarty M, Jones L, Kurganskiiy A, Rowney FM, Armitage C, Adams-Groom B, Ford CR, Petch GM, Creer S (2019) Temperate airborne grass pollen defined by spatio-temporal shifts in community composition. Nat Ecol Evol 3:750–754. https://doi.org/10.1038/s41559-019-0849-7

Buchholz, S, Egerer, MH (2020) Functional ecology of wild bees in cities: towards a better understanding of trait-urbanization relationships. Biodiver Conserv, 29: 2779-2801. https://doi.org/10.1007/s10531-020-02003-8

Buchholz S, Gathof AK, Grossmann AJ, Kowarik I, Fischer LK (2020) Wild bees in urban grasslands: Urbanisation, functional diversity and species traits. Landsc Urban Plan 196:103731. https://doi.org/10.1016/j.landurbplan.2019.103731

Buchholz S, Kowarik I (2019) Urbanisation modulates plant-pollinator interactions in invasive vs. native plant species. Sci Rep 9:6375. https://doi.org/10.1038/s41598-019-42884-6

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583. https://doi.org/10.1038/nmeth.3869

Cane JH (1987) Estimation of Bee Size Using Intertegular Span (Apoidea). Journal of the Kansas Entomological Society 60:145–147

Cane JH, Sipes S (2006) Characterizing floral specialization by bees: analytical methods and a revised lexicon for oligolecty. In Waser N, Ollerton J (eds) *Plant-pollinator interactions: from specialization to generalization*. University of Chicago Press, pp 99-122

Carrier, JA, Beebee, TJ (2003) Recent, substantial, and unexplained declines of the common toad *Bufo bufo* in lowland England. Biol Conserv 111: 395-399. https://doi.org/10.1016/S0006-3207(02)00308-7

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y, Leon C (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5:e8613. https://doi.org/10.1371/journal.pone.0008613

Clarke KR, Gorley R (2015) PRIMER version 7: User manual/tutorial. PRIMER-E, Plymouth

Corbet, SA, Bee, J, Dasmahapatra, K, Gale, S, Gorringe, E, La Ferla, B, Moorhouse, T, Trevail, A, Van Bergen, Y, Vorontsova, M (2001) Native or exotic? Double or single? Evaluating plants for pollinator-friendly gardens. Ann Bot-London 87:219–232. https://doi.org/10.1006/anbo.2000.1322

Davies ZG, Fuller RA, Loram A, Irvine KN, Sims V, Gaston KJ (2009) A national scale inventory of resource provision for biodiversity within domestic gardens. Biol Conserv 142:761–771. https://doi.org/10.1016/j.biocon.2008.12.016

De Palma A, Kuhlmann M, Roberts SPM, Potts SG, Börger L, Hudson LN, Lysenko I, Newbold T, Purvis A (2015) Ecological traits affect the sensitivity of bees to land-use pressures in European agricultural landscapes. J Appl Ecol 52:1567–1577. https://doi.org/10.1111/1365-2664.12524

De Vere N, Jones LE, Gilmore T, Moscrop J, Lowe A, Smith D, Hegarty MJ, Creer S, Ford CR (2017) Using DNA metabarcoding to investigate honey bee foraging reveals limited flower use despite high floral availability. Sci Rep 7:42838. https://doi.org/10.1038/srep42838

Devictor V, Clavel J, Julliard R, Lavergne S, Mouillot D, Thuiller W, Venail P, Villéger S, Mouquet N (2010) Defining and measuring ecological specialization. J Appl Ecol 47:15–25. https://doi.org/10.1111/j.1365-2664.2009.01744.x

Dormontt EE, van Dijk K-J, Bell KL, Bigffin E, Breed MF, Byrne M, Caddy-Retalic S, Encinas-Viso F, Nevill PG, Shapcott A, Young JM, Waycott M, Lowe AJ (2018) Advancing DNA Barcoding and Metabarcoding Applications for Plants Requires Systematic Analysis of Herbarium Collections—An Australian Perspective. Front Ecol Evol 6:134. https://doi.org/10.3389/fevo.2018.00134

Filipiak ZM, Filipiak M (2020) The Scarcity of Specific Nutrients in Wild Bee Larval Food Negatively Influences Certain Life History Traits. Biology 9:462 https://doi.org/10.3390/biology9120462

Fitch, G, Glaum, P, Simao, MC, Vaidya, C, Matthijs, J, Iuliano, B, Perfecto, I (2019) Changes in adult sex ratio in wild bee communities are linked to urbanization. Sci Reports 9:1-10. https://doi.org/10.1038/s41598-019-39601-8

Fox LR, Morrow PA (1981) Specialization: species property or local phenomenon? Science 211:887–893. https://doi.org/10.1126/science.211.4485.887

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nat Commun 8:1188. https://doi.org/10.1038/s41467-017-01312-x

Galimberti A, De Mattia F, Bruni I, Scaccabarozzi D, Sandionigi A, Barbuto M, Casiraghi M, Labra M (2014) A DNA barcoding approach to characterize pollen collected by honeybees. PLoS One 9:e109363. https://doi.org/10.1371/journal.pone.0109363

Gálvez-Reyes, N, Arribas, P, Andujar, C, Emerson, B. Piñero, D, Mastretta-Yanes, A (2020) Local-scale dispersal constraints promote spatial structure and arthropod diversity within a tropical sky-island. Authorea Preprints. doi:10.22541/au.160193334.45224582

Gao J, O'Neill BC (2020) Mapping global urban land for the 21st century with data-driven simulations and Shared Socioeconomic Pathways. Nat Commun 11:2302. https://doi.org/10.1038/s41467-020-15788-7

Goddard, MA, Dougill, AJ, Benton, TG (2010) Scaling up from gardens: biodiversity conservation in urban environments. Trends Ecol Evol 25: 90-98. https://doi.org/10.1016/j.tree.2009.07.016

Gous A, Eardley CD, Johnson SD, Swanevelder DZH, Willows-Munro S (2021) Floral hosts of leaf-cutter bees (Megachilidae) in a biodiversity hotspot revealed by pollen DNA metabarcoding of historic specimens. PLoS One 16:e0244973. https://doi.org/10.1371/journal.pone.0244973

Greenleaf, SS, Williams, NM, Winfree, R, Kremen, C (2007) Bee foraging ranges and their relationship to body size. Oecologia, 153: 589-596.

Gresty CEA, Clare E, Devey DS, Cowan RS, Csiba L, Malakasi P, Lewis OT, Willis KJ (2018) Flower preferences and pollen transport networks for cavity-nesting solitary bees: Implications for the design of agri-environment schemes. Ecol Evol 8:7574–7587. https://doi.org/10.1002/ece3.4234

Harrison T, Winfree R (2015) Urban drivers of plant-pollinator interactions. Funct Ecol

Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLoS One 6:e19254. https://doi.org/10.1371/journal.pone.0019254

Houston T (2000) Native bees on wildflowers in Western Australia : a synopsis of native bee visitation of wildflowers in Western Australia based on the bee collection of the Western Australian Museum. Western Australian Museum, Western Australia.

Hung K-LJ, Ascher JS, Davids JA, Holway DA (2019) Ecological filtering in scrub fragments restructures the taxonomic and functional composition of native bee assemblages. Ecology 100:e02654. https://doi.org/10.1002/ecy.2654

Huson D (2018) User Manual for MEGAN V6.13.5. Version 6.13.5. https://www.researchgate.net/profile/Daniel_Huson/publication/251816417_User_Manual_for_MEGAN_V35internal6/links/546cb2b10cf284dbf190e96f/User-Manual-for-MEGAN-V35internal6.pdf

Jarrett, C, Powell, LL, McDevitt, H, Helm, B, Welch, AJ (2020) Bitter fruits of hard labour: diet metabarcoding and telemetry reveal that urban songbirds travel further for lower-quality food. Oecologia 193:377–388. https://doi.org/10.1007/s00442-020-04678-w

Johnson SD (2010) The pollination niche and its role in the diversification and maintenance of the southern African flora. Philos Trans R Soc Lond B Biol Sci 365:499–516. https://doi.org/10.1098/rstb.2009.0243

Keller A, Danner N, Grimmer G, Ankenbrand M, von der Ohe K, von der Ohe W, Rost S, Härtel S, Steffan-Dewenter I (2015) Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. Plant Biol 17:. https://doi.org/10.1111/plb.12251

Krombein KV (1967) Trap-nesting wasps and bees. Washington, DC Smithsonian Inst Press, pp 570.

Lowry, H, Lill, Wong, BBM (2013) Behavioural responses of wildlife to urban environments. Biol Rev, 88: 537-549. https://doi.org/10.1111/brv.12012

Luck, GW, Smallbone, LT (2010) Species diversity and urbanisation: patterns, drivers and implications. In Gaston, KJ (eds) *Urban Ecology*. Cambridge University Press, pp 88-117.

MacIvor JS (2017) Cavity-nest boxes for solitary bees: a century of design and research. Apidologie 48:311–327. https://doi.org/10.1007/s13592-016-0477-z

McDonald RI, Colbert M, Hamann M, Simkin R, Walsh B (2018) Nature in the urban century: a global assessment of where and how to conserve nature for biodiversity and human wellbeing. The Nature Conservancy, Stockholm.

McFrederick, QS, LeBuhn, G (2006) Are urban parks refuges for bumble bees *Bombus spp.* (Hymenoptera: Apidae)? Biol Conserv 129: 372-382. https://doi.org/10.1016/j.biocon.2005.11.004

McKinney ML (2002) Urbanization, Biodiversity, and Conservation: The impacts of urbanization on native species are poorly studied, but educating a highly urbanized human population about these impacts can greatly improve species conservation in all ecosystems. Bioscience 52:883-890 https://doi.org/10.1641/0006-3568(2002)052[0883:UBAC]2.0.CO;2

McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217. https://doi.org/10.1371/journal.pone.0061217

Menz MHM, Phillips RD, Winfree R, Kremen C, Aizen MA, Johnson SD, Dixon KW (2011) Reconnecting plants and pollinators: challenges in the restoration of pollination mutualisms. Trends Plant Sci 16:4–12. https://doi.org/10.1016/j.tplants.2010.09.006

Moorhouse-Gann RJ, Dunn JC, de Vere N, Goder M, Cole N, Hipperson J, Symondson WOC (2018) New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. Sci Rep 8:8542. https://doi.org/10.1038/s41598-018-26648-2

Murray DC, Coghlan ML, Bunce M (2015) From benchtop to desktop: important considerations when designing amplicon sequencing workflows. PLoS One 10:e0124671. https://doi.org/10.1371/journal.pone.0124671

O'Connor RS, Kunin WE, Garratt MPD, Potts SG, Roy HE, Andrews C, Jones CM, Peyton Jm, Savage J, Harvey MC, Morris RK, Roberts SP, Wright I, Vanbergen AJ, Carvell C (2019) Monitoring insect pollinators and flower visitation: The effectiveness and feasibility of different survey methods. Methods Ecol Evol 10:2129–2140. https://doi.org/10.1111/2041-210x.13292

Oksanen J, Guillaume Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens HH, Szoecs E, Wagner H (2019) vegan: Community Ecology Package. Version 2.5-6 https://cran.ism.ac.jp/web/packages/vegan/vegan.pdf

Persson, AS, Ekroos, J, Olsson, P, Smith, HG (2020) Wild bees and hoverflies respond differently to urbanisation, human population density and urban form. Landsc Urban Plan, 204:103901.

Phillips RD, Hopper SD, Dixon KW (2010) Pollination ecology and the possible impacts of environmental change in the Southwest Australian Biodiversity Hotspot. Philos Trans R Soc Lond B Biol Sci 365:517–528. https://doi.org/10.1098/rstb.2009.0238

Pitts-Singer, TL (2015) Resource effects on solitary bee reproduction in a managed crop pollination system. Environ Entomol 44:11251138. https://doi.org/10.1093/ee/nvv088

Pornon A, Andalo C, Burrus M, Escaravage N (2017) DNA metabarcoding data unveils invisible pollination networks. Sci Rep 7:16828. https://doi.org/10.1038/s41598-017-16785-5

Pornon A, Escaravage N, Burrus M, Holota H, Khimoun A, Mariette J, Pelizzari C, Iribar A, Etienne R, Taberlet P, Vidal M, Winterton P, Zinger L, Andalo C (2016) Using metabarcoding to reveal and quantify plant-pollinator interactions. Sci Rep 6:27282. https://doi.org/10.1038/srep27282

Prendergast K (2020) Plant-pollinator network interaction matrices, and flowering plant species composition, in urban bushland remnants and residential gardens in the southwest Western Australian biodiversity hotspot [Data set]. *Plant-pollinator networks in Australian urban bushland remnants are not structurally equivalent to those in residential gardens*. doi:10.25917/5f3a0aa235fda

Prendergast K, Dixon KW, Bateman PW (2021) Interactions between the introduced European honey bee and native bees in urban areas varies by year, habitat type and native bee guild. Biol J Linn Soc Lond. https://doi.org/10.1093/biolinnean/blab024

Prendergast K, Menz MHM, Dixon KW, Bateman PW (2020) The relative performance of sampling methods for native bees: an empirical test and review of the literature. Ecosphere 11:206. https://doi.org/10.1002/ecs2.3076

Prendergast K, Ollerton J (2021) Plant-pollinator networks in Australian urban bushland remnants are not structurally equivalent to those in residential gardens. Urban Ecosyst. https://doi.org/10.1007/s11252-020-01089-w

Prodan A, Tremaroli V, Brolin H, Zwinderman AH, Nieuwdorp M, Levin E (2020) Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. PLoS One 15:e0227434. https://doi.org/10.1371/journal.pone.0227434

R Core Team (2019) R: A Language and Environment for Statistical Computing. Version 3.6.1. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rader R, Bartomeus I, Tylianakis JM, Laliberté E (2014) The winners and losers of land use intensification: pollinator community disassembly is non-random and alters functional diversity. Divers Distrib 20:908–917. https://doi.org/10.1111/ddi.12221

Ritchie AD, Ruppel R, Jha S (2016) Generalist Behavior Describes Pollen Foraging for Perceived Oligolectic and Polylectic Bees. Environ Entomol 45:909–919. https://doi.org/10.1093/ee/nvw032

Roulston TH, Cane JH (2000) The Effect of Diet Breadth and Nesting Ecology on Body Size Variation in Bees (Apiformes). J Kans Entomol Soc 73:129–142

Ryan, E, Bateman, P, Fernandes, K, van der Heyde, M, Nevill, P (2022) eDNA metabarcoding of log hollow sediments and soils highlights the importance of substrate type, frequency of sampling and animal size, for vertebrate species detection. Environ. DNA. https://doi.org/10.1002/edn3.306

Sánchez-Bayo F, Wyckhuys KAG (2019) Worldwide decline of the entomofauna: A review of its drivers. Biol Conserv 232:8–27. https://doi.org/10.1016/j.biocon.2019.01.020

Santini, L, González-Suárez, M, Russo, D, Gonzalez-Voyer, A, von Hardenberg, A, Ancillotto, L (2019) One strategy does not fit all: determinants of urban adaptation in mammals. Ecol Lett, 22: 365-376. https://doi.org/10.1111/ele.13199

Seidelmann, K, Ulbrich, K, Mielenz, N (2010) Conditional sex allocation in the Red Mason bee, *Osmia rufa*. Behav Ecol Sociobiol 64: 337–347. https://doi.org/10.1007/s00265-009-0850-2

Staab M, Pufal G, Tscharntke T, Klein A-M (2018) Trap nests for bees and wasps to analyse trophic interactions in changing environments—A systematic overview and user guide. Methods Ecol Evol 9:2226–2239. https://doi.org/10.1111/2041-210X.13070

Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012) Environ. DNA. Mol Ecol 21:1789–1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vernat T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. Nucleic Acids Res 35:e14–e14. https://doi.org/10.1093/nar/gkl938

Theodorou, P (2022) The effects of urbanisation on ecological interactions. Curr Opin Insect Sci 100922. https://doi.org/10.1016/j.cois.2022.100922

Threlfall CG, Walker K, Williams NSG, Hahs AK, Mata L, Stork N, Livesley SJ (2015) The conservation value of urban green space habitats for Australian native bee communities. Biol Conserv 187:240–248. https://doi.org/10.1016/J.BIOCON.2015.05.003

Torchio PF, Tepedino VJ (1980) Sex ratio, body size and seasonality in a solitary bee, *Osmia lignaria propinqua cresson* (Hymenoptera: Megachilidae). Evolution 34:993–1003. https://doi.org/10.1111/j.1558-5646.1980.tb04037.x

Tscharntke T, Gathmann A, Steffan-Dewenter I (1998) Bioindication using trap-nesting bees and wasps and their natural enemies: community structure and interactions. J Appl Ecol 35:708–719.

van der Heyde, M, Bunce, M, Wardell-Johnson, G, Fernandes, K, White, NE, Nevill, P (2020) Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. Mol Ecol Resour 20: 732-745. https://doi.org/10.1111/1755-0998.13148

Voulgari-Kokota A, Ankenbrand MJ (2019) Linking pollen foraging of megachilid bees to their nest bacterial microbiota. Ecol Evol 9:10788-1800. https://doi.org/10.1002/ece3.5599

Wcislo WT, Cane JH (1996) Floral resource utilization by solitary bees (Hymenoptera: Apoidea) and exploitation of their stored foods by natural enemies. Annu Rev Entomol 41:257–286. https://doi.org/10.1146/annurev.en.41.010196.001353

Wenzel A, Grass I, Belavadi VV, Tscharntke T (2020) How urbanization is driving pollinator diversity and pollination--A systematic review. Biol Conserv 241:108321

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York. https://ggplot2.tidyverse.org

Wilkinson SP, Davy SK, Bunce M, Stat M (2018) Taxonomic identification of environmental DNA with informatic sequence classification trees. PeerJ Preprints doi:10.7287/peerj.preprints.26812v1

Wilson RS, Keller A, Shapcott A, Leonhardt SD, Sickel W, Hardwick JL, Heard TA, Kaluza BF, Wallace HM (2021) Many small rather than few large sources identified in long-term bee pollen diets in agroecosystems. Agric Ecosyst Environ 310:107296. https://doi.org/10.1016/j.agee.2020.107296

Winfree R, Williams NM, Gaines H, Ascher JS, Kremen C (2008) Wild Bee Pollinators Provide the Majority of Crop Visitation across Land-Use Gradients in New Jersey and Pennsylvania, USA. J Appl Ecol 45:793–802

Wright IR, Roberts SPM, Collins BE (2015) Evidence of forage distance limitations for small bees (Hymenoptera: Apidae). Eur J Entomol 112:303–310. https://doi.org/10.14411/eje.2015.028

Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao P, Chen S (2010) Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS One https://doi.org/10.1371/journal.pone.0013102

Zurbuchen A, Cheesman S, Klaiber J, Müller A, Hein S, Dorn S (2010a) Long foraging distances impose high costs on offspring production in solitary bees. J Anim Ecol 79:674–681. https://doi.org/10.1111/j.1365-2656.2010.01675.x

Zurbuchen A, Landert L, Klaiber J, Müller A, Hein S, Dorn S (2010b) Maximum foraging ranges in solitary bees: only few individuals have the capability to cover long foraging distances. Biol Conserv 143:669–676. https://doi.org/10.1016/j.biocon.2009.12.003

4.8 Supplementary Tables

Table S 4-4-1 Floral host and intertegular span of the eight species of bee included in this study, including number of nesting tubes sampled from each habitat type. Floral hosts were observed through pollinator surveys in both residential gardens and bushland remnant habitats (Prendergast and Ollerton 2021) and surveys by K.S. Prendergast (unpub) in southwest WA. Intertegular span was measured from female body size. Lecty is in terms of known specialisation on pollen in terms of foraging from flower species from a single family for oligolectic bees.

Species	Intertegular span (mm)	Lecty (oligolectic (specialist) vs polylectic (generalist))	Floral host(s)	Residential garden	Bushland remnant
Megachile (Hackeriapis) canifrons (Smith, 1853)	3.17	Oligolectic	*Jacksonia sternbergiana, J. furcellata,* and *Verticordia plumosa*†	10	10
Megachile (Mitchellapis) fabricator (Smith, 1868)	3.57	Oligolectic	*Jacksonia sternbergiana* and *J. furcellata*	6	9
Rozenapis ignita (Smith, 1853)	3.25	Oligolectic	*Jacksonia sternbergiana*, *J. furcellata,* and *Corymbia calophylla*†	10	10
Hylaeus (Euprosopis) violaceus (Smith, 1853)	1.76	Polylectic	*Melaleuca preissiana, M. huegelii, M. teretifolia*, *Eucalyptus spp.,* and *Banksia sessilis*	9	8
Megachile aurifrons (Smith, 1853)	2.58	Polylectic	*Scaevola spp.*, *Plectranthus spp.*, *Jacksonia furcellata*, *J.sternbergiana, Melaleuca quinquenervia,* and *Babingtonia camphorosmae*	10	10
Megachile erythropyga (Smith., 1853)	3.2	Polylectic	*Scaevola spp.*, *Plectranthus spp., Corymbia calophylla*, and *Callistemon spp.*	10	10
Megachile (Hackeriapis) oblonga (Smith, 1879)	2.42	Polylectic	*Corymbia callophylla, Goodenia filiformis, Jacksonia sericea, Melaleuca huegelii, Babingtonia camphorosmae, Cassytha racemosa, Stylidium brunonianum, Asartea scoparia, Baekea spp., Thryptomene saxicola,* and *Mentha pulegium*	10	10
Megachile (Hackeriapis) tosticauda (Cockerell, 1912)	2.37	Polylectic	*Babingtonia camphorosmae, Corymbia callophylla, Billardiera variifolia, Baekea spp.,* and *Cassytha racemosa*	7	9

Table S 4-4-2 Abbreviations used for habitat characteristics

Abbreviation	Description
Area	Area of site
Builtspace	Percentage of built space within area
Dist.Bushland	Distance to nearest bushland
FloralN	Total number of flowers
FloralR	Species Richness of flowering plants
NativeFloraN	Number of native flowers
NativeFloraR	Species richness of native flowering plants
PropNnativeflora	Proportion of native flowers in total flowers
PropRnativeflora	Proportion of native flowering plant species richness in total species richness
Bare.ground	Bare ground cover
Woody.plant	Number of trees and shrubs
HoneybeeN	Number of honeybees observed
NativeBeeN	Number of native bees observed
TotalBeeN	Total number of bees observed
PropNativeBees	Proportion of native bees in total area
NativeBeeR	Native bee species richness caught during sampling period

Table S 4-4-3 List of plant detections from nesting tubes for bee species in this study

Oligolectic Bee Species				
Megachile canifrons				
Class	**Order**	**Family**	***Genus***	***Species***
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Asterales	Asteraceae		
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
Pinidae	Araucariales	Podocarpaceae	*Podocarpus*	
Pinidae	Cupressales	Cupressaceae		
Gunneridae	Myrtales	Lythraceae	*Lagerstroemia*	
Gunneridae	Brassicales	Brassicaceae		
Gunneridae	Gentianales	Rubiaceae	*Galium*	
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
Liliopsida	Poales	Poaceae		
Gunneridae	Fabales	Fabaceae		
Gunneridae	Dipsacales	Adoxaceae	*Viburnum*	
Gunneridae	Lamiales			
Gunneridae	Fagales	Juglandaceae		
Magnoliidae	Laurales	Lauraceae		
Gunneridae	Fabales	Fabaceae	*Daviesia*	
Gunneridae	Pentapetalae			
Megachile fabricator				
Class	**Order**	**Family**	**Genus**	**Species**
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Petrosaviidae	Poales	Poaceae	*Avena*	
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Asterales	Asteraceae		
Gunneridae	Fabales	Fabaceae		
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
Gunneridae	Santalales	Loranthaceae	*Nuytsia*	*Nuytsia floribunda*
Gunneridae	Dipsacales	Adoxaceae	*Viburnum*	
Gunneridae	Fagales	Juglandaceae		
Magnoliidae	Laurales	Lauraceae		
Gunneridae	Fabales	Fabaceae	*Daviesia*	
Gunneridae	Pentapetalae			
Liliopsida	Poales	Poaceae		
Rozenapis ignita				
Class	**Order**	**Family**	**Genus**	**Species**
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Pinidae	Cupressales	Cupressaceae	*Juniperus*	
Mesangiospermae				
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Asterales	Asteraceae		
Gunneridae	Asterales	Menyanthaceae		
Pinidae	Araucariales	Podocarpaceae	*Podocarpus*	
Pinidae	Cupressales	Cupressaceae		
Gunneridae	Fabales	Fabaceae		
Gunneridae	Fabales	Fabaceae	*Hardenbergia*	
Gunneridae	Brassicales	Brassicaceae		
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
Liliopsida	Poales	Poaceae	*Avena*	
Gunneridae	Fabales	Fabaceae	*Daviesia*	
Gunneridae	Solanales	Solanaceae		
Gunneridae	Gentianales	Rubiaceae	*Galium*	
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
Liliopsida	Poales	Poaceae		
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
Gunneridae	Sapindales	Anacardiaceae	*Schinus*	*Schinus terebinthifolia*
eudicotyledons	Proteales	Proteaceae	*Grevillea*	
Gunneridae	Fabales	Fabaceae	*Lotus*	*Lotus unifoliolatus*
Gunneridae	Apiales	Apiaceae		
Pinidae	Cupressales	Cupressaceae	*Callitris*	
Cycadidae	Cycadales	Zamiaceae	*Macrozamia*	

Liliopsida	Poales	Poaceae	*Ehrharta*	
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
Liliopsida	Zingiberales	Musaceae		
Gunneridae	Dipsacales	Adoxaceae	*Viburnum*	
Gunneridae	Asterales			
Gunneridae	Lamiales			
Gunneridae	Fagales	Juglandaceae		
Magnoliidae	Laurales	Lauraceae		
Gunneridae	Lamiales	Oleaceae	*Ligustrum*	
Gunneridae	Pentapetalae			

Polylectic Bee Species

Hylaeus violaceus

Class	**Order**	**Family**	**Genus**	**Species**
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Asterales	Asteraceae		
Pinidae	Araucariales	Podocarpaceae	*Podocarpus*	
Pinidae	Cupressales	Cupressaceae		
Gunneridae	Fabales	Fabaceae		
Gunneridae	Brassicales	Brassicaceae		
Gunneridae	Gentianales	Rubiaceae	*Galium*	
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
Gunneridae	Fabales	Fabaceae	*Grazielodendron*	
Liliopsida	Poales	Poaceae		
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
eudicotyledons	Proteales	Proteaceae	*Grevillea*	
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
Gunneridae	Dipsacales	Adoxaceae	*Viburnum*	
Gunneridae	Lamiales			
Gunneridae	Fagales	Juglandaceae		
Magnoliidae	Laurales	Lauraceae		
Gunneridae	Pentapetalae			
Pinidae	Araucariales			

Megachile aurifrons

Class	**Order**	**Family**	**Genus**	**Species**
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Gunneridae	Malvales	Malvaceae	*Taliparitи*	*Taliparitи tiliaceum*
Gunneridae	Asterales	Asteraceae	*Cotula*	*Cotula australis*
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Asterales	Asteraceae		
Gunneridae	Asterales	Menyanthaceae		
Pinidae	Cupressales	Cupressaceae		
Gunneridae	Fabales	Fabaceae		
Gunneridae	Myrtales	Lythraceae	*Lagerstroemia*	
Gunneridae	Fabales	Fabaceae	*Hardenbergia*	
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
Gunneridae	Solanales	Solanaceae		
Gunneridae	Gentianales	Rubiaceae	*Galium*	
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
Gunneridae	Fabales	Fabaceae	*Grazielodendron*	
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
eudicotyledons	Proteales	Proteaceae	*Grevillea*	
Gunneridae	Apiales	Apiaceae		
Gunneridae	Lamiales	Verbenaceae		
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
Gunneridae	Brassicales	Brassicaceae		
Gunneridae	Malvales	Malvaceae	*Taliparitи*	
Gunneridae	Lamiales			
Gunneridae	Fagales	Juglandaceae		
Magnoliidae	Laurales	Lauraceae		
Gunneridae	Lamiales	Oleaceae	*Ligustrum*	
Gunneridae	Fabales	Fabaceae	*Daviesia*	
Gunneridae	Pentapetalae			
Liliopsida	Poales	Poaceae		
Gunneridae	Solanales	Convolvulaceae		

Megachile erythropyga

Class	**Order**	**Family**	**Genus**	**Species**
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Gunneridae	Myrtales	Myrtaceae	*Syzygium*	
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Asterales	Asteraceae		
Pinidae	Araucariales	Podocarpaceae	*Podocarpus*	
Pinidae	Cupressales	Cupressaceae		
Gunneridae	Fabales	Fabaceae		

Liliopsida	Poales	Poaceae	*Avena*	
Gunneridae	Rosales	Moraceae	*Morus*	
Gunneridae	Solanales	Solanaceae		
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
Gunneridae	Sapindales	Anacardiaceae	*Schinus*	*Schinus terebinthifolia*
Gunneridae	Gentianales	Apocynaceae		
Gunneridae	Lamiales	Verbenaceae		
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
Liliopsida	Poales	Poaceae		
Cycadidae	Cycadales	Zamiaceae	*Macrozamia*	
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
Gunneridae	Lamiales			
Gunneridae	Fagales	Juglandaceae		
Gunneridae	Pentapetalae			
Gunneridae	Solanales	Convolvulaceae		

Megachile oblonga

Class	**Order**	**Family**	**Genus**	**Species**
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Asterales	Asteraceae		
Gunneridae	Asterales	Menyanthaceae		
Gunneridae	Fabales	Fabaceae		
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
Gunneridae	Gentianales	Rubiaceae	*Galium*	
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
Gunneridae	Fabales	Fabaceae	*Grazielodendron*	
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
Gunneridae	Fabales	Fabaceae	*Lotus*	*Lotus unifoliolatus*
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
Gunneridae	Santalales	Loranthaceae	*Nuytsia*	*Nuytsia floribunda*
Gunneridae	Dipsacales	Adoxaceae	*Viburnum*	
Gunneridae	Asterales			
Liliopsida	Poales	Poaceae		
Gunneridae	Lamiales			
Gunneridae	Fagales	Juglandaceae		
Gunneridae	Fabales	Fabaceae	*Daviesia*	
Gunneridae	Pentapetalae			

