**Molecular and Life Sciences**

# Assessing the use of eDNA metabarcoding to monitor mine site restoration

**Mieke Elisabeth van der Heyde**
**0000-0002-1658-9927**

**This thesis is presented for the Degree of**
**Doctor of Philosophy**
**of**
**Curtin University**

**January 2021**

# 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:** Mieke van der Heyde

**DATE:** 25 September, 2020

# Abstract

Globally, ecological restoration is increasingly recognized as the primary method to reverse biodiversity losses and reinstate ecosystems and their associated services. Monitoring restoration is key to assessing the success or failure of methodologies, to improve future restoration efforts, and indicate when further intervention may be necessary. Most restoration monitoring is centered on vegetation, and frequent monitoring is costly and labor intensive. This thesis explores the use of DNA metabarcoding, a technology that has become increasingly available over recent years, to monitor restoration.

DNA Metabarcoding involves the use of high throughput sequencing to sequence barcode regions of the genome to determine the community composition of an environmental sample. The first stage of this PhD involves testing multiple substrates (soil, feces, bulk plant material, bulk arthropods) to determine which sample material is suitable for restoration monitoring. I found that some substrates detected more diversity (scat) than others (soil) and community composition varied significantly between substrates. If the aim is to broadly capture all biota then multiple substrates will be required. This information is then applied to chronosequences of mine site restoration in three ecologically different locations in Western Australia to investigate the recovery of soil microbes, vertebrates, and invertebrates.

Soil microbial communities (SMC) are ubiquitous, respond rapidly to changes in the environment, and are the functional basis of ecosystems, making them excellent indicators for restoration monitoring. I found that microbial communities showed patterns of recovery, with communities becoming more similar to reference communities over time. However, these patterns were not consistent between locations or the target microbes (bacteria or fungi), and influenced by the addition of fresh topsoil during site preparation.

Similarly to SMC, fauna recovery in restoration is rarely studied, but nevertheless important as return of plant communities does not necessarily indicate the recovery of associated fauna. Here I use wide-scale testing in different ecosystems to demonstrate the use of a novel substrate (pooled scat samples) to assess bird and mammal diversity, and show the limitations in certain environments. This method is was able to differentiate between restored and reference sites, but is most suitable to environments with low vegetative cover with higher scat detectability and persistence.

Finally, invertebrates are ideal for restoration monitoring because they respond quickly to disturbances, indicate various ecosystem functions, and have been identified a key indicators of restoration success. They are also abundant, diverse, and require many expert hours to identify, making them an appealing target for DNA metabarcoding which allows rapid, cost-effective identification of the invertebrates and their associated plant communities. I found that ground dwelling invertebrates showed the strongest patterns of community recovery, while airborne invertebrate had less local fidelity because of their high dispersal abilities. Assessing plant diversity provides additional functional information about the interactions between invertebrate and plant communities, and indicated that invertebrates are foraging locally in restored sites.

The studies presented here illustrate the viability of this tool to improve biological monitoring of ecological restoration by expanding the range of what can be monitored, and testing its potential and limitations across a range of ecosystems. With further refinement to experimental design, these approaches may help guide future restoration efforts and interventions, and expand beyond mining restoration to other applications of biological monitoring.

# Acknowledgements

This thesis could not have been completed without the support and encouragement of an army of supervisors, collaborators, colleagues, friends, and family. I would like to thank my supervisory panel Dr. Paul Nevill, Prof. Michael Bunce, Dr. Nicole White, Prof. Grant Wardell-Johnson, and Prof. Kingsley Dixon for their guidance throughout this PhD. With your support I have learned so much about performing good science, the importance of collaboration, and the skills I will need to forge my own career.

I would not have enjoyed the last three and a half years without my friends and colleagues in the Trace and Environmental DNA Laboratory (TrEnD) and the Centre for Mine Site Restoration (CMSR). I would especially like to thank Matt Power and for introducing me to the laboratory skills I needed for much of the PhD, Sheree Walters, Kristen Fernandes, and Katrina West for help with field work, and my friends in the office Adam Koziol, Kristen Fernandes, Katrina West, Matt Heydenreych, Mahsa Mousavi-Derazmahalleh, Miwa Takahashi, Arne Adam, Georgia Nestor, Sophie Preston, Jackie Jones, Alishum Ali, David Juszkiewicz, Jason Alexander, Tina Berry, Sarah Crump, Sarah Luxton, Frederik Seersholm. Thank you for the good times, the support during tough times and the company throughout.

Most of my research could not have been conducted without the support of key people in BHP, Hanson Construction Materials, and South32. Thank you for helping me access sites, and for answering my endless questions and emails.

Finally I would like to thank my family who have always supported me and encouraged me to push my boundaries and see how far I could go.

The financial support for my PhD came from Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration (ICI150100041). I was

personally supported by the Curtin International Postgraduate Research Scholarship (CIPRS) and Research Stipend Scholarship.

# Table of Contents

# Statement of Contributions

**Chapter 2:** van der Heyde, M.,Bunce, M., 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*, 00: 1-14. https://doi.org/10.1111/1755-0998.13148.

**Author Contributions:** MvH. conducted the study and wrote the manuscript. PN, MB., NW and GW-J. were involved in the experimental design; samples were collected by MvH, PN, KF and GW-J, and processed by MvH and KF; molecular and bioinformatics work was performed by MvH.; all data were analysed and processed by MvH.; statistical analysis was done by MvH.; the manuscript was edited by all authors

**Chapter 3:** van der Heyde, M**.,** Bunce, M., Dixon, K., Wardell-Johnson, G., White, N.E., Nevill, P. 2020. Changes in soil microbial communities in post mine ecological restoration: implications for monitoring using high throughput DNA sequencing. *Science of the Total Environment.* https://doi.org/10.1016/j.scitotenv.2020.142262

**Author Contributions:** MvH conducted the study and wrote the manuscript. PN, MB, NW and GW-J were involved in the experimental design; samples were collected and processed by MvH; molecular and bioinformatics work was performed by MvH; all data were analysed and processed by MvH; statistical analysis was performed by MvH; the manuscript was edited by all authors

**Chapter 4:** van der Heyde, M**.,** Bunce, M., Bateman, P.W., Wardell-Johnson, G., White, N.E., Nevill, P. [Under review] Scat DNA provides important data for effective monitoring of bird and mammal diversity. *Biodiversity and Conservation*

**Author Contributions:** MvH conducted the study and wrote the manuscript. MvH, PN, MB, NW, and GW-J were involved in the experimental design. Samples were collected and processed by MvH; molecular and bioinformatics work was performed by MvH; all data was analyzed and processed by MvH; statistical analysis was done by MvH; the manuscript was edited by all authors.

**Chapter 5:** van der Heyde, M**.,** Bunce, M., Dixon, K., Wardell-Johnson, G., White, N.E., Nevill, P. Evaluating restoration trajectories using DNA metabarcoding of invertebrates and their associated plant communities. *In preparation*

**Author Contributions:** MvH conducted the study and wrote the manuscript. PN, MB, NW and GW-J were involved in the experimental design; samples were collected and processed by MvH; molecular and bioinformatics work was performed by MvH; all data were analysed and processed by MvH; statistical analysis was performed by MvH; the manuscript was edited by all authors

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| 16S V4 | Fourth hypervariable region of the 16S r RNA gene |
| ANOVA | Analysis of variance |
| ARC | Australian Research Council |
| ASV | Amplicon sequence variant |
| AUD | Australian Dollar |
| USD | United States Dollar |
| BLAST | Basic Local Alignment Search Tool |
| BOLD | Barcode of Life Database |
| BSA | Bovine Serum Albumin |
| bp | Base pairs |
| CMSR | Centre for Mine Site Restoration |
| CI | Confidence Interval |
| CIPRS | Curtin International Postgraduate Research Scholarship and Research Stipend Scholarship. |
| COI | Mitochondrial cytochrome c oxidase I |
| CT | Cycle threshold |
| dbRDA | Distance Based Redundancy Analysis |
| DNA | Deoxyribonucleic acid |
| DTPA | diethylene-triamine-penta-acetic acid |

| | |
|---|---|
| dNTPs | Deoxynucleotide triphosphates |
| eDNA | Environmental DNA |
| HSD | Honestly Significant Difference |
| HTS | High-throughput Sequencing |
| iDNA | Invertebrate derived DNA |
| IUCN | International Union for Conservation of Nature |
| JF | Jarrah Forest |
| MEGAN | Metagenome Analyzer |
| MgCl2 | Magnesium Chloride |
| MID | Multiplex identifier |
| N | North |
| NGS | Next Generation Sequencing |
| NMDS | Non-metric multidimensional scaling |
| OTU | Operational Taxonomic Unit |
| PB | Pilbara |
| PCR | Polymerase chain reaction |
| PERMANOVA | Permutational multivariate analysis of variance |
| qPCR | Quantitative PCR |
| RNA | Ribonucleic acid |
| rRNA | Ribosomal RNA |
| SCP | Swan Coastal Plain |
| SIMPROF | Similarity profile test |
| TGS | Third Generation Sequencing |
| TrEnD | Trace and Environmental DNA Laboratory |
| UN | United Nations |

| | |
|---|---|
| WA | Western Australia |
| ZOTU | Zero-radius operational taxonomic unit |

# CHAPTER 1

## INTRODUCTION

## 1.1 Environmental impact of mining

Mining in Australia began over 30,000 years ago when the Traditional Owners (First Nations of Australia) dug ochre for pigment (Paterson & Lampert, 1985). This activity expanded rapidly in the mid 1800's to a $205 billion industry today, which accounts for over 60% of Australia's exports and directly employs over 235,000 people (Resources and Energy Quarterly 2017). Mining is one of the many anthropogenic activities that have led to large scale land degradation and biodiversity losses (Chaudhary, Pfister, & Hellweg, 2016). While the footprint of mining is small compared to urbanization, agriculture and forestry (<1%) (Hodges, 1995), a majority (75%) of mines are in areas considered of high conservation value (Bridge, 2004; Miranda et al., 2003). Also, though small relative to other disturbances, the cumulative 'footprint' of mining is considerable. For example, in China alone the area of land degraded by mining is 3.7 million hectares, roughly equivalent to the size of Switzerland (Li et al. 2006). Landforming and restoring such large areas is costly (Menz, Dixon, & Hobbs, 2013), with a per hectare cost in Australia of up to AU$34,000 (Gardner & Bell, 2007).

Rehabilitation efforts have historically been inadequate; estimates put the number of abandoned mines in Australia, for example, at 50,000 (Unger et al. 2012). Abandoning mines may have been acceptable in the early days of the mining industry, but communities now expect that mine sites will be restored (Burton et al. 2012). As a result, mining companies are now legally responsible for the rehabilitation and/or restoration of mine sites in Australia (Commonwealth of Australia 2006; WA EPA 2006). Depending on the site-context restoration techniques can be applied including: landforming, to minimize erosion and promote revegetation; establishing plant growth medium, topsoil if available or various organic amendments; propagating plant species through seeds, and if necessary transplants (Australian Government, 2016). However, to date, very few sites have been confirmed as restored and officially closed (Campbell, Lindqvist, Browne, Swann, & Grudnoff, 2017).

Ecological restoration is the primary tool to combat losses of biodiversity and ecosystem services around the world (Cardinale et al., 2012; Mcdonald, Gann, Jonson, & Dixon, 2016). Ecological restoration has been defined as 'the process of

assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed' (Gann et al., 2019). The objective of restoration is to be on a pathway to a restored state (intact ecosystems); however, there can be abiotic and biotic constraints limiting efforts to replicate natural systems (EPA, 2006; Young et al., 2019). Rehabilitation is a return of an acceptable level of ecosystem functioning within the constraints of the site, which provide goods and ecosystem services (Burton et al., 2012; Cross et al., 2018).

## 1.2 Monitoring ecological restoration

Monitoring of restoration is conducted for several interconnected reasons. First, without monitoring, it is impossible to assess whether or not restoration efforts are successful (S. A. Thompson & Thompson, 2004). Evaluating restoration success helps researchers and practitioners determine the effectiveness of various remediation methods and amendments, improving future restoration outcomes (Collen & Nichols, 2012; Gann et al., 2019; Kupschus, Schratzberger, Righton, & Blanchard, 2016). This is important to maximize the efficacy of restoration methods and get the best 'bang' for your restoration buck. Second, monitoring can indicate whether further interventions may be necessary to reach the target, allowing the possibility of adaptive management (Murray & Marmorek, 2003; Thom, 2000). Adaptive management can help shorten the monitoring period by indicating when completion criteria are unlikely to be met and further remediation action is necessary (Murray & Marmorek, 2003; Thom, 2000). Third, restoration monitoring in a mining context must demonstrate to regulators that targets are being met. In Australia, mining proposals submitted to the Department of Mines and Petroleum must include Mine Closure Plans that detail completion criteria and plans for post closure monitoring (Government of Western Australia 1988). The minimum period of monitoring required is usually 10 years, but this period can be extended as long as completion criteria remain unmet. Monitoring must provide a high level of certainty that a tenement will meet the completion criteria before the tenement can be relinquished by the mining company (Government of Western Australia 2019).

Unfortunately, monitoring efforts are frequently limited in size and scope, and extensive monitoring programs are often prohibitively expensive (Thompson & Thompson, 2004). To date, assessments of restoration have emphasized flora recovery, while the recovery of animals, insects, and soil microbial communities is largely ignored (Cross, Bateman, & Cross, 2020; Cross, Tomlinson, Craig, Dixon, & Bateman, 2019) even though they can provide valuable indicators of success (Andersen & Sparling, 1997; Cross, Craig, Tomlinson, Dixon, & Bateman, 2020; Harris, 2003; Nichols & Nichols, 2003). For example, fauna are responsible for ecosystem functions such as pollination, seed dispersal, and nutrient cycling (Fleming & Muchhala, 2008; Herrera, 1995). They can act as facilitators of restoration through various plant-animal interactions such as seed dispersal and

seedling herbivory (Catterall, 2018). Restoring plant cover does not necessarily mean the restoration of other trophic levels (Cristescu et al. 2013) or ecosystem processes (Ruiz-jaen & Aide, 2005). Other little monitored groups include soil microbes, which form the functional basis for ecosystems and play key roles in nutrient cycling and decomposition (Meena et al., 2017), plant performance and community composition (Yang et al., 2018).

Currently, monitoring restoration involves comparisons between the restored areas and remnant or reference ecosystems (Gann et al., 2019; Gellie, Mills, Breed, & Lowe, 2017; Wallace, Laughlin, & Clarkson, 2017). The biotic components used to monitor ecosystem recovery must first demonstrate a measurable difference between restoration and reference ecosystem (Wallace et al., 2017). For successful restoration, the trajectories of these ecosystem properties are expected to converge towards the reference ecosystem (Gann et al., 2019; Hobbs & Harris, 2001; Hobbs & Norton, 1996; Suding, Gross, & Houseman, 2004). Wallace et al. (2017) describes successful restoration trajectories as having rapid, linear, lagged or threshold responses (Figure 1.1), while unsuccessful restoration trajectories fail to reach the level of reference ecosystems. The type of restoration trajectory will differ depending on the ecosystem property in question (Wallace et al., 2017). For example, richness of various communities may recovery rapidly, but community composition remains different between restored and reference sites for much longer (Andersen, Hoffmann, & Somes, 2003; Gellie et al., 2017).