Megachile tosticauda

Class	**Order**	**Family**	**Genus**	**Species**
Gunneridae	Myrtales	Myrtaceae	*Eucalyptus*	
Petrosaviidae	Poales	Poaceae	*Avena*	
Gunneridae	Myrtales	Myrtaceae		
Gunneridae	Fabales	Fabaceae		
Gunneridae	Myrtales	Myrtaceae	*Ugni*	*Ugni candollei*
Gunneridae	Myrtales	Myrtaceae	*Melaleuca*	*Melaleuca nodosa*
eudicotyledons	Proteales	Proteaceae	*Persoonia*	
Gunneridae	Fabales	Fabaceae	*Jacksonia*	*Jacksonia horrida*
Gunneridae	Dipsacales	Adoxaceae	*Viburnum*	
Gunneridae	Lamiales			
Gunneridae	Fagales	Juglandaceae		
Magnoliidae	Laurales	Lauraceae		
Gunneridae	Pentapetalae			
Gunneridae	Solanales	Convolvulaceae		

Table S 4-4-4 SIMPER analysis showing plant families ranked according to average Bray-Curtis dissimilarity between habitat types for overall native bee community and each species of bee. Frequency is based on abundance of ASVs identified to plant family found within those sites, a proxy for relative abundance

Overall Bee Community			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 60%**		
Myrtaceae	4.41	62.01	62.01
Fabaceae	2.1	24.5	86.51
Residential	**Average Similarity: 57.8%**		
Myrtaceae	5.02	63.28	63.28
Fabaceae	2.02	22.22	85.5
Both	**Average dissmilarity: 41.23%**		
Myrtaceae	4.41	21.06	21.06
Fabaceae	2.1	13.31	34.37
Poaceae	0.53	7.54	41.91
Asteraceae	0.43	6.17	48.09
Proteaceae	0.4	5.99	54.08
Juglandaceae	0.54	5.87	59.95
Cupressaceae	0.26	4.9	64.85
Adoxaceae	0.22	3.26	68.11
Lauraceae	0.17	2.89	71
Oligolectic Bees			
Megachile canifrons			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 60.6%**		
Myrtaceae	4.23	63.56	63.56
Fabaceae	1.92	28.26	91.81
Residential	**Average Similarity: 68.9%**		
Myrtaceae	6.69	72.06	72.06
Both	**Average dissimilarity: 38%**		
Myrtaceae	4.23	30.19	30.19
Fabaceae	1.92	8.2	38.39
Poaceae	0.5	7.18	45.57
Juglandaceae	0.5	6.42	52
Cupressaceae	0.24	5.77	57.76
Asteraceae	0.4	5.71	63.47
Proteaceae	0.17	5.11	68.59
Brassicaceae	0.4	4.87	73.46
Megachile fabricator			
Species	**Frequency**	**Contribution%**	**Cumulative%**

Bushland	**Average Similarity: 73.9%**		
Myrtaceae	5.21	66.78	66.78
Fabaceae	1.9	21.18	87.96
Residential	**Average Similarity: 78.9%**		
Myrtaceae	7.29	72.96	72.96
Both	**Average dissimilarity: 28.7%**		
Myrtaceae	5.21	33.03	33.03
Fabaceae	1.9	9.38	42.41
Asteraceae	0.11	7.56	49.97
Loranthaceae	0.11	6.82	56.79
Poaceae	0.33	6.21	63.01
Proteaceae	0.11	4.04	67.05
Lauraceae	0.22	3.24	70.29
Rozenapis ignita			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 58.1%**		
Fabaceae	2.89	37.95	37.95
Myrtaceae	2.71	33.31	71.26
Residential	**Average Similarity: 65.4%**		
Myrtaceae	4.03	37.81	37.81
Fabaceae	2.98	33.56	71.37
Both	**Average dissmilarity: 39.2%**		
Myrtaceae	4.03	17.93	17.93
Poaceae	1.11	9.32	27.25
Asteraceae	0.44	8.82	36.08
Fabaceae	2.98	6.88	42.96
Cupressaceae	0.61	6.69	49.65
Brassicaceae	0.3	5.06	54.71
Menyanthaceae	0.3	4.77	59.48
Juglandaceae	0.5	4.56	64.04
Rubiaceae	0.3	4.21	68.25
Polylectic Bee Species			
Hylaeus violaceus			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 61%**		
Myrtaceae	4.28	54.26	54.26
Fabaceae	2.29	26.63	80.89
Residential	**Average Similarity: 56.6%**		
Myrtaceae	5.43	80.65	80.65
Both	**Average dissimilarity: 39.2%**		
Myrtaceae	5.43	22.04	22.04
Fabaceae	1.42	16.88	38.92

Poaceae	0.74	8.28	47.2
Juglandaceae	0.29	7.43	54.63
Cupressaceae	0.29	5.52	60.15
Proteaceae	0.29	5.17	65.31
Adoxaceae	0.29	4.42	69.74
Lauraceae	0.29	4.27	74.01
Megachile aurifrons			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 62%**		
Myrtaceae	4.2	63.15	63.15
Fabaceae	2.3	25.12	88.26
Residential	**Average Similarity: 57.6%**		
Myrtaceae	4.14	48.72	48.72
Fabaceae	2.86	39.38	88.1
Both	**Average dissimilarity: 41%**		
Myrtaceae	4.2	16.72	16.72
Fabaceae	2.3	11.98	28.7
Asteraceae	0.73	7.82	36.52
Proteaceae	0.6	7.52	44.04
Malvaceae	0	6.29	50.33
Poaceae	0.3	5.87	56.2
Juglandaceae	0.38	5.59	61.79
Cupressaceae	0.43	4.89	66.68
Lythraceae	0	3.71	70.39
Megachile erythropyga			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 65.6%**		
Myrtaceae	5.09	69.25	69.25
Fabaceae	1.49	12.72	81.97
Residential	**Average Similarity: 55.7%**		
Myrtaceae	4.32	58.54	58.54
Cupressaceae	1.17	16.87	75.41
Both	**Average dissimilarity: 42.9%**		
Myrtaceae	4.32	13.19	13.19
Fabaceae	1.2	10.13	23.32
Cupressaceae	1.17	9.61	32.92
Poaceae	0.4	9.2	42.12
Asteraceae	0.63	7.23	49.35
Juglandaceae	0.6	5.38	54.73
Verbenaceae	0.6	5.35	60.08
Solanaceae	0.48	4.74	64.82
Anacardiaceae	0.4	4.1	68.91

Podocarpaceae	0.2	3.94	72.85
Megachile oblonga			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 62.3%**		
Myrtaceae	4.94	65.83	65.83
Fabaceae	2.52	27.47	93.29
Residential	**Average Similarity: 63.5%**		
Myrtaceae	4.57	69.64	69.64
Fabaceae	1.26	19.11	88.75
Both	**Average dissimilarity: 38.4%**		
Myrtaceae	4.57	22.8	22.8
Fabaceae	1.26	21.8	44.6
Poaceae	0.34	8.92	53.52
Juglandaceae	0.71	7.65	61.17
Asteraceae	0.14	5.63	66.8
Rubiaceae	0.14	5.18	71.98
Megachile tosticauda			
Species	**Frequency**	**Contribution%**	**Cumulative%**
Bushland	**Average Similarity: 66.4%**		
Myrtaceae	5.25	78.78	78.78
Residential	**Average Similarity: 59%**		
Myrtaceae	4.93	67.3	67.3
Fabaceae	1.54	20.9	88.2
Both	**Average dissimilarity: 34.4%**		
Myrtaceae	4.93	22.07	22.07
Fabaceae	1.54	13.04	35.11
Adoxaceae	0.5	7.76	42.87
Poaceae	0.25	7.7	50.57
Juglandaceae	0.75	6.79	57.36
Asteraceae	0.25	5.6	62.96
Proteaceae	0.25	5.45	68.41
Lythraceae	0.25	4.82	73.23

Table S 4-4-5 Distance based linear model (DistLM) analysis of plant ASV composition from brooding tubes across all sites. Constructed using a BEST selection procedure and the AICc selection criterion. The overall best solution is presented, and possible alternate models within 2 AICc of the best. * indicates significance at α = 0.05.

	Selections	Number of variables	AICc	R^2
Overall BEST solution	Habitat. Species, Lecty	3	1063.1	0.166
Alternate solutions within 2 AICc of BEST solution	Habitat, Species	2	1063.1	0.166
	Habitat. Species, Lecty, NativeFloraR	4	1063.3	0.180
	Habitat. Species, Lecty, Floral Richness, PropRnativeFlora	5	1063.8	0.192
	Habitat. Species, Lecty, Dist Bushland, Floral Richness, PropRnativeFlora	6	1064.3	0.203

Marginal tests for each variable

Variable	Characteristic	Pseudo-F	Proportion	P value
1	Habitat	4.214	0.031	0.0001*
2	Species	2.784	0.135	0.0001*
3	Lecty	2.37	0.017	0.0022*
4	Area	2.848	0.021	0.0002*
5	Builtspace	3.163	0.023	0.0001*
6	Dist.Bushland	2.633	0.019	0.0005*
7	FloralN	1.05	0.007	0.3615
8	FloralR	2.829	0.02	0.0002*
9	NativeFloraN	1.082	0.008	0.3249
10	NativeFloraR	2.515	0.018	0.0008*
11	PropNnativeflora	1.13	0.008	0.2676
12	PropRnativeflora	2.466	0.018	0.0014*
13	Bare.ground	2.249	0.017	0.0046*
14	Woody.plant	3.449	0.026	0.0001*

4.9 Supplementary Figures

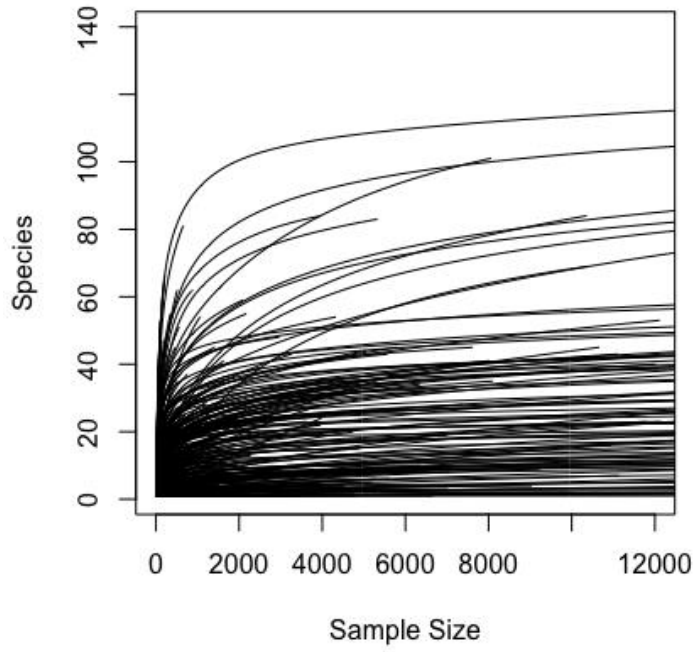


Figure S 4-1 Rarefaction curve of raw reads (Sample Size) versus ASV richness (Species)

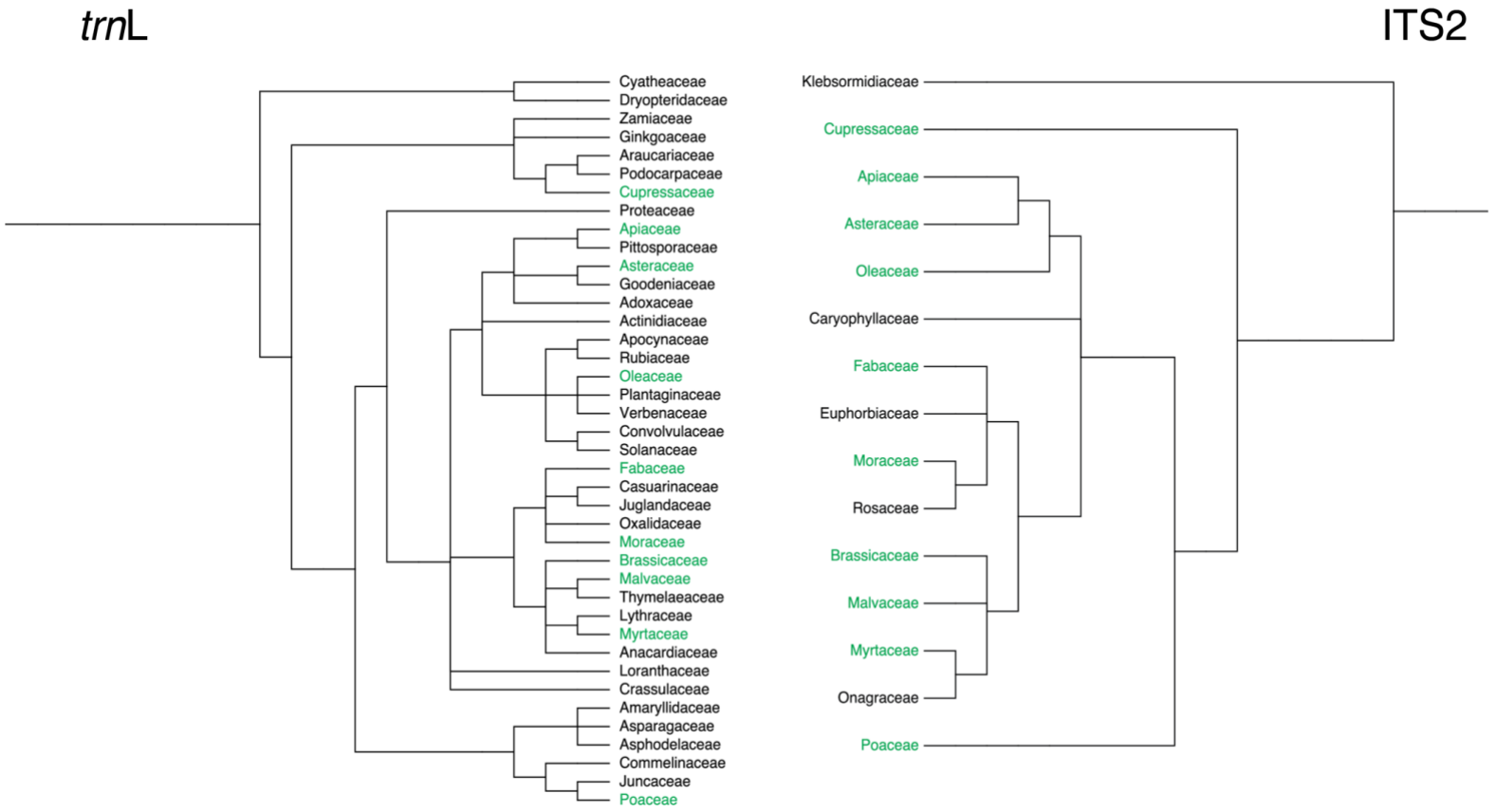


Figure S 4-2 Cladograms of trees generated from taxa assignments to plant Family level for *trn*L and ITS2 primers. Green indicates shared taxa between the two primer sets.

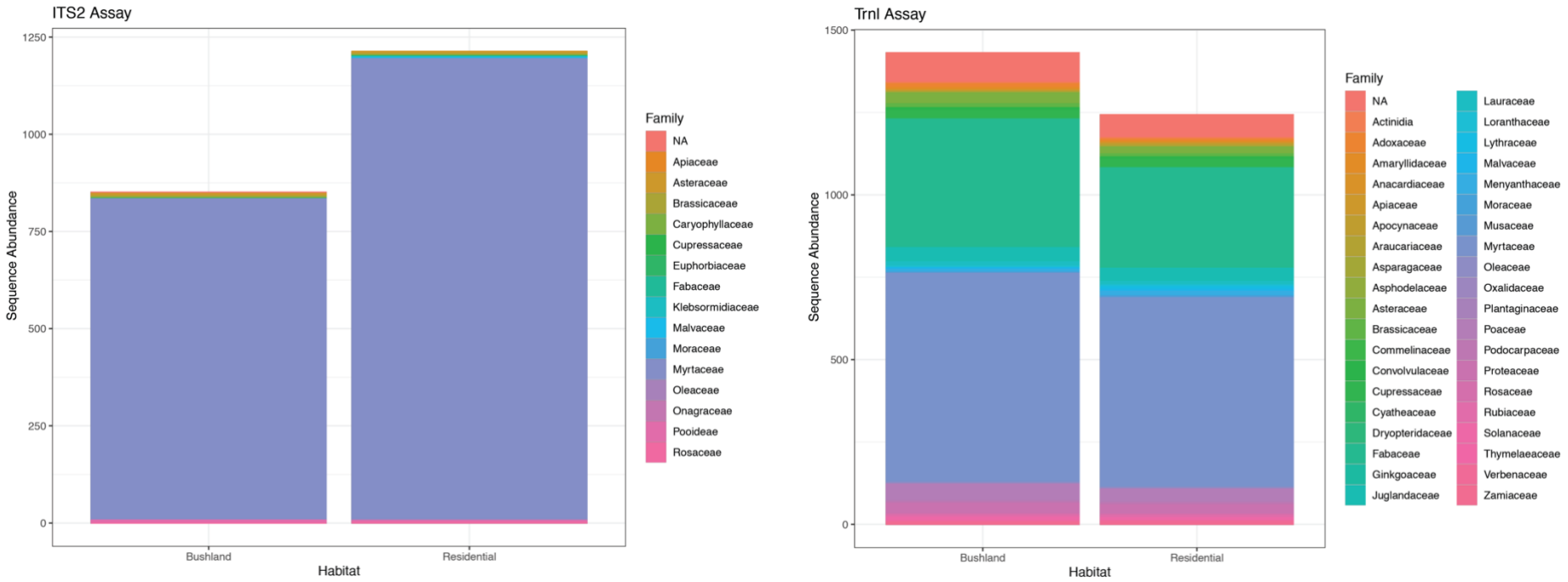


Figure S 4-3 Relative sequence abundances (from presence-absence detections) of the two PCR assays (ITS2 and *trn*L) from all samples showing the proportions of ASVs that could be identified to Family level between Bushland Remnants and Residential Garden habitat types.

This page has been left intentionally blank

CHAPTER 5

EXAMINING THE EFFECTS OF DISTANCE, TAXON, AND ANIMAL BIOMASS, ON VERTEBRATE DETECTIONS FROM CARRION FLY iDNA

5.1 Preface

This chapter is one of the first examples of the use of invertebrate derived DNA (iDNA) within an Australian context. Here, the parameters surrounding the effectiveness of carrion fly iDNA for detecting vertebrates are explored by considering how distance, taxonomic group, and biomass impacts the detectability of vertebrates from carrion flies. Currently there is very little known about how these different factors impact vertebrate detection from carrion flies, therefore the limitations and strengths of this method as a monitoring tool remain untested. Further, as carrion flies are more abundant and easier to sample in most environments than leeches (another common iDNA sampler used in tropical environments), this study paves the way for future work in this area. The results of this chapter informed the methods and the interpretations in Chapter 6.

This chapter consists of a manuscript that is currently being prepared for submission to the peer-reviewed journal *Royal Society Open Science*.

5.1.1 Acknowledgements

We acknowledge the Whadjuk people of the Noongar nation, their Elders, past, present, and emerging, as the Traditional Custodians of this country, on whose land this research was carried out. This work was supported by computational resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. We would like to thank Perth Zoo, especially Dr Peter Mawson and Caroline Lawrence, for their support. KF was partially supported by funding from Food Agility CRC Ltd, funded under the Commonwealth Government CRC Program. Funding was awarded for this project from the Holsworth Wildlife Endowment Fund in association with the Ecological Society of Australia.

5.1.2 Data Accessibility

R scripts and the accompanying data will be made available here: 10.5281/zenodo.6648515

5.1.3 Author Contributions

KF, PWB, BJS, MG, KB and PN conceptualised the study. KF collected the samples and processed them. KF conducted all the analysis. BJS assisted with statistical analysis. KF wrote the original draft of the manuscript. All authors contributed to the review and editing of the manuscript.

Examining the effects of distance, taxon, and animal biomass, on vertebrate detections from carrion fly iDNA

Kristen Fernandes [1,2,3*], Philip W Bateman [4], Benjamin J Saunders [5], Mark Gibberd [6], Michael Bunce [1,7], Kristine Bohmann [2], Paul Nevill [1,8]

1 Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia
2 Section for Molecular Ecology and Evolution, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen K, Denmark
3 Food Agility CRC Ltd, 175 Pitt St Sydney, NSW 2000, Australia
4 Behavioural Ecology Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia
5 School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia
6 Centre for Crop and Disease Management, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia
7 Environmental Science and Research (ESR), Kenepuru, Porirua 5022, New Zealand.
8 ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

5.2 Abstract

Invertebrate-derived DNA (iDNA) from ‘samplers’ including leeches, carrion flies, and mosquitoes, is a powerful, non-invasive tool that is increasingly used to assess terrestrial vertebrate diversity. However, for carrion flies, unknowns exist about the factors that influence vertebrate detections, such as spatial limits to iDNA signals, whether there are preferential detections of different taxonomic classes, or how the total biomass of a species in an area will influence their detection through iDNA. We examined these questions at a study site with a highly diverse population of exotic species (Perth Zoo), and along transects extending 4 km away from this location. From metabarcoding of 920 collected flies, we detected 28 vertebrate species, including zoo residents, domestic, and wild animals. iDNA detections were highly geographically localised, and only a few zoo animals were detected outside of the zoo setting with the furthest detections ca. 4 km away. The 28 detected species included eight mammal and one bird species out of the 43 mammal species and 41 bird species within the zoo that had access to outdoor enclosures during the sampling period. We detected none of the 14 reptile species kept at the zoo with access to outdoor enclosures. We found no influence of the taxonomic group of animals on their detectability from iDNA. Additionally, we found no relationship between biomass and the detection of vertebrates. Our data suggest that carrion fly detections in urban settings are predominantly determined by geographic proximity to the sampling location rather than the biomass of species in an area. This study provides an important step forward in understanding the limitations of this technique in terrestrial biodiversity monitoring.

5.3 Introduction

Monitoring species diversity and abundance are crucial to managing and informing efforts to prevent biodiversity loss (Legge et al., 2018). Information obtained from monitoring can improve understanding of population trajectories and extinction risk, the response of species to threatening processes, and evaluate management effectiveness (Balmford et al., 2005; Legge et al., 2018; Marsh & Trenham, 2008). Most current monitoring methods for terrestrial vertebrates rely on the direct observation of species in an area or the traces they may leave behind, such as hair, footprints, or scat (Kouakou et al., 2009). This can be time and resource-intensive and require the mobilisation of experts over long periods (Campbell et al., 2011). Various methods (e.g. camera traps, acoustic loggers) have emerged to monitor ecosystems to reduce the financial, expertise, and effort costs and, and enhance monitoring performance (Calvignac-Spencer et al., 2013). Recently, environmental DNA (eDNA) metabarcoding of environmental substrates, such as soil (Andersen et al., 2012; Ryan et al., 2022), water (Rodgers & Mock, 2015), scat (van der Heyde et al., 2021), tree hollow sediment (Newton et al., 2022), and air (Clare et al., 2022; Lynggaard et al., 2022) have grown in popularity to monitor vertebrate biodiversity. An extension of this is the use of invertebrate-derived DNA (iDNA), whereby vertebrate diversity is monitored using invertebrates that feed on the blood, flesh, or scat of vertebrates and therefore carry their DNA signal (Abrams et al., 2019; Calvignac-Spencer et al., 2013; Lynggaard et al., 2019; Schnell et al., 2015).

iDNA metabarcoding species detection works by extracting DNA from invertebrate samples, amplifying the DNA using nucleotide-labelled primers that target vertebrate sequences, sequencing the DNA on high-throughput platforms, and identifying the resulting sequencing against reference sequence databases (Calvignac-Spencer et al., 2013; Taberlet et al., 2012). This method is an efficient and cost-effective way to monitor vertebrate diversity in remote or difficult terrain (Massey et al., 2021). Unlike camera traps, which are biased toward detecting larger-bodied terrestrial mammals as their size can trigger the cameras more easily (Burton et al., 2015), iDNA can be used to detect smaller-bodied taxa, including arboreal or non-mammal species (Gogarten et al., 2019). Studies using iDNA have focused on invertebrate groups such as leeches (Schnell et al., 2015), sand flies (Kocher et al., 2017), mosquitoes (Massey et al., 2021), and carrion flies (Calvignac-Spencer et al., 2013) as 'samplers.' Compared to mosquitoes and sandflies, carrion flies (families Calliphoridae and Sarcophagidae) have been found to be the most effective iDNA sampler at capturing vertebrate species richness (Massey et al., 2021). Unlike leeches, which can retain vertebrate DNA in their guts for months at a time (Schnell et al., 2012), the short-lived persistence of DNA transported in carrion flies (24-96 hours) (Lee et al., 2015) makes them ideal candidates for measuring current vertebrate diversity and distributions. Furthermore, carrion flies are widespread in many different environments and are easy to sample (Norris, 1965). Studies using carrion flies for vertebrate detection have so far been undertaken on tropical mammal communities in Africa (Calvignac-Spencer et al., 2013; Schubert et al., 2015), South America (Rodgers et al., 2017), and Asia (Bagnall, 2017), and vertebrate communities in urbanised areas in Europe (Hoffmann et al., 2018) and North America (Owings et al., 2019).

Despite the utility of carrion fly iDNA for vertebrate monitoring, some aspects of its use are yet to be explored. Firstly, carrion flies are very mobile, with some species of calliphorid flies having been documented to disperse up to 63.5 km from a source (Braack & Retief, 1986). However, no studies have reported the spatial distribution of iDNA signals of vertebrate species from carrion flies. Furthermore, carrion flies are opportunistic feeders, and, unlike hematophagous invertebrates or mosquitoes, they display no species preference for feeding (Calvignac-Spencer et al., 2013; Massey et al., 2021). However, some classes of vertebrates, such as reptiles, can have poor DNA shedding rates and a lower detection in environmental DNA studies (Adams et al., 2019) although it is yet to be tested has yet to be tested whether reptiles have a relatively lower detection probability with iDNA than other vertebrate taxa. It is also unknown whether the biomass of a species in a given area will influence its detection via iDNA, i.e. if there are more individuals of species, or if the mean mass of individuals of a particular species is greater than that of others, whether that would provide more readily available food sources for flies and increase detections.

In this study, we aimed to use iDNA metabarcoding from carrion flies to detect vertebrate species and examine several factors that could impact their detectability. We used a controlled environment of a zoo with known vertebrate species diversity, abundance, and position. Firstly, we aimed to determine the approximate geographic range of an iDNA signal. We predicted that the detections of zoo taxa would decline with distance from their enclosures. We then set out to assess the influence of taxon and biomass on the detectability of vertebrates through carrion fly iDNA. Based on previous studies, we predicted that taxon would influence detection, based on assumed differential DNA shedding or degradation rates of different classes of vertebrates, i.e., reptiles would be less likely to be detected than mammals. We also predicted that animal populations representing higher biomass will be detected more frequently and at farther distances than animals from populations with lower biomass.

5.4 Methods

5.4.1 Study site

This study was conducted within the Perth metropolitan region in south-west Western Australia, using baited plastic bottle traps to capture flies (Fig 5-1B, inset) during the Spring/Summer season (November-December 2020). The sampling areas were distributed along the Swan River/Derbarl Yerrigan (Fig 5-1B). This area is predominantly residential, interspersed with medium and high-density housing, commercial buildings, and parklands. To measure the distance of iDNA detections from a source point, the Perth Zoo was chosen as the source of a known diversity and abundance of exotic animals not native to the local area.

Perth Zoo is a 17-ha zoological park located in the Perth metropolitan area; it houses 164 animal species. The zoo has created 'animal zones' that group the displayed animals according to the geographical region of origin. There are three main open-air animal zones, the 'African Savannah' that houses animals found on the African continent or surrounding islands, 'Australian Animals' that houses native Australian animals from Australia, and the 'Asian Rainforest' that houses animals found on the Asian continent (Fig 5-1A). While some species are kept in indoor enclosures, 101 are displayed with access to an outside enclosure. These animals can thus more readily interact with flies. They include three amphibians, 41 birds, 43 mammals, and 14 reptiles.

5.4.2 Sample collection

Traps were made from 2 L plastic bottles which had the top and nozzle cut off and then inverted into the bottle (Fig 5-1B inset). Before deployment, traps were sterilised for 30 minutes with a 20% bleach solution. Gloves were used to handle all traps. Flies could easily enter the trap but could not fly vertically upwards to leave. Approximately 20-30 mL of a commercial, organic, fly bait ("Magna Fly Bait", GEPRO, Australia) was put into each trap.

Six trapping locations were identified inside the zoo near 'target' taxa, comprising reptiles, birds and mammals (Fig 5-1A). Two 'target' species of each taxon were chosen, and traps were placed less than 5 m away from their enclosure. The two reptiles were the radiated tortoise (*Astrochelys radiata*) and the Galapagos tortoise (*Chelonoidis nigra*), the two birds were the emu (*Dromaius novaehollandiae*) and Southern Cassowary (*Casuarius casuarius*), and the two mammals were the numbat (*Myrmecobius fasciatus*), and lion (*Panthera leo*). These target species were chosen because they had access to open-air enclosures, which meant that they were accessible to flies. At each of the six trapping locations in the zoo, three traps were deployed.

Outside the zoo, fly traps were set up along two transects away from the zoo. Fly traps were set every 250 m up to four km away from the zoo (Fig 5-1B). Here, four km is assumed to be the average daily flight distance of most calliphorid fly species (Braack & Retief, 1986; MacLeod & Donnelly, 1963; Norris, 1965). The transects followed the coast of the peninsula of south Perth in two directions to account for prevailing wind directions, which can be a key factor for carrion fly dispersal (Spivak et al., 1991). During the sampling period, Perth experienced a strong easterly during the day with a south-westerly sea breeze in the afternoons.

Traps were checked both within and outside the zoo, and flies were collected every 72 hours over 21 days. This yielded seven possible sampling time points for each of the 18 deployed traps in the zoo and 32 deployed traps outside the zoo. Following each collection, traps were rinsed with bottled water before the next sampling. As not every sample contained flies at every collection point, in total we collected 123 fly samples across the sampling period. Samples were transported chilled on ice, and then frozen at -18˚ C upon arrival at the laboratory.

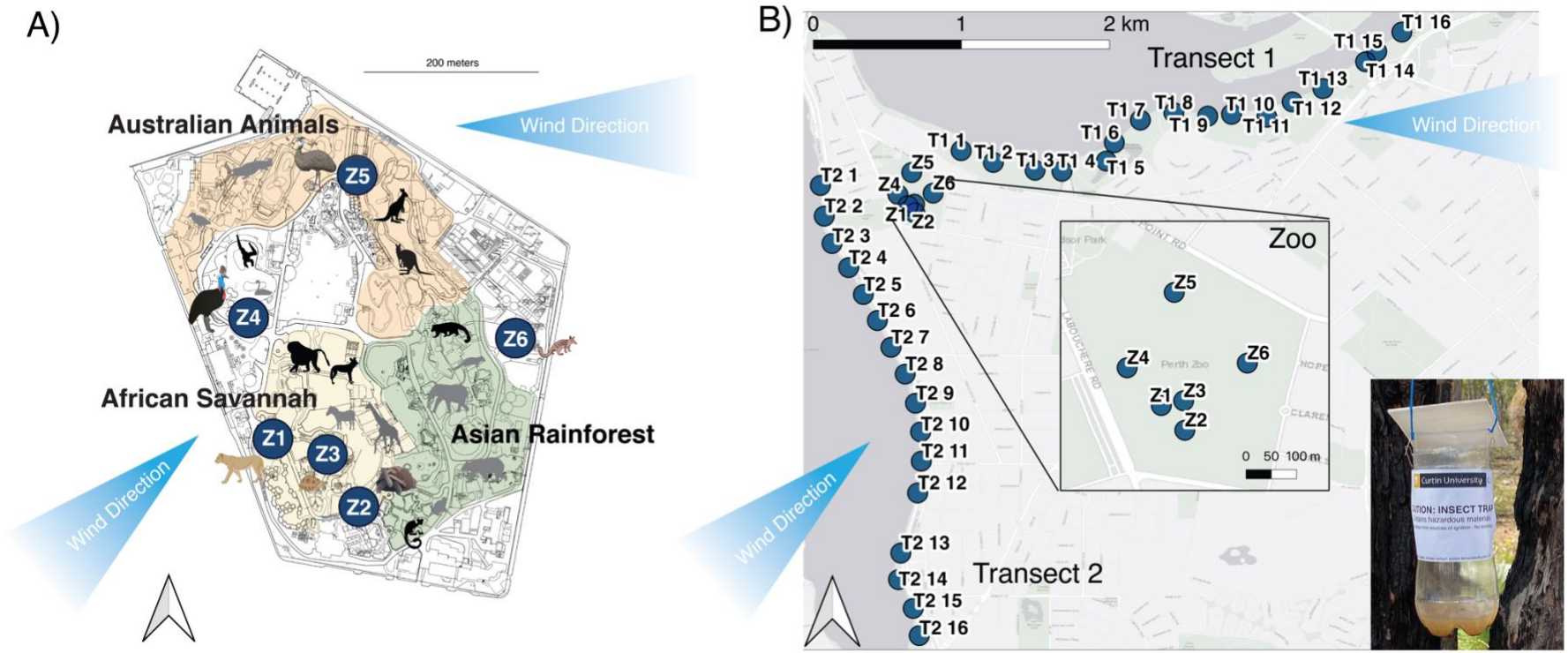


Figure 5-1 Map of Perth Zoo showing the locations of carrion fly traps within the zoo (Z1-Z6). Major 'animal zones' designated by Perth Zoo that corresponds to the groups of animals kept within the zoo are illustrated in colour with corresponding labels; African Savannah, Asian Rainforest, and Australian Animals. 'Target' zoo animal enclosures are included in colour, other zoo animals that were detected in this study are indicated as black silhouettes, zoo animals located in the same animal zone but not detected are shown as grey silhouettes. Prevailing wind directions are marked as blue arrows. B) Map showing the sampling locations (dark blue circles) spaced approximately 250 m apart along 4 km transects away from Perth Zoo (16 sampling locations per transect). Inset map shows the location of the fly traps within the zoo. Prevailing wind directions are marked as blue arrows. Carrion fly samplers inset image. These are made from plastic bottles and filled with commercial fly bait.

5.4.3 Sample processing and DNA extraction

All flies collected from a single trap at a single time point were considered as one sample. Each sample contained no more than 54 flies. Samples were thawed and flies were gently washed with ultrapure water to remove any traces of fly bait before being stored in 70% ethanol at -18 ˚C, prior to DNA extraction. DNA was extracted using a non-destructive DNA digestion protocol (Nielsen et al., 2019) that lyses and releases cellular DNA without damaging the exoskeleton of the flies. This was followed by purification and concentration using a DNEasy Blood and Tissue kit following the manufacturer's instructions (Qiagen, Netherlands) modified for 400 µL of digestion fluid and a 200 µL elution volume in EB buffer. Negative extraction controls were included for every 24 samples. DNA extracts were stored at -18 ˚C in 'Lo-bind' Eppendorf tubes.

5.4.4 PCR amplification and sequencing.

Two metabarcoding primer sets were used to target two taxonomically informative markers. These were chosen based on in-silico assessment for affinity for the target zoo species to ensure that between the two primer sets the six target species - Astrochelys radiata, Chelonoidis nigra, Dromaius novaehollandiae, Casuarius casuarius, Myrmecobius fasciatus, and Panthera leo - would be detected. Additionally, relatively short markers were needed to ensure PCR amplification despite potential degradation due to digestion within the fly digestive tract. Based on these criteria, a vertebrate-specific primer set was selected (vertebrate 12s), which targets the mitochondrial 12s rRNA region (F: 5'-TAGAACAGGCTCCTCTAG-3'; R: 5'-TTAGATACCCCACTATGC-3'; 98 bp; (Riaz et al., 2011)) and a mammal specific primer set was selected (mammal 16s), which targets the mitochondrial 16s rRNA region (F:5'-CGGTTGGGGTGACCTCGGA-3'; R:5'-GCTGTTATCCCTAGGGTAACT-3'; 130 bp; (Taylor, 1996)).

Prior to metabarcoding, quantitative PCR (qPCR) was carried out on all 123 sample extracts to assess the amplification efficiency and the presence of PCR inhibitors in the samples using a StepOnePlus Real-Time PCR System (Applied Biosystems). This was carried out on dilution series for each sample extract (undiluted, 1:2 and 1:5 dilutions). qPCRs were carried out in 12.5 µl reactions containing: 1 U of AmpliTaq gold, 1 x PCR Gold Buffer and 2 mM MgCl2 (all from Applied Biosystems, USA), 0.4 mg/mL bovine serum albumin (Fisher Biotec, Australia), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4 µM of each forward and reverse primer, 0.6 µL of 1/1000 SYBR Green (Invitrogen, USA), and 1 µL of template DNA. The qPCR program for the vertebrate 12s assay was as follows: 95˚ C for 5 minutes, followed by 45 cycles of denaturation at 95˚ C for 30s, 60˚ C for 30s, and 72˚C for 45s, with a final elongation at 72˚C for 10 minutes. For the mammal 16s assay, the qPCR program was as follows: 95˚C for 5 minutes, followed by 45 cycles of 95˚C for 30s, 57˚C for 30s, and 72˚C for 45s, with a final elongation at 72˚C for 10 minutes. Negative extraction, qPCR, 'fly bait', and positive (quenda; *Isoodon fusciventer*) controls were included alongside the sample extracts. Positive controls were chosen as a species that is known to amplify well using laboratory workflows, although as this species was also known to inhabit the local area, positive control samples used in the qPCR screening were not processed any further to minimise possible cross-contamination.

Following qPCR screening, metabarcoding was carried out. For each sample, sample extract dilutions that had shown the highest level of amplification and minimal inhibition were amplified using qPCR with 'fusion primers', called a one-step approach (Bohmann et al., 2021). The fusion primers consisted of the metabarcoding primers, labelled on the 5' end of both the forward and reverse primer with 6-8 nucleotide long molecular identification (MID) tags followed by Illumina sequencing adaptors. Amplifications were carried out so that each sample was tagged with a unique combination of forward and reverse MID tags. All MID tag combinations were not previously used in the laboratory. To minimise the risk of contamination, qPCR reactions were prepared in an ultra-clean pre-PCR laboratory free from extracted or amplified DNA and DNA was added in another DNA extraction laboratory free from amplicons.

All sample extracts, extraction and qPCR negative controls, and a ‘fly bait’ control were amplified in duplicate using the qPCR conditions mentioned above. Duplicates carried the same MID tags. PCR products were pooled in approximate equimolar concentrations into 20 sample pools informed by the qPCR amplification curves. DNA in the pools were then quantified using a QIAxcel Advanced System (Qiagen) with the QIAxcel DNA High-Resolution Kit. Sample pools were then pooled into one sequencing library in equimolar ratios. The library was size-selected using a Pippin Prep 2% agarose Marker B cassette (Sage Science, USA) for 150-600 bp fragments to eliminate primer dimer. The library pool was then purified using a QIAquick PCR purification kit (Qiagen) as per the manufacturer's instructions with the addition of a 5-minute incubation at room temperature before elution. The purified library was eluted in 40 µL and quantified with a QuBit (Invitrogen, USA) using double-stranded DNA high-sensitivity reagents. The library was sequenced on an Illumina MiSeq (Illumina, USA) flowcell using a single-end 300 cycle V2 kit as per the manufacturer's directions.

5.4.5 Bioinformatics and sequence processing

Sequence data for each primer set was processed separately. Sequences were demultiplexed using 'obitools' (Boyer et al., 2016), with no mismatches in the MID/primer sequence allowed. Sequences were then length filtered for a minimum length of 50 bp. The 'DADA2' package (Callahan et al., 2015) in R v. 3.6.3 (R Core Team, 2019) was then used to further quality filter sequences. Sequences were quality filtered with a max expected error of 2, and those identified as chimeras were removed. Sequences were then dereplicated to produce amplicon sequence variants (ASVs). ASVs were matched to the NCBI GenBank reference database (www.ncbi.nlm.nih.gov/genbank/) using the Basic Local Alignment Search Tool (BLAST) for taxonomic assignment using a high-performance cluster computer (Pawsey Computing Centre, Perth, Western Australia). Taxonomic assignments were made to the lowest common ancestor (LCA) using the LCA script from eDNAFlow (Mousavi-Derazmahalleh et al., 2021) with a minimum query coverage of 100% and identity threshold of 95%. Where the absolute value of the difference between % identity of ASVs was <1, species taxonomy was not returned and the ASV was assigned to the closed common ancestor. ASVs identified in the fly bait; cow (*Bos taurus*), pig (*Sus scrofa*), and sheep (*Ovis aries*); negative controls, cow and chicken (*Gallus gallus*), both common laboratory contaminants; and human (*Homo sapiens*), were removed from the data set using the 'phyloseq' package (McMurdie & Holmes, 2013). ASVs were then agglomerated at a species level, retaining ASVs identified at a higher taxonomic level. For analysis requiring ASV tables from both assays to be combined, ASV tables were merged whereby if the vertebrate species identity was identical between both assays the ASVs were agglomerated. The ASV tables from each primer set were kept separate for other analyses.

5.4.6 Statistical analysis

All statistical tests were run on R v. 3.6.3 (R Core Team, 2019). All maps were created using QGIS v. 3.2.3-'Bonn' (QGIS Development Team, 2018). To investigate the effectiveness of our sampling effort, the ASVs from both assays were combined. Species accumulation curves for only zoo taxa and overall taxa detected were created using the 'specaccum' function from the 'vegan' package (Oksanen et al., 2019) using the exact method with 999 permutations. From the species that were detected at the zoo and along the transects (n=4), the number of sequences found at each distance (metres) was assessed using a Spearman rank correlation. Sequences are not compared between species, but rather together from sequences of zoo animals located in the zoo (0 m) to 4 km away.

We then assessed the influence of taxonomic group and biomass on the detectability of zoo taxa. First, for each species present at the zoo, biomass was calculated based on the number of individuals kept at the zoo multiplied by the average mass of an individual of that species. The average mass was calculated based on a literature search. Then, two logistic regression models with binomial distribution were used to test the probability of detection of a species kept at the zoo from carrion fly iDNA (based on presence/absence of the species), the first with the factor of taxon (four levels: amphibian, bird, mammal, and reptile) was created using the species detections from the general vertebrate 12s primer set (Riaz et al., 2011), so as to avoid any confounding effects of a mammal specific primer set. The second model was created with the total biomass of each species as a continuous explanatory variable, and this was conducted on the combined species detections from both primers (vertebrate 12s and mammal 16s).

5.5 Results

In total, 920 flies were collected. Over 13,000,000 reads were generated from these flies, with an average per sample of 52,546 (± 4,003 s.e.) from the mammal 16s assay and 61,254 (± 7,265 s.e.) from the vertebrate 12s assay. The use of passive traps, where flies can interact with fly bait, meant that the bait signal (cow, pig, and sheep) was high across samples. After removing ASVs associated with bait, there was a significant reduction in the number of sequences, with an average of 7.5% of sequences retained overall from samples across both primer sets. Within these, 50 ASVs not associated with the fly bait were detected. These included vertebrates across four taxonomic classes spanning species kept in the zoo, animals used as food sources for zoo animals and humans, domestic, and wild animals known to occur in the study area (Table 5-1).

The following are results from the combination of both assays. Out of the 123 samples, 87 had vertebrate detections; 64 out of 90 samples along the transects and 23 out of 35 samples from the zoo. On average, there were 5 (± 0.4 s.e.) vertebrate ASVs per sample. There were 19 samples in total that contained zoo taxa, eight samples from the zoo and 11 samples along the transects. Of the six target species, only the lion was detected from iDNA (from both assays). There were no detections of the other target mammal, the numbat. There were also no detections of the two target bird species, emu and southern cassowary, and the two target reptile species, the radiated tortoise, and Galapagos tortoise. The other taxa that were detected using both assays included two amphibians, one reptile, one bird, and two mammal species that were wild species. The nine other remaining detections were of domestic species classed as pets: domestic dogs (*Canis lupus familiaris*) and cats (*Felis catus*); or food sources for humans or zoo animals: deer (*Cervus sp.*), goat (*Capra hircus*), and barramundi (*Lates calcifer*) (Table 5-1). Species accumulation curves calculated from both assays at all sites (Fig 5-2A) and within the zoo (Fig 5-2B) did not reach asymptote indicating that more sampling was needed to capture the vertebrate diversity that carrion fly iDNA was likely to sample.

Within the 64 bulk carrion fly samples collected along the two 4 km transects, only four of the taxa present in the zoo were detected along the transects (from both assays), these were all mammal species: lion, African painted dog (*Lycaon pictus*), Javan gibbon (*Hylobates moloch*), and cotton-top tamarin (*Saguinus oedipus*) (Fig 5-3). The number of sequencing reads from these species decreased as samples were collected at greater distances from the zoo.

Table 5-1 The vertebrate species detected through metabarcoding of carrion fly iDNA, not including humans and species that are found in commercial fly bait. Common names are included where appropriate. The detection location of each animal is noted as follows: Zoo (red circle, Z), Transect 1 (yellow circle, 1), and Transect 2 (green circle, 2). Further notation is included for animals other than the detected zoo taxa to indicate whether the animal was a domestic pet (†) or was a food source animal for humans or zoo animals (*).

Class	Order	Family	Genus	Species	Location
Zoo animals					
Aves	Pelecaniformes	Ardeidae	*Ardea*		Z
Mammalia	Primates	Cebidae	*Saguinus*	*Saguinus oedipus* (Cotton-top tamarin)	Z 2
Mammalia	Primates	Cebidae	*Saguinus*		Z
Mammalia	Primates	Cercopithecidae	*Papio*	*Papio hamadryas* (Hamadryas baboon)	Z
Mammalia	Primates	Hylobatidae	*Hylobates*	*Hylobates moloch* (Javan gibbon)	Z 1 2
Mammalia	Carnivora	Felidae	*Panthera*	*Panthera leo* (Lion)	Z 1 2
Mammalia	Carnivora	Viverridae	*Arctictis*	*Arctictis binturong* (Binturong)	Z
Mammalia	Diprotodontia	Macropodidae	*Macropus*	*Macropus fuliginosus* (Western grey kangaroo)	2
Mammalia	Diprotodontia	Macropodidae	*Osphranter*	*Osphranter rufus* (Red kangaroo)	Z
Mammalia	Carnivora	Canidae	*Lycaon*	*Lycaon pictus* (African painted dog)	Z 2
Wild, domestic†, or food source animals*					
Actinopteri	Perciformes	Centropomidae	*Lates*	*Lates calcarifer* (Barramundi)*	2
Actinopteri	Scombriformes	Scombridae*			Z 1 2
Amphibia	Anura	Hylidae	*Litoria*	*Litoria moorei* (Motorbike frog)	Z
Amphibia	Anura	Myobatrachidae	*Pseudophryne*	*Pseudophryne guentheri* (Gunther's toadlet)	Z
Aves	Pelecaniformes	Ardeidae	*Nycticorax*		Z
Lepidosauria	Squamata	Scincidae	*Hemiergis*		1
Mammalia	Artiodactyla	Bovidae	*Capra*	*Capra hircus* (Goat)*	Z 1 2
Mammalia	Artiodactyla	Cervidae	*Cervus**		1 2
Mammalia	Carnivora	Canidae	*Canis*	*Canis lupus familiaris* (Domestic dog) †	Z 1 2
Mammalia	Carnivora	Canidae	*Canis†*		Z 1 2
Mammalia	Carnivora	Felidae	*Felis*	*Felis catus* (Domestic cat)†	Z 1 2
Mammalia	Diprotodontia	Phalangeridae	*Trichosurus*	*Trichosurus vulpecula* (Brush tailed possum)	Z 1 2
Mammalia	Rodentia	Muridae	*Rattus*		Z 1 2

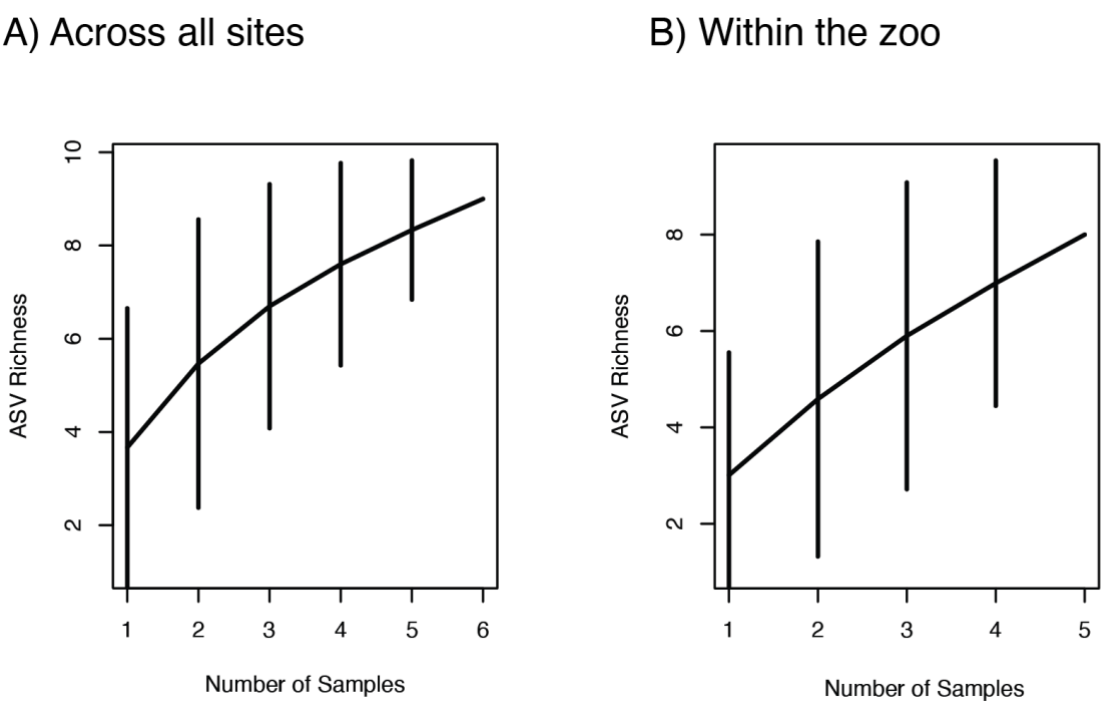


Figure 5-2 Species accumulation curve for zoo animals detected from carrion flies caught on seven sampling days. A) The species accumulation across all sample collection sites and B) The species accumulation from only the zoo sample collection sites.

This decrease was approaching statistical significance at α = 0.05 (Spearman rank correlation R = -0.33, p=0.069, n= 17). For example, within the samples from within the zoo, a total of 8,889 sequences of African painted dog were detected, while only 16 sequences were detected in samples collected 1.25 km from the zoo along Transect 2 (Fig 5-3B). The furthest detection of a zoo animal was in a sample collected 4 km from the zoo in Transect 2 where a single sequence of lion was detected (Fig 5-3A).

From the vertebrate 12s assay only, four species kept at the zoo were detected: one bird species of the 40 bird species with access to outdoor enclosures and three mammal species of the 44 mammal species with access to outdoor enclosures. With the addition of the mammal 16s assay, a further four mammal species found within the zoo could be detected. Using the data from the vertebrate 12s assay only, the likelihood of being detected did not differ between taxon (Deviance explained= 0.0008; d.f.= 3,97; p=0.28).

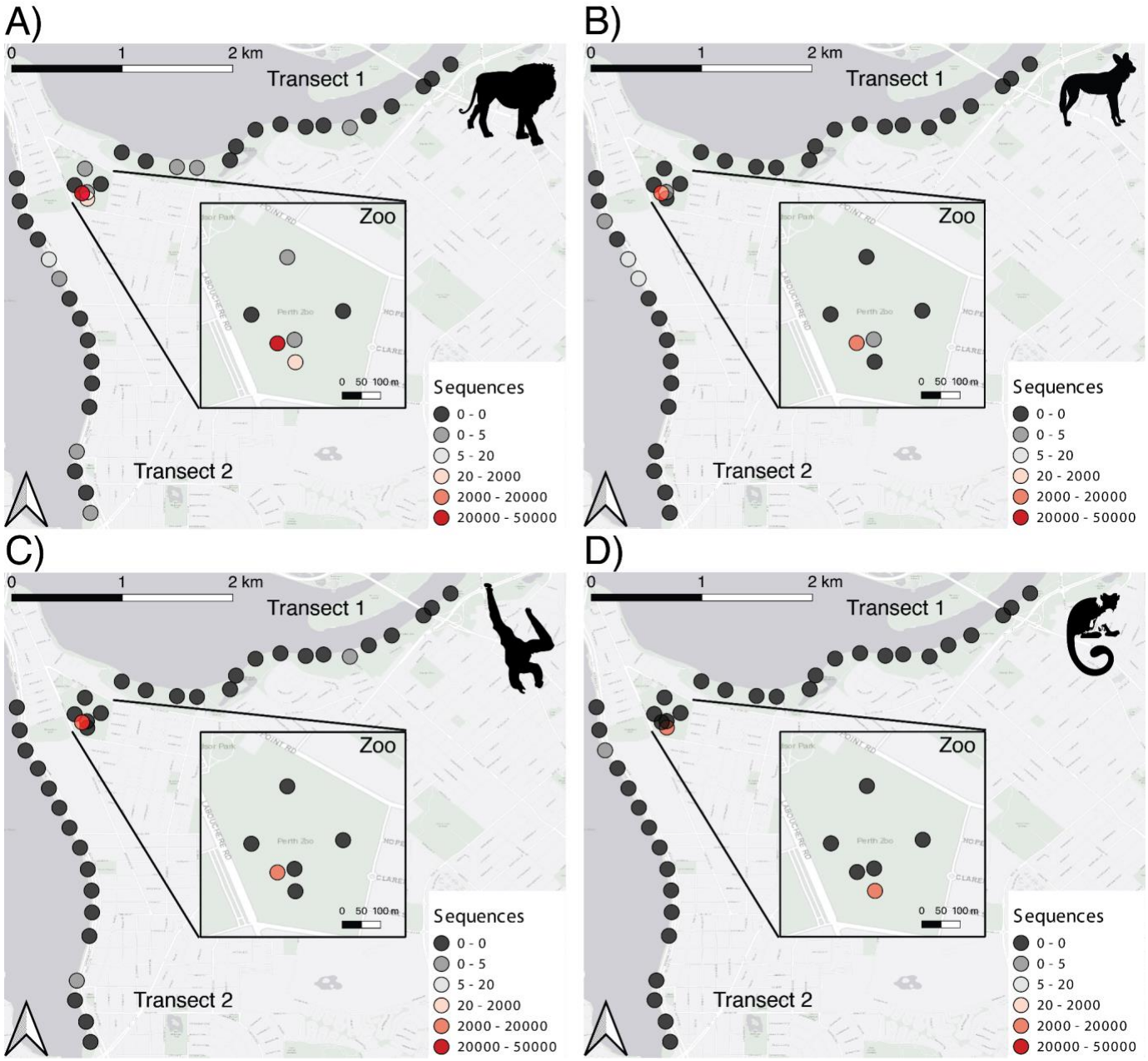


Figure 5-3 Number of sequences of the four zoo mammal species detected both at the zoo (inset) and along the transects A) *Panthera leo* (Lion), B) *Lycaon pictus* (African painted dog) C) *Hyloblates moloch* (Javan gibbon), and D) *Saguinus oedipus* (Cotton-top tamarin). Colour of circles indicates the total number of sequences detected from samples at the transect point across all sampling days. Silhouettes of the detected animals are included.

From the 104 species residing at the zoo, mammals made up 87.6% of the total biomass, followed by reptiles at 9.7% and birds at 2.7%. Some of the mammals contributing to the high total biomass such as the white rhinoceros (*Ceratotherium simum*) (~4000 kg) or giraffes (*Giraffa sp.*) (~4125 kg) were not detected, and neither were mammals with a very small total biomass, such as the black-capped squirrel monkey (*Saimiri boliviensis*) (~1.25kg). The total biomass of the detected zoo animals ranged from ~4.3 kg (cotton-top tamarin) to ~440 kg (African hunting dog). The likelihood of being detected indicated that there was no significant relationship with biomass (Deviance explained= 0.0006; d.f. = 1,99, p= 0.85).

5.6 Discussion

Invertebrate-derived DNA (iDNA) is increasingly used to assess vertebrate species diversity (e.g., Hoffmann et al., 2016; Massey et al., 2021; Rodgers et al., 2017; Schubert et al., 2015). Therefore, it is vital to examine the factors that influence detection of species using this technique. Here, we used two metabarcoding assays to detect vertebrate taxa spanning zoo, domestic and wild vertebrate species, from carrion flies collected from within the Perth Zoo, and two transects extending away from the zoo. Despite carrion flies' assumed mobility, we found that the signals from iDNA detections were predominantly geographically localised to the zoo, with a maximum detection distance of 4 km away from the zoo. We found no association between the detection of an animal and its taxon group (i.e., mammal, reptile, and bird). Further, there was no relationship between the biomass of zoo animals and their detection from iDNA. This indicates that the detections of vertebrate species from carrion fly iDNA are largely indiscriminate and restricted to a fine spatial scale.

5.6.1 Distance from a source of DNA

Detections of zoo vertebrate species from carrion fly iDNA were highly localised to the location of the vertebrates in the zoo, even though previous studies have shown that carrion flies can travel 2.2 - 6.4 km over a 24-hr period depending on the species (Braack & Retief, 1986; Norris, 1959, 1965). However, these studies aimed to discover the maximum dispersal distance from a release event and not the distance a fly travels from a food source. In the presence of a food source, it is possible to distinguish carrion fly communities within 300 m (Perez et al. 2016). Further, these studies have not shown how structural components of the environment may affect movement (e.g., in an urban setting, the impacts of roads or buildings). Within the samples collected from the zoo, zoo animal species were detected with relatively high numbers of sequences, indicating a highly localised signal that weakens with distance from the DNA source. Recent studies using air eDNA also found that detection probability improved with closer proximity to the vertebrate DNA source (Lynggaard et al., 2022). It should be noted that the number of sequences cannot be interpreted without the caveat that PCR and sequencing conditions can skew sequence abundance (Fonseca, 2018; Murray et al., 2015), especially in samples of mixed taxa. Thus, we did not compare sequence numbers between species. However, from the overall data, there is an observed (but

statistically insignificant) pattern of decline in the number of sequences from zoo taxa further away from the zoo. Detections at the max distance of 4 km suggest that the sampled distance was too small, and the potential to detect species beyond this scale is unknown. However, whilst zoo animals could be detected up to 4 km away from the zoo, these were represented by only very few sequences (<20), and more stringent filtering requirements may have removed these detections completely from the dataset. Investigating the geographic detection range of species in metabarcoding studies represents an important avenue for future research. The movement of eDNA in terrestrial systems has been poorly explored, and future studies are needed that examine the likelihood of eDNA movement by different forces including water and air.

This localisation effect is also similar to a previous vertebrate survey using iDNA from leeches, which have very specific microhabitats and limited dispersal abilities (Schnell et al., 2015). While we assumed that carrion flies would transport iDNA much further than leeches because of the latter's poor dispersal ability, our findings suggest that this may not be the case, and that the distances are comparative (Ji et al., 2022). This is beneficial, as leeches are not present in many environment types as carrion flies, and sampling carrion flies using baited traps is relatively easier. A narrow detection distance can be helpful in fine-scale studies, such as tracking threatened vertebrate species, determining species occupancy, or tracking the invasion fronts of invasive species. However, this is with the caveat that in heterogenous habitats, the scale of detection within ~1 km may be too coarse. Additionally, the distances required to measure occupancy is heavily dependent on the ecological questions being asked. Especially for rare species with narrow habitat ranges, where a broader detection within a range of 1-2 km could introduce a false positive detection and impact occupancy models. Further investigation using this technique to look at distance in different habitat types, in different climatic conditions, or during different seasonal conditions, which can all impact fly dispersal patterns (MacLeod & Donnelly, 1963; Norris, 1965), is required to test the broader applicability of our findings.