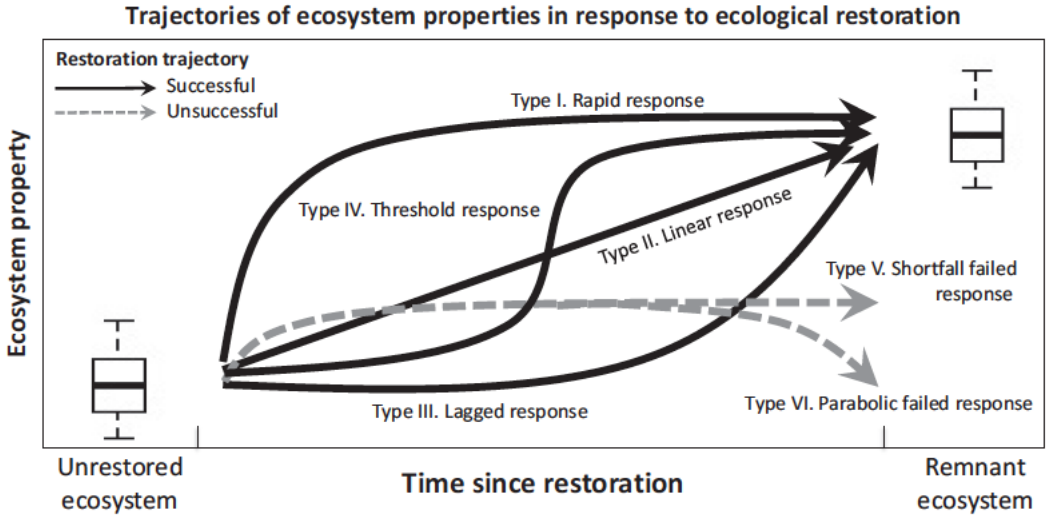


**Figure 1.1** **"**Conceptual diagram illustrating multiple possible trajectories of ecosystem properties to restoration efforts over time since restoration (Hobbs and Norton 1996, Hobbs and Harris 2001, Suding et al. 2004). The four solid arrows represent successful trajectories that a property may follow from an unrestored level to reach the target level of remnant ecosystems: (I) rapid response, (II) linear response, (III) lagged response, or (IV) threshold response. The response shape may depend on the nature of the property itself or could be affected by management actions. All solid response curves eventually reach values found in remnant ecosystems, but I, II, and IV display marked thresholds where the rate of change is drastically altered. The two dashed arrows represent unsuccessful trajectories where target levels are never attained: (V) shortfall failed response and (VI) parabolic failed response, where restoration efforts were only temporarily effective. This conceptual diagram displays hypothetical unrestored and remnant ecosystem means and distributions using boxplots (on the left and right, respectively). Unrestored ecosystem values are arbitrarily shown as low and remnant ecosystem values as high, but the inverse may be true depending on the ecosystem property." Reprinted from 'Exotic weeds and fluctuating microclimate can constrain native plant regeneration in urban forest restoration,' by Wallace et al. 2017, Ecological Applications, 27(4), pg 2. Reprinted with permission.

Methods of performing biodiversity assessments vary depending on the target taxa. While plants can be readily observed and assessed using a variety of plant survey method (Rochefort, Isselin-Nondedeu, Boudreau, & Poulin, 2013), accurate identification can depend on the season and age of the plant (Thompson & Newmaster, 2014). Animals may require traps, cameras, and invasive habitat searches (Environmental Protection Authority, 2010). Smaller organisms such as arthropods may be trapped and removed for identification under a microscope, and some soil microorganisms may be identified using a combination of microscopy and

culturing (Kirk et al., 2004). Attempting to monitor multiple taxa may therefore require multiple methods and hundreds of identification hours. These types of field based visual surveys are often time-consuming and expensive (Thompson & Newmaster, 2014; Thomsen & Willerslev, 2015) and not standardized across projects or observers (Milberg, Bergstedt, Fridman, Odell, & Westerberg, 2008). As a result, there is a need to develop tools that can expand our monitoring capabilities. One of the tools that shows a lot of potential in the field of biodiversity assessment is eDNA metabarcoding.

## 1.3 eDNA Metabarcoding

### 1.3.1 Recent advances sequencing technologies

When the Human Genome project began in the late 1980s it stimulated innovation and the development of new sequencing technologies that reduced the cost of sequencing at an unprecedented rate (Hagen 2014). It took over a decade to sequence a complete human genome, but the introduction of high-throughput (HTS) or 'Next Generation' sequencing (NGS) allowed millions of sequencing reads to be generated in parallel, making it possible to sequence an entire human genome in a single day (Behjati & Tarpey, 2013). This new accessibility of sequencing technologies paved the way for new applications, developing the field of metagenomics: using sequencing to determine the identity of organisms in a sample. Metabarcoding is a type of metagenomics that involves the identification of organisms in a mixed sample (more than one organism present) through the use of certain 'barcoding' regions (Taberlet, Coissac, Hajibabaei, & Rieseberg, 2012). Using this tool, DNA from environmental sources can be used for biodiversity assessment (Taberlet, Coissac, et al., 2012)

### 1.3.2 Environmental DNA

Environmental DNA (eDNA) is DNA that is collected from an environmental sample, such as soil or water, rather than sampling from individual organisms. Organisms shed DNA into the environment through defecation, shedding, skin cells, mucus, saliva, etc., and this DNA can be extracted and sequenced from a sample to provide taxonomic identification of organisms (Bohmann et al., 2014; Taberlet, Coissac, et al., 2012). DNA can also be extracted from 'bulk' samples. These are samples that contain pieces of organisms, for example bulk insect samples in which insects in a trap are homogenized together and the DNA is extracted from the homogenate (Yu et al., 2012). Traditionally 'bulk' and 'environmental' samples have been considered as somewhat different categories, although the terminology in this field is growing and changing rapidly along with the potential sample types. For example, the term 'iDNA' has been proposed for invertebrate derived DNA, that when invertebrate samples are used to extract non invertebrate DNA (e.g. vertebrates) (Schnell et al., 2015).

### 1.3.3 Metabarcoding workflow

In a metabarcoding analyses samples are collected (e.g. soil, feces, insects, etc.), DNA is then extracted from a sample, a short 'barcoding' gene region is amplified using Polymerase Chain Reaction (PCR). The gene region targeted is sequenced and then 'identified' or queried against a reference biological database. This approach is cost-effective, and produces comprehensive datasets more quickly than standard monitoring methods (Ji et al., 2013). Additionally, it allows collection of information on groups, such as soil microbes, that are otherwise impossible to monitor.

#### *1.3.3.1 Sample collection*

The source of DNA for metabarcoding (sample material) greatly influences what organisms can be detected. Water, soil, air and faeces contain traces of eDNA that can be extracted, sequenced and processed to use for biological monitoring (Bohmann et al., 2014). For targeted monitoring (i.e. diet of a predator) sample choice is relatively simple (faeces of the predator). However, for broad biodiversity assessment sampling is more complicated. Most metabarcoding studies have been conducted aquatic systems (reviewed in Thomsen & Willerslev, 2015), where species can be detected from DNA floating in water (Ficetola et al. 2008; Goldberg et al. 2011; Jerde et al. 2011), or in aquatic sediments (Turner et al. 2015). There have been several reviews on all aspects of aquatic eDNA metabarcoding, including number and size of samples (Machler et al. 2016; Rees et al. 2014), different methods for capturing and extracting eDNA (Deiner et al. 2015; Eichmiller et al. 2016; Minamoto et al. 2016), persistence of eDNA (Barnes et al., 2014; Díaz-Ferguson & Moyer, 2014; Rees et al., 2014) choice of molecular makers (Freeland, 2017) and more (see Freeland 2017). However, terrestrial metabarcoding is considerably less developed.

#### *1.3.3.2 DNA extraction*

Sample processing depends on the type of sample, but for DNA extraction, it is common to use a small amount (~100mg-1g) of homogenized sample. The DNA is then extracted using various commercial kits, often with modifictaions, such as the DNeasy Blood and Tissue kit (Qiagen) (Beng et al., 2016), the PowerSoil kit (Previously MoBio Laboratories, Carlsbad, California now Qiagen) (Drummond et

al., 2015) or similar. Although it is possible to extract from greater volumes (up to 10g) using specialized, and more expensive, kits like the PowerMax Soil kit (Qiagen) (Yoccoz et al., 2012) there are also new innovations for capturing extracellular DNA from large amounts of starting material more cheaply using saturated sodium phosphate buffers (Taberlet, Homme, et al., 2012). It is recommended to perform any extraction in sterile conditions in a laboratory physically separate from any PCR products to prevent cross contamination (Goldberg et al., 2016).

#### *1.3.3.3 DNA amplification*

Once the DNA has been extracted, the next step is the amplification of target barcoding region using PCR. A pair of primers, short DNA fragments that anneal and amplify target DNA, is typically called an assay. An assay can be used to target a species or a group of organisms (Epp et al., 2012; Vamos, Elbrecht, & Leese, 2017). There is no universal barcode (i.e. one gene region for all biodiversity) that contains the resolution power to identify every source of DNA in a sample. Primers are chosen depending on the target organisms to amplify barcodes that are theoretically similar within species but contain enough variation to separate different species (Cannon et al., 2016; Riaz et al., 2011). Assays are continually being developed, refined and tested. In selecting markers, there are trade-offs between the size, breadth, resolution of barcodes, and the availability of reference sequences for taxonomic identification. Larger barcodes provide greater taxonomic resolution but are likely degraded in environmental DNA. For example, Lahaye et al. (2008) recommend *mat*K as a universal barcode for plants, while Fahner et al. (2016) recommend using *rbc*L and ITS2, partly because of existing databases for taxonomic identification. The P6 loop of the trnL intron is suggested for plants because while it has low resolution, it is short and highly length variable (10-143bp), and more likely to be found in degraded environmental DNA than longer barcodes (Taberlet et al. 2007). Barcodes also have certain taxa that they cannot detect reliably, creating biases in the dataset (Clarke et al. 2014; Deagle et al. 2014). Ultimately, using multiple primer sets to amplify multiple barcodes will help control the biases of any one barcode on its own.

A typical PCR reaction includes an aqueous buffer, deoxynucleoside triphosphates (dNTPs) which consist of the four basic nucleotides (adenine, thymine,

cytosine and guanine), a thermostable DNA polymerase, a pair of primers, and the template DNA extracted from the sample. At high temperatures the double stranded template DNA is denatured (separates), then the temperature is reduced so the primers can anneal and the polymerase extends the primer using the dNTPs and copying the template DNA. After 35-40 cycles of these three steps (denaturing, annealing, extending), this process can make millions of copies of the target DNA, the sequence between the two primers also referred to as an 'amplicon'. PCR amplification typically plateaus around 35-40 cycles; however, certain molecules such as tannins, can inhibit amplification and result in biased sequencing results (Murray, Coghlan, & Bunce, 2015). Using quantitative PCR on a Real-Time PCR instrument also allows users to screen for inhibition by amplifying various dilutions of the DNA template (Murray et al., 2015). They can then select the working dilution with the greatest DNA copy number that showed uninhibited amplification, resulting in more accurate relative sequencing in terms of relative sequence abundance and richness (Murray et al., 2015).

In order to allow the simultaneous sequencing of multiple samples, samples are multiplexed by adding unique "tags" or "indexes" to the sequences from each sample. PCR is used to add a single index to the sequencing adaptors, but double-indexing can reduce both the rates of false assignments and the costs (Kircher, Sawyer, & Meyer, 2012). These samples can then be pooled together and sequenced in parallel using NGS, most commonly on Illumina instruments which dominate the short-read sequencing industry and can sequence 100-150 samples in a single run (Goodwin, McPherson, & McCombie, 2016).

#### *1.3.3.4 Data Processing post sequencing*

Post-sequencing, there are various software's and programs (DADA2 Callahan et al., 2016; QIIME Caporaso et al., 2010; USEARCH Edgar and Flyvbjerg, 2015) to perform the bioinformatics necessary to go from raw sequences to ecologically informative data, and these are constantly evolving. Typically sequences are clustered together in operational taxonomic units (OTUs). The 97% threshold for OTUs was initially proposed for the 16S ribosomal RNA gene to characterize bacteria in 1994 and is considered to represent approximately species-level diversity (Stackebrant & Goebel, 1994). This assumption has been repeatedly

challenged (Edgar, 2018), resulting in increased adoption of Amplicon sequence variants (ASVs) or zero-radius OTUs (ZOTUS) (Callahan et al., 2016; Edgar, 2016), which are similar to OTUs clustered at 100%.

### 1.3.4 Benefits of metabarcoding

Perhaps the greatest benefit of using high throughput sequencing in a restoration context is the increased breadth of biodiversity that can be monitored. Restoration monitoring traditionally has a botanical bias (Cross et al. 2018) because it is assumed that fauna will return with the restoration of plant communities. This is not always the case and recently there has been increased interest in broader biodiversity monitoring (Cross et al., 2019; Majer, 2009). Environmental DNA metabarcoding can be used to monitor vertebrates (Andersen et al., 2012), invertebrates (Yu et al., 2012), plants (Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012), and soil microbial communities (Creer et al., 2016), and is effective at identifying cryptic or undescribed taxa that are difficult to detect using standard methods (Biggs et al., 2015; Yoccoz et al., 2012). Several review articles have argued that eDNA metabarcoding should be used for biodiversity monitoring (Deiner, Bik, Elvira, et al., 2017; Hajibabaei, Baird, Fahner, Beiko, & Golding, 2016; Holdaway et al., 2017), which could supplement baseline surveys (Garris, Baldwin, Van Hamme, Gardner, & Fraser, 2016), and increase the taxonomic breadth of post-closure monitoring (Fernandes et al., 2018). This is not to say that sequencing technology is without its challenges (see Deiner et al. 2017), but it has great potential and many of the limitations involved may be addressed in the near future as the discipline grows.

DNA metabarcoding is dramatically changing our approach to biodiversity assessment. One of the benefits of this technique is the potential cost-effectiveness. Estimates of costs for metabarcoding vary depending on the technology and sample type from a few dollars (de Mattia et al., 2012) to a few hundred dollars (Ji et al., 2013) per sample. Because the costs tend to increase per sample rather than per specimen, metabarcoding would be most cost-effective in highly biodiverse systems, and targeting highly diverse taxa such as arthropods (Ji et al., 2013). As the technology improves, costs are projected to decrease allowing the use of more frequent sampling, which can be used to track changes in restoration and identify

trajectories in the biological changes throughout the restoration process. Depending on the type of sample material used, eDNA can also be a non-invasive tool detecting diversity from traces of the organism rather than direct handling (Deiner, Bik, Mächler, et al., 2017). This method is also less reliant on specialized taxonomic expertise, which is dwindling worldwide (Pearson, Hamilton, & Erwin, 2011). Finally, metabarcoding data can be audited by regulators and third-parties (Yu et al., 2012), which can help minimize the effect of observer bias (Milberg et al., 2008). Unlike standard morphological methods, with metabarcoding the raw sequence files that provide all the biological data can be stored and re-analyzed if there is doubt on the validity of a survey.

## 1.4 Methodological testing

### 1.4.1 Selection of appropriate eDNA metabarcoding substrates

There is limited research on appropriate sampling methods for terrestrial biodiversity assessment using metabarcoding. Researchers have sampled tropical arthropods, and used metabarcoding to characterize communities faster and cheaper than traditional taxonomic identification (Beng et al., 2016; Gibson, Shokralla, Porter, King, van Konynenburg, et al., 2014; Ji et al., 2013; Yu et al., 2012). Metabarcoding has also been applied to feces for diet analysis (De Barba et al., 2014; Kowalczyk et al., 2011; Rayé et al., 2011; Valentini et al., 2009), which may also give indicators of biodiversity using the animal as an environmental sampler. Previous studies have also sampled soil and successfully sequenced vertebrates (K. Andersen et al., 2012), earthworms (Bienert et al., 2012; Taberlet et al., 2012) and plants (Fahner et al. 2016; Taberlet et al., 2012; Yoccoz et al., 2012). Ultimately, there are many possible substrate types. Hence the most appropriate substrate to monitor restoration of mine sites remains unknown.