5.6.2 Detection of different animal taxon

We found there was no relationship between the taxonomic group of an animal and its detectability. Even when traps were located in enclosures containing our target reptile and bird species, detection success was poor. These results are comparable to those of other studies using different eDNA substrates to detect vertebrates, such as air (Clare et al., 2022; Lynggaard et al., 2022), water (Mas-Carrió et al., 2022), and soil (Ryan et al., 2022), where mammals were most commonly represented in detections, and reptile species were least represented or completely absent. However, unlike these studies, detection of birds from carrion fly iDNA was poor, with only two bird taxa detected, and neither of our target species. As we assume that the detections from carrion flies are predominantly related to feeding on faeces (Owings et al., 2019), and especially in a zoo environment, the lack of bird detections from carrion fly iDNA could be related to the presence of free DNases in avian faeces (Regnaut et al., 2006), which is known to rapidly degrade DNA.

The difficulties associated with detecting reptiles through metabarcoding could possibly be explained by the relatively lower deposition rate of reptile waste compared to other classes of vertebrates (Adams et al., 2019); however, tortoise faeces, which would theoretically provide a food source for carrion flies, was observed in the enclosures of both species. The absence of both target reptile species from our results may indicate either avoidance of reptile faeces as a food source for carrion flies or rapid and extensive degradation of reptile DNA in waste. Whilst it is possible to obtain DNA from reptile faeces to conduct monitoring and analysis (Jones et al., 2008; Pearson et al., 2015; Ratsch et al., 2020), this is usually dependent on the collection of fresh samples because of the high rates of DNA degradation by faecal microbes in reptiles (Jones et al., 2008). Additionally, as iDNA degrades quickly within the digestive tract of invertebrates (Lee et al., 2015; Wilting et al., 2021), it could be possible that DNA from reptiles degraded to the point where it was unidentifiable.

Primer bias may also have affected the detection of these target species. In this study, we used the same metabarcoding primer sets as used in previous carrion fly iDNA studies (Calvignac-Spencer et al., 2013; Hoffmann et al., 2018; Kocher et al., 2017; Schubert et al., 2015), which included a general vertebrate primer (Riaz et al., 2011) and a mammal specific primer (Taylor, 1996). While the general vertebrate primer set (Riaz et al., 2011) can detect mammals, birds, and amphibians and, in-silico, matched against our target taxa from reptiles and birds, they may exhibit preferential binding towards mammals in a mix of DNA and may have preferentially bound to the fly bait DNA. Whilst we detected more mammals with a mammal-specific primer set, we could only detect a fraction of the total mammal biodiversity present at the zoo. Therefore, whilst the inclusion of a more specialised taxon-specific reptile or bird primer set may increase the biodiversity detected from these taxon groups, our results show that more intensive sampling is required to increase the detections overall. Increasing sampling intensity could be accomplished through a longer sampling regime or installing more traps. Also, increasing the sequencing depth of samples (Singer et al., 2019) or using an inorganic bait solution may have improved the identification of zoo vertebrate species.

Like other studies on carrion fly iDNA, we found that vertebrate signals were primarily overwhelmed by the DNA associated with fly bait (Calvignac-Spencer et al., 2013). Another study that used a mammalian bait to attract flies also found that only 12% of their sequences were non-bait taxa (Calvignac-Spencer et al., 2013), similar to the 7.5% of sequences in our study that were non-bait. To overcome this issue, some iDNA studies have used blocking primers, enabling the detection of more 'target' species to reveal more 'wild' diversity by minimising the detection of domestic animal signals or the animal used as a fly bait (Hoffmann et al., 2018; Schubert et al., 2015). Blocking primers work by selectively blocking the DNA of a species whose amplification can potentially mask the presence of DNA from other species, for example, blocking the amplification of human DNA (Rojahn et al., 2021). Blocking primers are common in dietary studies when the amplification of the host DNA can overwhelm the signal from the dietary prey items (Shehzad et al., 2012) or for avoiding human contaminant sequences (Andersen et al., 2012). Hence, having prior knowledge of the content of fly bait can be vital. However, blocking primers can also block more taxa than just their intended target, thereby introducing bias in survey results (Vestheim & Jarman, 2008). Therefore, improving trap design to avoid the interaction between flies and the bait/create a barrier between the bait and collection fluid may be more beneficial than blocking primers to reduce the signal.

5.6.3 Impact of biomass on detection

In contrast to the findings of studies using water or air as sampler of vertebrate diversity (Carvalho et al., 2021; Lynggaard et al., 2022), we found no correlation between the detections of our species and increasing biomass. Our data also suggests that larger animals are not more likely to be detected than smaller ones. This is important as many smaller-bodied, arboreal, or volant mammals, such as rodents or bats, are difficult to detect using approaches such as camera trapping (Gogarten et al., 2019). An explanation for this finding is that flies do not move far from a food source, so a fly's proximity to a sampling point would be the main driving factor of iDNA detection rather than an animal's total biomass. Physical barriers may also impede movement or odour cues that would attract carrion flies to a potential food source (Benbow et al., 2015). In our study, many of the mammals with the largest total biomass, such as the elephants (*Elephas maximus*), giraffes (*Giraffa sp.*) and zebras (*Equus quagga*), which were all undetected, were located in sunken enclosures that might physically restrict the movement of flies from this area to the fly traps in other locations.

5.7 Conclusions

We examined the impact of various organism characteristics on vertebrate species detection using iDNA metabarcoding. We found detections from carrion flies to be spatially localised, and context-specific, and that taxon and biomass had no effect on vertebrate species detection likelihood. While no monitoring method is unbiased, and detectability between different monitoring methods is variable across taxonomic classes and sizes of animals, understanding these biases is important for interpreting results in an ecological context. Here, we answer some critical questions needed for interpreting iDNA-based detections. As such, we provide an important steppingstone for improving the uptake of iDNA techniques in terrestrial ecological monitoring.

5.8 References

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged

Abrams, J. F., Hörig, L. A., Brozovic, R., Axtner, J., Crampton-Platt, A., Mohamed, A., Wong, S. T., Sollmann, R., Yu, D. W., & Wilting, A. (2019). Shifting up a gear with iDNA : From mammal detection events to standardised surveys. *The Journal of Applied Ecology*, *56*(7), 1637–1648.

Adams, C. I. M., Hoekstra, L. A., Muell, M. R., & Janzen, F. J. (2019). A Brief Review of Non-Avian Reptile Environmental DNA (eDNA), with a Case Study of Painted Turtle (Chrysemys picta) eDNA Under Field Conditions. *Diversity*, *11*(4), 50.

Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjaer, K. H., Orlando, L., Gilbert, M. T. P., & Willerslev, E. (2012). Meta-barcoding of "dirt" DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, *21*(8), 1966–1979.

Bagnall, A. J. (2017). *Use of iDNA in detecting elusive mammals: a case study on the Indonesian island of Seram* [PhD Thesis, The University of Edinburgh]. https://era.ed.ac.uk/handle/1842/35558

Balmford, A., Bennun, L., Brink, B. T., Cooper, D., Côte, I. M., Crane, P., Dobson, A., Dudley, N., Dutton, I., Green, R. E., Gregory, R. D., Harrison, J., Kennedy, E. T., Kremen, C., Leader-Williams, N., Lovejoy, T. E., Mace, G., May, R., Mayaux, P., … Walther, B. A. (2005). Ecology: The Convention on Biological Diversity's 2010 target. *Science*, *307*(5707), 212–213.

Benbow, E. M., Tomberlin, J. K., & Tarone, A. M. (2015). *Carrion Ecology, Evolution, and Their Applications*. CRC Press.

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A., & Creer, S. (2021). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231–1246.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12428

Braack, L. E., & Retief, P. F. (1986). Dispersal, density and habitat preference of the blow-flies *Chrysomyia albiceps* (Wd.) and *Chrysomyia marginalis* (Wd.) (Diptera: Calliphoridae). *The Onderstepoort Journal of Veterinary Research*, *53*(1), 13–18.

Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., & Boutin, S. (2015). Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. *The Journal of Applied Ecology*, *52*(3), 675–685.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., & Han, A. W. (2015). DADA2: High resolution sample inference from amplicon data. *bioRxiv*. https://www.biorxiv.org/content/10.1101/024034v1.abstract

Calvignac-Spencer, S., Leendertz, F. H., Gilbert, M. T. P., & Schubert, G. (2013). An invertebrate stomach's view on vertebrate ecology: certain invertebrates could be used as "vertebrate samplers" and deliver DNA-based information on many aspects of vertebrate ecology. *BioEssays*, *35*(11), 1004–1013.

Calvignac-Spencer, S., Merkel, K., & Kutzner, N. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Molecular Ecology, 22*(4), 915-924. https://doi.org/10.1111/mec.12183

Campbell, G., Kuehl, H., Diarrassouba, A., N'Goran, P. K., & Boesch, C. (2011). Long-term research sites as refugia for threatened and over-harvested species. *Biology Letters*, *7*(5), 723–726.

Carvalho, C. S., de Oliveira, M. E., Rodriguez-Castro, K. G., Saranholi, B. H., & Galetti, P. M., Jr. (2021). Efficiency of eDNA and iDNA in assessing vertebrate diversity and its abundance. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13543

Clare, E. L., Economou, C. K., Bennett, F. J., Dyer, C. E., Adams, K., McRobie, B., Drinkwater, R., & Littlefair, J. E. (2022). Measuring biodiversity from DNA in the air. *Current Biology*, *32*(3), 693–700.e5.

Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. *Molecular Ecology Resources*, *18*(5), 923–926.

Gogarten, J. F., Hoffmann, C., Arandjelovic, M., Sachse, A., Merkel, K., Dieguez, P., Agbor, A., Angedakin, S., Brazzola, G., Jones, S., Langergraber, K. E., Lee, K., Marrocoli, S., Murai, M., Sommer, V., Kühl, H., Leendertz, F. H., & Calvignac-Spencer, S. (2019). Fly-derived DNA and camera traps are complementary tools for assessing mammalian biodiversity. *Environmental DNA*, *9*, 79.

Hoffmann, C., Merkel, K., Sachse, A., Rodríguez, P., Leendertz, F. H., & Calvignac-Spencer, S. (2018). Blow flies as urban wildlife sensors. *Molecular Ecology Resources*, *18*(3), 502–510.

Hoffmann, C., Stockhausen, M., Merkel, K., Calvignac-Spencer, S., & Leendertz, F. H. (2016). Assessing the feasibility of fly based surveillance of wildlife infectious diseases. *Scientific Reports*, *6*, 37952.

Ji, Y., Baker, C. C. M., Popescu, V. D., Wang, J., Wu, C., Zhengyang, W., Li, Y., Wang, L., Hua, C., Yang, Z., Yang, C., Xu, C. C. Y., Diana, A., Wen, C., Pierce., N. E., Yu., D. W. (2022). Measuring protected-area effectiveness using vertebrate distributions from leech iDNA. *Nature Communications*, *13*, 1555

Jones, R., Cable, J., & Bruford, M. W. (2008). An evaluation of non-invasive sampling for genetic analysis in northern European reptiles. *The Herpetological Journal*, *18*(1), 32–39.

Kocher, A., de Thoisy, B., Catzeflis, F., Valière, S., Bañuls, A.-L., & Murienne, J. (2017). iDNA screening: Disease vectors as vertebrate samplers. *Molecular Ecology*, *26*(22), 6478–6486.

Kouakou, C. Y., Boesch, C., & Kuehl, H. (2009). Estimating chimpanzee population size with nest counts: validating methods in Taï National Park. *American Journal of Primatology*, *71*(6), 447–457.

Lee, P.-S., Sing, K.-W., & Wilson, J.-J. (2015). Reading Mammal Diversity from Flies: The Persistence Period of Amplifiable Mammal mtDNA in Blowfly Guts (*Chrysomya megacephala*) and a New DNA Mini-Barcode Target. *PloS One*, *10*(4), e0123871.

Legge, S., Robinson, N., Lindenmayer, D., Scheele, B., Southwell, D., & Wintle, B. (2018). *Monitoring Threatened Species and Ecological Communities*. CSIRO Publishing.

Lynggaard, C., Bertelsen, M. F., Jensen, C. V., Johnson, M. S., Frøslev, T. G., Olsen, M. T., & Bohmann, K. (2022). Airborne environmental DNA for terrestrial vertebrate community monitoring. *Current Biology*, *32*(3), 701–707.e5.

Lynggaard, C., Nielsen, M., Santos-Bay, L., Gastauer, M., Oliveira, G., & Bohmann, K. (2019). Vertebrate diversity revealed by metabarcoding of bulk arthropod samples from tropical forests. *Environmental DNA*, *1*(4), 329–341.

MacLeod, J., & Donnelly, J. (1963). Dispersal and Interspersal of Blowfly Populations. *The Journal of Animal Ecology*, *32*(1), 1–32.

Marsh, D. M., & Trenham, P. C. (2008). Current trends in plant and animal population monitoring. *Conservation Biology*, *22*(3), 647–655.

Mas-Carrió, E., Schneider, J., Nasanbat, B., Ravchig, S., Buxton, M., Nyamukondiwa, C., Stoffel, C., Augugliaro, C., Ceacero, F., Taberlet, P., Glaizot, O., Christe, P., & Fumagalli, L. (2022). Assessing environmental DNA metabarcoding and camera trap surveys as complementary tools for biomonitoring of remote desert water bodies. *Environmental DNA*, *4*(3), 580–595.

Massey, A. L., Bronzoni, R. V. de M., da Silva, D. J. F., Allen, J., de Lazari, P. R., Dos Santos Filho, M., Canale, G. R., Bernardo, C. S. S., Peres, C. A., & Levi, T. (2021). Invertebrates for vertebrate biodiversity monitoring: comparisons using three insect taxa as iDNA samplers. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13525

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One*, *8*(4), e61217.

Mousavi-Derazmahalleh, M., Stott, A., Lines, R., Peverley, G., Nester, G., Simpson, T., Zawierta, M., De La Pierre, M., Bunce, M., & Christophersen, C. T. (2021). eDNAFlow, an automated, reproducible and scalable workflow for analysis of environmental DNA (eDNA) sequences exploiting Nextflow and Singularity. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13356

Murray, D. C., Coghlan, M. L., & Bunce, M. (2015). From benchtop to desktop: important considerations when designing amplicon sequencing workflows. *PloS One*, *10*(4), e0124671.

Newton, J. P., Bateman, P. W., Heydenrych, M. J., Mousavi-Derazmahalleh, M., & Nevill, P. (2022). Home is where the hollow is: Revealing vertebrate tree hollow user biodiversity with eDNA metabarcoding. *Environmental DNA*. https://doi.org/10.1002/edn3.313

Nielsen, M., Gilbert, M. T. P., Pape, T., & Bohmann, K. (2019). A simplified DNA extraction protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity. *Environmental DNA*, edn3.16.

Norris, K. R. (1959). The Ecology of Sheep Blowflies in Australia. In A. Keast, R. L. Crocker, & C. S. Christian (Eds.), *Biogeography and Ecology in Australia* (pp. 514–544). Springer Netherlands.

Norris, K. R. (1965). The Bionomics of Blow Flies. *Annual Review of Entomology*, *10*(1), 47–68.

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (Version 2.5-6) [Computer software]. https://cran.ism.ac.jp/web/packages/vegan/vegan.pdf

Owings, C. G., Banerjee, A., Asher, T. M. D., Gilhooly, W. P., 3rd, Tuceryan, A., Huffine, M., Skaggs, C. L., Adebowale, I. M., Manicke, N. E., & Picard, C. J. (2019). Female Blow Flies As Vertebrate Resource Indicators. *Scientific Reports*, *9*(1), 10594.

Pearson, S. K., Tobe, S. S., Fusco, D. A., Bull, C. M., & Gardner, M. G. (2015). Piles of scats for piles of DNA: deriving DNA of lizards from their faeces. *Australian Journal of Zoology*, *62*(6), 507–514.

Perez, A. E., Haskell, N. H., & Wells, J. D. (2016). Commonly used intercarcass distances appear to be sufficient to ensure independence of carrion insect succession pattern. *Annals of the Entomological Society of America*, *109*(1), 72-80.

QGIS Development Team. (2018). *QGIS Geographic Information System* (3.2.3-Bonn) [Computer software]. http://qgis.osgeo.org.

Ratsch, R., Kingsbury, B. A., & Jordan, M. A. (2020). Exploration of Environmental DNA (eDNA) to Detect Kirtland's Snake (*Clonophis kirtlandii*). *Animals*, *10*(6). https://doi.org/10.3390/ani10061057

R Core Team. (2019). *R: A Language and Environment for Statistical Computing* (Version 3.6.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Regnaut, S., Lucas, F. S., & Fumagalli, L. (2006). DNA degradation in avian faecal samples and feasibility of non-invasive genetic studies of threatened capercaillie populations. *Conservation Genetics*, *7*(3), 449–453.

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, *39*(21), e145.

Rodgers, T. W., & Mock, K. E. (2015). Drinking water as a source of environmental DNA for the detection of terrestrial wildlife species. *Conservation Genetics Resources*, *7*(3), 693–696.

Rodgers, T. W., Xu, C. C. Y., Giacalone, J., Kapheim, K. M., Saltonstall, K., Vargas, M., Yu, D. W., Somervuo, P., McMillan, W. O., & Jansen, P. A. (2017). Carrion fly-derived DNA metabarcoding is an effective tool for mammal surveys: Evidence from a known tropical mammal community. *Molecular Ecology Resources*, *17*(6), e133–e145.

Rojahn, J., Gleeson, D. M., Furlan, E., Haeusler, T., & Bylemans, J. (2021). Improving the detection of rare native fish species in environmental DNA metabarcoding surveys. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *31*(4), 990–997.

Ryan, E., Bateman, P., Fernandes, K., Heyde, M., & Nevill, P. (2022). eDNA metabarcoding of log hollow sediments and soils highlights the importance of substrate type, frequency of sampling and animal size, for vertebrate species detection. *Environmental DNA*. https://doi.org/10.1002/edn3.306

Schnell, I. B., Sollmann, R., Calvignac-Spencer, S., Siddall, M. E., Yu, D. W., Wilting, A., & Gilbert, M. T. P. (2015). iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool - prospects, pitfalls and avenues to be developed. *Frontiers in Zoology*, *12*, 24.

Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R. D., Willerslev, E., Bertelsen, M. F., & Gilbert, M. T. P. (2012). Screening mammal biodiversity using DNA from leeches. *Current Biology*, *22*(8), R262–R263.

Schubert, G., Stockhausen, M., Hoffmann, C., Merkel, K., Vigilant, L., Leendertz, F. H., & Calvignac-Spencer, S. (2015). Targeted detection of mammalian species using carrion fly-derived DNA. *Molecular Ecology Resources*, *15*(2), 285–294.

Shehzad, W., Riaz, T., Nawaz, M. A., Miquel, C., Poillot, C., Shah, S. A., Pompanon, F., Coissac, E., & Taberlet, P. (2012). Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology*, *21*(8), 1951–1965.

Singer, G. A. C., Fahner, N. A., Barnes, J. G., McCarthy, A., & Hajibabaei, M. (2019). Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Scientific Reports*, *9*(1), 5991.

Spivak, M., Conlon, D., & Bell, W. J. (1991). Wind-Guided Landing and Search Behavior in Fleshflies and Blowflies Exploiting a Resource Patch (Diptera: Sarcophagidae, Calliphoridae). *Annals of the Entomological Society of America*, *84*(4), 447–452.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050.

Taylor, P. G. (1996). Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Molecular Biology and Evolution*, *13*(1), 283–285.

van der Heyde, M., Bateman, P. W., Bunce, M., Wardell-Johnson, G., White, N. E., & Nevill, P. (2021). Scat DNA provides important data for effective monitoring of mammal and bird biodiversity. *Biodiversity and Conservation*, *30*(12), 3585–3602.

Vestheim, H., & Jarman, S. N. (2008). Blocking primers to enhance PCR amplification of rare sequences in mixed samples - a case study on prey DNA in Antarctic krill stomachs. *Frontiers in Zoology*, *5*(1), 12.

Wilting, A., Nguyen, T. V., Axtner, J., Nguyen, A., Schmidt, A., Le, M., Nguyen, A. H. Q., Rawson, B. M., Tilker, A., & Fickel, J. (2021). Creating genetic reference datasets: Indirect sampling of target species using terrestrial leeches as sample “collectors.” *Environmental DNA*. https://doi.org/10.1002/edn3.256

This page has been left intentionally blank

CHAPTER 6

CARRION FLY iDNA METABARCODING TO MONITOR MAMMALS IN A FRAGMENTED TERRESTRIAL ECOSYSTEM

6.1 Preface

This chapter uses the results from Chapter 5 to inform the methods and interpretations of data gathered from carrion fly iDNA. Here, the focus is implementation of iDNA metabarcoding for conservation management with the wheatbelt region of Western Australia as a focal study. This region is highly fragmented, with much of the natural vegetation cleared for cereal cropping (hence the name, wheatbelt). However, in small pockets of remnant vegetation there lies highly biodiverse communities of mammals. Many of which are classed as 'Threatened' and protected under the Australian Government's *Environment Protection and Biodiversity Conservation Act 1999* (EPBC Act). Therefore, monitoring of these species is essential for management in this region. In areas where little is known about species assemblages, or where access to conservation reserves is restricted by terrain, iDNA metabarcoding can be a viable solution to mitigate these issues. Thus, this chapter explores the utility of carrion fly iDNA metabarcoding in a conservation context, specifically looking at whether this technique can determine the movement of conservation critical species and invasive predator species through a fragmented landscape.

This chapter consists of a manuscript that is currently being prepared for submission to the peer reviewed journal *Molecular Ecology*.

6.1.1 Acknowledgements

This work was supported by computational resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. KF was partially supported by funding from Food Agility CRC Ltd, funded under the Commonwealth Government CRC Program

6.1.2 Data Accessibility

R scripts and the accompanying data will be made available here: 10.5281/zenodo.6767690

6.1.3 Author Contributions

KF, PWB, BJS, KB and PN conceptualised the study. KF collected the samples and processed them. KF conducted all the analysis. BJS assisted with statistical analysis. KF wrote the original draft of the manuscript. All authors contributed to the review and editing of the manuscript.

Monitoring biodiversity on-the-fly: Carrion fly iDNA metabarcoding to monitor mammals in a fragmented terrestrial ecosystem.

Kristen Fernandes [1,2,3*], Philip W Bateman [4], Benjamin J Saunders [5], Kristine Bohmann [2], Paul Nevill [1,6]

1 Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

2 Section for Molecular Ecology and Evolution, Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen K, Denmark

3 Food Agility CRC Ltd, 175 Pitt St Sydney, NSW 2000, Australia

4 Behavioural Ecology Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

5 School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

6 ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

6.2 Abstract

Predation from invasive predators, habitat clearing, and inappropriate fire regimes have been attributed as the primary causes of mammal declines in Australia. In severely fragmented habitats, risks to native mammal populations are higher due to the isolation of reserve areas and the increase in edge effects. Therefore, monitoring these reserves, and the anthropogenically-altered landscapes between, can be valuable in tracking mammal population movement through landscapes. Here we use metabarcoding of invertebrate-derived DNA (iDNA) from carrion flies to track mammal populations in the wheatbelt region of Western Australia, where widespread clearing for agriculture has removed most of the native perennial vegetation and replaced it with an agricultural system. We aimed to investigate if the localisation of the iDNA signal reflected the expected distribution of four native species - echidna *(Tachyglossus aculeatus)*, numbat (*Myrmecobius fasciatus)*, woylie (*Bettongia penicillata)* and chuditch (*Dasyurus geoffroii*) - and two non-native mammal species - red fox (*Vulpes vulpes*) and feral cat (*Felis catus*) - in a fragmented landscape. Further, we aimed to assess the differences in diversity and composition of mammals within and outside conservation reserves. We collected iDNA samples from three conservation reserves and road edges in between them. We detected 14 of the 40 known mammal species (native and non-native) to inhabit the region, including all our target species. We found that detections of our target native taxa were centred around conservation reserve areas, but there were also some detections on road edges nearby. We detected the presence of foxes and feral cats, throughout the study area, including all conservation reserves, despite targeted management of these species. The diversity and composition of species on road edges and conservation reserves were significantly different, with the conservation reserves hosting more native biodiversity than road edges. Here we demonstrate the utility of iDNA metabarcoding for mammal monitoring, with our data suggesting that the signals from iDNA are highly localised and reflective of the known distribution of mammals. The development of these methods shows promise for the future conservation management of mammals.

6.3 Introduction

The current accelerated loss of vertebrate species across the globe has heralded the likely beginning of the Earth's sixth mass extinction event (Ceballos et al., 2017; Grooten & Almond, 2018; Tilman et al., 2017). Extinctions and decline of ecosystem functionality have been mainly attributed to human activities causing habitat loss (Brodie et al., 2021). However, evidence suggests that the actual scale of species decline is underestimated because the focus has been on the complete loss of species rather than the decline and loss of populations (Ceballos et al., 2017). Some regions face more extinction threats than others. For example, in Australia, the total decline of mammal species represents about one-third of all mammal species that have gone extinct in the last 500 years (Baillie & Groombridge, 1996). Since European colonisation in the 1700s, 22 native Australian mammal species have become extinct (Burbidge & McKenzie, 2009; McKenzie et al., 2007; Woinarski et al., 2011), comprising 6% of Australia's marsupial species and 14% of its rodent species (Baillie & Groombridge, 1996). Another eight species of mammals that previously occurred widely throughout Australia now only exist as populations on a small number of islands or as translocated populations within intensively managed fenced reserves (Burbidge & McKenzie, 2009). However, extinctions and declines have not been spread evenly across the Australian continent - the two most impacted areas are part of the southern arid zone and the Wheatbelt region of Western Australia (Short & Smith, 1994).

Mammal declines have primarily been attributed to ecological disturbances from human activities: habitat clearing and fragmentation for agricultural use, the introduction of domestic animals, changes in burning regimes, hunting and culling of native fauna, and increased predation from introduced predators (Short & Smith, 1994; Woinarski et al., 2019). Attempts have been made to curb species decline and protect existing populations of native animals in the southern arid zone and the Wheatbelt region of Western Australia by creating conservation reserves. However, the establishment of conservation reserves alone is insufficient to maintain biodiversity within an area (Woinarski et al., 2011), and their effectiveness is determined by the appropriate management of these protected areas (Watson et al., 2014), including predator control, adequate fire regimes, and sufficient funding to enact and enforce protection of conservation areas. Furthermore, conservation reserves are often faced with the 'island dilemma' whereby if the surrounding area is unsuitable, the reserve can only ever hold the number of species within it an equilibrium on its size and degree of isolation (Diamond, 1975). In some cases, road reserves, the area of public land between private property and roads, can act as dispersal corridors for plant and animal species in agricultural landscapes (Dean et al., 2018; Ding & Eldridge, 2022). Wildlife can use these corridors to move through disturbed landscapes (Bennett, 1991), decreasing the isolation level between conservation reserves. However, in the highly fragmented native bush remaining in the Wheatbelt region of Western Australia, little is known about the use of road reserves by wildlife as a means of moving between conservation areas (Abensperg-Traun, 1991).

Tracking wildlife use and movement through environments requires direct observation of the species or the traces they may leave behind (Kouakou et al., 2009). However, this can be time and resource-intensive to conduct over large areas or in remote locations (Campbell et al., 2011). Further, some sampling procedures, such as trapping, can be harmful or stressful to animals (Putman, 1995). Therefore, to reduce the costs and risks associated with monitoring wildlife, non-invasive sampling techniques, such as DNA metabarcoding, can be applied for a fraction of the person-hours as traditional observation-based surveys (Calvignac-Spencer et al., 2013). DNA metabarcoding techniques can detect vertebrates through fragments of DNA that they shed into the environment. Different environmental substrates, such as water (Harper et al., 2019; Mas-Carrió et al., 2022), scat (van der Heyde et al., 2021), soil (Ryan et al., 2022), sediment from tree hollows (Newton et al., 2022), and air (Clare et al., 2022; Lynggaard et al., 2022) have been used to detect vertebrate diversity successfully. An extension of this method is detecting vertebrates from invertebrates that interact with them, known as invertebrate-derived DNA (iDNA) metabarcoding. iDNA metabarcoding has been used to document vertebrate diversity across the globe (Abrams et al., 2019; Calvignac-Spencer et al., 2013; Lynggaard et al., 2019; Schnell et al., 2015). Invertebrates such as leeches, mosquitoes, sand flies or carrion flies have been used as sources of vertebrate DNA (Massey et al., 2021) as they interact with the flesh, blood, or scat of vertebrates. Briefly, iDNA metabarcoding works by extracting the total DNA from invertebrate samples and amplifying only vertebrate DNA using nucleotide-labelled, vertebrate-specific primers (Bohmann et al., 2021). The vertebrate DNA is then sequenced, and the resulting sequences are matched against reference databases to identify the species detected (Calvignac-Spencer et al., 2013; Taberlet et al., 2012).

Here we aimed to use iDNA from carrion flies to detect native Australian mammals in the wheatbelt region of Western Australia. We investigated if the localisation of the iDNA signal reflects the expected distribution of four native and two non-native target mammal species in a fragmented landscape. We specifically aimed to assess the composition and richness of mammals within and outside conservation reserves. We predicted that the target native marsupial species - the carnivorous chuditch (*Dasyurus geoffroii*), the myrmecophagous numbat (*Myrmecobius fasciatus*) and the herbivorous and fungivorous woylie (*Bettongia penicillata*) would be restricted to conservation reserves, with little evidence of dispersal across the spaces between the reserves. We predicted that the myrmecophagous and termitophagous monotreme, the echidna (*Tachyglossus aculeatus*), would be most frequently detected in conservation reserves. However, as they have wide home ranges in cropping and pastoral areas (Sprent & Nicol, 2012), we would also detect them between the reserves. Finally, we predicted that invasive predator species - red fox (*Vulpes vulpes*) and feral cat (*Felis catus*) - would be equally frequently detected across the landscape.

6.4 Methods

6.4.1 Site descriptions

The Wheatbelt region is located in the south-west of Western Australia and describes a 140,000 km^2 area of land predominantly used for cereal growing. It has a Mediterranean climate characterised by hot, dry summers and cool winters. Average maximum temperatures range from 31 ˚C in summer (January) to 15.4 ˚C in winter (July). Annual rainfall varies through the region, but within the sampling area, rainfall averages 444.4 mm, with the most rainfall occurring in July. More than 93% of the natural vegetation in this region has been replaced with exotic grasses and cereal crops. The remaining native vegetation is scattered through the region in thousands of remnants of varying size, shape, and degree of isolation (Saunders, 1989).