In addition to the detectable diversity, for accurate biodiversity assessment, there must be some certainty that the organisms detected are present in the area. In soil DNA can be preserved for thousands of years under the right (i.e. very cold) conditions (Epp et al., 2012; Pedersen et al., 2014). In surface soils, DNA has been shown to remain detectable up to 77 days (Widmer et al. 1997) or even 6 years (Andersen et al. 2012) after the organisms' removal in Denmark. This presents a problem for mine site monitoring, where the aim is to determine current biodiversity. Sampling only the soil surface may lessen the risk of detecting past diversity, as the DNA is leached (Andersen et al. 2012) or degraded (Lindahl, 1993) over time. The temperatures in Australia are much higher than may be expected in Denmark (Andersen et al., 2012) or Oregon (Widmer, Seidler, Donegan, & Reed, 1997), and DNA persistence would likely be lower as a result. In addition to temperature, soil chemistry and texture can influence the distribution and persistence of eDNA (K. Andersen et al., 2012; Levy-booth et al., 2007). For example, in sandy soil, DNA tends to leach into the soil profile, while on clay soil DNA is limited to the surface 2 cm (Andersen et al. 2012). Assessing DNA of soil samples from hotter climates will

determine if they can be useful for biodiversity surveys, or if the DNA is degraded too quickly.

### 1.4.2 Established targets of eDNA metabarcoding applied to monitor mine site restoration

In terrestrial systems, metabarcoding has mostly been used to describe microbial communities (Hartmann et al. 2015; Lesaulnier et al., 2008; Schmidt et al., 2013). These communities are relatively simple to sample as they are easily extracted from soil. Most microorganisms are unculturable (Kim & Pham, 2012) which means they are difficult to monitor without molecular techniques such as eDNA metabarcoding. Microbial communities are important in soil formation (Schulz et al., 2013) and nutrient cycling (Chen, Zhu, & Zhang, 2003), and can be used to assess degradation and restoration success (Harris, 2003). They have high growth rates and respond quickly to environmental changes such as soil chemistry (Leff et al., 2015; Šmejkalová, Mikanová, & Borůvka, 2003), physical soil disturbance (Dong et al., 2017; Kabiri, Raiesi, & Ghazavi, 2016), and plant communities (Burns et al. 2015). As such they may be good indicators of restoration success. Previous studies have suggested that SMC may be used to monitor revegetation (Gellie, Mills, Breed, & Lowe, 2017; Yan et al., 2019; Yan et al., 2018), but these were conducted at only one location. To assess if high throughput sequencing of SMC is appropriate for restoration monitoring, it is important to test for consistency across multiple locations (see Section 1.4.3).

### 1.4.3 Wide scale testing of novel substrates

When testing novel substrates, or using a particular substrate in a novel context, wide scale testing is important to determine if the substrate and organisms detected are limited to certain environments or more broadly applicable. Climate variables such as UV radiation, temperature, and rainfall can affect DNA degradation rates (Barnes et al., 2014; Levy-booth et al., 2007; Sirois & Buckley, 2019). Environmental DNA will not be useful for current biodiversity assessment in either extremes: in cold dark conditions eDNA metabarcoding may pick up ancient DNA from extinct animals (Pedersen et al., 2014; Willerslev et al., 2003), while in hot areas with high UV radiation the degradation rates may be too high to detect

diversity. In addition to DNA degradation rates, variability in ecosystems may affect the availability of a particular substrate. For example, leeches can been used to generate vertebrate diversity (Schnell et al., 2015), but finding leeches is dependent on the presence of suitable wetland habitat (Kasparek et al., 2000). For restoration monitoring, substrates need to be available within restoration site, and able to detect organisms that can show differences between restored and reference sites. Testing and applying substrates to multiple ecosystems can determine if patterns of community recovery are consistent, or if other environmental conditions need to be considered.

### 1.4.4 Applying eDNA metabarcoding to established indicators of restoration success with poor reference databases

Invertebrates are established indicators of restoration success (Andersen, Hoffmann, Müller, & Griffiths, 2002; Andersen & Sparling, 1997; Majer, 2009) and ecosystem function (Folgarait, 1998; Rosenberg, Danks, & Lehmkuhl, 1986). They play an important roles in soil formation, nutrient cycling, seed dispersal and pollination (Bronstein, Alarcón, & Geber, 2006; Catterall, 2018; Hunter, 2001; Majer, 1989, 1997) and provide food sources for higher trophic organisms (S. L. Cross, Craig, Tomlinson, & Bateman, 2020). Invertebrates are numerous, diverse and easy to capture (Gaston, 1991; Yeates, Harvey, & Austin, 2003), but expensive and time consuming to identify because of their high diversity and the lack of specialized taxonomic expertise (Obrist & Duelli, 2010; Whitehead, 1990). Many species are difficult to identify or are are still undescribed, especially in Australia which has a high degree of endemism because of its geologic history and isolation (Mummery & Hardy, 1994; Yeates et al., 2003). eDNA metabarcoding has been shown to reliably identify mixed samples of invertebrates faster and cheaper than morphological identification (Ji et al., 2013; Yu et al., 2012). As a result, eDNA metabarcoding of invertebrate communities may represent a rapid, cost-effective survey method to monitor an important aspect of ecosystem functioning.

One of the most important challenge in the use of metabarcoding is the poor state of reference sequence databases. Taxonomic reference databases contain sequence data from taxonomically identified specimens, and are key to identifying

the metabarcoding sequences. Initiatives such as the Barcode of Life Data Systems (BOLD) are aimed at producing high quality reference libraries, improving on databases like GenBank by having permanent voucher specimens, minimum sequence length of 500 base pairs, and limit to certain barcoding regions (Ratnasingham & Hebert, 2007). At present, BOLD accepts only the cytochrome c oxidase subunit I (COI) gene for fauna barcoding, though the lack of conserved regions makes this gene unsuitable for a lot of amplicon based metabarcoding (Deagle et al. 2014). Without conserved regions within the gene, creating primers to amplify smaller sections is unreliable and leads to biases in the amplified taxa which affect the biodiversity estimates (Yu et al. 2012; Deagle et al. 2014). For geographic regions where there is little development of barcodes for species, there is a limiting factor present in species level identification (Gibson, Shokralla, Porter, King, Konynenburg, et al., 2014). In Australia, up to 75% of invertebrate fauna remains undescribed, so species level identification is not within the means of even morphological identification (Austin et al., 2004; Majer, Brennan, & Moir, 2007). There is no doubt that reference databases will grow, and more barcoding regions will likely be developed, but in the meantime it may be useful to determine if monitoring can also be conducted without verified taxonomic identification using the sequence variants.

## 1.5 Thesis overview

The primary question of this thesis is; “Is eDNA metabarcoding a viable tool to monitor mine site restoration”. This will be done by addressing the specific questions and aims across four data chapters (Figure 1.2). The data chapters in this thesis are written and formatted as four journal articles and include two articles that are published (Chapters 2 and 3), one that is currently under review (Chapter 4), and one in preparation (Chapter 5). Each chapter will have an abstract, introduction, methods, results, discussion, conclusion, and references section. For clarity and continuity between chapters, each chapter begins with a preface to give readers a brief description of the purpose of the chapter within the context of the thesis. For consistency, formatting of published papers has been modified to enable referencing between chapters. The final discussion (Chapter 6) synthesizes the main findings of the thesis and the critical study design considerations for terrestrial eDNA surveys. This chapter also considers the limitations and future work necessary to develop eDNA metabarcoding as a valid monitoring tool.

In Chapter 2, I test multiple substrates at two location in Western Australia to see what taxa can be generated from each substrate and where they overlap. The aim of this chapter is to assess the potential of each substrate for use in terrestrial biodiversity surveys, and it is the first study comparing multiple terrestrial substrates using the same eDNA assays. The results of this chapter were used to select appropriate substrates for the following chapters where eDNA metabarcoding was applied to mine site restoration monitoring.

In Chapter 3, I investigate the recovery of soil microbial communities following mine site restoration. SMC are a well-established target of eDNA metabarcoding, and researchers have applied eDNA sequencing to characterize microbial communities for over a decade. However, they rarely examine changes in mine site restoration and those that to typically do not; study both bacteria and fungi, use functional annotation to make inferences, or assess more than one location. This chapter examines the community and functional changes in both bacteria and fungi across three locations in Western Australia to determine their potential for monitoring mine site restoration.

In chapter 4, I trial a novel methodology for assessing vertebrate diversity using bulk scat samples. This methodology is tested in three locations in Western Australia to investigate if environmental conditions affect the suitability of this method for monitoring. I sampled sites of mine site restoration and reference to determine if scat collection could distinguish fine scale differences between restoration and reference sites.

In Chapter 5, I investigate the return of invertebrate communities, well-known indicators of ecological restoration. This chapter also functions as an example of monitoring using eDNA metabarcoding where reference databases are depauperate. Here I examine ground dwelling and airborne invertebrates to determine if dispersal ability affects the patterns of community recovery. I also test whether invertebrate recovery (increasing community similarity to reference community) is consistent across different ecosystems. Lastly, this chapter uses a plant assay on invertebrate samples to examine the interactions between the invertebrates and plant community, providing indication of function by determining if invertebrates are using restoration sites or simply passing through.

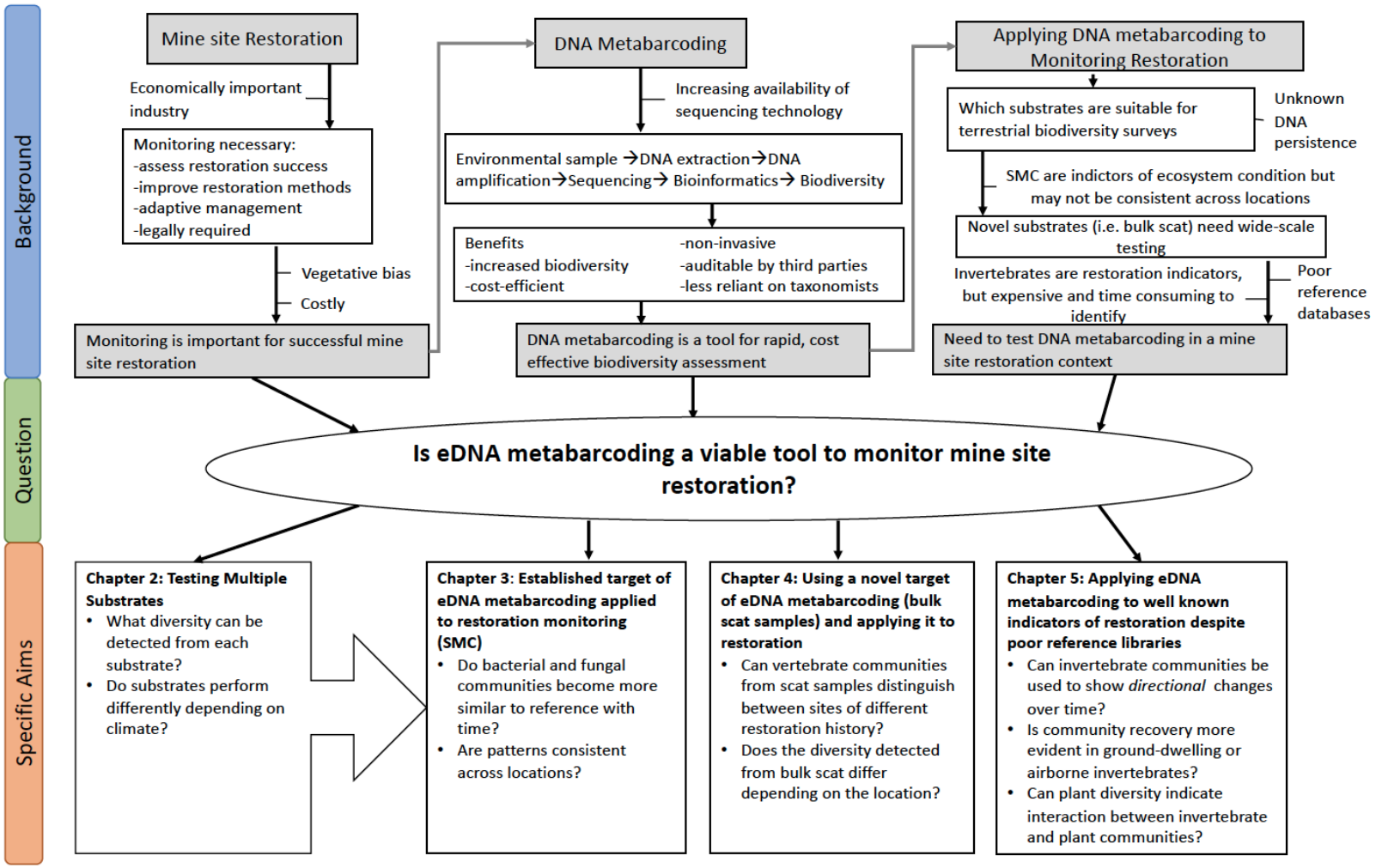


**Figure 1.2** Conceptual framework of thesis structure**.**

## 1.6 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.


Andersen, A. N., Hoffmann, B. D., Müller, W. J., & Griffiths, A. D. (2002). Using ants as bioindicators in land management: Simplifying assessment of ant community responses. *Journal of Applied Ecology*, *39*(1), 8–17. doi: 10.1046/j.1365-2664.2002.00704.x

Andersen, A. N., Hoffmann, B. D., & Somes, J. (2003). Ants as indicators of minesite restoration: Community recovery at one of eight rehabilitation sites in central Queensland. *Ecological Management and Restoration*, *4*, S12–S19. doi: 10.1046/j.1442-8903.4.s.2.x

Andersen, A. N., & Sparling, G. P. (1997). Ants as Indicators of Restoration Success : Relationship with Soil Microbial Biomass in the Australian Seasonal Tropics. *Society for Ecological Restoration*, *5*(2), 109–114.

Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjær, K. H., … Willerslev, E. (2012). Meta-barcoding of "dirt" DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, *21*(8), 1966–1979. doi: 10.1111/j.1365-294X.2011.05261.x

Austin, A. D., Yeates, D. K., Cassis, G., Fletcher, M. J., La Salle, J., Lawrence, J. F., … Taylor, G. S. (2004). Insects "Down Under" - Diversity, endemism and evolution of the Australian insect fauna: Examples from select orders. *Australian Journal of Entomology*, *43*(3), 216–234. doi: 10.1111/j.1326-6756.2004.00448.x

Australian Government. (2016). Rehabilitation Implementation. In *Mining Rehabilitation - Leading Practice Sustainable Development Program for the Mining Industry* (pp. 27–44). doi: 10.1002/9781119389668.ch22

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science and Technology*, *48*(3), 1819–1827. doi: 10.1021/es404734p

Behjati, S., & Tarpey, P. S. (2013). What is next generation sequencing? *Archives of Disease in Childhood: Education and Practice Edition*, *98*(6), 236–238. doi: 10.1136/archdischild-2013-304340

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

Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J.-J., & Taberlet, P. (2012). Tracking earthworm communities from soil DNA. *Molecular Ecology*, *21*, 2017–2030. doi: 10.1111/j.1365-294X.2011.05407.x

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., … Dunn, F. (2015). Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt ( Triturus cristatus ). *Biological Conservation*, *183*, 19–28. doi: 10.1016/j.biocon.2014.11.029

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., … de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, *29*(6), 358–367. doi: 10.1016/j.tree.2014.04.003

Bridge, G. (2004). Contested terrain: mining and the environment. *Annual Review of Environment and Resources*, *29*(1), 205–259. doi: 10.1146/annurev.energy.28.011503.163434

Bronstein, J. L., Alarcón, R., & Geber, M. (2006). The evolution of plant-insect mutualisms. *New Phytologist*, *172*(3), 412–428. doi: 10.1111/j.1469-8137.2006.01864.x

Burton, M., Jasmine Zahedi, S., & White, B. (2012). Public preferences for timeliness and quality of mine site rehabilitation. The case of bauxite mining in Western Australia. *Resources Policy*. doi: 10.1016/j.resourpol.2012.01.004

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. doi: 10.1038/nmeth.3869

Campbell, R., Lindqvist, J., Browne, B., Swann, T., & Grudnoff, M. (2017). The Dark Side of the Boom. *The Australia Institute, Canberra*. doi: 10.1093/acprof:oso/9780190215132.003.0002

Cannon, M. V., Hester, J., Shalkhauser, A., Chan, E. R., Logue, K., Small, S. T., & Serre, D. (2016). In silico assessment of primers for eDNA studies using PrimerTree and application to characterize the biodiversity surrounding the Cuyahoga River. *Scientific Reports*, *6*(1), 22908. doi: 10.1038/srep22908

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Publishing Group*, *7*(5), 335–336. doi: 10.1038/nmeth0510-335

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., … Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67. doi: 10.1038/nature11148

Catterall, C. P. (2018). Fauna as passengers and drivers in vegetation restoration: A synthesis of processes and evidence. *Ecological Management and Restoration*, *19*, 54–62. doi: 10.1111/emr.12306

Chaudhary, A., Pfister, S., & Hellweg, S. (2016). Spatially Explicit Analysis of Biodiversity Loss Due to Global Agriculture, Pasture and Forest Land Use from a Producer and Consumer Perspective. *Environmental Science and Technology*, *50*(7), 3928–3936. doi: 10.1021/acs.est.5b06153

Chen, G., Zhu, H., & Zhang, Y. (2003). Soil microbial activities and carbon and nitrogen fixation. *Microbiology*, *154*, 393–398. doi: 10.1016/S0923-2508(03)00082-2

Clarke, L. J., Soubrier, J., Weyrich, L. S., & Cooper, A. (2014). Environmental metabarcodes for insects: In silico PCR reveals potential for taxonomic bias. *Molecular Ecology Resources*, *14*(6). doi: 10.1111/1755-0998.12265

Collen, B., & Nichols, E. (2012). Taking the measure of change. *Science*, *346*(6206), 166–167. doi: 10.1126/science.337.6092.285

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., … Bik, H. M. (2016). The ecologist ' s field guide to sequence-based identification of

biodiversity. *Methods in Ecology Adn Evolution*, *7*, 1008–1018. doi: 10.1111/2041-210X.12574

Cristescu, R. H., Rhodes, J., Frére, C., & Banks, P. B. (2013). Is restoring flora the same as restoring fauna? Lessons learned from koalas and mining rehabilitation. *Journal of Applied Ecology*, *50*(2), 423–431. doi: 10.1111/1365-2664.12046

Cross, A. T., Young, R., Nevill, P., McDonald, T., Prach, K., Aronson, J., … Dixon, K. W. (2018). Appropriate aspirations for effective post-mining restoration and rehabilitation: a response to Kaźmierczak et al. *Environmental Earth Sciences*, *77*(6), 1–6. doi: 10.1007/s12665-018-7437-z

Cross, S. L., Craig, M. D., Tomlinson, S., & Bateman, P. W. (2020). I don't like crickets, I love them: invertebrates are an important prey source for varanid lizards. *Journal of Zoology*, *310*(4), 323–333. doi: 10.1111/jzo.12750

Cross, Sophie 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 and Restoration*, *21*(1), 4–8. doi: 10.1111/emr.12393

Cross, Sophie L., Craig, M. D., Tomlinson, S., Dixon, K. W., & Bateman, P. W. (2020). Using monitors to monitor ecological restoration: Presence may not indicate persistence. *Austral Ecology*, 1–12. doi: 10.1111/aec.12905

Cross, Sophie L., Tomlinson, S., Craig, M. D., Dixon, K. W., & Bateman, P. W. (2019). Overlooked and undervalued: The neglected role of fauna and a global bias in ecological restoration assessments. *Pacific Conservation Biology*, *25*(4), 331–341. doi: 10.1071/PC18079

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources*, *14*(2), 306–323. doi: 10.1111/1755-0998.12188

de Mattia, F. D. E., Gentili, R., Bruni, I., Galimberti, A., Sgorbati, S., Casiraghi, M., & Labra, M. (2012). A multi-marker DNA barcoding approach to save time and resources in vegetation surveys. *Botanical Journal of Linnean Society*,

*169*, 518–529.

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), 1789–1793. doi: 10.1098/rsbl.2014.0562

Deiner, K., Bik, H. M., Elvira, M., Seymour, M., Creer, S., Bista, I., … Vere, N. De. (2017). Environmental DNA metabarcoding : Transforming how we survey animal and plant communities. *Molecular Ecology*, (September), 1–24. doi: 10.1111/mec.14350

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, Vol. 26. doi: 10.1111/mec.14350

Deiner, K., Walser, J. C., Mächler, E., & Altermatt, F. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biological Conservation*, *183*, 53–63. doi: 10.1016/j.biocon.2014.11.018

Díaz-Ferguson, E. E., & Moyer, G. R. (2014). History, applications, methodological issues and perspectives for the use environmental DNA (eDNA) in marine and freshwater environments. *Revista de Biología Tropical*, *62*(4), 1273–1284. doi: 10.15517/rbt.v62i4.13231

Dong, W., Liu, E., Yan, C., Tian, J., Zhang, H., & Zhang, Y. (2017). Impact of no tillage vs. conventional tillage on the soil bacterial community structure in a winter wheat cropping succession in northern China. *European Journal of Soil Biology*, *80*, 35–42. doi: 10.1016/j.ejsobi.2017.03.001

Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C. M., … Nelson, N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience*, *4*(46). doi: 10.1186/s13742-015-0086-1

Edgar, R. C. (2016). UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, 081257. doi: 10.1101/081257

Edgar, R. C. (2018). Updating the 97% identity threshold for 16S ribosomal RNA

OTUs. *Bioinformatics*, *34*(14), 2371–2375. doi: 10.1093/bioinformatics/bty113

Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476–3482. doi: 10.1093/bioinformatics/btv401

Eichmiller, J. J., Miller, L. M., & Sorensen, P. W. (2016). Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. *Molecular Ecology*, *16*, 56–68. doi: 10.1111/1755-0998.12421

Environmental Protection Authority and Department of Environment and Conservation. (2010). Technical Guidance - Sampling Methods for Terrestrial Vertebrate Fauna (eds B.M. Hyder, J. Dell and M.A Cowan). *Perth, Western Australia*.

EPA. (2006). Guidance Statement No. 6: Guidance for the assessment of environmental factors: rehabilitation of terrestrial ecosystems, Environment Protection Authority, Perth.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., … Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology*, *21*(8), 1821–1833. doi: 10.1111/j.1365-294X.2012.05537.x

Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-scale monitoring of plants through environmental DNA metabarcoding of soil: Recovery, resolution, and annotation of four DNA markers. *PLoS ONE*, *11*(6), 1–16. doi: 10.1371/journal.pone.0157505

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

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, *4*(4), 423–425. doi: 10.1098/rsbl.2008.0118

Fleming, T. H., & Muchhala, N. (2008). Nectar-feeding bird and bat niches in two

worlds: pantropical comparisons of vertebrate pollination systems. *Journal of Biogeography*, *35*(5), 764–780. doi: 10.1111/j.1365-2699.2007.01833.x

Folgarait, P. J. (1998). Ant biodiversity and its relationship to ecosystem functioning : a review. *Biodiversity and Conservation, 7,* 1224-1244.

Freeland, J. R. (2017). The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA. *Genome*, *60*(4), 358–374.

Gann, G. D., McDonald, T., Walder, B., Aronson, J., Nelson, C. R., Jonson, J., … Dixon, K. W. (2019). International principles and standards for the practice of ecological restoration. Second edition. *Restoration Ecology*, *27*(S1), S1–S46. doi: 10.1111/rec.13035

Gardner, J. H., & Bell, D. T. (2007). Bauxite mining restoration by Alcoa World Alumina Australia in Western Australia: Social, political, historical, and environmental contexts. *Restoration Ecology*, *15*(SUPPL. 4), 3–10. doi: 10.1111/j.1526-100X.2007.00287.x

Garris, H. W., Baldwin, S. A., Van Hamme, J. D., Gardner, W. C., & Fraser, L. H. (2016). Genomics to assist mine reclamation: A review. *Restoration Ecology*, *24*(2), 165–173. doi: 10.1111/rec.12322

Gaston, K. J. (1991). The Magnitude of Global Insect Species Richness. *Conservation Biology*, *5*(3), 283–296. doi: 10.1111/j.1523-1739.1991.tb00140.x

Gellie, N. J. C., Mills, J. G., Breed, M. F., & Lowe, A. J. (2017). Revegetation rewilds the soil bacterial microbiome of an old field. *Molecular Ecology*, *26*(11), 2895–2904. doi: 10.1111/mec.14081

Gibson, J., Shokralla, S., Porter, T. M., King, I., Konynenburg, S. Van, & Janzen, D. H. (2014). *Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics*. *111*(22), 8007–8012. doi: 10.1073/pnas.1406468111

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). doi: 10.1073/pnas.1406468111

Goldberg, C. S., Pilliod, D. S., Arkle, R. S., & Waits, L. P. (2011). Molecular detection of vertebrates in stream water: A demonstration using rocky mountain tailed frogs and Idaho giant salamanders. *PLoS ONE*, *6*(7). doi: 10.1371/journal.pone.0022746

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., … Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, *7*(11), 1299–1307. doi: 10.1111/2041-210X.12595

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351. doi: 10.1038/nrg.2016.49

Government of Western Australia Department of Mines Industry Regulation and safety. (2019). *Guidance - How to Prepare a Mine Closure Plan in accordance with Part 1 of Guidelines for Mine Closure Plans in Western Australia*. (September).

Government of Western Australia Department of Mines Industry Regulation and Safety. (1988). *Guidelines for Mine Closure Plans*. (March), 3–5.

Hajibabaei, M., Baird, D. J., Fahner, N. A., Beiko, R., & Golding, G. B. (2016). A new way to contemplate Darwin's tangled bank: how DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *371*, 20150330. doi: 10.1098/rstb.2015.0330

Harris, J. A. (2003). Measurements of the soil microbial community for estimating the success of restoration. *European Journal of Soil Science*, *54*(4), 801–808. doi: 10.1046/j.1351-0754.2003.0559.x

Hartmann, M., Frey, B., Mayer, J., & Ma, P. (2015). Distinct soil microbial diversity under long-term organic and conventional farming. *ISME*, *9*, 1177–1194. doi: 10.1038/ismej.2014.210

Herrera, C. M. (1995). Plant-vertebrate seed dispersal systems in the Mediterranean:

Ecological, evolutionary, and historical determinants. *Annual Review of Ecology and Systematics*, *26*(1), 705–727. doi: 10.1146/annurev.es.26.110195.003421

Hobbs, R. J., & Harris, J. A. (2001). Restoration ecology: Repairing the earth's ecosystems in the new millennium. *Restoration Ecology*, *9*(2), 239–246. doi: 10.1046/j.1526-100X.2001.009002239.x

Hobbs, R., & Norton, D. A. (1996). Towards a Conceptual Framework for Restoration Ecology. *Restoration Ecology*, *4*(2), 93–110.

Hodges, C. A. (1995). Mineral resources, environmental issues, and land use. *Science*, *268*(5215), 1305–1312. doi: 10.1126/science.268.5215.1305

Holdaway, R. J., Wood, J. R., Dickie, I. A., Orwin, K. H., Bellingham, P. J., Richardson, S. J., … Buckley, T. R. (2017). Using DNA metabarcoding to assess New Zealand ' s terrestrial biodiversity. *New Zealand Journal of Ecology*, *41*, 251–262. doi: 10.20417/nzjecol.41.28

Hunter, M. D. (2001). Insect population dynamics meets ecosystem ecology: Effects of herbivory on soil nutrient dynamics. *Agricultural and Forest Entomology*, *3*(2), 77–84. doi: 10.1046/j.1461-9563.2001.00100.x

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, *4*(2), 150–157. doi: 10.1111/j.1755-263X.2010.00158.x

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

Kabiri, V., Raiesi, F., & Ghazavi, M. A. (2016). Tillage effects on soil microbial biomass, SOM mineralization and enzyme activity in a semi-arid Calcixerepts. *Agriculture, Ecosystems and Environment*, *232*, 73–84. doi: 10.1016/j.agee.2016.07.022

Kasparek, M., Demirsoy, A., Akbulut, A., Akbulut, N., Çalişkan, M., & Durmuş, Y. (2000). Distribution and status of the medicinal leech (Hirudo medicinalis L.) in Turkey. *Hydrobiologia*, *441*, 37–44. doi: 10.1023/A:1017555322002

Kim, J., & Pham, V. H. T. (2012). Cultivation of unculturable soil bacteria Cultivation of unculturable soil bacteria. *Trends in Biotechnology*, *30*(9), 475–484. doi: 10.1016/j.tibtech.2012.05.007

Kircher, M., Sawyer, S., & Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Research*, *40*(1), 1–8. doi: 10.1093/nar/gkr771

Kirk, J. L., Beaudette, L. A., Hart, M., Moutoglis, P., Klironomos, J. N., Lee, H., & Trevors, J. T. (2004). Methods of studying soil microbial diversity. *Journal of Microbiological Methods*, *58*(2), 169–188. doi: 10.1016/j.mimet.2004.04.006

Kowalczyk, R., Taberlet, P., Coissac, E., Valentini, A., Miquel, C., & Wójcik, J. M. (2011). *Forest Ecology and Management Influence of management practices on large herbivore diet — Case of European ˙ Primeval Forest ( Poland ) bison in Białowie za*. *261*, 821–828. doi: 10.1016/j.foreco.2010.11.026

Kupschus, S., Schratzberger, M., Righton, D., & Blanchard, J. (2016). Practical implementation of ecosystem monitoring for the ecosystem approach to management. *Journal of Applied Ecology*, *53*(4), 1236–1247. doi: 10.1111/1365-2664.12648

Lahaye, R., Bank, M. Van Der, Bogarin, D., Warner, J., Pupulin, F., Gigot, G., … Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots. *PNAS*, *105*(8), 2923–2928.

Leff, J. W., Jones, S. E., Prober, S. M., Barberán, A., Borer, E. T., Firn, J. L., … Fierer, N. (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences*, *112*(35), 10967–10972. doi: 10.1073/pnas.1508382112

Lesaulnier, C., Papamichail, D., Mccorkle, S., Ollivier, B., Skiena, S., Taghavi, S., … Lelie, D. Van Der. (2008). Elevated atmospheric CO 2 affects soil microbial diversity associated with trembling aspen. *Environmental Microbiology*, *10*(4), 926–941. doi: 10.1111/j.1462-2920.2007.01512.x

Levy-booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, R., Klironomos, J. N., … Dunfield, K. E. (2007). Cycling of extracellular DNA

in the soil environment. *Soil Biology & Biochemistry*, *39*, 2977–2991. doi: 10.1016/j.soilbio.2007.06.020

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, *362*, 709–715.

Machler, E., Deiner, K., Spahn, F., & Altermatt, F. (2016). Fishing in the Water: Effect of Sampled Water Volume on Environmental DNA-Based Detection of Macroinvertebrates. *Environmental Science and Technology*, *50*(1), 305–312. doi: 10.1021/acs.est.5b04188

Majer, J. D. (1989). *Animals in Primary Succession: the Role of Fauna in Reclaimed Lands. (Cambridge University Press: Cambridge.)*.