Three conservation reserves surrounded by cropping and pastoral land were chosen as representing 'islands' of native vegetation within the farmland of the Wheatbelt (Fig 6-1). These reserves were Boyagin Nature Reserve (Boyagin NR), Dryandra Woodland National Park (Dryandra NP), and Lupton Conservation Park (Lupton CP).

Boyagin NR, in the north of our study area, is a 49 km^2 reserve located beyond the eastern edge of the Jarrah Forest bioregion (Schut et al., 2012). The predominant vegetation communities within Boyagin NR are kwongan scrubland and *Eucalyptus* woodland (Dames & Moore, 1985). Boyagin NR consists of eastern and western halves separated by a 500m wide strip of farmland. The trapping from our study was in the western half, which has had fox control in place since 1989. Woylies and numbats have been re-introduced into Boyagin NR, with populations established in both halves of the reserve. In addition, 18 other mammal species have been recorded at Boyagin NR, including our other native and non-native target species.

Dryandra NP, in the south of our study area, is the largest remnant of natural vegetation within the Wheatbelt, comprising 280 km^2 of 17 blocks of natural bushland, mostly woodland of powderbark wandoo (*Eucalyptus accedens*) and white wandoo (*E. wandoo*). It is one of the most diverse reserves in the region. There are 29 species of native mammal species found in Dryandra NP, including all our target species. Dryandra NP plays a crucial role in conserving threatened species. For example, the numbat population within Dryandra NP represents half the total number of individuals of the species found in the wild (Friend et al., 1995). There is extensive predator control within the reserve and surrounding areas, with regular fox baiting and feral cat management; however, both species are still recorded in the reserve (Marlow et al., 2015a).

Lupton CP, on the western edge of our study area, is part of the northern jarrah forest. The park area encompasses 87 km^2 but is contiguous with the rest of the forest, which comprises mainly jarrah (*E. marginata*) and marri (*Corymbia calophylla*). Several common mammal species are known to inhabit the region, although limited records are available for the specific species that occur in Lupton CP (*Northern Forest Management Plan 1987-1997*, 1987). We assume that while numbat and woylies do not occur in Lupton CP or at very low numbers, our other target species (both native and non-native) do. In addition, feral cats and foxes are known to be frequently detected within the jarrah forest, and predator control through baiting is a regular part of management throughout both the northern and southern areas of the jarrah forest (Dundas et al., 2014).

Our other study site was made up of the road edges between these conservation areas, representing the farm matrix in which the conservation areas are embedded. Road edges were chosen based on their accessibility. Most of these were <5 m wide with sparse vegetation, usually consisting of trees with little to no natural understorey and a ground cover of exotic grasses.

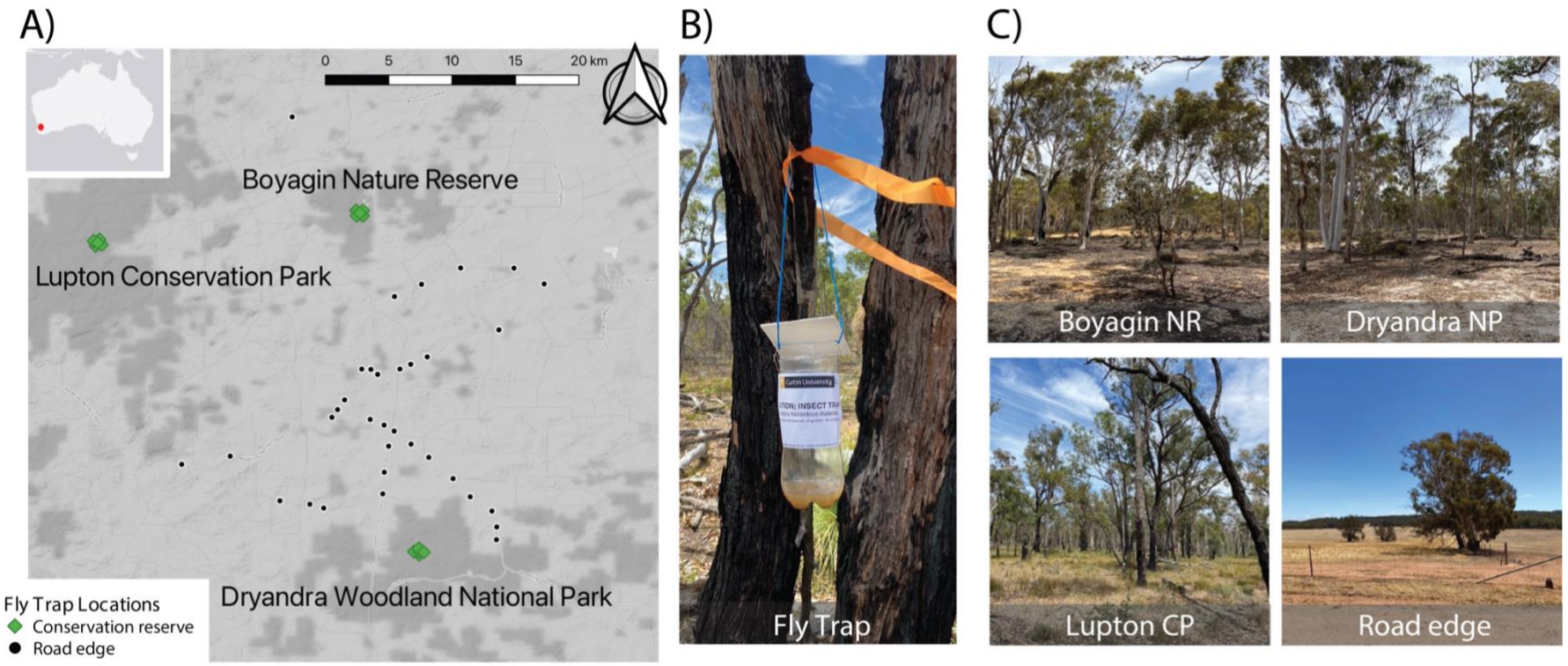


Figure 6-1 A) Map showing the trapping locations alongside 35 road edges and at the three conservation reserves, Boyagin Nature Reserve (Boyagin NR), Dryandra Woodland National Park (Dryandra NP), and Lupton Conservation Park (Lupton CP), in Western Australia. Locations along road edges are noted as black circles, and conservation reserves are green diamonds. Within each conservation reserve, five traps are placed approximately 500 m from each other in an 'X' pattern. B) Image showing a fly trap made from plastic bottles with the top and nozzle cut off and then inverted and inserted into the bottle. C) Images showing typical habitat at Boyagin NR, Dryandra NP, Lupton CP, and road edges.

6.4.2 Sample collection

Traps were made from plastic bottles with the top and nozzle cut off and then inverted and inserted into the bottle. Traps were sterilised for 30 minutes with a 20% bleach solution before deployment, and gloves were used to handle the sterilised traps. Approximately 20-30 mL of a commercial, organic fly bait ("Magna Fly Bait", GEPRO, Australia) was put into each trap as an attractant and sampling fluid (Fig 6-1B).

Samples were set up and collected during January 2021 (summer). All traps were handled with gloves. Gloves were replaced between every trap. Five fly traps were placed approximately 500 m apart at each of the three conservation reserves in an 'X' formation. These five traps were kept as separate samples. A single fly trap was set up at each of the 35 road edges through the region. The 35 road edge sampling sites were based on accessibility to public land and the availability of trees on which to hang fly traps. All fly traps were left out for 72 hours. Sampling fluid (and flies) were decanted into falcon tubes. Traps were then thoroughly cleaned, sampling fluid replaced, and set up again. The same trap was used at each sampling location. Fly sampling at these locations occurred at three time points over 14 days. Samples were then stored on ice during transport. Samples were frozen on arrival at the laboratory at -18 ˚C.

6.4.3 Sample processing and DNA extraction

Processing of samples occurred in a dedicated pre-PCR lab. Each trap and time point was considered a sample and extracted separately to increase the chance of detecting vertebrate species. There were 150 samples extracted in total. The number of flies extracted in a sample ranged from 1 to 300, with an average of 24 (± 3 s.e.) flies. Samples were gently washed with ultrapure water to remove any traces of fly bait before being stored in 70% ethanol at -18 ˚C before DNA extraction. Prior to DNA extraction, the ethanol was decanted and evaporated off the samples. A sample of the bait used during this study was also processed alongside fly samples. DNA was extracted using a non-destructive DNA digestion protocol (Nielsen et al., 2019) that lyses and releases cellular DNA without damaging the exoskeleton of the flies. Flies within samples that contained more than 40 individuals were divided into separate

tubes of 40 flies or less, and these separate tubes were digested individually to ensure even coverage of digestion fluid across the whole sample. Samples were digested overnight at 56 ˚C. The digest fluid from these separate tubes was then combined equally for the purification and concentration of DNA. Purification and concentration of DNA were conducted using a DNEasy Blood and Tissue kit following the manufacturer's instructions (Qiagen, Netherlands) modified to include a 400 µL volume of digestion fluid and a 100 µL double elution in EB buffer. DNA extraction controls were included every 24 samples. DNA extracts were stored at -18 ˚C.

6.4.4 PCR amplification and sequencing.

Two metabarcoding assays were used to analyse the iDNA from carrion fly samples across two gene regions. Assays were chosen based on *in-silico* assessment against the target mammal species, whereby between the two assays all species were detected. Additionally, short amplicon sizes were needed to account for any degradation due to digestion from flies. Based on these criteria, a vertebrate-specific assay was selected targeting the 12s rRNA region (F: 5'-TAGAACAGGCTCCTCTAG-3'; R: 5'-TTAGATACCCCACTATGC-3'; 98 bp excluding primers; (Riaz et al., 2011)) (vertebrate 12s) and a mammal specific assay was selected targeting the 16s rRNA region (F: 5'-CGGTTGGGGTGACCTCGGA-3'; R:5'-GCTGTTATCCCTAGGGTAACT-3'; 130 bp excluding primers; (Taylor, 1996)) (mammal 16s).

SYBR Green quantitative PCR (qPCR) was carried out on all samples to assess the amplification efficiency and the presence of PCR inhibitors in the samples. This was done using undiluted samples and 1:2 and 1:5 dilutions. qPCR reactions were carried out in 12.5 µl reactions containing: 1 U of AmpliTaq gold, 1 x PCR Gold Buffer and 2 mM $MgCl_2$ (all from Applied Biosystems, USA), 0.4 mg/mL bovine serum albumin (Fisher Biotec, Australia), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4 µM of each forward and reverse primer, 0.6 µL of 1/1000 SYBR Green (Invitrogen, USA), and 1 µL of template DNA. The qPCR program for the vertebrate 12s assay was as follows: 95˚C for 5 minutes, followed by 45 cycles of denaturation at 95˚C for 30s, 60˚C for 30s, and 72˚C for 45s, with a final elongation at 72˚C for 10 minutes. For the mammal 16s assay, the qPCR program was as follows: 95˚C for 5 minutes, followed by 45 cycles of 95˚C for 30s, 57˚C for 30s, and 72˚C for 45s, with a final elongation at 72˚C for 10 minutes. Negative extraction controls, a 'fly bait' control, qPCR negative controls and positive (quenda; *Isoodon fusciventer*) controls were included alongside the sample. Positive controls are used as a basis for comparing the quality and quantity of DNA from other samples. The positive control is a proven sample that will be amplified using laboratory workflows. As there are members of this species inhabiting the study sites, positive controls were not sequenced to minimise any possible cross-contamination.

Following qPCR, dilutions that showed the highest level of amplification and minimal inhibition were selected for amplification using qPCR with ‘fusion primers’, which are gene-specific primers labelled on both the forward and reverse with 6-8bp molecular identification (MID) tags and Illumina sequencing adaptors (Bohmann et al., 2021). The qPCR reactions were prepared in an ultra-clean laboratory free from extracted or amplified DNA to minimise the possibility of contamination. The DNA was then added to reactions in a separate pre-PCR laboratory, absent from amplified DNA. Each sample was tagged with a unique combination of forward and reverse MID tags not previously used within the laboratory. Using the above qPCR conditions, 121 samples were amplified using the mammal 16s primer, and 75 samples were amplified using the vertebrate 12s primer. All respective samples were amplified in duplicate, using the same tag combinations for each sample, including extraction and qPCR negative controls, and the ‘fly bait’ control, but not the positive control. Following qPCR amplification, PCR products were pooled in approximate equimolar concentration pools containing no more than 10 samples each based on the qPCR amplification curves. These pools were then quantified using a QIAxcel Advanced System (Qiagen) with the QIAxcel DNA High-Resolution Kit. As per the quantification results, they were then pooled in approximate equimolar ratios to create a single amplicon library pool for sequencing. The library pool was size-selected for fragments between 150-600 bp long using a Pippin Prep 2% agarose Marker B cassette (Sage Science, USA). The library pool was then purified using a QIAquick PCR purification kit (Qiagen) as per the manufacturer's instructions with the addition of a 5-minute incubation at room temperature before elution. The purified library was eluted in 40 µl and quantified with a QuBit (Invitrogen, USA) using double-stranded DNA high-sensitivity reagents. The library pool was sequenced on an Illumina MiSeq (Illumina, USA) using a single-end 300 cycle V2 kit as per the manufacturer's directions.

6.4.5 Bioinformatics and sequence processing

Sequence data for each primer set was processed separately. Sequences were demultiplexed using 'obitools' (Boyer et al., 2016), with no mismatches in the MID/primer sequence allowed. Sequences were then length filtered for a minimum length of 50 bp. The 'DADA2' package (Callahan et al., 2015) in R v. 3.6.3 (R Core Team, 2019) was then used to further quality filter sequences. Sequences were quality filtered with a max expected error of 2, and those identified as chimeras were removed. Sequences were then dereplicated to produce amplicon sequence variants (ASVs). ASVs were matched to the NCBI GenBank reference database (www.ncbi.nlm.nih.gov/genbank/) using the Basic Local Alignment Search Tool (BLAST) for taxonomic assignment using a high-performance cluster computer (Pawsey Computing Centre, Perth, Western Australia). Taxonomic assignments were made to the lowest common ancestor (LCA) using the LCA script from eDNAFlow (Mousavi-Derazmahalleh et al., 2021) with a minimum query coverage of 100%, identity threshold of 95%. Where the absolute value of the difference between % identity of ASVs was <1, species taxonomy was not returned and the ASV was assigned to the closed common ancestor. ASVs identified in the fly bait: cow (*Bos taurus*), pig (*Sus scrofa*), and sheep (*Ovis aries*); negative controls, cow and chicken (*Gallus gallus*), both common laboratory contaminants; and human (*Homo sapiens*), were removed from the data set using the 'phyloseq' package (McMurdie & Holmes, 2013). ASVs from both metabarcoding assays were combined for further analysis. ASVs were agglomerated at a species level and per trap to account for the use of the same fly-trap between the three sample collection times. The ASV table was transformed to presence-absence.

6.4.6 Statistics

All statistical tests were run on R v. 3.6.3 (R Core Team, 2019). All maps were created using QGIS v. 3.2.3-'Bonn' (QGIS Development Team, 2018)**.** Only ASVs identified as mammals were used for statistical analysis. The ASVs were categorised into either native or non-native taxa. ASV richness (observed number of ASVs) was then calculated for these two groups at the sampling points using the 'phyloseq' package (McMurdie & Holmes, 2013) on R v. 3.6.3 (R Core Team, 2019). ASV richness was square-root transformed to meet the assumption of homogeneity of variance and normality. A two-factor analysis of variance (ANOVA) was then used to compare the square-root transformed ASV richness of the groups of mammals (Factor Native vs Non-native, two levels; native and non-native), between the reserves (Factor Site, four levels; Boyagin NR, Dryandra NP, Lupton CP, and the road edges). A Tukey-HSD post-hoc test was then run using the 'agricolae' (de Mendiburu & de Mendiburu, 2019) package to assess the significant differences between the groups. A one-way permutational MANOVA (PERMANOVA) was conducted to compare the ASV composition between the different reserves (Factor Site) using the 'vegan' package (Oksanen et al., 2019) with Jaccard similarity and 9999 permutations. Pairwise comparisons were performed using the 'PairwiseAdonis' (Arbizu, 2020) package with a Benjamini-Hochberg correction for multiple comparisons.

6.5 Results

6.5.1 Sequencing results

In total, 3,428 flies were captured during the sampling period. After removing the 'bait' signal, 1,248,049 sequences were retained across both assays, with a mean of 13,287 (± 3686 s.e.) sequences per sample. Across all samples, 45 ASVs were detected, of which seven ASVs belonged to the class Aves, and 36 belonged to the class Mammalia. Two ASVs could not be identified beyond the phylum Chordata and were removed from the data set prior to further analysis. In total, we detected 14 mammal taxa of the 40 mammal species (both native and non-native) and three species of bird of the 174 recorded in the region (Table 6-1).

We detected 12 non-native taxa (including species that were also detected in the fly bait) (Table 6-1). As the Wheatbelt region where the sampling took place is predominantly agricultural, detections of these taxa (cow, pig, sheep, and chicken) would not be completely unexpected. Therefore, while we have removed the specific ASVs associated with the fly bait from our analysis, we have kept other ASVs identified as these species for diversity and species composition assessments.

We detected 15 native taxa, including our four target species (Table 6-1). We also detected the bilby (*Macrotis lagotis*) from samples collected at Dryandra NP. The only bilby population in this region is located in the Barna Mia animal sanctuary, a fenced area of the Dryandra NP that houses several other protected marsupial species. Fly traps in Dryandra NP were set up less than 200 m away from the border of this sanctuary.

Table 6-1 All ASVs identified to at least family level that were detected via carrion fly DNA. Location where each taxon is detected is noted with letters: Boyagin Nature Reserve (B), Dryandra Woodland National Park (D), Lupton Conservation Park (L), and road edges (R). Taxa are divided into non-native taxa and native taxa to the region. Species with * are also taxa that were detected in the 'fly bait' used in this study. Taxa highlighted in bold are the target taxa investigated in this study.

Non-native Taxa						
Class	**Order**	**Family**	**Genus**	**Species**	**Common Name**	**Location Detected**
Aves	Galliformes	Phasianidae	*Gallus*	*Gallus gallus**	Chicken	B, D, L, R
Mammalia	Artiodactyla	Bovidae	*Bos*	*Bos taurus**	Cow	B, D, L, R
Mammalia	Artiodactyla	Bovidae	*Ovis*	*Ovis aries**	Sheep	B, D, L, R
Mammalia	Artiodactyla	Suidae	*Sus*	*Sus scrofa**	Pig	B, D, L, R
Mammalia	Carnivora	Canidae	*Canis*			B, D, R
Mammalia	**Carnivora**	**Canidae**	***Vulpes***	***Vulpes vulpes***	**Red Fox**	B, D, L, R
Mammalia	**Carnivora**	**Felidae**	***Felis***	***Felis catus***	**Feral cat**	B, D, R
Mammalia	Lagomorpha	Leporidae	*Oryctolagus*	*Oryctolagus cuniculus*	European rabbit	B, L, R
Mammalia	Perissodactyla	Equidae	*Equus*	*Equus caballus*	Horse	B, D, R
Mammalia	Rodentia	Muridae	*Mus*			B, D, R
Mammalia	Rodentia	Muridae	*Mus*	*Mus musculus*	House mouse	B, R
Mammalia	Rodentia	Muridae	*Rattus*			R
Native Taxa						
Class	**Order**	**Family**	**Genus**	**Species**	**Common Name**	**Location Detected**
Aves	Anseriformes	Anatidae				L
Aves	Coraciiformes	Meropidae	*Merops*		Rainbow bee-eater	R
Aves	Galliformes	Phasianidae	*Coturnix*	*Coturnix pectoralis*	Stubble Quail	B, D, L, R
Aves	Passeriformes	Corvidae	*Corvus*			D, L, R
Aves	Passeriformes	Corvidae				D, L, R
Mammalia	Chiroptera	Vespertilionidae	*Eptesicus*	*Eptesicus regulus*	Southern Forest Bat	R
Mammalia	**Dasyuromorphia**	**Dasyuridae**	***Dasyurus***	***Dasyurus geoffroii***	**Chuditch**	B
Mammalia	**Dasyuromorphia**	**Myrmecobiidae**	***Myrmecobius***	***Myrmecobius fasciatus***	**Numbat**	D, R
Mammalia	Diprotodontia	Macropodidae	*Notamacropus*	*Notamacropus irma*	Western Brush Wallaby	L, R
Mammalia	Diprotodontia	Macropodidae	*Osphranter*			D, L, R
Mammalia	Diprotodontia	Macropodidae				D, L, R
Mammalia	Diprotodontia	Phalangeridae	*Trichosurus*	*Trichosurus vulpecula*	Brush-tail Possum	D, R
Mammalia	**Diprotodontia**	**Potoroidae**	***Bettongia***	***Bettongia penicillata***	**Woylie**	B, D, R
Mammalia	Diprotodontia	Potoroidae	*Bettongia*			D, R
Mammalia	**Monotremata**	**Tachyglossidae**	***Tachyglossus***	***Tachyglossus aculeatus***	**Echidna**	D, L, R
Mammalia	Peramelemorphia	Peramelidae	*Macrotis*	*Macrotis lagotis*	Bilby	D

6.5.2 Spatial distribution of six target mammals.

We detected all four native target species from carrion fly iDNA: chuditch, numbat, woylie, and echidna. Numbat detections were restricted to Dryandra NP (3 out of 5 traps) despite the species being recorded at Boyagin NR. Woylie was detected at both Dryandra NP (1 out of 5 traps) and at Boyagin NR (1 out of 5 traps), and echidna was detected in Dryandra NP (1 out of 5 traps) and Lupton CP (2 out of 5 traps) but not Boyagin NR despite being frequently observed there. Chuditch were only detected in Boyagin NR (1 out of 5 traps) (Fig 6-2). We also detected woylies (1 out of 35 traps), numbats (1 out of 35 traps), and echidnas (2 out of 35 traps) from road edges adjacent to Dryandra, with woylies detected up to 10 km away (Fig 6-2). Foxes were detected in all conservation reserves (Boyagin NR, 3 out of 5 traps; Dryandra NP, 4 out of 5 traps; Lupton CP, 1 out of 5 traps), and feral cats were detected in Boyagin NR (1 out of 5 traps) and Dryandra NP (5 out of 5 traps) but not Lupton CP. These predators were also detected at many road edges (foxes, 18 out of 35 traps: cats, 10 out of 35 traps) (Fig 6-2).

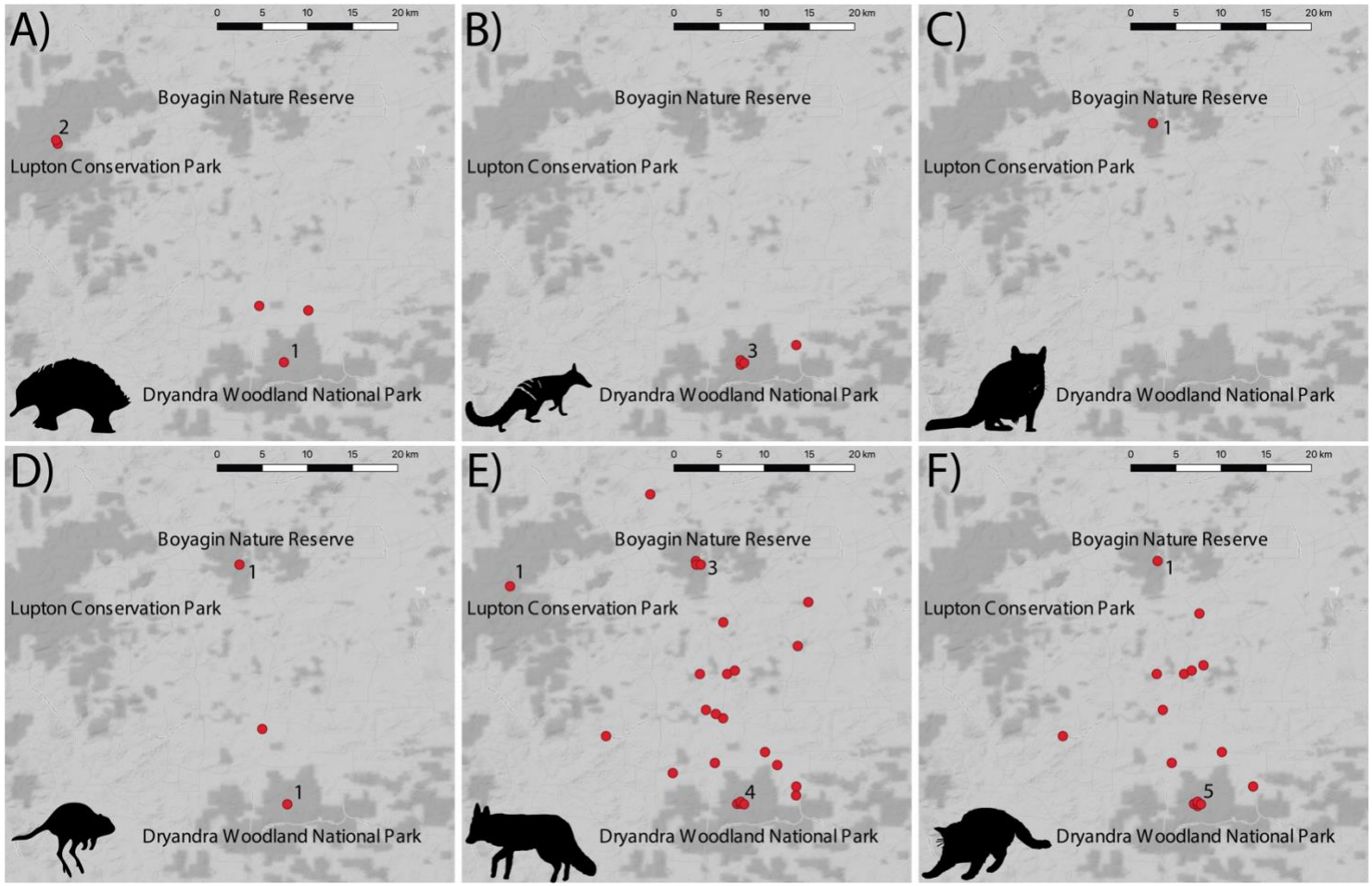
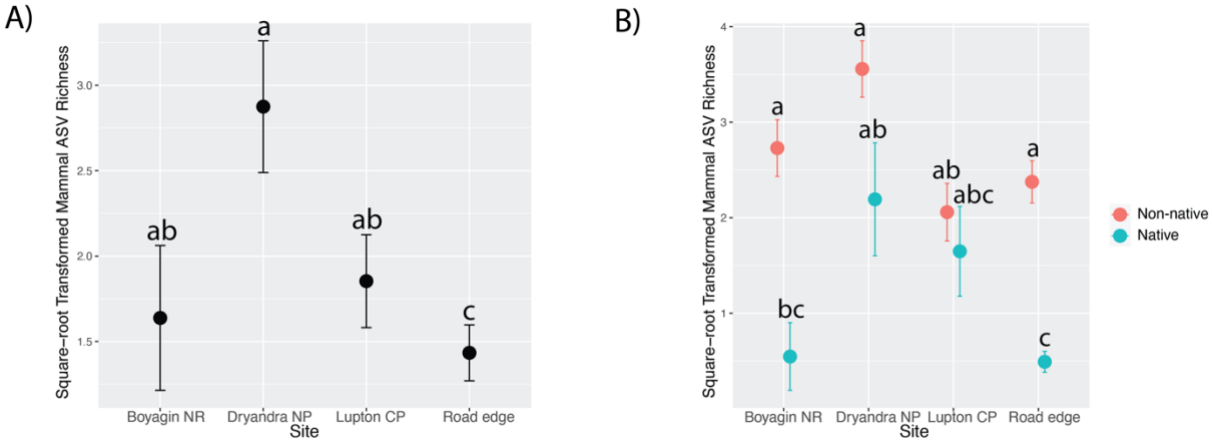


Figure 6-2 Detection locations of the six target species from the carrion fly iDNA sampling in 35 road edges and three conservation reserves, Boyagin Nature Reserve, Dryandra Woodland National Park, and Lupton Conservation Park, in Western Australia. Black circles indicate the locations of the fly traps through the region, red circles indicate positive detections of the target species. Inset are silhouettes of the target species. Within conservation reserves, numbers indicate how many traps out of the five these species were detected in during the sampling period. A) echidna (*Tachyglossus aculeatus*), B) numbat (*Myrmecobius fasciatus*), C) chuditch (*Dasyurus geoffroii*), D) woylie (*Bettongia penicillata*), E) red fox (*Vulpes vulpes*), and F) feral cat (*Felis catus*).

6.5.3 Comparison between conservation reserves and road edges.

The fully factorial ANOVA test on ASV richness found that there was no significant interaction between the factors Native vs Non-native and Site ($F_{(3, 98)} = 1.77$, $p=0.159$), a significant effect of Site ($F_{(3, 98)} = 5.91$, $p<0.001$), and a significant main effect of the factor Native vs Non-native ($F_{(1, 98)} = 78.75$, $p<0.0001$). The total mammal ASV richness was highest in Dryandra NP (Fig 6-3A, Fig 6-4A), with no significant differences between the ASV richness at Lupton CP and Boyagin NR conservation reserves. ASV richness was lowest at the road edges (Fig 6-3A). However, when looking specifically at the differences between native and non-native mammal ASV richness, non-native mammal ASV richness was higher than native ASV richness (Fig 6-3B). Boyagin NR and Dryandra NP had similar non-native ASV richness to the road edges, but Lupton CP has significantly lower non-native ASV richness (Fig 6-3B). Dryandra NP and Lupton CP had higher native mammal ASV richness than Boyagin NR and the road edges, which were similar (Fig 6-3B). Native mammal ASV richness centred around the conservation reserves (Fig 6-4B), and this richness decreased at sites further away from the conservation reserves. Non-native mammal ASV richness was high throughout the region, including the road edges furthest away from conservation reserves (Fig 6-4C).