Majer, J. D. (1997). Invertebrates assist in the restoration process: an Australian perspective. In *(eds. K. M. Urbanska, N. R. Webb & P. J. Edwards) Restoration Ecology and Sustainable Development. Cambridge University Press, Cambridge.* (pp. 212-237.).

Majer, J. D. (2009). Animals in the Restoration Process — Progressing the Trends. *Restoration Ecology*, *17*(4), 315–319. doi: 10.1111/j.1526-100X.2009.00528.x

Majer, J. D., Brennan, K. E. C., & Moir, M. L. (2007). Invertebrates and the restoration of a forest ecosystem: 30 years of research following bauxite mining in Western Australia. *Restoration Ecology*, *15*(SUPPL. 4), 104–115. doi: 10.1111/j.1526-100X.2007.00298.x

Mcdonald, T., Gann, G. D., Jonson, J., & Dixon, K. W. (2016). International standards for the practice of ecological restoration - including principles and key concepts. *Communications Standards*, (December), 427–437. doi: 10.1016/b978-0-08-034092-0.50030-2

Menz, M. H. M., Dixon, K. W., & Hobbs, R. J. (2013). Hurdles and opportunities for landscape-scale restoration. *Science*, *339*(6119), 526–527. doi: 10.1126/science.1228334

Milberg, P., Bergstedt, J., Fridman, J., Odell, G., & Westerberg, L. (2008). Observer bias and random variation in vegetation monitoring data. *Journal of Vegetation Science*, *19*(5), 633–644. doi: 10.3170/2008-8-18423

Minamoto, T., Naka, T., Moji, K., & Maruyama, A. (2016). Techniques for the practical collection of environmental DNA : filter selection , preservation , and extraction. *Limnology*, *17*(1), 23–32. doi: 10.1007/s10201-015-0457-4

Miranda, M., Burris, P., Froy Bincang, J., Shearman, P., Oliver Briones, J., La Viña, A., & Menard, S. (2003). *Mining and Critical Ecosystems: Mapping the Risks*.

Mummery, J., & Hardy, N. (1994). Australia's biodiversity : an overview of selected significant components. *Dept. of the Environment, Sport and Territories, Biodiversity Unit Canberra*.

Murray, C., & Marmorek, D. (2003). Adaptive management and ecological restoration. In *Ecological restoration of southwestern ponderosa pine forests Vol. 2* (pp. 417–428).

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

Nichols, O. G., & Nichols, F. M. (2003). Long-term trends in faunal recolonization after bauxite mining in the Jarrah forest of Southwestern Australia. *Restoration Ecology*, *11*(3), 261–272. doi: 10.1046/j.1526-100X.2003.00190.x

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

Paterson, N., & Lampert, R. J. (1985). A Central Australian ochre mine. *Records of the Australian Museum*, *37*(1), 1–9. doi: 10.3853/j.0067-1975.37.1985.333

Pearson, D. L., Hamilton, A. L., & Erwin, T. L. (2011). Recovery Plan for the Endangered Taxonomy Profession. *BioScience*, *61*(1), 58–63. doi: 10.1525/bio.2011.61.1.11

Pedersen, M. W., Overballe-petersen, S., Ermini, L., Sarkissian, C. Der, Haile, J., Hellstrom, M., … Willerslev, E. (2014). Ancient and modern environmental DNA. *Philosophical Transactions of the Royal Society B: Biological*

*Sciences*, *370*, 20130383.

Ratnasingham, S., & Hebert, P. D. N. (2007). BARCODING BOLD : The Barcode of Life Data System. *Molecular Ecology Notes*, *7*, 355–364. doi: 10.1111/j.1471-8286.2006.01678.x

Rayé, G., Miquel, C., Coissac, E., Redjadj, C., Loison, A., & Taberlet, P. (2011). New insights on diet variability revealed by DNA barcoding and high-throughput pyrosequencing: Chamois diet in autumn as a case study. *Ecological Research*, *26*(2), 265–276. doi: 10.1007/s11284-010-0780-5

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, *51*, 1450–1459. doi: 10.1111/1365-2664.12306

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), 1–11. doi: 10.1093/nar/gkr732

Rochefort, L., Isselin-Nondedeu, F., Boudreau, S., & Poulin, M. (2013). Comparing survey methods for monitoring vegetation change through time in a restored peatland. *Wetlands Ecology and Management*, *21*(1), 71–85. doi: 10.1007/s11273-012-9280-4

Rosenberg, D. M., Danks, H. V., & Lehmkuhl, D. M. (1986). Importance of insects in environmental impact assessment. *Environmental Management*, *10*(6), 773–783. doi: 10.1007/BF01867730

Ruiz-jaen, M. C., & Aide, T. M. (2005). Restoration Success : How Is It Being Measured? *Restoration Ecology*, *13*(3), 569–577.

Schmidt, P., Bálint, M., Greshake, B., Bandow, C., Römbke, J., & Schmitt, I. (2013). Soil Biology & Biochemistry Illumina metabarcoding of a soil fungal community. *Soil Biology and Biochemistry*, *65*, 128–132. doi: 10.1016/j.soilbio.2013.05.014

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*(1), 1–14. doi: 10.1186/s12983-015-0115-z

Schulz, S., Brankatschk, R., Dumig, A., Kogel-Knabner, I., Schloter, M., & Zeyer, J. (2013). The role of microorganisms at different stages of ecosystem development for soil formation. *Biogeosciences*, *10*, 3983–3996. doi: 10.5194/bg-10-3983-2013

Sirois, S. H., & Buckley, D. H. (2019). Factors governing extracellular DNA degradation dynamics in soil. *Environmental Microbiology Reports*, *11*(2), 173–184. doi: 10.1111/1758-2229.12725

Šmejkalová, M., Mikanová, O., & Borůvka, L. (2003). Effects of heavy metal concentrations on biological activity of soil micro-organisms. *Plant, Soil and Environment*, *49*(7), 321–326.

Stackebrant, E., & Goebel, B. M. (1994). Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *International Journal of Systematic and Evolutionary Microbiology*, *44*(4), 846–849. doi: 10.1099/00207713-44-4-846

Suding, K. N., Gross, K. L., & Houseman, G. R. (2004). Alternative states and positive feedbacks in restoration ecology. *Trends in Ecology and Evolution*, *19*(1), 46–53. doi: 10.1016/j.tree.2003.10.005

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, Pierre, Coissac, E., Hajibabaei, M., & Rieseberg, L. H. L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8, SI), 1789–1793. doi: 10.1111/j.1365-294X.2012.05542.x

Taberlet, Pierre, Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., … Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3). doi:

10.1093/nar/gkl938

Taberlet, Pierre, Homme, S. M. P., Campione, E., Roy, J., Miquel, C., Shehzad, W., … Melodelima, C. (2012). Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, *21*, 1816–1820. doi: 10.1111/j.1365-294X.2011.05317.x

Thom, R. M. (2000). Adaptive management of coastal ecosystem restoration projects. *Ecological Engineering*, *15*(3–4), 365–372. doi: 10.1016/S0925-8574(00)00086-0

Thompson, K. A., & Newmaster, S. G. (2014). Molecular taxonomic tools provide more accurate estimates of species richness at less cost than traditional morphology-based taxonomic practices in a vegetation survey. *Biodiversity and Conservation*, *23*(6), 1411–1424. doi: 10.1007/s10531-014-0672-z

Thompson, S. A., & Thompson, G. G. (2004). Adequacy of rehabilitation monitoring practices in the Western Australian mining industry. *Ecological Management and Restoration*, *5*(1), 30–33. doi: 10.1111/j.1442-8903.2004.00172.x

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

Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, *183*, 93–102. doi: 10.1016/j.biocon.2014.11.017

Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E., Coissac, E., Pompanon, F., … Taberlet, P. (2009). New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: The trnL approach. *Molecular Ecology Resources*, *9*(1), 51–60. doi: 10.1111/j.1755-0998.2008.02352.x

Vamos, E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, *1*, e14625. doi: 10.3897/mbmg.1.14625

Wallace, K. J., Laughlin, D. C., & Clarkson, B. D. (2017). Exotic weeds and fluctuating microclimate can constrain native plant regeneration in urban forest restoration. *Ecological Applications*, *27*(4), 1268–1279. doi:

10.1002/eap.1520

Whitehead, P. (1990). Systematics : An Endangered Species. *Systematic Biology*, *39*(2), 179–184.

Widmer, F., Seidler, R. J., Donegan, K. K., & Reed, G. L. (1997). Quantification of transgenic plant marker gene persistence in the field. *Molecular Ecology*, *6*(1), 1–7. doi: 10.1046/j.1365-294X.1997.00145.x

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., … Cooper, A. (2003). Diverse plant and animal genetic records from holocene and pleistocene sediments. *Science*, *300*(5620), 791–795. doi: 10.1126/science.1084114

Yan, Dongfeng, Bissett, A., Gellie, N., Mills, J. G., Lowe, A. J., & Breed, M. F. (2019). Soil bacterial community differences along a coastal restoration chronosequence. *Plant Ecology*, *221*(9), 795–811. doi: 10.1007/s11258-019-00979-0

Yan, DongFeng, Mills, J. G., Gellie, N. J. C., Bissett, A., Lowe, A. J., & Breed, M. F. (2018). High-throughput eDNA monitoring of fungi to track functional recovery in ecological restoration. *Biological Conservation*, *217*, 113–120. doi: 10.1016/j.biocon.2017.10.035

Yeates, D. K., Harvey, M. S., & Austin, A. D. (2003). New estimates for terrestrial arthropod species-richness in Australia. *Records of the South Australian Museum Monograph Series*, *7*, 231–241.

Yoccoz, N. G., Brathen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., … Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, *21*, 3647–3655. doi: 10.1111/j.1365-294X.2012.05545.x

Young, R., Manero, A., Miller, B., Kragt, M. E., Standish, R. J., Jasper, D., & Boggs, G. (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. doi: 10.1111/j.2041-210X.2012.00198.x

# CHAPTER 2

## TESTING MULTIPLE SUBSTRATES FOR TERRESTRIAL BIODIVERSITY MONITORING USING ENVIRONMENTAL DNA (EDNA) METABARCODING

## 2.1 Preface

*This chapter consists of a published manuscript titled '**Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA (eDNA) metabarcoding**' [Molecular Ecology Resources 2020]. The content in section 2.2 is the same as the published manuscript with only minor changes in formatting to accommodate thesis referencing.*

This chapter describes a systematic test of common terrestrial sample types (substrates). While methodologies for assessing soil microbial communities have been established in the literature, when it comes to eukaryotic biodiversity there are a lot of unknowns in terms of what samples are appropriate. This is especially true in Australia, where the high temperatures and UV radiation may degrade DNA quickly, and make some substrates unsuitable for biodiversity assessment. The substrates (bulk arthropods, soil, plant material, scat) were collected from two sites in Western Australia with different climates and vegetation, to ensure the results would be applicable to more than one environment. The same four assays were applied to all substrates to compare the vertebrate, invertebrate, and plant diversity that could be detected in each substrate. It was necessary to test multiple substrates to determine which substrates would be suitable to apply to mine site restoration for the later chapters.

### 2.1.1 Acknowledgements

This work was supported by the Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration (ICI150100041) and the Pawsey Supercomputing Centre. We would like to thank Katrina West and Khiraj Bhalsing for help with sample collection. We acknowledge the support of the members of the Trace and Environmental DNA (TrEnD) Laboratory with metabarcoding workflows and bioinformatics.

### 2.1.2 Data Accessibility

Sequencing data and DADA2 script is available at the Dryad Digital Repository: https://doi.org/10.5061/dryad.38100f6

### 2.1.3 Author Contributions

MvH conducted the study and wrote the manuscript. PN, MB, NW, and GW-J were involved in the experimental design; samples were collected by MvH, PN, KF, GW-J, and processed by MvH and KF; molecular and bioinformatics work was performed by MvH; all data was analyzed and processed by MvH; statistical analysis was done by MvH; the manuscript was edited by all authors.

## TESTING MULTIPLE SUBSTRATES FOR TERRESTRIAL BIODIVERSITY MONITORING USING ENVIRONMENTAL DNA (EDNA) METABARCODING

*Molecular Ecology Resources* (2020), Vol. 00, Pg. 1-14. doi: 10.1111/1755-0998.13148

**van der Heyde, M.[1,2*],** Bunce, M.[2], Wardell-Johnson, G.[1],Fernandes, K.[2], White, N.E.[2], Nevill, P.[1,2]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, GPP Box U1987, Perth, Western Australia, 6845

[2]Trace and Environmental DNA Laboratory, School of Life and Molecular Sciences, Curtin University, GPP Box U1987, Perth, Western Australia, 6845

*Corresponding author

**keywords:** DNA barcoding, metagenomics, biological audit, terrestrial, Biodiversity

## 2.2 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, likely 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.

## 2.3 Introduction

There is a growing need for effective biomonitoring with increasing pressure on ecological systems from human population growth, resource use and climate change (UNEP 2011; Pimm et al. 2014; Dirzo et al. 2014). Biomonitoring is necessary for effective ecosystem management including the early detection of invasive species (Epanchin-Niell et al. 2012), measurement of trajectories following ecological restoration (Herrick, Schuman, and 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 (i) phenotypic plasticity (Demes, Graham, and Suskiewicz 2009; Weigand et al. 2011), (ii) juveniles with ambiguous morphology (Richard et al. 2010; Ji et al. 2013), and (iii) taxa having different levels of detectability according to season and time (K. A. Thompson and Newmaster 2014; Fernandes et al. 2018). There has also been a worldwide decline in taxonomic expertise (Pearson, Hamilton, and Erwin 2011), which further limits traditional approaches. In addition, it is difficult to rely on morphology to monitor across a broad taxonomic range, as expertise and methods tend to be taxon-specific. With the demand for efficient biomonitoring, new technologies are being developed to expand the monitoring "toolkit" to complement traditional methods.

One such method is environmental DNA (eDNA) metabarcoding, a process of sequencing barcode regions from DNA that has been isolated from environmental samples including sediment, water, seawater, bulk arthropods and air (Bohmann et al. 2014). Several reviews have advocated the use of eDNA metabarcoding for biodiversity monitoring (Bohmann et al., 2014; Creer et al., 2016; Fernandes et al., 2018; Taberlet et al., 2012; Thomsen & Willerslev, 2015; Williams et al., 2014) as it has the potential to increase the range of biodiversity detected and to include a broader array of forms (i.e. immature specimens, cryptids, and phoretic individuals). The data are also readily auditable by third parties (Ji et al. 2013) and the cost may be calculated based on number of samples, rather than number of specimens, making it more cost-effective, especially in highly diverse systems (Ji et al. 2013).

The reduction in the cost of high-throughput sequencing has led to a rapid increase in the number of eDNA studies, as well as commercial interest (Supporting Information Koziol et al. 2018). Soil microbial researchers have been using eDNA metabarcoding for over two decades (Anderson and Cairney 2004) and there is now growing evidence that barcoding may be useful to monitor plant communities (de Mattia et al. 2012; K. A. Thompson and Newmaster 2014; Fahner et al. 2016), vertebrates (Andersen et al., 2012; Calvignac-spencer, Merkel, & Kutzner, 2013; Fernandes et al., 2019) and invertebrates (Ji et al. 2013; Yang et al. 2014). Researchers have successfully sequenced: top soil (K. Andersen et al. 2012; Fahner et al. 2016), scat (De Barba et al. 2014), ancient middens (Dáithí C. Murray et al. 2012), air (Kraaijeveld et al. 2015), bulk arthropods (Yu et al. 2012; Ji et al. 2013), leaf material (K. A. Thompson and Newmaster 2014), flowers (Thomsen and Sigsgaard 2019) and more. However, almost all these studies have employed only one eDNA substrate (Koziol et al. 2018). Yang et al. (2014) is one of the few that sampled multiple terrestrial substrates (soil, leaf litter, and insect traps) but they targeted the soil fauna for extraction (separated from the soil) rather than extracting DNA from the soil directly. Yang et al. (2014) also used different PCR assays for their bulk arthropod and soil/leaf litter samples, affecting their comparability. Thus far there has been no study that compares the same barcode across multiple terrestrial substrates. Most samples are able to detect multiple taxonomic groups. However, without a systematic, comparative substrate trial using multiple barcodes it is not possible to determine overlap.