Figure 6-3 Square-root transformed means (including standard error bars) of mammal ASV richness (number of mammal ASVs) from Boyagin Nature Reserve, Dryandra Woodland National Park, and Lupton Conservation Park, and Road edges in Western Australia. Letters (a,b,c) show the results from the Tukey-HSD test. Different letters indicate statistically different richness values. A) Overall mammal ASV richness and B) the ASV richness from non-native and native mammals at each site.



The ASV composition was significantly different between the different reserves (PERMANOVA; $F_{(3,\ 43)}$= 1.72, p=0.0014). Pairwise comparisons indicated that Dryandra NP's composition significantly differed from the other conservation reserves (F>2.49, p<0.025). Lupton CP and Boyagin NR did not have significantly different ASV compositions (F=1.57, p=0.067). The composition of Dryandra NP and Lupton CP was significantly different to the road edges (F>1.84, p<0.025), however, the ASV composition between Boyagin NR and road edges was not significantly different (F=0.73, p=0.78).

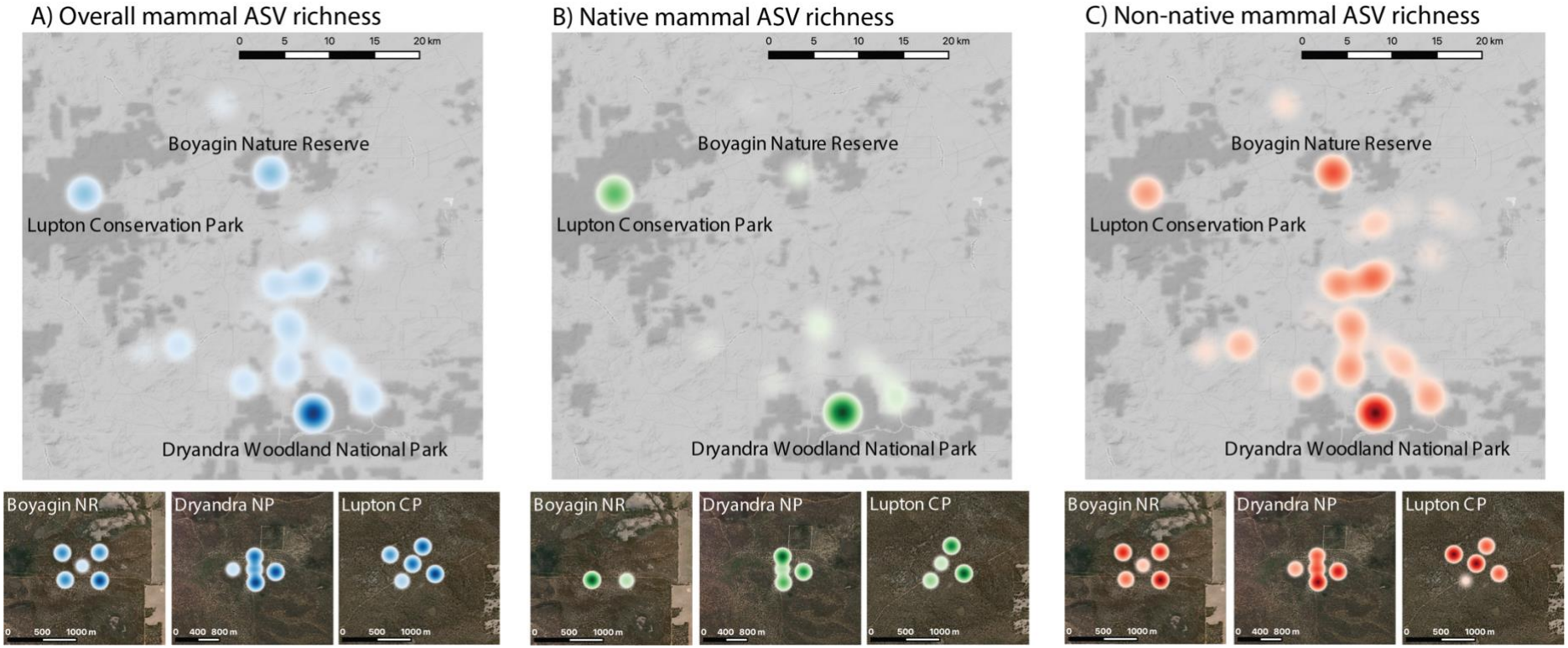


Figure 6-4 Heat map based on mammal ASV richness (Alpha diversity) values from fly iDNA sampling in 35 road edges and five traps at each of the three conservation reserves, Boyagin Nature Reserve (Boyagin NR), Dryandra Woodland National Park (Dryandra NP), and Lupton Conservation Park (Lupton CP), in Western Australia. A) Overall mammal ASV richness, B) Native mammal ASV richness and C) Non-native mammal ASV richness.

6.6 Discussion

We successfully detected 14 out of the 40 mammals recorded in our study area, including all our target species, native and non-native, from iDNA. Our predictions were broadly supported in that the detected distributions of all target native species were mostly restricted to conservation areas, while non-native feral cats and foxes were detected across almost the whole landscape. Further, the composition of mammals within conservation reserves and road edges reflected our predictions of the different species richness and composition at these sites. These results support the value of conservation reserves in the region and reveal the differential response of native and non-native mammals in fragmented landscapes.

Three of our target native species - numbat, woylie, and echidna - were detected within conservation reserves and alongside road edges, but only those roads close to the conservation reserves. The farthest detection from a reserve was woylie, on a road edge 10km north of Dryandra NP. The detection of the target species broadly reflects our prediction about their distribution. We found that the diversity of native mammals centred in the various conservation reserves and that, within the central road edges of our sampling area, there were no native mammals detected. This implies that mammal populations are isolated within each of these reserves, or that the sampling efforts within the middle of study area was inadequate to pick up the signals of native species. Potentially because within this area, native mammals may be more dispersed or spend less time in these central areas because of lack of suitable habitat.

Isolation can be quite detrimental for native mammals in already reduced habitat ranges. Whilst it is known that woylie populations within other conservation reserves in this region are genetically distinct due to restricted gene spread (Pacioni et al., 2020), this could also be true for other species living in these isolated conservation reserves. However, numbats and chuditch, particularly young ones, regularly disperse to new habitats (Soderquist & Serena, 2000; Thorn et al., 2022) and it seems likely that remnant vegetation along roadsides would act as corridors for this dispersal: our lack of detection of these species from outside the conservation areas probably reflects the length of time for which we took samples, or that no animals were dispersing at the time, or that animals disperse too rapidly to be detected at our level of sampling. Even so, it is evident that management strategies need to address the isolation of these conservation reserves and the effects on the long-term viability of mammal populations within them.

We expected that chuditch would be detected in Dryandra NP due to longstanding records of this species within this reserve (Orell 2004), but we did not detect it in our samples. However, we did detect chuditch in one sample at Boyagin NR. As the chuditch is a primarily solitary and nocturnal animal with a wide home range (Rayner et al., 2011), this lack of detection may be due to insufficient sampling in the conservation park rather than the absence of this species from the area. Furthermore, we expected that echidnas would be found widely through the region, both within and outside the conservation reserves, as it is known that echidnas can travel through pastoral areas (Sprent & Nicol, 2012). However, the detections of echidna also centred around conservation areas.

Cats and foxes, as predicted, were detected both in conservation areas and across the matrix of farmland between them. Predation by foxes and feral cats has been identified as one of the most significant causes of Australian mammal extinctions (Woinarski et al., 2015). Dryandra NP has carried out extensive fox eradication over the last 40 years (Marlow et al., 2015b); however, foxes were still detected within this reserve. Fox eradication within Dryandra NP is conducted through the regular deployment of 1080 poison baits (sodium fluoroacetate), a fast-acting and naturally occurring lethal toxin by which native animals in this region are unaffected (Marlow et al., 2015a; Thomson & Algar, 2000). Baiting occurred in the reserve four days before fly sampling began (Department of Biodiversity, Conservation, and Attractions, pers comms.), which might have influenced the detections of foxes from iDNA if there was more carrion for flies to feed on. However, feral cats are not as attracted to these baits, and declines in foxes have increased cat activity and predation rates (Marlow et al., 2015a), which can be detrimental to native mammal populations surviving in these protected regions. Cats were not detected in Lupton CP, despite being contiguous with the rest of the jarrah forest region. This might be due to cat control in the area or reflect dynamic, local cat movements.

We detected non-native animals, including fox and cat signals within the central road edges, where native animal diversity was generally absent. Predator species have been shown to use road edge habitats more frequently than macropods (Wysong et al., 2020), which implies that they could move through areas easily if they are well connected with roads. This indicates that managing foxes and feral cats must occur beyond just within reserves and into the surrounding pastoral areas to eradicate these predator populations effectively. This is supported by findings that fox home ranges can average 704 ha, and inadequate baiting within this range can decrease the effectiveness of baiting programs because of decreased chance of uptake (Carter et al., 2011). Further, there is evidence to suggest that the use of linear infrastructure (such as roads) by predator species can cause avoidance of these areas by prey animals because of the increased predation risk (Latham et al., 2011). If native mammals avoid road areas because of the risk of predators, this may increase the isolation of the conservation reserves even if the habitat is maintained between them to facilitate movement between the areas.

Our results suggest either a geographic limit for which carrion flies can transport an iDNA signal or that some native mammals, during our study, were also present in habitats close to but beyond the conservation reserves. While studies have shown that flies can disperse at great distances (MacLeod & Donnelly, 1963), iDNA signals appear to stay within a very close area to a vertebrate population (Fernandes et al., unpublished). Within Dryandra NP, a fenced conservation sanctuary site, Barna Mia, contains several species of marsupial that are otherwise locally extinct, including the bilby. We detected bilby iDNA within 200 m of the Barna Mia fence, from which we infer that, for temporal and spatial constraints of this study, flies detected from this trap were travelling no more than 500 m from the DNA source (the distance from the approximate centre of the Barna Mia sanctuary area to the trap where this species was detected). Efforts have been made in and around the Dryandra reserve to maintain 'wildlife corridors' - areas of retained bushland that allow for linkages between remnant areas in a predominantly agricultural landscape (Hobbs & Saunders, 1991). Whilst not every road edge, lined with remnant trees, can act as a wildlife corridor, the verges near conservation reserves may facilitate wildlife movement through these areas (Bennett, 1991). The highly localised iDNA signal to the target population distributions suggests that carrion fly iDNA would be viable for monitoring vertebrate diversity at a fine scale, especially in remote regions where trapping to assess species presence can be difficult to undertake.

6.6.1 Limitations and future directions

Understanding the strengths and limitations of carrion fly iDNA monitoring for species detection is essential for future use of this method in conservation management. Here, we demonstrated the utility of iDNA metabarcoding for monitoring the distributions of mammal populations. While we only detected 14 out of 40 mammal species known to occur in the Wheatbelt region, we detected all six of our target species. It is not known why certain species were absent from the record and could be reflective of the sampling intensity/timespan, local animal movements, or biases with the method. In this case, we cannot rule out an absence from an iDNA record as an absence in the area. In future studies, exploring the effects of increasing sampling intensity, either spatially or temporally, on mammal species detection will be worthwhile. However, it must be remembered that while metabarcoding studies can reduce the overall costs of conducting biodiversity surveys, molecular work can be expensive, and increasing sampling intensity can dramatically increase associated costs (Borrell et al., 2017). This is an important consideration for conservation studies where funding may be limited, and in this case, iDNA could be used to guide the design of more targeted surveying efforts (Abrams et al., 2019). The effects of sampling season and landscape structure on mammal species detection are also worth exploring. Our study was conducted during a single season and the effects of weather conditions on iDNA dispersion and animal detection are unknown and may enable more robust estimates of mammal occupancy. Also, it is not clear how the structure of the environment can influence the detections of mammals from carrion fly iDNA. For instance, it is known that odour or chemical signals of decaying matter that attract carrion flies can be obstructed by physical barriers (Benbow et al., 2015).

6.7 References

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged

Abensperg-Traun, M. (1991). A study of home range, movements and shelter use in adult and juvenile echidnas, Tachyglossus aculeatus (Monotremata: Tachyglossidae), in Western Australian wheatbelt reserves. *Australian Mammalogy*, *14*(1), 13–21.

Abrams, J. F., Hörig, L. A., Brozovic, R., Axtner, J., Crampton-Platt, A., Mohamed, A., Wong, S. T., Sollmann, R., Yu, D. W., & Wilting, A. (2019). Shifting up a gear with iDNA : From mammal detection events to standardised surveys. *The Journal of Applied Ecology*, *56*(7), 1637–1648.

Arbizu, M. (2020). *PairwiseAdonis: Pairwise Multilevel Comparison Using Adonis*. [Computer software] https://github. com/pmartinezarbizu/pairwiseAdonis.

Baillie, J., & Groombridge, B. (1996). *1996 IUCN red list of threatened animals* (No. 333.95 I61 1996). IUCN, Gland (Suiza). Species Survival Commission.

Benbow, E. M., Tomberlin, J. K., & Tarone, A. M. (2015). *Carrion Ecology, Evolution, and Their Applications*. CRC Press.

Bennett, A. F. (1991). Roads, roadsides and wildlife conservation: a review. In D. A. Saunders R. J. Hobbs (eds.), *Nature Conservation 2: The Role of Corridors* (pp. 99-117). Surrey Beatty.

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A., & Creer, S. (2021). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231–1246.

Borrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K., & Garcia-Vazquez, E. (2017). DNA in a bottle-Rapid metabarcoding survey for early alerts of invasive species in ports. *PloS One*, *12*(9), e0183347.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12428

Brodie, J. F., Williams, S., & Garner, B. (2021). The decline of mammal functional and evolutionary diversity worldwide. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(3). https://doi.org/10.1073/pnas.1921849118

Burbidge, A. A., & McKenzie, N. L. (2009). Conservation status and biogeography of Australia's terrestrial mammals. *Australian Journal of Zoology, 56,* 411-422. https://doi.org/10.1071/ZO08027

Callahan, B. J., McMurdie, P. J., Rosen, M. J., & Han, A. W. (2015). DADA2: High resolution sample inference from amplicon data. *bioRxiv*. https://www.biorxiv.org/content/10.1101/024034v1.abstract

Calvignac-Spencer, S., Leendertz, F. H., Gilbert, M. T. P., & Schubert, G. (2013). An invertebrate stomach's view on vertebrate ecology: certain invertebrates could be used as "vertebrate samplers" and deliver DNA-based information on many aspects of vertebrate ecology. *BioEssays*, *35*(11), 1004–1013.

Calvignac-Spencer, S., Merkel, K., & Kutzner, N. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Molecular Ecology, 22*(4), 915-924. https://doi.org/10.1111/mec.12183

Campbell, G., Kuehl, H., Diarrassouba, A., N'Goran, P. K., & Boesch, C. (2011). Long-term research sites as refugia for threatened and over-harvested species. *Biology Letters*, *7*(5), 723–726.

Carter, A., Luck, G. W., & McDonald, S. P. (2011). Fox-baiting in agricultural landscapes in south-eastern Australia: a case-study appraisal and suggestions for improvement. *Ecological Management & Restoration*, *12*(3), 214–223.

Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), E6089–E6096.

Clare, E. L., Economou, C. K., Bennett, F. J., Dyer, C. E., Adams, K., McRobie, B., Drinkwater, R., & Littlefair, J. E. (2022). Measuring biodiversity from DNA in the air. *Current Biology, 32*(3), 693–700.e5.

Dames, & Moore. (1985). *The vegetation of Boyagin Nature Reserve*. Department of Environment and Conservation.

Dean, W. R. J., Seymour, C. L., & Joseph, G. S. (2018). Linear structures in the Karoo, South Africa, and their impacts on biota. *African Journal of Range & Forage Science*, *35*(3-4), 223–232.

de Mendiburu, F., & de Mendiburu, M. F. (2019). *Package "agricolae."* [Computer software] https://cran.r-project.org/web/packages/agricolae/index.html

Diamond, J. M. (1975). The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation*, *7*(2), 129–146.

Ding, J., & Eldridge, D. J. (2022). Roadside verges support greater ecosystem functions than adjacent agricultural land in a grassy woodland. *Journal of Environmental Management*, *308*, 114625.

Dundas, S. J., Adams, P. J., & Fleming, P. A. (2014). First in, first served: uptake of 1080 poison fox baits in south-west Western Australia. *Wildlife Research*, *41*(2), 117–126.

Friend, T., Bowra, T., Gorton, S., Hilder, D., Mitchell, D., Moncrieff, D., & Sutton, A. (1995). *Dryandra Woodland Management Plan 1995-2005*. Department of Conservation and Land Management. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.623.1498&rep=rep1&type=pdf

Grooten, M., & Almond, R. E. A. (2018). *Living Planet Report-2018: Aiming Higher*. WWF International.

Harper, L. R., Lawson Handley, L., Carpenter, A. I., Ghazali, M., Di Muri, C., Macgregor, C. J., Logan, T. W., Law, A., Breithaupt, T., Read, D. S., McDevitt, A. D., & Hänfling, B. (2019). Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. *Biological Conservation*, *238*, 108225.

Hobbs, R., & Saunders, D. (1991). Wildlife corridors in the wheatbelt. *Journal of the Department of Agriculture, Western Australia, 32*(2). https://researchlibrary.agric.wa.gov.au/cgi/viewcontent.cgi?article=1352&context=journal_agriculture4

Kouakou, C. Y., Boesch, C., & Kuehl, H. (2009). Estimating chimpanzee population size with nest counts: validating methods in Taï National Park. *American Journal of Primatology*, *71*(6), 447–457.

Latham, A. D. M., Latham, M. C., Boyce, M. S., & Boutin, S. (2011). Movement responses by wolves to industrial linear features and their effect on woodland caribou in northeastern Alberta. *Ecological Applications*, *21*(8), 2854–2865.

Lynggaard, C., Bertelsen, M. F., Jensen, C. V., Johnson, M. S., Frøslev, T. G., Olsen, M. T., & Bohmann, K. (2022). Airborne environmental DNA for terrestrial vertebrate community monitoring. *Current Biology*, *32*(3), 701–707.e5.

Lynggaard, C., Nielsen, M., Santos-Bay, L., Gastauer, M., Oliveira, G., & Bohmann, K. (2019). Vertebrate diversity revealed by metabarcoding of bulk arthropod samples from tropical forests. *Environmental DNA*, *1*(4), 329–341.

MacLeod, J., & Donnelly, J. (1963). Dispersal and Interspersal of Blowfly Populations. *The Journal of Animal Ecology*, *32*(1), 1–32.

Marlow, N. J., Thomas, N. D., Williams, A. A., Macmahon, B., Lawson, J., Hitchen, Y., ... & Berry, O. (2015a). Lethal 1080 baiting continues to reduce European red fox (*Vulpes vulpes*) abundance after more than 25 years of continuous use in south-west Western Australia. *Ecological Management & Restoration*, 16(2), 131-141.

Marlow, N. J., Thomas, N. D., Williams, A. A. E., Macmahon, B., Lawson, J., Hitchen, Y., Angus, J., & Berry, O. (2015b). Cats (*Felis catus*) are more abundant and are the dominant predator of woylies (*Bettongia penicillata*) after sustained fox (*Vulpes vulpes*) control. *Australian Journal of Zoology*, *63*(1), 18–27.

Mas-Carrió, E., Schneider, J., Nasanbat, B., Ravchig, S., Buxton, M., Nyamukondiwa, C., Stoffel, C., Augugliaro, C., Ceacero, F., Taberlet, P., Glaizot, O., Christe, P., & Fumagalli, L. (2022). Assessing environmental DNA metabarcoding and camera trap surveys as complementary tools for biomonitoring of remote desert water bodies. *Environmental DNA*, *4*(3), 580–595.

Massey, A. L., Bronzoni, R. V. de M., da Silva, D. J. F., Allen, J., de Lazari, P. R., Dos Santos Filho, M., Canale, G. R., Bernardo, C. S. S., Peres, C. A., & Levi, T. (2021). Invertebrates for vertebrate biodiversity monitoring: comparisons using three insect taxa as iDNA samplers. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13525

McKenzie, N. L., Burbidge, A. A., Baynes, A., Brereton, R. N., Dickman, C. R., Gordon, G., Gibson, L. A., Menkhorst, P. W., Robinson, A. C., Williams, M. R., & Woinarski, J. C. Z. (2007). Analysis of factors implicated in the recent decline of Australia's mammal fauna. *Journal of Biogeography*, *34*(4), 597–611.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One*, *8*(4), e61217.

Mousavi-Derazmahalleh, M., Stott, A., Lines, R., Peverley, G., Nester, G., Simpson, T., Zawierta, M., De La Pierre, M., Bunce, M., & Christophersen, C. T. (2021). eDNAFlow, an automated, reproducible and scalable workflow for analysis of environmental DNA (eDNA) sequences exploiting Nextflow and Singularity. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13356

Newton, J. P., Bateman, P. W., Heydenrych, M. J., Mousavi-Derazmahalleh, M., & Nevill, P. (2022). Home is where the hollow is: Revealing vertebrate tree hollow user biodiversity with eDNA metabarcoding. *Environmental DNA*. https://doi.org/10.1002/edn3.313

Nielsen, M., Gilbert, M. T. P., Pape, T., & Bohmann, K. (2019). A simplified DNA extraction protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity. *Environmental DNA, 1*(2), 144-154. https://doi.org/10.1002/edn3.16

Northern Forest Management Plan 1987-1997. (1987). Department of Conservation and Land Management. https://library.dbca.wa.gov.au/static/FullTextFiles/006653.pdf

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (Version 2.5-6) [Computer software]. https://cran.ism.ac.jp/web/packages/vegan/vegan.pdf

Pacioni, C., Trocini, S., Wayne, A. F., Rafferty, C., & Page, M. (2020). Integrating population genetics in an adaptive management framework to inform management strategies. *Biodiversity and Conservation*, *29*(3), 947–966.

Putman, R. J. (1995). Ethical considerations and animal welfare in ecological field studies. *Biodiversity & Conservation*, *4*(8), 903–915.

QGIS Development Team. (2018). *QGIS Geographic Information System* (3.2.3-Bonn) [Computer software]. http://qgis.osgeo.org.

Rayner, K., Chambers, B., Johnson, B., Morris, K. D., & Mills, H. R. (2011). Spatial and dietary requirements of the chuditch (*Dasyurus geoffroii*) in a semiarid climatic zone. *Australian Mammalogy*, *34*(1), 59–67.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing* (Version 3.6.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, *39*(21), e145.

Ryan, E., Bateman, P., Fernandes, K., Heyde, M., & Nevill, P. (2022). eDNA metabarcoding of log hollow sediments and soils highlights the importance of substrate type, frequency of sampling and animal size, for vertebrate species detection. *Environmental DNA*. https://doi.org/10.1002/edn3.306

Saunders, D. A. (1989). Changes in the Avifauna of a region, district and remnant as a result of fragmentation of native vegetation: the wheatbelt of western Australia. A case study. *Biological Conservation*, *50*(1), 99–135.

Schnell, I. B., Sollmann, R., Calvignac-Spencer, S., Siddall, M. E., Yu, D. W., Wilting, A., & Gilbert, M. T. P. (2015). iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool - prospects, pitfalls and avenues to be developed. *Frontiers in Zoology*, *12*, 24.

Schut, A. G. T., Wardell-Johnson, G. W., & Baran, I. (2012). Canopy profiling for vegetation mapping in south-western Australian forested ecosystems. *Annals of Photogrammetry Remote Sensing and Spatial Information Sciences*, *I-7*, 365–370.

Short, J., & Smith, A. (1994). Mammal Decline and Recovery in Australia. *Journal of Mammalogy*, *75*(2), 288–297.

Soderquist, T. R., & Serena, M. (2000). Juvenile behaviour and dispersal of chuditch (*Dasyurus geoffroii*) (Marsupialia : Dasyuridae). *Australian Journal of Zoology*, *48*(5), 551–560.
Sprent, J., & Nicol, S. C. (2012). Influence of habitat on home-range size in the short-beaked echidna. *Australian Journal of Zoology*, *60*(1), 46–53.
Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050.
Taylor, P. G. (1996). Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Molecular Biology and Evolution*, *13*(1), 283–285.
Thomson, P. C., & Algar, D. (2000). The uptake of dried meat baits by foxes and investigations of baiting rates in Western Australia. *Wildlife Research*, *27*(5), 451–456.
Thorn, S., Maxwell, M., Ward, C., & Wayne, A. (2022). Remote sensor camera traps provide the first density estimate for the largest natural population of the numbat (Myrmecobius fasciatus). *Wildlife Research*. https://doi.org/10.1071/WR21115
Tilman, D., Clark, M., Williams, D. R., Kimmel, K., Polasky, S., & Packer, C. (2017). Future threats to biodiversity and pathways to their prevention. *Nature*, *546*(7656), 73–81.
van der Heyde, M., Bateman, P. W., Bunce, M., Wardell-Johnson, G., White, N. E., & Nevill, P. (2021). Scat DNA provides important data for effective monitoring of mammal and bird biodiversity. *Biodiversity and Conservation*, *30*(12), 3585–3602.
Watson, J. E. M., Dudley, N., Segan, D. B., & Hockings, M. (2014). The performance and potential of protected areas. *Nature*, *515*(7525), 67–73.
Woinarski, J. C., Braby, M. F., Burbidge, A. A., Coates, D., Garnett, S. T., Fensham, R. J., ... & Murphy, B. P. (2019). Reading the black book: The number, timing, distribution and causes of listed extinctions in Australia. *Biological Conservation*, *239*, 108261.
Woinarski, J. C. Z., Burbidge, A. A., & Harrison, P. L. (2015). Ongoing unraveling of a continental fauna: decline and extinction of Australian mammals since European settlement. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(15), 4531–4540.
Woinarski, J. C. Z., Legge, S., & Fitzsimons, J. A. (2011). The disappearing mammal fauna of northern Australia: context, cause, and response. *Conservation*. https://doi.org/10.1111/j.1755-263X.2011.00164.x
Wysong, M. L., Iacona, G. D., Valentine, L. E., Morris, K., & Ritchie, E. G. (2020). On the right track: placement of camera traps on roads improves detection of predators and shows non-target impacts of feral cat baiting. *Wildlife Research*, *47*(8), 557–569.

CHAPTER 7

GENERAL DISCUSSION

7.1 The uses and value of invertebrate DNA metabarcoding as a terrestrial biomonitoring tool

Invertebrates are routinely used in monitoring terrestrial ecosystems; they are highly diverse, essential for ecosystem functioning, and interact widely with other organisms in the ecosystem (Kim, 1993). Therefore, they have fundamental value as indicators to measure health and ecosystem functioning (McGeoch, 1998). Yu et al. (2012) was the first to demonstrate the capabilities of metabarcoding for detecting biodiversity when applied to diverse invertebrate communities. Since then, the applications of invertebrate DNA metabarcoding for terrestrial monitoring have appeared seemingly limitless. Within this thesis, I have shown that the application of metabarcoding to invertebrates can elevate the capabilities for terrestrial biomonitoring, not only as a tool to monitor invertebrate biodiversity but also for monitoring ecosystem connectivity and species ecology with a range of applications for conservation and land management. The scope of this thesis was to build knowledge on how DNA metabarcoding could be used to monitor terrestrial ecosystems. Using a series of case studies within my thesis, I demonstrate the wide variety of applications that metabarcoding methods can have in biomonitoring. The results from this body of work indicate that depending on the sources of DNA, it is possible to use invertebrates to monitor various aspects of terrestrial ecosystem biodiversity and functioning. This final general discussion chapter synthesises the main findings of the previous chapters and considers the value of this methodology for terrestrial biomonitoring. Additionally, I identify the limitations of this body of work, discuss knowledge gaps, and suggest what I believe to be important future avenues of research (Fig 7-1).

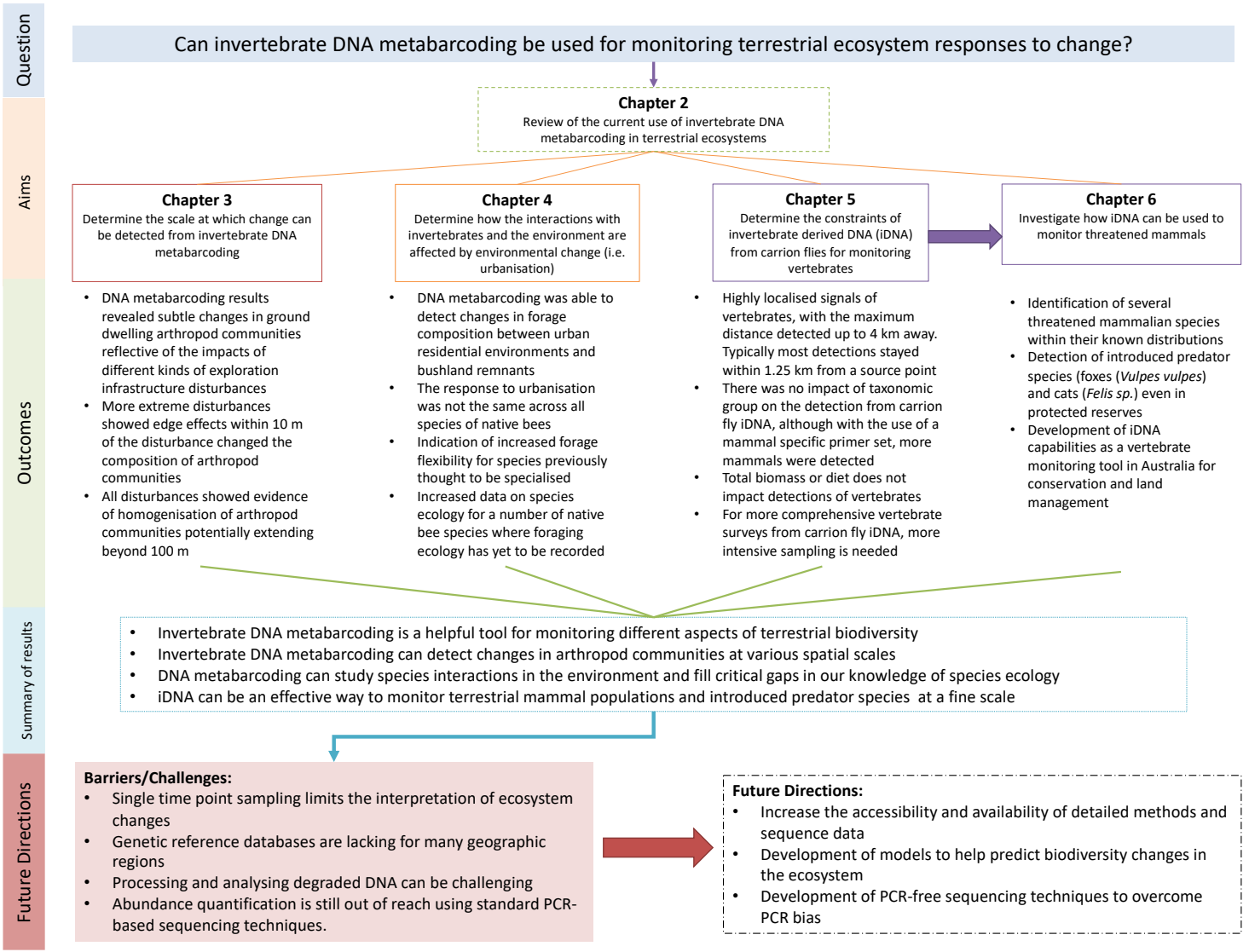


Figure 7-1 Conceptual flow diagram outlining the findings from this thesis and avenues for future research

7.2 Summary of main findings

7.2.1 Invertebrate DNA metabarcoding is a helpful tool for monitoring different aspects of terrestrial biodiversity.

A growing number of studies have recently used bulk arthropod sample metabarcoding to assess ecosystem biodiversity in multiple contexts. Chapter 2: '*Current and future applications of bulk arthropod sample metabarcoding in terrestrial ecological monitoring',* reviewed studies that applied DNA metabarcoding to bulk arthropod samples to document various aspects of terrestrial biodiversity. Whilst predominantly the focus for most bulk-arthropod metabarcoding studies has been on uncovering arthropod biodiversity, by using the DNA that arthropods have in them or on them, it is possible to elevate the capabilities beyond solely monitoring arthropod biodiversity and move towards studying terrestrial ecosystems and the interactions occurring within them with more nuance and depth. Hence, I explored multiple uses for this technique in habitat quality assessment, the study of plant-pollinator interactions, vertebrate monitoring, and the transmission of vector-borne pathogens. Here I also discussed some current limitations and suggestions for future areas of research that need to be conducted for bulk-arthropod metabarcoding to become more commonplace and useful in the biomonitoring toolkit, such as the capabilities for machine learning, long-term temporal studies, and the development of techniques to quantify abundance from metabarcoding data. Chapter 2 provides an in-depth examination of how terrestrial environments can be monitored using the DNA from arthropods and provides context for the other chapters presented in this thesis.

7.2.2 Invertebrate DNA Metabarcoding can detect changes in arthropod communities at various spatial scales.

In Chapter 3: '*Metabarcoding reveals the impacts of exploration infrastructure on arthropod communities at mine sites',* I use metabarcoding

to determine the scale at which metabarcoding can be used to detect changes in the environment. Here, I use the case study of mining exploration infrastructure on ground-dwelling arthropod communities at mine sites in the Midwest region of Western Australia. This chapter presents one of the most common uses of invertebrate DNA metabarcoding, uncovering arthropod biodiversity from bulk-arthropod samples. Applying DNA metabarcoding in this context provides a rapid and informative method to assess biodiversity in areas where cost or time would be prohibitive barriers to in-depth biodiversity assessment (Ji et al., 2013; Yu et al., 2012). Within mining operations, disturbances caused by exploration infrastructure are rarely characterised due to the comparatively small disturbance footprint they generate (McKenney & Kiesecker, 2010). Furthermore, the difficulty associated with monitoring arthropod communities to measure disturbance impacts means that the scale and extent of these effects are unknown. Therefore, metabarcoding can be used to provide valuable information on the degree and extent of disturbance caused by different exploration infrastructure types.

Here, I discovered that metabarcoding could be used to detect disturbances to arthropod communities at varying spatial scales depending on the degree of intensity of the disturbance. For all infrastructure types (drill pads, maintained tracks, and unmaintained tracks), the surrounding environment was impacted similarly, with an initial clearing of vegetation directly on the disturbance and an increase in vegetation cover and habitat structure at 100 metres away. However, each infrastructure type uniquely impacted arthropod communities. For more intense disturbances, i.e., maintained tracks, arthropod communities showed significant differences within 10 metres of the disturbance, and communities were still distinct in comparison to other arthropod communities at the same site at 100 m away. For the disturbances that were left to revegetate over time, i.e., drill pads and unmaintained tracks, the arthropod communities homogenised at each infrastructure type. This pattern is commonly observed in arthropods in response to vegetation disturbances (Majer et al., 2020). The results from

this study suggest a more pervasive impact of these exploration infrastructures on ecosystems than has been previously measured. Further, in the context of this thesis, this study provides evidence of the sensitivity of arthropods as indicators of disturbance in the environment and the ability for DNA metabarcoding to measure these arthropod communities effectively and at fine scales. Metabarcoding hence can be used to provide critical biodiversity information to develop more effective management strategies to conserve, restore, or mitigate risks from exploration disturbances in ecosystems.

7.2.3 DNA metabarcoding can study species interactions in the environment and fill critical gaps in our knowledge of species ecology.

In Chapter 4: '*DNA Metabarcoding identifies urban foraging patterns of oligolectic and polylectic cavity-nesting bees'*, I investigate the impacts of urbanisation on the foraging habits of solitary cavity-nesting bees in the Perth metropolitan region in southwestern Australia. This chapter presents a method to use metabarcoding to monitor the changes in species behaviour, moving beyond biodiversity audits to infer species interactions in disturbed environments by identifying the pollen and plant material gathered by native bees. Within this study, bees were selected with varying degrees of pollen specialisation, with oligolectic bees as specialists in a single family of plants and polylectic bees as generalists that could forage from a range of plant families. There is little known about many native bee species' foraging preferences and behaviour; therefore, using metabarcoding to identify forage sources provides an opportunity to fill knowledge gaps. Here, I found that oligolectic bees in residential gardens had a degree of adaptive forage flexibility. Despite their previously observed specialist behaviour, these species had a range of sources of plant material identified in their brooding tubes. Polylectic bees, however, did not change their foraging behaviour or preferences even in residential gardens. Using this information, it was possible to infer traits about behavioural plasticity in response to altered habitats that had not been observed in these species using traditional means

of monitoring. By tailoring the application of metabarcoding towards a specific source of DNA, this method could answer questions about species interactions, foraging preferences, and behaviour in response to changes in the environment. These results show positive connotations towards the available food sources for sustaining wild bee populations in urban areas and provide avenues for future research to uncover information about the energetic costs, health, and survival of native bees in urban environments. The ability to uncover this information is particularly beneficial where species may be difficult to observe in the wild due to rarity or cryptic behaviour and helps fill in knowledge gaps to aid in the conservation and management of these species.

7.2.4 iDNA can be an effective way to monitor terrestrial mammal populations and introduced predator species at a fine scale

By tailoring the application of DNA metabarcoding, it is possible to extend beyond studying arthropod ecology and investigate other aspects of ecosystem health and functioning. In Chapters 5: '*Examining the effects of distance, taxon, and animal biomass, on vertebrate detections from carrion fly iDNA'* and Chapter 6: '*Carrion fly iDNA metabarcoding to monitor mammals in a fragmented terrestrial ecosystem*', I apply metabarcoding to carrion flies to detect the vertebrate species in the environment. Many invertebrate species, such as carrion flies, will feed on and interact with vertebrates and sample their DNA. Using invertebrates as a source of vertebrate DNA has been coined 'invertebrate-derived DNA' or 'iDNA'. Many iDNA studies using carrion flies have emerged in recent years and have shown to be an effective tool for assessing the biodiversity of vertebrate populations (Calvignac-Spencer et al., 2013; Rodgers et al., 2017; Schubert & Stockhausen, 2015). However, the value of a monitoring tool is understanding the parameters surrounding its effectiveness. This is especially the case if used in a conservation or land management context for threatened species.

For carrion flies, some work has previously been conducted regarding the longevity of DNA signals within the guts of flies (Owings et al., 2019). Still, the distance from a vertebrate population, the effects of vertebrate biomass, or vertebrate taxonomic groups on iDNA detections remained to be tested. Therefore, Chapter 5 explored those specific factors into the impacts on iDNA detections from carrion flies using Perth Zoo, within the Perth metropolitan region in southwestern Australia, as a source of known exotic biodiversity.

Chapter 5 showed that iDNA signals from carrion flies are highly localised to vertebrate populations. With a mammal specific primer set, I detected more mammals than other groups of vertebrates. However, reptiles and birds were still underrepresented within the data set. Furthermore, these signals were not affected by biomass or the diet type of the detected mammals. As the detections of mammals were not influenced by biomass, this implied that even rare mammal species with smaller biomass within a region would have just a likely chance of being detected in carrion flies. These results suggest that carrion fly iDNA could be a valuable tool for mammalian monitoring – especially for detecting fine-scale differences in populations or locating threatened or rare species.

In Chapter 6, I implement carrion fly iDNA monitoring across the highly fragmented Wheatbelt region of southwestern Australia. Here, I aimed to uncover the effects of fragmentation on mammalian species by looking at populations of mammals within conservation reserve areas and the usage of road reserves through detections at these sites. I specifically looked at four native target species, the numbat (*Myrmecobius fasciatus*), chuditch (*Dasyurus geoffroii*), woylie (*Bettongia penicillata)*, and echidna (*Tachyglossus aculeatus*), and two introduced predators the red fox (*Vulpes vulpes*) and feral cat (*Felis catus*) and their movement through the region. In Chapter 5, I uncovered that the signals from carrion flies were highly localised to the area where the vertebrate populations were found; therefore, the results from Chapter 6 could be used to pinpoint the presence of these mammalian species through this region at a fine scale. Chapter 6 showed a

higher diversity of native mammals located within conservation reserve areas, while non-native species were present throughout the region. Whilst our key native taxa were detected at some road reserves, these road reserves were closely connected with conservation reserves. In the middle of the region, where there was little connection to main conservation reserve areas, there was lower native mammal diversity, and native mammals were completely absent from some road reserves. This result implies a degree of isolation between conservation reserves caused by the fragmentation of the habitat, as native mammals could not travel through the landscape easily.

Additionally, at all conservation reserves and road reserves, we were able to detect introduced predators, either foxes or feral cats. So, whilst native mammals were isolated between reserves, introduced predators could move within the region and utilise anthropogenically disturbed landscapes. Though extensive introduced predator management has been occurring in these conservation reserves (Marlow et al., 2015; Thomson & Algar, 2000), the widespread detections of these predator species throughout road reserves indicated that current management strategies have not eradicated them from the region entirely. This chapter demonstrates the value of carrion fly iDNA to complement conservation and land management practices, and could be especially useful in remote areas or rough terrain, where tracking mammal populations may be difficult to undertake.

7.3 Challenges and limitations

This thesis demonstrates the potential for applying metabarcoding to invertebrates as a monitoring tool across terrestrial ecosystems. Chapter 2 (Section 2.8) reviews the current barriers to using metabarcoding in terrestrial land management in depth. Therefore, within this section, I will discuss the limitations of the work in my thesis. While each of the chapters within this thesis touches on those elements and discusses its constraints and challenges, here, I will discuss four overarching themes that have emerged across all the chapters as limitations towards the results I was able to obtain; single time point sampling, poorly populated reference databases, the impacts of DNA degradation, and abundance quantification.

Firstly, Chapters 3-6 in this thesis examine biodiversity at a single time point. Whilst this does not diminish the significance of the findings from each of the studies, further data as a time series would increase the ability to interpret trends observed in the environment. Seasonal effects have been detected using metabarcoding across various organisms (Barsoum et al., 2019; Baselga et al., 2013; Chain et al., 2016; Stoeckle et al., 2017). Characterisation of the seasonal effects on biodiversity shifts detected from and by invertebrates may be essential to conducting terrestrial monitoring studies using the metabarcoding methodology. This is mainly because invertebrate communities are heavily impacted by seasonal changes (Davis et al., 2006; Westveer et al., 2018). Therefore, the application of DNA metabarcoding to studying invertebrate communities or the communities they are interacting with needs to be cognisant of surrounding seasonal effects. Previous work on environmental DNA (eDNA) from water sources has shown that time differences across a single week can demonstrate variability in detected communities that are equivalent to spatial replicates (Beentjes et al., 2019). This suggests that the turnover effects from eDNA are possibly more substantial than that of community heterogeneity in water bodies. While overall seasonal changes may be necessary when interpreting trends from organisms in bulk samples, short-term temporal heterogeneity may impact what is observed in iDNA results. Here, by using time points

across several days as sampling replicates (Chapters 5 and 6), I aimed to reduce the impacts of this temporal discrepancy. However, these effects have not been empirically characterised with iDNA within my studies and may impact how many vertebrate species were detected using iDNA.

The lack of complete genetic reference databases was another challenge faced by my studies and is common to many other metabarcoding datasets. Many taxonomic assignments were made across my chapters that could not be resolved to species-level. For fields of study where taxonomic sufficiency is routinely practised, such as monitoring indicator arthropods to study habitat quality, species-level identification, although helpful, is unnecessary and broader taxonomic classifications can be just as useful. However, species-level identification is essential for conservation, such as identifying specific diet items or tracking threatened species. For Chapter 4, identifications for some plant ASVs were only resolvable to order level. This limits the interpretation of the dataset and how the information from this study can be used in a conservation context. As the extent, quality, and coverage of genetic reference databases expand over time, this limitation will decrease. Techniques such as genome skimming may become useful to fill these database gaps (Trevisan et al., 2019).

The issues associated with taxonomic identification of metabarcoding ASVs can also be related to the short fragments of DNA. Digestion or physical environmental conditions can mean that any DNA gathered from invertebrates, or the environment is subject to degradation, shortening DNA fragments' length (Nielsen et al., 2007). In Chapter 4, the assay with the shorter amplicon length was able to identify more families of plants, presumably because of considerable DNA degradation through larval activity in the brooding tubes.

Shorter DNA fragments can also make it difficult to identify closely related species. For example, several Macropod ASVs identified in Chapter 6 could not be resolved beyond family level because of how genetically similar the ASVs were to other known species in the region. Some solutions to this can

be through using a multiple metabarcoding assays across different gene regions to increase the possibility of identifying organisms (Lynggaard et al., 2019; Yoccoz et al., 2012). Furthermore, assessing the efficacy of different DNA extraction methods, PCR polymerase choice, PCR additives and choice of molecular markers can also improve the quality and quantity of DNA retrieved from degraded samples (Särkinen et al., 2012). For every chapter in this thesis, there was considerable unpublished experimentation and optimisation that went into improving outcomes for DNA extraction, PCR, and sequencing. Whilst costs and time may be barriers to undertaking such investigation for every study, I believe it is important that this experimentation is undertaken when any new substrate or DNA source is put through laboratory workflows, as it improves the outcomes for downstream analysis.

Finally, for many traditional biomonitoring programs, abundance of animals is a critical component to help interpret patterns of change (van Klink et al., 2020). However, equating abundance with numbers of metabarcoding sequences is unreliable (Elbrecht & Leese, 2015). This is predominantly related to the reliance of PCR to generate sequences, whereby primer bias can skew sequence abundance information (Schloss et al., 2011). Therefore, relative abundance using the repeated occurrence or relative sequencing depth is often used in DNA metabarcoding studies as a proxy for abundance (Hänfling et al., 2016; Klymus et al., 2017; Richardson et al., 2015). In Section 2.8, I discuss how the estimation of species abundance is a barrier for bulk arthropod metabarcoding and the various developments or new technologies required to help quantify abundance from metabarcoding data. However, there is very little information on how iDNA detections are influenced by abundances of vertebrates (Carvalho et al., 2021; Schnell et al., 2015). In Chapter 5, I began to explore the impacts of biomass (which is influenced by animal abundance). Despite varying levels of biomass within an area, carrion fly iDNA detections were not impacted. Future work investigating the influence of population abundances using DNA mass from carrion fly samples or relative detection number could be valuable.

7.4 Significance of thesis

Across the world, faunal biodiversity is in decline (Butchart et al., 2010; Sánchez-Bayo & Wyckhuys, 2019; Wagner, 2020). Whilst there are many different solutions to solving this biodiversity crisis, an important part of the conservation and management of landscapes will always be monitoring biodiversity. There has been a considerable amount of research into the use of invertebrate metabarcoding in aquatic ecosystems (see for example, Bush et al., 2019; Elbrecht et al., 2017; Leese et al., 2020; Martins et al., 2019), however the research into applying metabarcoding to terrestrial invertebrates and invertebrate derived (iDNA) studies could expand. The research conducted in this thesis demonstrates the value of invertebrate DNA metabarcoding (and iDNA) for monitoring terrestrial ecosystems and adds important information regarding how this method can be used effectively, demonstrating both its strengths and limitations.

The main findings from this thesis have direct implications for the management of species, from arthropods to charismatic megafauna. Collectively, the chapters have used metabarcoding to uncover new aspects of ecosystem connectivity and biodiversity. Chapters 3 and 4 expand on the current applications of DNA metabarcoding, adding important information to understand how Western Australian arthropod communities respond to environmental disturbances. These chapters provide evidence of how arthropods respond to changes in the environment in a region where this information has been challenging to uncover. Chapters 5 and 6 demonstrate a critical new development for iDNA applications. Using well studied systems, the research within these chapters aimed to validate and confirm the effectiveness of iDNA as a technique to monitor vertebrates. In Chapter 5 I uncovered critical information regarding how iDNA signals move through ecosystems, how they are impacted by biomass and taxonomic group of vertebrates they are used to monitor. The gap in the research that Chapter 5 filled has been proposed by Schnell et al. (2015) as critical information required to make iDNA monitoring studies more effective in a conservation context. Understanding these limitations means that in Chapter

6 when this method was implemented in the field, the findings could give meaningful conservation outcomes for threatened mammalian species.

Whilst the results from my thesis have shown that metabarcoding cannot be used as a magic monitoring tool that will fix all problems associated with current methods of terrestrial biomonitoring, this thesis provides the guidelines for how best it can be implemented to complement and elevate existing monitoring tools. There is still plenty of future work that needs to be undertaken before DNA metabarcoding can become a regular part of the biomonitoring toolkit. Still, I believe that the results presented here indicate the priorities and directions for future research.

7.5 Future directions

More studies are now using DNA metabarcoding to uncover biodiversity than ever before. While completing this thesis, the field has already evolved, and there have been significant improvements to laboratory workflows, sequencing technologies, and bioinformatics pipelines. These improvements have meant that the reliability of DNA metabarcoding results has improved, errors have been reduced, and the costs associated with processing and sequencing samples have decreased. However, there is still a significant amount of research that needs to take place before we see metabarcoding being regularly deployed in monitoring programs. Here, I will describe some areas of future research that I consider to be important for improving the accuracy, utility, and effectiveness of metabarcoding studies using invertebrates.

7.5.1 Standardisation and automation of methods

The first critical area for future development will be making the information gained from DNA metabarcoding studies useful and informative for land managers. In biosecurity, for example, a large barrier to implementation of metabarcoding tools has been the lack of standardisation (Piper et al., 2019). This will be a difficult task to undertake in an emerging field where methodological development is required. For example, there is no consensus on the best way to sample the environment. Increasingly, studies are showing that different types of substrates are needed for detecting different aspects of biodiversity (Koziol et al., 2019; van der Heyde et al., 2020). Even with iDNA, the choice of invertebrate can impact what vertebrate diversity can be detected (Massey et al., 2021). Further, new discoveries, such as DNA from the air to detect vertebrates (Clare et al., 2022; Lynggaard et al., 2022), are currently emerging. Therefore, standardisation may be very difficult. Nevertheless, new methodologies need to be calibrated, and the benefits and limitations fully explored if they are to be of value in a management context.

With the development and refining of methodologies, automation may be the next frontier to explore. Already, automated robotics and liquid handling is becoming more common place in laboratory workflows. Utilising automated methods can improve the speed of obtaining results, reduce the risks of cross contamination and increases the reliability and comparability of results (Buchner et al., 2021). Recent research has also explored the implementation of machine learning, imagery-based monitoring techniques for species identification from mixed arthropod samples (Wührl et al., 2022). Implementing machine learning with molecular tools to potentially tag images with DNA barcodes has great potential to build capabilities for monitoring bulk samples and expediting the discovery and monitoring of cryptic species. Future research may see the improvement of the portability of these tools, and implementation for automated field-based monitoring, or even field-based sequencing. This has great utility in biosecurity for invasive species control, for example an in-field detection of a major agricultural pest species (e.g., Cotton bollworm (*Helicoverpa armigera*) or Tobacco whitefly (*Bemisia tabaci*)) could improve the speed of management response to mitigate any potential outbreaks.

7.5.2 Data accessibility

Part of improving the implementation of DNA metabarcoding-based methods will be the accessibility of data. This includes sharing and publishing detailed methodologies and raw data that include geographic and climatic information for every dataset. This data needs to be easy for researchers to use, but also accessible for policy and decision makers to understand. Developing user-friendly platforms has been identified as a method to improve the bioinformatic literacy involved in making informed decisions about molecular monitoring methods and analysis of results (Ruppert et al., 2019). Metricising biodiversity data to create a health-score or indicator metric in the same way pollution metrics are calculated (Odiete, 2020) may also be part of the future. Citizen science and community-based sampling and monitoring programs can benefit and build on these metrics

and data, adding to their value with tangible local impact. Storage of data is just as important, and stable storage of data will be essential for comparing seasonal and even multi-year variations. Temporal replication is extremely valuable in detecting the response of biodiversity to, for example, climatic shifts (Berry et al., 2019).

7.5.3 Modelling and future prediction

The availability of long-term temporal data will make it easier to develop models that can predict changes in biodiversity. However, this is contingent on our understanding the variables that impact biodiversity detection from metabarcoding. In Chapter 5, I explored the factors that could influence iDNA detections. However, there is a need to build on this research and to develop a mechanistic model that can identify the factors that underpin the likelihood that a species will be detected and that can be used predictively. This type of model would be essential to develop for any application of DNA metabarcoding where insects act as vectors for ecosystem biodiversity. Additionally, this has utility beyond invertebrates, and can be used to tackle whole-ecosystem level monitoring if calibrated with the most effective mix of eDNA substrates. To develop these models, considerations towards the biological, chemical, and physical factors that influence DNA detection in different environmental conditions is necessary. Mechanistic models to develop biodiversity predictions have been suggested as important in combatting the effects of climate change (Mokany & Ferrier, 2011; Urban et al., 2016). However, there have been very few models developed that have been implemented with success across multiple groups of organisms, with predominant focus on just one or two groups (Leidinger & Cabral, 2017; Shipley et al., 2006). With the ability to use invertebrate DNA metabarcoding to examine multiple trophic groups all at once, there is scope in this space to develop modelling tools that predict holistic biodiversity changes in ecosystems.

7.5.4 PCR-free Metabarcoding

A possible future direction is the potential to move away from PCR-based methods and develop new sequencing technologies that can overcome the issues associated with primer bias. As discussed previously, amplicon sequencing is heavily impacted by primer choice, PCR conditions, and reagents. These factors can make it challenging to use amplicon sequencing to effectively characterise biodiversity in the same way that traditional methods can (i.e., detect abundance) (Bohmann et al., 2021; Elbrecht & Leese, 2015). Therefore, the development of PCR-free sequencing methods could theoretically overcome some of these biases. One such method that is already readily implemented using current next-generation sequencing technologies is mitochondrial metagenomics (mitogenomics), where Illumina sequencing is conducted on the whole mitochondrial genomes of bulk samples (Crampton-Platt et al., 2016). This approach has been shown to be useful in estimating biomass for invertebrate species (Bista et al., 2018). Furthermore, because whole genomes are being sequenced, evolutionary information can also be acquired. However, the efficacy of mitogenomics is dependent on the availability of appropriate reference genomes for mapping (Gómez-Rodríguez et al., 2015), which is also a limitation for metabarcoding studies. Mitogenomics is also considerably more costly than amplicon sequencing. Additionally, the sequencing depth and computer processing requirements to analyse this data are immense and can present a limitation for research groups without extensive funding or bioinformatic resources (Cowart et al., 2018). Therefore, it may be a question of improving the quantitative capabilities of metabarcoding to inform when mitogenomics may be a valuable investment. Further developments into populating reference databases and improving sequencing technologies to lower costs and increase the availability of bioinformatic resources will also be necessary to increase the accessibility of this technique.

7.5.5 Advancing experimental design for eDNA/iDNA studies

Whilst it is important to design experiments around ecological questions and the limitations of the collection method, for many of the chapters within this thesis, I explored novel substrate types (i.e., bee nesting tubes, iDNA from carrion flies), as such making decisions around aspects of iDNA/eDNA experimental design challenging. In such exploratory studies, it is difficult to know *a-priori* whether the replication, controls, or spatial/temporal range is appropriate for the question(s) or powered for the statistical analysis. Furthermore, it is hard to go back once the data is collected to address these limitations. Within this thesis, there are instances in every chapter where experimental design could be improved to mitigate confounding variables and unknowns – many of which only became apparent as the data was analysed. It is easy to design eDNA/iDNA experiments that are logistically impossible or too expensive – the more considered approach is to maximise the cost/benefit so that budgets for biomonitoring can extend temporally and spatially. Therefore, moving forward, especially with novel substrates and eDNA, iterative study design that builds on previous work and improves data collection, from appropriate replicate number to parsing of data, is an important step towards standardising metabarcoding methods for terrestrial monitoring.

7.6 Thesis conclusion

Since the first study applying metabarcoding techniques to invertebrates a decade ago (Yu et al., 2012), there have been considerable advances in the technology and the analysis of metabarcoding data from invertebrates. This thesis demonstrates the value invertebrates have in ecosystems and how this importance and interconnectedness can be leveraged to help monitor and protect existing biodiversity via the use of DNA metabarcoding. The overarching question of this thesis was, "Can invertebrate DNA metabarcoding be used for monitoring terrestrial ecosystem responses to change?" To answer this question, I reviewed current literature regarding bulk arthropod metabarcoding for monitoring terrestrial ecosystems (Chapter 2). Using this knowledge, I developed experimental studies that would employ invertebrate DNA metabarcoding in ecosystems subject to disturbance and change and monitor how the diversity and ecology of the invertebrates in this environment responded (Chapters 3 and 4). Lastly, I demonstrated that invertebrate metabarcoding could be used for more than just invertebrate monitoring and developed studies that used carrion fly iDNA to monitor vertebrates in changed environments (Chapters 5 and 6). These results show just how versatile metabarcoding of invertebrates is and that it has immense strength as a tool to monitor changing and disturbed terrestrial ecosystems. However, invertebrate metabarcoding is not without its limitations, and understanding these limitations is essential to acquiring robust, accurate, and informative results. The results from this body of work suggest new research directions and questions that need to be answered to make the use of metabarcoding in terrestrial environments more effective. While the future of monitoring will always lie in a combination of techniques, there is mounting evidence that demonstrates the utility of DNA metabarcoding and other molecular tools for the global monitoring of terrestrial ecosystems.

7.7 References

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged

Barsoum, N., Bruce, C., Forster, J., Ji, Y.-Q., & Yu, D. W. (2019). The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity. *Ecological Indicators*, *101*, 313–323.

Baselga, A., Fujisawa, T., Crampton-Platt, A., Bergsten, J., Foster, P. G., Monaghan, M. T., & Vogler, A. P. (2013). Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels. *Nature Communications*, *4*, 1892.

Beentjes, K. K., Speksnijder, A. G. C. L., Schilthuizen, M., Hoogeveen, M., & van der Hoorn, B. B. (2019). The effects of spatial and temporal replicate sampling on eDNA metabarcoding. *PeerJ*, *7*, e7335.

Berry, T. E., Saunders, B. J., Coghlan, M. L., Stat, M., Jarman, S., Richardson, A. J., Davies, C. H., Berry, O., Harvey, E. S., & Bunce, M. (2019). Marine environmental DNA biomonitoring reveals seasonal patterns in biodiversity and identifies ecosystem responses to anomalous climatic events. *PLoS Genetics*, *15*(2), e1007943.

Bista, I., Carvalho, G. R., Tang, M., Walsh, K., Zhou, X., Hajibabaei, M., Shokralla, S., Seymour, M., Bradley, D., Liu, S., Christmas, M., & Creer, S. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12888

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A., & Creer, S. (2021). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231–1246.

Buchner, D., Macher, T.-H., Beermann, A. J., Werner, M.-T., & Leese, F. (2021). Standardized high-throughput biomonitoring using DNA metabarcoding: Strategies for the adoption of automated liquid handlers. *Environmental Science and Ecotechnology*, *8*, 100122.

Bush, A., Compson, Z. G., Monk, W. A., Porter, T. M., Steeves, R., Emilson, E., Gagne, N., Hajibabaei, M., Roy, M., & Baird, D. J. (2019). Studying Ecosystems With DNA Metabarcoding: Lessons From Biomonitoring of Aquatic Macroinvertebrates. *Frontiers in Ecology and Evolution*, *7*, 434.

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., … Watson, R. (2010). Global biodiversity: indicators of recent declines. *Science*, *328*(5982), 1164–1168.

Calvignac-Spencer, S., Merkel, K., Kutzner, N., Kühl, H., Boesch, C., Kappeler, P. M., Metzger, S., Schubert, G., & Leendertz, F. H. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Molecular Ecology*, *22*(4), 915–924.

Carvalho, C. S., de Oliveira, M. E., Rodriguez-Castro, K. G., Saranholi, B. H., & Galetti, P. M., Jr. (2021). Efficiency of eDNA and iDNA in assessing vertebrate diversity and its abundance. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13543

Chain, F. J. J., Brown, E. A., MacIsaac, H. J., & Cristescu, M. E. (2016). Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. *Diversity & Distributions*, *22*(5), 493–504.

Clare, E. L., Economou, C. K., Bennett, F. J., Dyer, C. E., Adams, K., McRobie, B., Drinkwater, R., & Littlefair, J. E. (2022). Measuring biodiversity from DNA in the air. *Current Biology*, *32*(3), 693-700.e5.

Cowart, D. A., Murphy, K. R., & Cheng, C.-H. C. (2018). Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. *Marine Genomics*, *37*, 148–160.