Generally, eDNA studies have occurred in temperate regions or colder (K. Andersen et al. 2012; Fahner et al. 2016), where DNA preservation is more optimal for metabarcoding. Few if any non-microbial barcoding studies have been performed on soils from hotter climates such as those found in parts of Australia. This is true of most biodiversity research; hotter climates are critically understudied (Titley, Snaddon, and Turner 2017), despite having the highest extinction rates (Wiens 2016). As the climate continues to warm and dry (Huang et al. 2016), developing tools to monitor these regions globally has become increasingly important.

In this study, we tested five common terrestrial substrates (soil, bulk scat, bulk plant material and bulk arthropods from pitfall traps and vane traps) with four

eDNA barcoding assays to detect a wide range of plants, vertebrates and arthropods. Two well characterized study sites were chosen in Western Australia to examine the feasibility of metabarcoding from substrates collected from hot desert and Mediterranean climates. We aim to improve decision making for terrestrial eDNA surveys by:

1) Examining the diversity within and overlap between commonly sampled substrates. Ideal substrates for monitoring should detect both the greatest richness per sample and greatest overall diversity. The degree of overlap in diversity detected will indicate if multiple substrates are necessary for broad biological surveys.

2) Examining differences between sites as some substrates may perform poorly in certain climates.

3) Comparing DNA sequence and traditional biodiversity survey methods. In particular we compare a DNA sequence based approach with vegetation surveys based on plant morphology to understand the extent to which they complement one another.

## 2.4 Materials and methods

### 2.4.1 Study sites

Samples (n = 200) were collected from two study sites 1000 km apart in Western Australia; one in the Pilbara, and the other in the south-western Swan Coastal Plain (SCP) (Figure 2.1). The Pilbara site (-23.19°, 119.24°) is a valley bottom of red clay and strewn pebbles (Burbidge, Johnstone, and Pearson 2010), has an arid climate with hot summers, mild winters, more than 10 hours sunshine a day, and low but variable rainfall (Sudmeyer 2016). It is classified as hot desert (Bwh, Beck et al. 2018) and the dominant vegetation consists of *Acacia* shrublands with hummock grasses (*Triodia*) (Burbidge, Johnstone, and Pearson 2010). The Swan Coastal Plain (SCP) site (-31.76°, 115.95°) is in a highly diverse Banksia Woodland on sandy soils. The region has a hot Mediterranean (Csa, Beck et al. 2018) climate with hot, dry summers and cool, wet winters. Study sites with different climates and soil types were chosen because they may affect DNA preservation. Both sites have a broad array of taxa, and the SCP is located in one of Australia's two biodiversity hotspots. The Pilbara also hosts a globally significant resources industry (Argent 2013) where current and proposed projects require regular biomonitoring or collection of baseline data.

### 2.4.2 Sample collection

At each site, 5 different substrates were collected; soil, scat, plant material, arthropods from pitfall traps, and arthropods from vane traps. Samples were collected in October and November 2017 (early summer) totally 200 (2 sites x 20 samples per substrate x 5 substrates).

Sample points were 50-60m apart in a grid like pattern across the 15 ha sites. At each sample point, 5 soil subsamples were collected from the surface 5 cm using sterilized equipment and gloves that were re-sterilized with bleach between each sample point. The subsamples were collected randomly within a 10 m x 10 m plot and mixed in the field to form one sample. Scat was sampled by collecting any visible scat approximately 200 m around each sample point and collecting any scat that was visible. Soil and scat samples were kept cool, and frozen within a few hours.

Plant material was collected by a non-specialist but with some training in flora surveys. A leaf was collected from each plant species within a quadrat (50 m x 50 m for the Pilbara, 10 m x 10 m for the SCP) and stored in envelopes in silica gel. Size of quadrats was based on standard monitoring plots used by mining companies in the area. Each sample point also had four pitfall traps (12 cm deep, 4 cm diameter) combined to form one sample, and one yellow vane trap. The traps were left out for 7 days to catch arthropods and contained ethylene glycol in the form of concentrated auto coolant as a capture fluid.

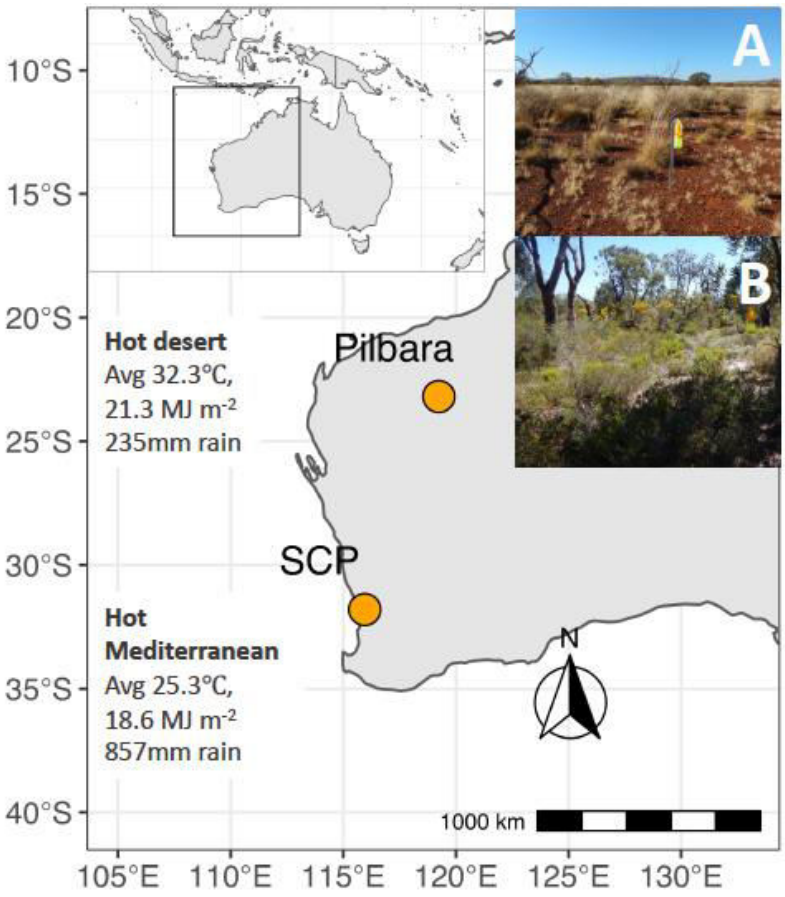


**Figure 2.1** Map of the two study sites in Western Australia. The Pilbara (A) and the Swan Coastal Plain (B) are shown in the photos. Information on the left contains the climate type (Koppen Classification), the average daily maximum temperature, average daily solar exposure, and the total rainfall for the sample year.

### 2.4.3 Sample Processing and DNA extraction

All samples were extracted using the Qiacube extraction platform (Qiagen, Germany). Soil samples were manually homogenized and DNA was extracted from 300 mg using the Qiagen DNeasy PowerLyser Powersoil kit (Qiagen, Germany). The Qiagen PowerFecal DNA kit (Qiagen, Germany) was used to extract DNA from 250 mg of each of the scat samples. For the plant samples small sections of each leaf were homogenized dry using Precellys 7ML Hard Tissue Homogenizing Ceramic Beads kit for 8 minutes, then again for 2 minutes with 3 mL of AP1 buffer from the Qiagen DNeasy Plant Mini kit (Qiagen, Germany). They were digested overnight and the DNA extracted using the Plant Mini kit. The arthropod samples were rinsed with de-ionized water using 20 micron sieves that were sterilized in bleach and UV between every sample. They were then homogenized using a hand-held blender (OMNI Tip homogenizer, Kennesaw, GA, USA) and the DNA extracted with a Qiagen QIAmp DNA Mini Kit modified with a starting volume of 400 µL of digest fluid and a 100 µL elution. DNA extraction controls (blanks) were carried out for every 20 samples using the extraction reagents only.

### 2.4.4 Assessment of DNA extracts

Quantitative polymerase chain reaction (qPCR) was used to assess the quality and quantity of DNA in the extract, as well as determine the optimal level of DNA input for metabarcoding (Dáithí C. Murray, Coghlan, and Bunce 2015). Four qPCR assays (described below) were run on all samples and all substrates, to determine if there was sufficient amplification to attempt sequencing. Due to the degraded nature of eDNA, all primers used targeted short amplicons (72bp to 157bp) to improve amplification success from samples. ZBJ-ArtF1c/ZBJ-ArtR2c (~157bp, Zeale, Butlin, Barker, Lees, & Jones, 2011) was chosen as a general arthropod primer, with an addition of Ant236/361 (~72bp, Fernandes et al., 2019) to target arthropod orders such as Hymenoptera, which ZBJ-ArtF1c/ZBJ-ArtR2c has shown some bias against (Fernandes, van der Heyde, et al. 2019; Clarke et al. 2014). Primer bias may differentially affect sites with different community composition, so the combination of the two invertebrate primers were chosen to control for this bias. Both target sections of cytochrome c oxidase subunit 1 (COI) have extensive reference databases available online to improve taxonomic assignment compared to gene regions with

smaller databases. Two plant primer sets were used that target the chloroplast genome: 1) trnL-g/h primers (Taberlet et al., 2007), which produces a PCR amplicon of variable length (10-143 bp) from the P6 loop of the trnL (UAA) intron; 2) rbcL-h1aF/h2aR primers (96bp, Poinar et al., 1998) was used only on plant material samples for comparison between the plant sequences and the traditional plant survey.

The vertebrate primers targeted the 12S gene 12SV5-F/R (98bp, Riaz et al., 2011). The PCR mix for quantitation contained: 2.5 mM MgCl2 (Applied Biosystems, USA), 1× PCR Gold buffer (Applied Biosystems), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4 mg/ml bovine serum albumin (Fisher Biotec, Australia), 0.4 µmol/L forward and reverse primer, 1 U AmpliTaq Gold DNA polymerase (Applied Biosystems) and 0.6 µl of a 1:10,000 solution of SYBR Green dye (Life Technologies, USA). They were run on a StepOne Plus (Applied BioSystems) real-time qPCR instrument with the following conditions: 5 min at 95°C, 40 cycles of 95°C for 30s, 30s at the annealing temperature (52°C for trnL and rbcL, 53°C for ZBJ-Art, 50°C for Ant236, 60°C for 12SV5) and 45s at 72°C, ending with 10 min elongation at 72°C. Contamination was minimised by preparing the PCR mixes in a dedicated clean room and then adding sample in a separate laboratory in specialized UV cabinets.

### 2.4.5 DNA Amplification and Sequencing

Samples that yielded sufficient amplifiable DNA, as determined by the qPCR screening, were assigned a unique combination of fusion tag primers that contained a unique multiplex identifier (MID) tag between 6-8 bp in length, the gene-specific primer (described above) and Illumina's sequencing adaptors (i.e. P5 and P7). These MID-tag (fusion) primers were then used in qPCR with the same reagents and cycling conditions described above. A single-step fusion protocol was employed with no reuse of index combinations. The MID-tag amplicons were generated in duplicate and then pooled together. Pooled amplicons were cleaned using the QIAquick PCR Purification Kit (Qiagen, Germany) and quantified using the QIAxcel Advanced System (Qiagen, Germany). Pools were combined in approximate equimolar ratios based on this quantitation to create a DNA library for sequencing. Amplicons in this library were size selected using a Pippin Prep (Sage Science), cleaned using the QIAquick PCR Purification Kit (Qiagen, Germany) and

eluted into 50µl. The final DNA library was quantified using Qubit Fluorometric Quantitation (Thermo Fisher Scientific) and sequenced as per Illumina sequencing protocols for single-end sequencing.

### 2.4.6 Sequence Analysis

Raw sequence reads were demultiplexed (i.e. assigned back to sample using MID-tag primer combos) using 'obitools' (Boyer et al. 2016a), then sequences were quality filtered, the errors and chimeras removed using DADA2 (Callahan et al. 2016) on R 3.5.1 (R Core Team 2018) (script available on https://doi.org/10.5061/dryad.38100f6). DADA2 denoises sequences using error rates estimated from the sequencing run, producing amplicon sequence variants (ASVs) that are not clustered like traditional operational taxonomic units (OTUs), but are analyzed in much the same way. The resulting ASV tables of were then analyzed in R 3.5.1 (R Core Team 2018). We ignored sequence counts less than 5, and removed samples with less than 200 seq/sample. Any ASVs that were present in the extraction (i.e. laboratory) controls were then removed from the dataset. ASV sequences were matched to a reference database using the Basic Local Alignment Search Tool (BLASTn) on a high performance cluster computer (Pawsey Supercomputing Centre; Perth, WA, Australia) against the online reference database Genbank (https://www.ncbi.nlm.nih.gov/genbank/) for taxonomic assignment. The ZBJ-Art and Ant236 assays were searched against both Genbank and Arthropod COI sequences extracted from the Barcode of Life Database (BOLD: https://www.barcodeoflife.org), because there are arthropod sequences on this database not present in Genbank. BLASTn results returned the top 10 hits with a minimum query coverage of 80% and minimum percent identity or 80%. Taxonomic identification was assigned to the lowest common ancestor with MEGAN (Huson et al. 2007) with minimum support of 140 (ZBJ-Art), 60 (Ant236), 90 (12SV5), 50, (trnL), and 90 (rbcL).

### 2.4.7 Vegetation Surveys

Vegetation surveys were conducted by an expert botanist, for each sample point on the SCP in November 2017. Due to time constraints while working remotely, the Pilbara vegetation surveys include 8 sample points conducted in

September 2018. The vegetation survey quadrat areas matched those of the sample collection of plant material (10 m x 10 m in SCP and 50 m x 50 m in Pilbara), with the intent of providing a morphological comparison for the plant samples. For the comparison between sequencing data and the morphological surveys, only sample points with both vegetation surveys and sequencing data were included. Identifications of all sampled specimens were confirmed by botanists at the Western Australian Herbarium but no voucher specimens were lodged.

### 2.4.8 Statistical analysis

Statistical analysis was performed on R 3.5.1 (R Core Team 2018). For the four assays that were tested on all substrates (ZBJ-Art, Ant236, trnL, 12SV5), we calculated the ASV richness for each substrate (soil, scat, plant material, arthropods from pitfall traps, and arthropods from vane traps) at both sites (SCP and Pilbara) and tested the differences between substrates and site using an two-way analysis of variance (ANOVA) where the data met the assumptions (12SV5). For most assays (trnL, ZBJ-Art, and Ant236), the groups did not have equal variance and so we used permutational analysis of variance (PERMANOVA) from the R package 'vegan' (J. Oksanen et al. 2019) with 999 permutations. We tested the differences in community composition between sites and substrates by first combining the four assays into a presence absence matrix that included taxonomic families and the samples in which they were detected. Then the Bray-Curtis similarity was calculated between samples and the differences in community composition were tested using the PERMANOVA with 999 permutations. This was visualized using nonmetric multidimensional scaling (NMDS) from the same package (J. Oksanen et al. 2019). This matrix was also used to determine the families associated with each substrate by using the R package 'indicspecies' (De Cáceres and Legendre 2009). Plant families detected from plant material samples were also compared to the families identified in the plant survey.