Crampton-Platt, A., Yu, D. W., Zhou, X., & Vogler, A. P. (2016). Mitochondrial metagenomics: letting the genes out of the bottle. *GigaScience*, *5*, 15.

Davis, C. A., Austin, J. E., & Buhl, D. A. (2006). Factors influencing soil invertebrate communities in riparian grasslands of the central Platte River floodplain. *Wetlands*, *26*(2), 438–454.

Elbrecht, V., & Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. *PloS One*, *10*(7), e0130324.

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, *8*(10), 1265–1275.

Gómez-Rodríguez, C., Crampton-Platt, A., Timmermans, M. J. T. N., Baselga, A., & Vogler, A. P. (2015). Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution*, *6*(8), 883–894.

Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., Blackman, R. C., Oliver, A., & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, *25*(13), 3101–3119.

Henniges, M. C., Powell, R. F., Mian, S., Stace, C. A., Walker, K. J., Gornall, R. J., Christenhusz, M. J. M., Brown, M. R., Twyford, A. D., Hollingsworth, P. M., Jones, L., de Vere, N., Antonelli, A., Leitch, A. R., & Leitch, I. J. (2022). A taxonomic, genetic and ecological data resource for the vascular plants of Britain and Ireland. *Scientific Data*, *9*(1), 1.

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P. M., Woodcock, P., Edwards, F. A., Larsen, T. H., Hsu, W. W., Benedick, S., Hamer, K. C., Wilcove, D. S., Bruce, C., Wang, X., Levi, T., Lott, M., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257.

Kim, K. C. (1993). Biodiversity, conservation and inventory: why insects matter. *Biodiversity and Conservation*, *2*(3), 191–214.

Klymus, K. E., Marshall, N. T., & Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PloS One*, *12*(5), e0177643.

Koziol, A., Stat, M., Simpson, T., Jarman, S., DiBattista, J. D., Harvey, E. S., Marnane, M., McDonald, J., & Bunce, M. (2019). Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources*, *19*(2), 366–376.

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., & Zizka, V. M. A. (2020). Improved freshwater macroinvertebrate detection from eDNA through minimized non-target amplification. *bioRxiv*. https://doi.org/10.1101/2020.04.27.063545

Leidinger, L., & Cabral, J. S. (2017). Biodiversity Dynamics on Islands: Explicitly Accounting for Causality in Mechanistic Models. *Diversity*, *9*(3), 30.

Lynggaard, C., Bertelsen, M. F., Jensen, C. V., Johnson, M. S., Frøslev, T. G., Olsen, M. T., & Bohmann, K. (2022). Airborne environmental DNA for terrestrial vertebrate community monitoring. *Current Biology*, *32*(3), 701-707.e5.

Lynggaard, C., Nielsen, M., Santos-Bay, L., Gastauer, M., Oliveira, G., & Bohmann, K. (2019). Vertebrate diversity revealed by metabarcoding of bulk arthropod samples from tropical forests. *Environmental DNA*, *1*(4), 329–341.

Majer, J. D., de Sousa-Majer, M. J., & Heterick, B. E. (2020). Partial clearing of a road corridor leads to homogenisation of the invertebrate fauna. *Pacific Conservation Biology*. https://doi.org/10.1071/PC20039

Marlow, N. J., Thomas, N. D., Williams, A. A. E., Macmahon, B., Lawson, J., Hitchen, Y., Angus, J., & Berry, O. (2015). Cats (*Felis catus*) are more abundant and are the dominant predator of woylies (*Bettongia penicillata*) after sustained fox (*Vulpes vulpes*) control. *Australian Journal of Zoology*, *63*(1), 18–27.

Martins, F. M. S., Galhardo, M., Filipe, A. F., Teixeira, A., Pinheiro, P., Paupério, J., Alves, P. C., & Beja, P. (2019). Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. *Molecular Ecology Resources*, *19*(4), 863–876.

Massey, A. L., Bronzoni, R. V. de M., da Silva, D. J. F., Allen, J., de Lazari, P. R., Dos Santos Filho, M., Canale, G. R., Bernardo, C. S. S., Peres, C. A., & Levi, T. (2021). Invertebrates for vertebrate biodiversity monitoring: comparisons using three insect taxa as iDNA samplers. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13525

McGeoch, M. A. (1998). The selection, testing and application of terrestrial insects as bioindicators. *Biological Reviews of the Cambridge Philosophical Society*, *73*(2), 181–201.

McKenney, B. A., & Kiesecker, J. M. (2010). Policy development for biodiversity offsets: a review of offset frameworks. *Environmental Management*, *45*(1), 165–176.

Mokany, K., & Ferrier, S. (2011). Predicting impacts of climate change on biodiversity: a role for semi-mechanistic community-level modelling. *Diversity & Distributions*, *17*(2), 374–380.

Nielsen, K. M., Johnsen, P. J., Bensasson, D., & Daffonchio, D. (2007). Release and persistence of extracellular DNA in the environment. *Environmental Biosafety Research*, *6*(1–2), 37–53.

Odiete, W. E. (2020). Novel metric for managing the protection of humanity and the environment against pollution and its adverse effects. *Heliyon*, *6*(11), e05555.

Owings, C. G., Banerjee, A., Asher, T. M. D., Gilhooly, W. P., 3rd, Tuceryan, A., Huffine, M., Skaggs, C. L., Adebowale, I. M., Manicke, N. E., & Picard, C. J. (2019). Female Blow Flies As Vertebrate Resource Indicators. *Scientific Reports*, *9*(1), 10594.

Piper, A. M., Batovska, J., Cogan, N. O. I., Weiss, J., Cunningham, J. P., Rodoni, B. C., & Blacket, M. J. (2019). Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *GigaScience*, *8*(8). https://doi.org/10.1093/gigascience/giz092

Richardson, R. T., Lin, C., Quijia, J. O., Riusech, N. S., Goodell, K., & Johnson, R. M. (2015). *Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach* [Data set]. https://doi.org/10.3732/apps.1500043

Rodgers, T. W., Xu, C. C. Y., Giacalone, J., Kapheim, K. M., Saltonstall, K., Vargas, M., Yu, D. W., Somervuo, P., McMillan, W. O., & Jansen, P. A. (2017). Carrion fly-derived DNA metabarcoding is an effective tool for mammal surveys: Evidence from a known tropical mammal community. *Molecular Ecology Resources*, *17*(6), e133–e145.

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, *17*, e00547.

Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, *232*, 8–27.

Särkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., & Bakker, F. T. (2012). How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PloS One*, *7*(8), e43808.

Schloss, P. D., Gevers, D., & Westcott, S. L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PloS One*, *6*(12), e27310.

Schnell, I. B., Sollmann, R., Calvignac-Spencer, S., Siddall, M. E., Yu, D. W., Wilting, A., & Gilbert, M. T. P. (2015). iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool - prospects, pitfalls and avenues to be developed. *Frontiers in Zoology*, *12*, 24.

Schubert, G., & Stockhausen, M. (2015). Targeted detection of mammalian species using carrion fly-derived DNA. *Molecular Ecology*. doi:10.1111/1755-0998.12306

Shipley, B., Vile, D., & Garnier, E. (2006). From plant traits to plant communities: a statistical mechanistic approach to biodiversity. *Science*, *314*(5800), 812–814.

Stoeckle, M. Y., Soboleva, L., & Charlop-Powers, Z. (2017). Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PloS One*, *12*(4), e0175186.

Thomson, P. C., & Algar, D. (2000). The uptake of dried meat baits by foxes and investigations of baiting rates in Western Australia. *Wildlife Research*, *27*(5), 451–456.

Trevisan, B., Alcantara, D. M. C., Machado, D. J., Marques, F. P. L., & Lahr, D. J. G. (2019). Genome skimming is a low-cost and robust strategy to assemble complete mitochondrial genomes from ethanol preserved specimens in biodiversity studies. *PeerJ*, *7*, e7543.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J.-B., Pe'er, G., Singer, A., Bridle, J. R., Crozier, L. G., De Meester, L., Godsoe, W., Gonzalez, A., Hellmann, J. J., Holt, R. D., Huth, A., Johst, K., Krug, C. B., Leadley, P. W., Palmer, S. C. F., Pantel, J. H., … Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. *Science*, *353*(6304). https://doi.org/10.1126/science.aad8466

van der Heyde, M., Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., & Nevill, P. (2020). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Molecular Ecology Resources*, *20*(3). https://doi.org/10.1111/1755-0998.13148

van Klink, R., Bowler, D. E., Gongalsky, K. B., Swengel, A. B., Gentile, A., & Chase, J. M. (2020). Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science*, *368*(6489), 417–420.

Wagner, D. L. (2020). Insect Declines in the Anthropocene. *Annual Review of Entomology*, *65*, 457–480.

Westveer, J. J., van der Geest, H. G., van Loon, E. E., & Verdonschot, P. F. M. (2018). Connectivity and seasonality cause rapid taxonomic and functional trait succession within an invertebrate community after stream restoration. *PloS One*, *13*(5), e0197182.

Wührl, L., Pylatiuk, C., Giersch, M., Lapp, F., von Rintelen, T., Balke, M., Schmidt, S., Cerretti, P., & Meier, R. (2022). DiversityScanner: Robotic handling of small invertebrates with machine learning methods. *Molecular Ecology Resources*, *22*(4), 1626–1638.

Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., Von Stedingk, H., Brysting, A. K., Coissac, E., Pompanon, F., Sønstebø, J. H., Miquel, C., Valentini, A., De Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., … Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, *21*(15), 3647–3655.

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, *3*(4), 613–623.

This page has been left intentionally blank

APPENDIX A: CO-AUTHORSHIP STATEMENTS

Chapter 2: In preparation for submission to the journal *Environmental DNA*.

Fernandes, K, Lynggaard, C, Nielsen, M, Chua, P, Nevill, PG, Bunce, M, Bohmann, K. (2022) Current and future applications of bulk arthropod sample metabarcoding in terrestrial ecological monitoring. *In preparation.*

To Whom It May Concern, I, Kristen Marie Rose Fernandes, contributed to the conceptualisation, writing, and editing of the manuscript. I compiled the final version of the manuscript.

Signed: Kristen Fernandes

I, as Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

Christina Lynggaard

Martin Nielsen

Physilia Chua

Paul Nevill

Michael Bunce

Kristine Bohmann

Chapter 3: In preparation for submission.

Fernandes, K, Clark-Ioannou, S, Saunders, BJ, Bateman, PW, Bunce, M, Nevill, PG. (2022) Metabarcoding reveals the impacts of exploration infrastructure on ground-dwelling arthropod communities at mine sites. *In preparation.*

To Whom It May Concern, I, Kristen Marie Rose Fernandes, contributed to processing the samples, analysed the data, and wrote and edited the resulting manuscript.

Signed: Kristen Fernandes

I, as Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

Sophia Clark-Ioannou

Benjamin J Saunders

Philip W Bateman

Michael Bunce

Paul Nevill

Chapter 4: Accepted with revisions to *Oecologia.* As the first author, permission is automatically granted to reproduce this copyrighted material to include the article in my thesis.

Fernandes, K, Prendergast, K, Bateman, PW, Saunders BJ, Gibberd, M, Bunce, M, Nevill, PG. (2022). DNA Metabarcoding identifies urban foraging patterns of oligolectic and polylectic cavity-nesting bees. *In review (Oecologia)*. doi: 10.21203/rs.3.rs-1256295/v1

To Whom It May Concern, I, Kristen Marie Rose Fernandes, contributed to the conceptualisation of the study, processed the samples, analysed the data, and wrote and edited the resulting manuscript.

Signed: Kristen Fernandes

I, as Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

Kit Prendergast

Philip W Bateman

Benjamin J Saunders

Mark Gibberd

Michael Bunce

Paul Nevill

Chapter 5: In preparation for submission to the journal *Royal Society Open Science*.

Fernandes, K, Bateman, PW, Saunders, BJ, Gibberd, M, Bunce, M, Bohmann, K, Nevill, PG. (2022). Examining the effects of distance, taxon, animal biomass, and diet on vertebrate detections from carrion fly iDNA. *In preparation.*

To Whom It May Concern, I, Kristen Marie Rose Fernandes, contributed to the conceptualisation of the study, collecting, and processing the samples, analysed the data, and wrote and edited the resulting manuscript.

Signed: Kristen Fernandes

I, as Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

Philip W Bateman

Benjamin J Saunders

Mark Gibberd

Michael Bunce

Kristine Bohmann

Paul Nevill

Chapter 6: In preparation for submission to the journal *Molecular Ecology*.

Fernandes, K, Bateman, PW, Saunders, BJ, Bohmann, K, Nevill, PG. (2022). Monitoring biodiversity on-the-fly: Carrion fly iDNA metabarcoding can monitor threatened mammals in fragmented terrestrial ecosystems. *In preparation.*

To Whom It May Concern, I, Kristen Marie Rose Fernandes, contributed to the conceptualisation of the study, collecting, and processing the samples, analysed the data, and wrote and edited the resulting manuscript.

Signed: Kristen Marie Rose Fernandes

I, as Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

Philip W Bateman

Benjamin J Saunders

Kristine Bohmann

Paul Nevill

APPENDIX B: ADDITIONAL PUBLICATIONS

There were five other articles published during this PhD that were not included in the thesis. As an author, permission is automatically granted to reproduce this material.

Fernandes, K., van der Heyde, M., Bunce, M., Dixon, K., Harris, R. J., Wardell-Johnson, G., & Nevill, P. G. (2018). DNA metabarcoding—a new approach to fauna monitoring in mine site restoration. *Restoration Ecology*, *26*(6), 1098–1107. doi: 10.1111/rec.12868

Fernandes, K., van der Heyde, M., Coghlan, M., Wardell-Johnson, G., Bunce, M., Harris, R., & Nevill, P. (2019). Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences. *Restoration Ecology*, *27*(5), 1177-1186–10. doi: 10.1111/rec.12976

Ryan, E., Bateman, P., Fernandes, K., van der Heyde, M., & Nevill, P. (2022). eDNA metabarcoding of log hollow sediments and soils highlights the importance of substrate type, frequency of sampling and animal size, for vertebrate species detection. *Environmental DNA*. doi:10.1002/edn3.306

van der Heyde, M., Bunce, M., Dixon, K. W., Fernandes, K., Majer, J., Wardell-Johnson, G., White, N., & Nevill, P. (2022). Evaluating restoration trajectories using DNA metabarcoding of ground-dwelling and airborne invertebrates and associated plant communities. *Molecular Ecology*, *31*(7), 2172-2188. doi: 10.1111/mec.16375

van der Heyde, M., Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., & Nevill, P. (2020). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Molecular Ecology Resources*, *20*(3), 732-745. doi:10.1111/1755-0998.13148

Restoration Ecology
THE JOURNAL OF THE SOCIETY FOR ECOLOGICAL RESTORATION

REVIEW ARTICLE

DNA metabarcoding—a new approach to fauna monitoring in mine site restoration

Kristen Fernandes[1,2], Mieke van der Heyde[1,2], Michael Bunce[2], Kingsley Dixon[1], Richard J. Harris[3], Grant Wardell-Johnson[1], Paul G. Nevill[1,4]

Ecological restoration of landscapes is an integral part of the mining process. However, restoration is often constrained by a lack of consistent monitoring approaches. For example, the need for specialist techniques and trapping approaches often limits monitoring of fauna recovery. Application of molecular tools has made important contributions to understanding factors influencing restoration success. Here, we outline advances in next-generation sequencing methods, especially metabarcoding of environmental DNA. These have potential to revolutionize the practical contribution of genetics to the monitoring of fauna in a restoration context. DNA metabarcoding involves the simultaneous characterization of biota using DNA barcodes. It is a powerful method to assess the biodiversity contained within environmental samples (e.g. scats, bulk arthropods, soil, water, and sediment). This review outlines the challenges associated with current approaches to monitoring faunal biodiversity throughout ecological restoration. We also demonstrate how the emergence of DNA metabarcoding could recast monitoring capacity for improved ecological restoration outcomes, while discussing current limitations of a DNA-based approach to biodiversity assessment.

Key words: ecological restoration, environmental DNA, fauna, metabarcoding, mine site, monitoring

Implications for Practice

- Following mining, there is a lack of monitoring of fauna in ecological restoration, and a reliance on vegetation monitoring to indicate whole ecosystem recovery.
- DNA metabarcoding makes fauna biodiversity surveys easier to conduct by using DNA extracted from environmental samples (e.g. scats, bulk arthropods, soil, or water) and sequenced using next-generation platforms.
- Current challenges with DNA metabarcoding for ecological restoration monitoring include DNA persistence, barcode choice, taxonomic reference databases, inability to survey abundance, and vagrant DNA.
- As the DNA metabarcoding methodology develops and the challenges are addressed through further research, this technique has potential to become a key component of best-practice fauna monitoring as well as broader biodiversity monitoring, ultimately improving ecological restoration outcomes.

Mine Site Restoration

Mining is the basis of many global economies. In the last 60 years, this importance has grown as demand and production of metals has increased (Cooke & Johnson 2002). In 2016, a global analysis of the top 40 mining companies estimated the value of mining to be $US 748 billion, with profits of $US 20 billion (PWC 2017). However, this high economic value is accompanied by a resource footprint in the range of 1,000 km^2 per company operating globally (Environmental Protection Authority 2014; Stevens & Dixon 2017). Environmental considerations have now become a key legislative requirement of mining projects in most developed countries (Mudd 2007). Therefore, ecological restoration is increasingly an integral part of the mining process (Mchaina 2001; Bridge 2004; Cross et al. 2018). However, mine closure and ecological restoration is increasingly expensive (Costanza et al. 2014). For example, it is estimated that in one mining domain in Australia, the total rehabilitation, restoration, and closure costs for all mines will be $US 3–4.5 billion ($A 4–6 billion) (Gorey et al. 2016).

Common goals set for mine sites following closure include restoration of structure, diversity, and function of the disturbed ecosystems (Majer & Nichols 1998; Cristescu et al. 2012; Miller et al. 2017). In most cases, returning lost biodiversity is focused on restoring plant communities (Majer & Nichols 1998; Bisevac & Majer 1999; Longcore 2003; Cristescu et al. 2012; Miller et al. 2017). However, the restoration of other elements of biodiversity, including fauna, is relatively poorly understood.

Author contributions: PGN, KF formulated the original idea; KF created the figure and the table; all authors contributed to writing the manuscript.

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, GPO Box U1987, Perth, WA 6102, Australia
[2]Trace and Environmental DNA Laboratory, School of Molecular and Life Sciences, Curtin University, GPO Box U1987, Perth, WA 6102, Australia
[3]School of Molecular and Life Sciences, Curtin University, GPO Box U1987, Perth, WA 6102, Australia
[4]Address correspondence to P. G. Nevill, email paul.nevill@curtin.edu.au

doi: 10.1111/rec.12868

RESTORATION ECOLOGY
The Journal of the Society for Ecological Restoration
SER SOCIETY FOR ECOLOGICAL RESTORATION

RESEARCH ARTICLE

Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences

Kristen Fernandes[1,2,3], Mieke van der Heyde[1,2], Megan Coghlan[2], Grant Wardell-Johnson[1], Michael Bunce[2], Richard Harris[4], Paul Nevill[1]

Invertebrate biomonitoring can reveal crucial information about the status of restoration projects; however, it is routinely underused because of the high level of taxonomic expertise and resources required. Invertebrate DNA metabarcoding has been used to characterize invertebrate biodiversity but its application in restoration remains untested. We use DNA metabarcoding, a new approach for restoration assessment, to explore the invertebrate composition from pitfall traps at two mine site restoration chronosequences in southwestern Australia. Invertebrates were profiled using two cytochrome oxidase subunit 1 assays to investigate invertebrate biodiversity. The data revealed differences between invertebrate communities at the two mines and between the different age plots of the chronosequences. Several characteristic taxa were identified for each age within the chronosequence, including springtails within the youngest sites (Order: Collembola) and millipedes within the oldest and reference sites (Order: Julida). This study facilitates development of a molecular "toolkit" for the monitoring of ecological restoration projects. We suggest that a metabarcoding approach shows promise in complementing current monitoring practices that rely on alpha taxonomy.

Key words: biomonitoring, chronosequence, metabarcoding, mine site restoration, next-generation sequencing

Implications for Practice

- Invertebrate monitoring is underused in restoration monitoring because it is resource intensive and requires taxonomic expertise.
- The DNA metabarcoding approach has potential to add considerable value to fauna monitoring programs.
- Using the invertebrate DNA metabarcoding technique it was possible to identify differences in invertebrate communities across restoration chronosequences.
- With further research and development of this technique, DNA metabarcoding has the potential to become a key component of the 'toolkit' for restoration monitoring.

Introduction

Ecosystem restoration is driven by a desire to address continued global declines in land quality, biodiversity, and ecosystem services (Bullock et al. 2011). The ultimate goals of restoration are to return a predisturbance level of condition and function to degraded systems and to enable the restored environment to persist without further intervention from land managers (Hobbs & Suding 2009). However, these goals require an understanding of the complex ecological processes underpinning the environment, as well as the role of biotic and abiotic factors in those processes. Adequate baseline monitoring is therefore crucial to determine ecosystem functioning and composition prior to planned environmental degradation (Pereira et al. 2013). In a restoration context, sites also require consistent monitoring to improve restoration practices through adaptive management interventions (Palmer & Filoso 2009; Wortley et al. 2013). Furthermore, consistent routine monitoring allows regulators and stakeholders to determine whether restoration goals are being achieved (McDonald et al. 2016; Miller et al. 2017). Monitoring is therefore critical to "best-practice" ecological restoration.

Historically, monitoring in restoration has focused on plant species (Young 2000; Ruiz-Jaen & Mitchell Aide 2005). However, plant community monitoring alone cannot ascertain the return of essential ecosystem functions (e.g. nutrient cycling; Herrick et al. 2006). Invertebrates, for example, are essential to

Author contributions: KF conducted the study and wrote the manuscript; PN, RH, MB were involved in the experimental design; PN, RH, MVH were responsible for sample collection; GWJ conducted flora surveys; MC, MB were involved in metabarcoding workflows; KF, MVH conducted the statistical analysis; all authors contributed to editing the manuscript.

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia
[2]Trace and Environmental DNA Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia
[3]Address correspondence to K. Fernandes, email kristen.fernandes@student.curtin.edu.au
[4]School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

doi: 10.1111/rec.12976
Supporting information at:
http://onlinelibrary.wiley.com/doi/10.1111/rec.12976/suppinfo

Received: 26 May 2021 | Revised: 29 March 2022 | Accepted: 4 April 2022

DOI: 10.1002/edn3.306

ORIGINAL ARTICLE

Environmental DNA

WILEY

eDNA metabarcoding of log hollow sediments and soils highlights the importance of substrate type, frequency of sampling and animal size, for vertebrate species detection

Ethan Ryan[1,2] | Philip Bateman[3] | Kristen Fernandes[2] | Mieke van der Heyde[1,2] | Paul Nevill[1,2]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Perth, Western Australia, Australia

[2]Trace and Environmental DNA Laboratory, School of Life and Molecular Sciences, Curtin University, Perth, Western Australia, Australia

[3]Behavioural Ecology Lab, School of Molecular and Life Sciences, Curtin University, Perth, Western Australia, Australia

Correspondence
Paul Nevill, ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences Curtin University, GPO box U19787, Perth, WA 6845, Australia.
Email: paul.nevill@curtin.edu.au

Funding information
IGO Limited Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration, Grant/Award Number: ICI150100041

Abstract

Fauna monitoring often relies on visual monitoring techniques such as camera trapping, which have biases leading to underestimates of vertebrate species diversity. Environmental DNA (eDNA) metabarcoding has emerged as a new source of biodiversity data that may improve biomonitoring; however, eDNA-based assessments of species richness remain relatively untested in terrestrial environments. We investigated the suitability of fallen log hollow sediment as a source of vertebrate eDNA, across two sites in southwestern Australia—one with a Mediterranean climate and the other semi-arid. We compared two different approaches (camera trapping and eDNA metabarcoding) for monitoring of vertebrate species, and investigated the effect of other factors (frequency of species, timing of visits, frequency of sampling, and body size) on vertebrate species detectability. Metabarcoding of hollow sediments resulted in the detection of higher species richness in comparison (29 taxa: six birds, three reptiles, and 20 mammals) to metabarcoding of soil at the entrance of the hollow (13 taxa: three birds, two reptiles, and eight mammals). We detected 31 taxa in total with eDNA metabarcoding and 47 with camera traps, with 14 taxa detected by both (12 mammals and two birds). By comparing camera trap data with eDNA read abundance, we were able to detect vertebrates through eDNA metabarcoding that had visited the area up to two months prior to sample collection. Larger animals were more likely to be detected, and so were vertebrates that were identified multiple times in the camera traps. These findings demonstrate the importance of substrate selection, frequency of sampling, and animal size, on eDNA-based monitoring. Future eDNA experimental design should consider all these factors as they affect detection of target taxa.

KEYWORDS
biodiversity, biomonitoring, camera traps, environmental DNA, fauna surveys, metabarcoding, tree hollows, vertebrates

Received: 29 January 2021 | Revised: 5 December 2021 | Accepted: 19 January 2022

DOI: 10.1111/mec.16375

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Evaluating restoration trajectories using DNA metabarcoding of ground-dwelling and airborne invertebrates and associated plant communities

Mieke van der Heyde[1,2] | Michael Bunce[2,3] | Kingsley W. Dixon[1] | Kristen Fernandes[2] | Jonathan Majer[1] | Grant Wardell-Johnson[1] | Nicole E. White[2] | Paul Nevill[1,2]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Perth, Western Australia, Australia

[2]Trace and Environmental DNA Laboratory, School of Life and Molecular Sciences, Curtin University, Perth, Western Australia, Australia

[3]Institute of Environmental Science and Research (ESR), Porirua, New Zealand

Correspondence
Mieke van der Heyde, ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, Perth, Western Australia.
Email: Mieke.vanderheyde@curtin.edu.au

Funding information
Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration, Grant/Award Number: ICI150100041

Handling Editor: Carla Lopes

Abstract

Invertebrates are important for restoration processes as they are key drivers of many landscape-scale ecosystem functions; including pollination, nutrient cycling and soil formation. However, invertebrates are often overlooked in restoration monitoring because they are highly diverse, poorly described, and time-consuming to survey, and require increasingly scarce taxonomic expertise to enable identification. DNA metabarcoding is a relatively new tool for rapid survey that is able to address some of these concerns, and provide information about the taxa with which invertebrates are interacting via food webs and habitat. Here, we evaluate how invertebrate communities may be used to determine ecosystem trajectories during restoration. We collected ground-dwelling and airborne invertebrates across chronosequences of mine-site restoration in three ecologically disparate locations in Western Australia and identified invertebrate and plant communities using DNA metabarcoding. Ground-dwelling invertebrates showed the clearest restoration signals, with communities becoming more similar to reference communities over time. These patterns were weaker in airborne invertebrates, which have higher dispersal abilities and therefore less local fidelity to environmental conditions. Although we detected directional changes in community composition indicative of invertebrate recovery, patterns observed were inconsistent between study locations. The inclusion of plant assays allowed identification of plant species, as well as potential food sources and habitat. We demonstrate that DNA metabarcoding of invertebrate communities can be used to evaluate restoration trajectories. Testing and incorporating new monitoring techniques such as DNA metabarcoding is critical to improving restoration outcomes.

KEYWORDS
DNA metabarcoding, environmental DNA, Invertebrates, monitoring, restoration, trajectory

Received: 16 August 2019 | Revised: 30 December 2019 | Accepted: 10 February 2020

DOI: 10.1111/1755-0998.13148

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding

Mieke van der Heyde[1,2] | Michael Bunce[2] | Grant Wardell-Johnson[1] | Kristen Fernandes[2] | Nicole E. White[2] | Paul Nevill[1,2]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Perth, WA, Australia

[2]Trace and Environmental DNA Laboratory, School of Life and Molecular Sciences, Curtin University, Perth, WA, Australia

Correspondence
Mieke van der Heyde, ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, Perth, WA, Australia.
Email: mieke.vanderheyde@postgrad.curtin.edu.au

Funding information
ARC Centre for Mine Site Restoration, Grant/Award Number: ICI150100041

Abstract

Biological surveys based on visual identification of the biota are challenging, expensive and time consuming, yet crucial for effective biomonitoring. DNA metabarcoding is a rapidly developing technology that can also facilitate biological surveys. This method involves the use of next generation sequencing technology to determine the community composition of a sample. However, it is uncertain as to what biological substrate should be the primary focus of metabarcoding surveys. This study aims to test multiple sample substrates (soil, scat, plant material and bulk arthropods) to determine what organisms can be detected from each and where they overlap. Samples (n = 200) were collected in the Pilbara (hot desert climate) and Swan Coastal Plain (hot Mediterranean climate) regions of Western Australia. Soil samples yielded little plant or animal DNA, especially in the Pilbara, probably due to conditions not conducive to long-term preservation. In contrast, scat samples contained the highest overall diversity with 131 plant, vertebrate and invertebrate families detected. Invertebrate and plant sequences were detected in the plant (86 families), pitfall (127 families) and vane trap (126 families) samples. In total, 278 families were recovered from the survey, 217 in the Swan Coastal Plain and 156 in the Pilbara. Aside from soil, 22%–43% of the families detected were unique to the particular substrate, and community composition varied significantly between substrates. These results demonstrate the importance of selecting appropriate metabarcoding substrates when undertaking terrestrial surveys. If the aim is to broadly capture all biota then multiple substrates will be required.

KEYWORDS

biodiversity, biological audit, DNA barcoding, metagenomics, terrestrial

1 | INTRODUCTION

There is a growing need for effective biomonitoring with increasing pressure on ecological systems from human population growth, resource use and climate change (Dirzo et al., 2014; Pimm et al., 2014; UNEP, 2011). Biomonitoring is necessary for effective ecosystem management, including the early detection of invasive species (Epanchin-Niell, Haight, Berec, Kean, & Liebhold, 2012), measurement of trajectories following ecological restoration (Herrick, Schuman, & Rango, 2006), and the conservation of threatened or endangered species and ecological communities (Campbell et al., 2002). Traditionally, biomonitoring has relied on visual surveys and traps with species identification based on morphology. However, this presents challenges in some groups due to (a) phenotypic