## 2.5 Results

A total of 26 589 497 metabarcoding sequences were generated from the four PCR assays (See Table 2.1 for summary). Small quantities of ASVs were present in extraction controls, highest in the trnL assay (5.7%) and lowest in the ZBJ-Art Assay (0.7%). A fish sequence variant was removed as a likely contaminant, it was found in 0.01% of 12SV5 sequences (See Supplementary information for details). In total, there were 278 taxonomic families detected from 87 orders.

**Table 2.1** Summary of sequencing results for each assay

| Assay | Sequences | Samples (>200 reads) | ASVs | Substrates |
|---|---|---|---|---|
| 12SV5 | 268,038 | 40 | 221 | Scat |
| ZBJ-Art | 5,676,094 | 160 | 2,253 | Scat, plant, pitfall, vane, soil |
| Ant236 | 4,491,408 | 119 | 2,145 | Scat, plant, pitfall, vane |
| trnL | 11,660,421 | 164 | 546 | Scat, plant, pitfall, vane, soil |
| rbcL | 7,080,330 | 40 | 1,064 | Plant[a] |
| Total | 29,176,291 | 523 | 6,229 | |

[a]*rbcL* was only applied to plant samples and was not tested for the other substrates.

### 2.5.1 Substrate Diversity and Richness

Substrates varied significantly in diversity and ASV richness. Pitfall traps detected the greatest number of ASVs overall (1792 ASVs), while vane traps ranked third overall with 1208 ASVs in total. Approximately the same number of families were detected from both traps across the two sites (127 in pitfall traps and 126 in vane traps), and they had similar proportions of families unique to that substrate (pitfall traps 35.4%; vane traps 33.3%).

Scat samples were found to have fewer ASVs overall (1333 ASVs) than pitfall traps, but had the highest ASV richness per sample for most assays (Figure 2.2). There are differences between the per sample richness and the accumulated richness; for example, although scat had higher average ZBJ-Art ASV richness (32.1± 4.4se) than vane samples (18.5 ± 2.0se), there were more ZBJ-Art ASVs in the vane substrate than in scat (542 vs 470). The differences between per sample

richness and accumulated richness are related to the overlap between samples. Scat required fewer samples to achieve the same proportion of cumulative diversity than pitfall and vane traps, which have more variation between samples. Overall, scat samples detected the most families overall (131 families), had the most families unique to scat samples (56 families, 42.7%), and was the only substrate that showed successful amplification with the vertebrate 12SV5 assay (Figure 2.2; Table S2.9.1).

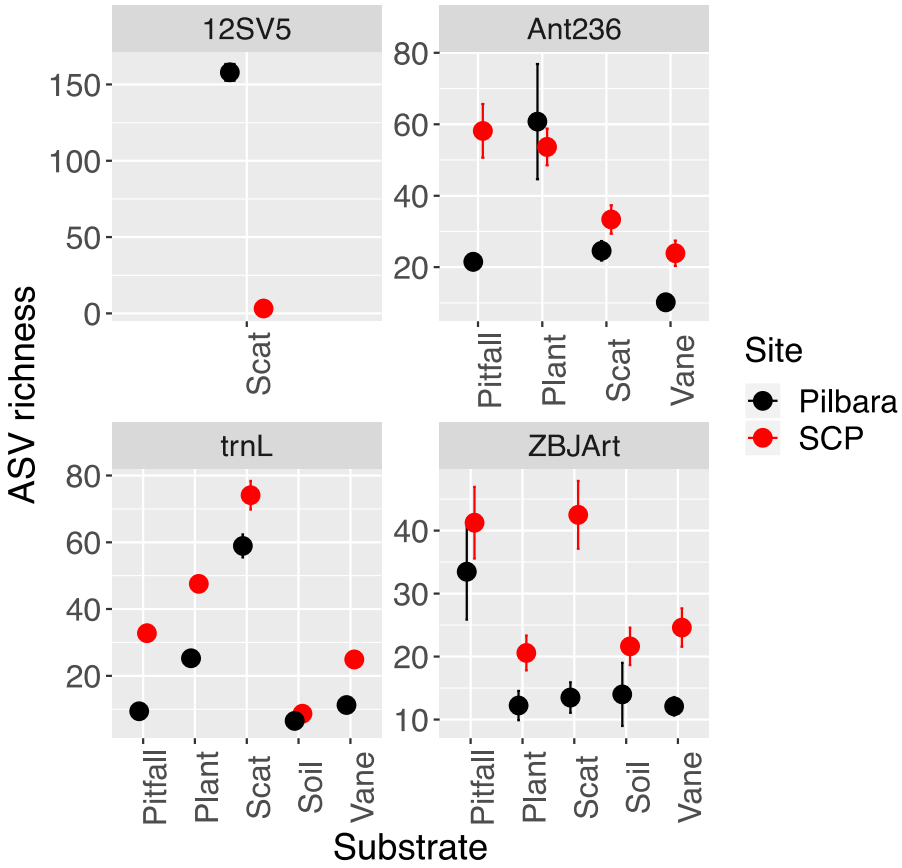


**Figure 2.2** ASV richness of all four barcoding assays in all the substrates. Error bars indicate standard error, red points are SCP samples and black points are from the Pilbara.

Plant samples were found to have similar overall ASV richness (1326 ASVs) to scat (1333 ASVs) and vane traps (1208 ASVs), but fewer families (86 families). The degenerate nature of the Ant236 primers resulted in plant DNA being sequenced, which was confirmed with the trnL assay. Overall, 22.1% of the families detected in the plant samples were unique to this substrate.

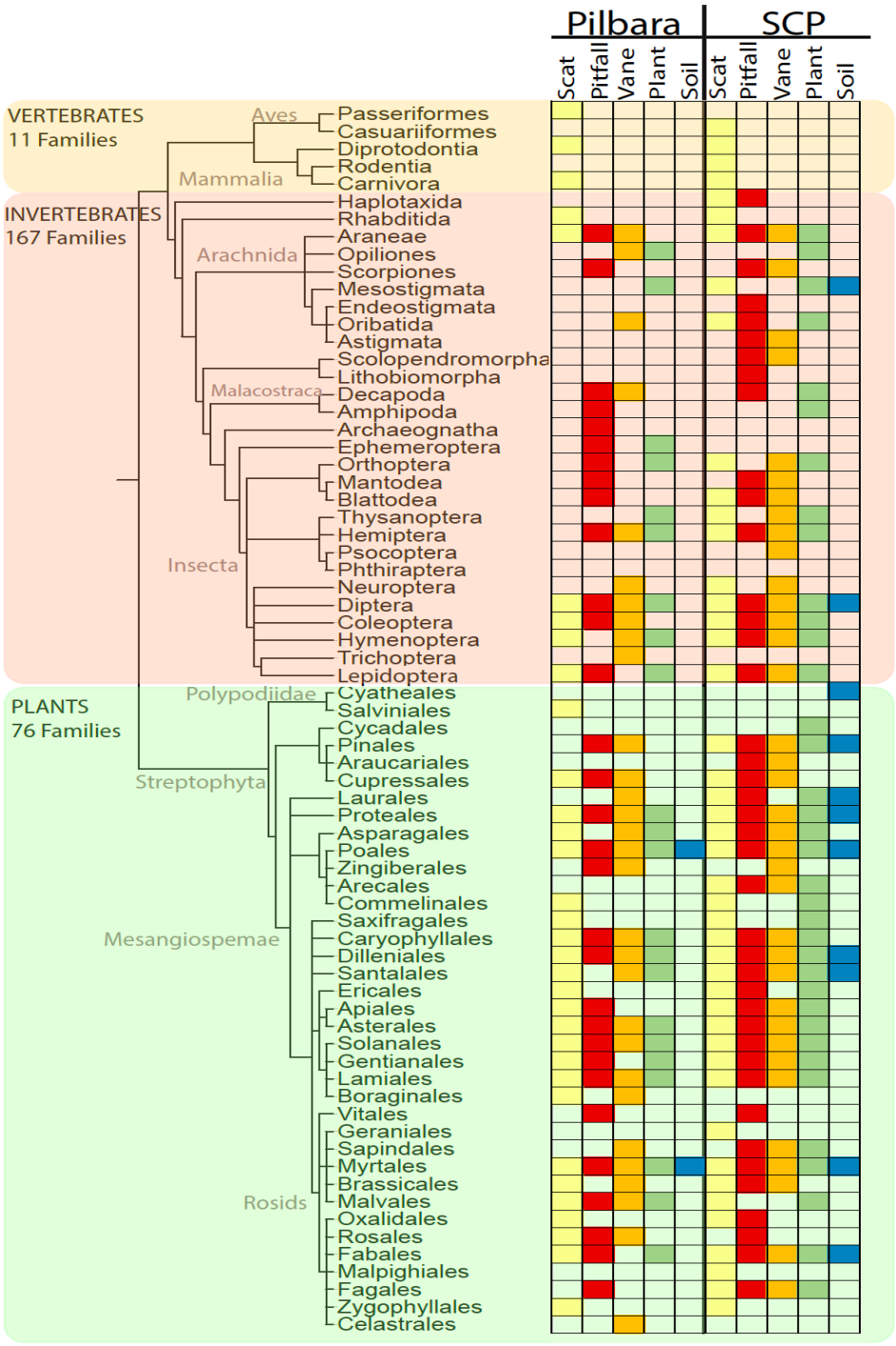


**Figure 2.3** Taxonomic orders detected in each substrate. Fungal and algal orders were removed as the assays are not equipped to properly detect fungal diversity. Orders were chosen for this figure because there were too many families to fit in one figure.

Despite numerous optimizations and extraction attempts (modified bead bashing time, pelleting time, alternate extraction method), less than half of the soil

samples successfully amplified (Table S2.9.1). For trnL and ZBJ-Art, 11/40 soil samples (27.5%) and 18/40 (47%) soil samples amplified respectively. Soil samples were the lowest in the number of total ASVs detected (224 ASV), and the lowest taxonomic diversity identified (18 families; Figure 2.3 & 2.4). Most (63%) of families identified from soil samples were detected in at least two other substrates and only 11.1% of families were unique to soil.

Only 3% of families were detected in all substrates, 10% if soil is excluded from the analysis. All 8 of the families found in all substrates including soil were plant families (Table 2.2). The results of our indicator analyses showed that 23 out of 157 families in the Pilbara, and 86 out of 217 families in the SCP were significantly associated with one or more substrates ($p < 0.05$, see Table S2.9.4 for details). Over half (59%) of the families were detected in only one substrate (130 in SCP, 109 Pilbara), and of the families significantly associated with substrates ($p < 0.05$ indicator analysis, Table S2.9.4), the majority (63% SCP, 74% Pilbara) were associated with a single substrate. For example, vertebrate families such as Macropodidae (wallabies and kangaroos), and Dromaiidae (emus) were associated with scat samples, while Formicidae (ants) and Termitidae (termites) were associated with pitfall traps. Other families were associated with multiple substrates; i.e. Noctuidae (Owlet moths) in plant and scat, Poaceae (grasses) in plant and scat and soil, and Acrididae (Short-horned grasshopper) in pitfall and vane trap samples. In the SCP, five plant families were significantly associated with all substrates except soil (Myrtaceae, Fabaceae, Asteraceae, Proteaceae, and Dilleniaceae).

### 2.5.2 Site differences

ASV richness was greater in the SCP than in the Pilbara (Figure 2.2). There was an interaction between site and substrate for both the trnL and Ant236 assay, and site and substrate were significant terms for all assays ($p < 0.001$, see Supplementary Material for details). The vertebrate assay was the only one where the Pilbara had significantly ($p < 0.001$) greater ASV richness, and 92% of the ASVs were assigned to the Macropodidae family. The remaining assays showed SCP having higher ASV richness than the Pilbara (Figure 2.2). Taxonomic diversity was also higher in the SCP (217 families) compared to the Pilbara (156 families)(Figure 2.4). Additionally, more samples from the SCP were able to be amplified and sequenced; in the trnL

assay we were only able to sequence 2 Pilbara soil samples and 9 SCP soil samples. Similarly, in the ZBJ-Art assay, we were able to sequence 16 SCP soil samples and only 2 Pilbara soil samples.

Site influenced the families found in each substrate ($F_{4,168}$ = 10.694, R2 = 0.112, p < 0.001, Figure 4A), and site ($F_{1,168}$ = 38.4, R2 = 0.101, p < 0.001) and substrate ($F_{4,168}$ = 32.5, R2 = 0.343, p<0.001) were also highly significant (See Table S2.9.3 for details)

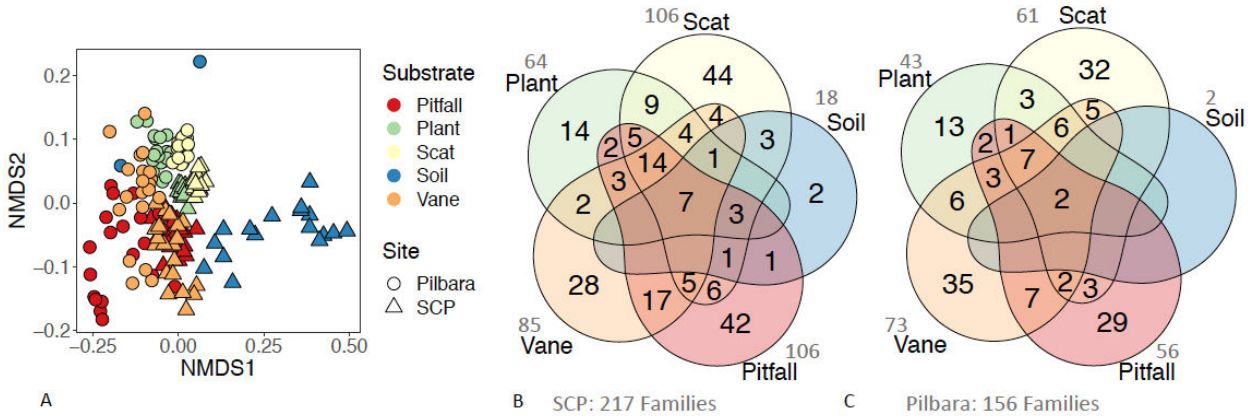


**Figure 2.4** Ordination and venn diagrams of the families detected in the various substrates. A) Nonmetric multidimensional scaling (NMDS) ordination of a presence-absence matrix of families detected in each sample (stress = 0.166, similarity = Bray-Curtis). Venn diagram of families shared between substrates at B) SCP and C) Pilbara site. Gray numbers indicate the total number of families in that substrate.

### 2.5.3 Comparison between plant surveys

The morphology based survey identified 51 families in total (Table S2.9.6). Several plant families found in the sequencing (e.g. Mazaceae and Hyacinthaceae) had been reassigned or renamed, and this was accounted for in our comparison with the morphological survey. The Pilbara has the lowest proportion of families found in both the sequencing and morphological survey (48.1%, Figure 2.5) compared to the SCP (61.7% Figure 2.5). In total, of the 15 plant families identified in the morphological survey and not found in the sequencing data, 8 were identified in only one sample. Families detected in the SCP sequencing data and not in the morphological survey include Pinaceae, a family of plants not present in the study site but in high abundance in an adjacent pine plantation. Four of the families found

in sequencing only were detected at a single sample point. Seven families had less than 700 reads, and the average number of reads per family was 28 580 for the trnL assay and 45 867 for the rbcL assay. In the SCP, almost all sequencing reads (99%) and morphological plant species (93%) were from families identified in both the sequencing and morphological survey; in the Pilbara, a smaller proportion of reads (81%) and sequences (83%) were from families shared between methods (Figure 2.5).

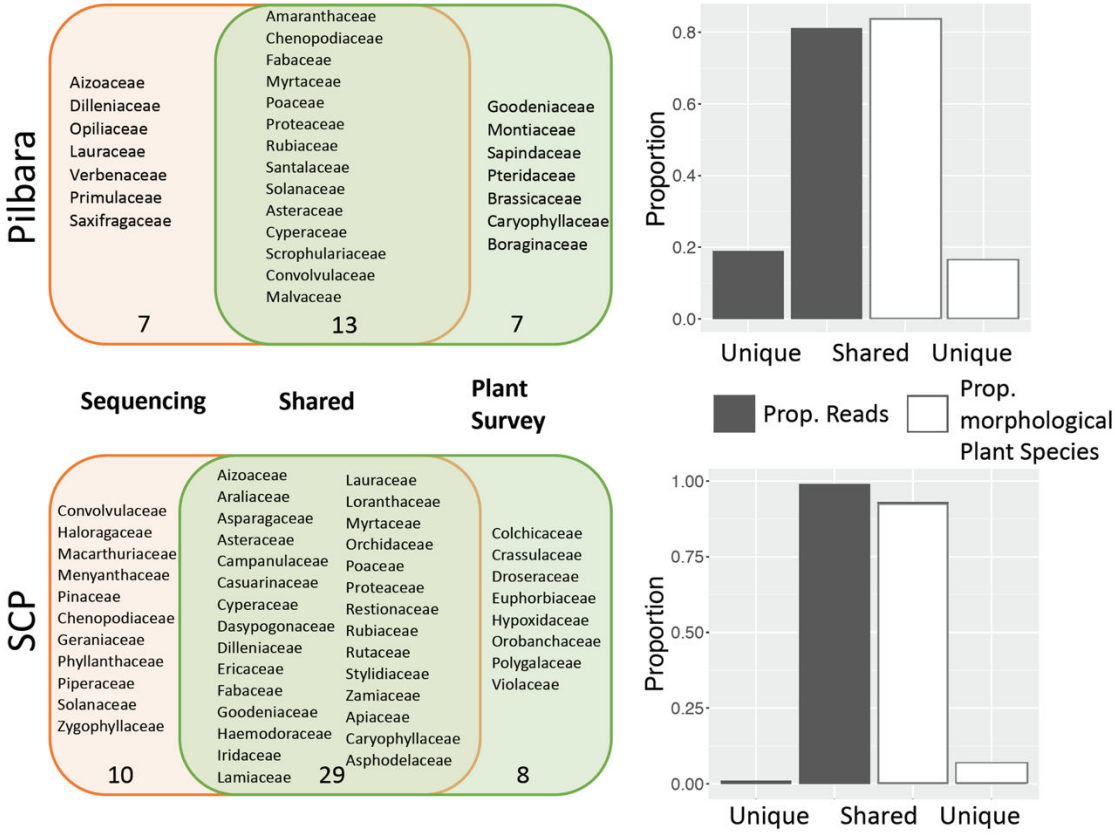


**Figure 2.5** Comparison between plant families found in the morphological plant survey and using sequencing. Numbers indicate the number of families in each category. Bar plots show the proportion (reads for sequencing data, or plant species for morphological survey) from families that were common between sequencing or the morphological plant survey.

## 2.6 Discussion

Use of DNA barcoding and eDNA substrates to monitor biodiversity is on the rise (Koziol et al., 2018). In terrestrial systems researchers are expanding both the range (e.g. Yang et al., 2014) and purpose (Thomsen and Sigsgaard 2019) of substrates collected. In this study, we tested multiple terrestrial substrates with four barcoding assays and found that terrestrial substrates can detect a broad range of taxonomic groups (Figure 2.2). Invertebrate and plant DNA was found in all substrates, although scat was the only substrate to consistently yield vertebrate sequences. Some cosmopolitan taxa were shared with multiple substrates. However, many taxa are specific to a particular substrate, and no one substrate was able to detect all taxa.

### 2.6.1 Substrate Diversity and Richness

Each substrate identified a different biological community in the five substrates and four assays used. Even the most diverse substrate (scat) only comprised half of the total diversity detected (47%). The combination of pitfall trap and scat samples increased detection to 76% of total and with the addition of vane traps, 92% of families detected from all five substrates. For terrestrial biodiversity monitoring, the more substrates chosen, the greater the range of biodiversity that can be detected. This result mirrors a substrate comparison in marine environments and further demonstrates that the manner in which environments are sampled strongly influences both the sensitivity of detection and the assemblages recovered (Koziol et al. 2018).

Generally, the literature agrees with what we were able to detect in our substrates. Like Ji et al. (2013) and Yu et al. (2012) we found invertebrate sequences in the pitfall and vane trap samples (Figure 2.3). Families of flying insects were associated with vane traps (Phoridae, Cicadellidae, and Crambidae) but pitfall traps caught both ground dwellers (Formicidae, Lycosidae) and flying insects (Empididae). While Calvignac-Spencer et al. (2013) were able to retrieve vertebrate DNA from their carrion flies, we were not able to successfully sequence our bulk arthropod samples with the vertebrate assay. Perhaps not enough carnivorous invertebrates were trapped, or the ones trapped had not recently fed. Plant sequences

were detected in bulk arthropod samples and may have come from pollinating insects (Pornon et al. 2016), ingested plants (Jurado-Rivera et al. 2009), or plant material that fell into the traps. The traps differed in the biota that they detected, as we expected based on entomology studies that show the importance of trap type to the biodiversity captured (Prasifka et al. 2007; Santos, Cabanas, and Pereira 2007). The traps are complementary, and when used together can detect the majority of families (Figure 2.4B).

In the plant samples, we found both plant and invertebrate sequences. In plant samples, the source of plant sequences is self-evident, while invertebrate DNA may come from larvae, or cells and fecal matter left on plants (Thomsen and Sigsgaard 2019), spider webs (Blake et al. 2016), and empty leaf mines (Derocles et al. 2015). As evidence, the arthropod families strongly associated with plant samples are all those that feed and reproduce on living plants (i.e. Gelechiidae, Aphididae, Cecidomyiidae), giving them ample opportunity to leave DNA on the plant. Some plant taxa were commonly found in multiple substrates. These include orders such as Poales, Myrtales, and Proteales, which were found in all five substrates (including soil samples - Figure 2.3). The cosmopolitan nature of these taxa suggest they may have airborne dispersal, and Kraaijeveld et al. (2015) were able to detect several taxa in these orders from airborne pollen traps.

The most taxonomically diverse substrate was scat, which picked up vertebrates, invertebrates and plants (Figure 2.3). Typically, scat is used to study the diet of a particular organism and researchers choose their barcoding assays accordingly; insectivores' scat are barcoded with invertebrate primers (Zeale et al. 2011), herbivores' scat is barcoded using plant primers (Valentini et al. 2009), and carnivores' scat is barcoded using vertebrate primers (Arteaga Claramunt et al. 2018). Our scat samples were dominated by Macropididae (Kangaroos and Wallabies), likely as a result of sampling bias. Rather than target a specific organism for scat collection, we collected scat along a transect, and Macropodidae scat is both numerous and easy to see, resulting in an overabundance of this family. This sampling bias is exacerbated by site differences, for example, Passeriformes scat is relatively easy to see in the red clay of the Pilbara, but almost impossible to find in the sand of the SCP. In addition, although many samples were from only herbivorous

animals (Macropodidae and Dromaiidae) they still picked up invertebrate sequences. This could be from contact with invertebrates on the ground such as beetles (Coleoptera Figure 2.3). Furthermore, the presence of families of moth within scat samples (Noctuidae, Oecophoridae) can be explained through biological material that could have been eaten (larvae, fecal matter, and other sources) along with plant material and survived to be sequenced from the scat samples. The richness per sample (Figure 2.2) and overall diversity suggest that scat samples are appropriate for broad biodiversity surveys. Nevertheless, caution should be applied as there is no guarantee that the diversity detected was not transported from outside the study area. Diet analysis of targeted organisms may be more informative than scats along a transect; for example, analyzing restoration success through diet changes in frugivorous bats (Galimberti et al. 2016).

**Table 2.2** Taxa of Interest

| Taxa | Common names | Substrate | Interest |
|---|---|---|---|
| Dilleniaceae, Fabaceae, Loranthaceae, Myrtaceae, Pinaceae, Poaceae, Proteaceae, Lauraceae | Plant families | All | These families were detected in all substrates |
| Noctuidae, Geometridae, and Pyralidae | Owlet moths, Geometer moths, Grass moths | Scat | Invertebrate families associated with scat samples ($p<0.05$) that reproduce on plant tissue |
| Aphididae, Cecidomyiidae, Noctuidae, Clubionidae, Curculionidae, Galumnidae, and Zyganidae. | Aphids, Gallmidges, Owlet moths, Club spiders, Weevils, Mites/ticks, Burnet moths | Plant | Invertebrate families associated with plant samples ($p<0.05$) that all live and reproduce on plant tissue |
| *Macathuria* | | Plant | Cryptid, often difficult to find because of growth habit. This taxa was found in sequencing data from plant material samples, but not the plant survey results |
| *Vulpes vulpes* | Red Fox | Scat | Invasive species that has led to the decline of native fauna |
| Poaceae, Fabaceae, Asteraceae | Grasses, Legumes, Daisies | Scat | Found in every scat sample in both SCP and the Pilbara, likely common food sources for Macropodidae (Kangaroos and Wallabies) |
| Pinaceae | Pines | Plant | Detected in plant material sequences but not morphological survey, likely sourced from pine plantations near the SCP site |

### 2.6.2 Site differences

The SCP generated higher biodiversity and greater ASV richness than the Pilbara, as expected based on known biodiversity in these regions (Rix et al. 2015). Patterns of diversity were consistent across site; substrates generated similar levels of unique and overlapping diversity. The greatest difference between sites, which performed poorly overall, but particularly so in the Pilbara where only two samples were successfully amplified in two assays. While other studies using soil samples as a substrate were able to reflect the above ground diversity of plants (Yoccoz et al.

2012; Fahner et al. 2016), vertebrates (K. Andersen et al. 2012), and other metazoan (Drummond et al. 2015), less than half of our soil samples successfully amplified. This may be a result of our sampling method and the environment (i.e. hot desert, hot mediterranean) from which the samples were sourced. This is the first study investigating non-microbial soil eDNA in hot Mediterranean and desert climates. DNA at our study sites would be relatively more degraded by the heat (Sirois and Buckley 2019) and high UV radiation (M. A. Barnes et al. 2014) than Denmark (K. Andersen et al. 2012) or New Zealand (Drummond et al. 2015). This might explain why the Pilbara, which is hotter than the SCP and less shaded, had fewer soil samples that amplified successfully. Despite these results, we feel further testing is needed before dismissing soil as a worthwhile substrate for non-microbial eDNA studies in similar environments. We sampled in a relatively hot period and only the surface 0-5cm. It is possible that better results may be achieved by sampling deeper in the soil profile, in cooler weather, and perhaps with more subsamples.

### 2.6.3 Comparison between plant surveys

The majority (83-93%) of morphological plant species identified were from families also detected in the sequencing data. In total, 7 families in the Pilbara and 8 families in the SCP were identified in the morphological survey and not by metabarcoding (Figure 2.5). One possible explanation is that assigning taxonomy to a DNA sequence (i.e. barcode) is dependent on the quality of the reference database. Thompson and Newmaster (2014) found that metabarcoding (rbcL and ITS2 gene regions) was more accurate than their morphology based plant survey. However, they had access to a comprehensive and fully referenced barcoding database for the Boreal forest in Canada. The Pilbara and SCP have a much more diverse flora than the Boreal forests of Canada (Rix et al. 2015), and much of the biodiversity is just beginning to enter barcoding libraries (Dormontt et al., 2018; Nevill et al. 2020). Fortunately, reference databases are continuously growing at an impressive rate often through large-scale initiatives (Hendrich et al. 2015; Costa and Carvalho 2017).

Many (7/12) of the families unique to the morphological surveys were found in very low abundance at only one sample point, and may have been missed in the collection of plant material for metabarcoding. However, another consideration is that DNA is more easily extracted from some plant species than others based on the

amount of secondary metabolites and variation in leaf structure (Khanuja, Shansany, and Kumar 1999; Friar 2005). Some recalcitrant plant species with low DNA concentrations or poor quality extracts may have been drowned out by the more easily processed plants in the mix, resulting in plants not showing up in sequencing data, despite having been collected.

### 2.6.4 Limitations

Our results highlight some important limitations in using metabarcoding for terrestrial biodiversity monitoring. We discussed above the importance of reference databases and extraction bias, but our results also bring into focus the difficulty in determining when DNA might have originated from outside the study area (Figure 2.5). Several plant orders found in the scat and invertebrate samples (Rosales, Vitales, Brassicales) are not present in the plant samples or the plant surveys, indicating they were either missed in surveys or likely originated from nearby suburban gardens. Depending on the study objectives, proponents may want to limit assays and substrates to those that target taxa that are more likely to occur at the sample point (e.g. plant material from plants in a quadrat, invertebrate sequences from bulk invertebrate samples). However, the presence of DNA, even if not proven to be from an organism inhabiting a particular area, does indicate connection between the sample area and those organisms.

## 2.7 Conclusion

We tested five terrestrial eDNA substrates with four metabarcoding assays for a total of 523 sequenced samples, to examine the impact of substrate on eDNA based biodiversity assessment. Our results are consistent with many other eDNA studies by demonstrating the ability of eDNA metabarcoding as a powerful tool for terrestrial biomonitoring, providing a broad survey of terrestrial environments. While we cannot equivocally state that these results would transfer across all biomes, the consistency in patterns across two sites is highly suggestive of strong spatial fidelity. We showed in our systematic comparison of substrates that the choice of substrate heavily dictates what taxa will be detected and that each additional substrate will increase the number of taxa detected. Therefore, substrates should be selected with care based on the purpose of monitoring and available funding (see Table 2.3 for recommendations based on target taxa and survey limitations). For example, soil and plant samples identified relatively few unique families. If these substrates were excluded from this study, 92% of the overall diversity would have still been detected. In contrast, scat and pitfall samples had higher per sample richness, making them suitable for surveys of total biodiversity where funding is limited. We cannot make any specific recommendations on which assays to use, as this is beyond the scope of this paper. Instead we can make broad recommendations for appropriate assay targets based on survey target. For example, where there are concerns about the DNA source, proponents may need to limit surveys to substrates and assay combinations more likely to be locally present (invertebrate assays on bulk arthropods, vertebrate assays on scat etc.)

We also show the limitations of metabarcoding where reference databases are depauperate, and that caution should be exercised with regard to the source of DNA in each substrate, which might originate from organisms outside the immediate study area. Our study highlights the utility of eDNA as biomonitoring tool but also cautions that, like other survey methods, its utility, sensitivity and efficacy will be influenced by how studies are designed and executed.

**Table 2.3** Substrates recommended for eDNA surveys based on target taxa and survey limitations.

| | **Target taxa** | | | |
|---|---|---|---|---|
| **Limitations** | *Invertebrates* | *Plants* | *Vertebrates* | *Total Biodiversity* |
| *None* | Pitfall+Vane | Scat+Plant | Scat | Pitfall+Vane+Scat+Plant |
| *DNA Source* | Pitfall+Vane | Plant | Scat | Pitfall+Vane+Scat+Plant |
| *Funding* | Pitfall/Vane | Scat | Scat | Scat+Pitfall |

* Colour of substrate recommended indicates appropriate assay: Invertebrate assay, Plant assay, Vertebrate Assay, Multiple assays

## 2.8 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.


Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjær, K. H., … Willerslev, E. (2012). Meta-barcoding of "dirt" DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, *21*(8), 1966–1979. doi: 10.1111/j.1365-294X.2011.05261.x

Anderson, I. C., & Cairney, J. W. G. (2004). Diversity and ecology of soil fungal communities: Increased understanding through the application of molecular techniques. *Environmental Microbiology*, *6*(8), 769–779. doi: 10.1111/j.1462-2920.2004.00675.x

Argent, N. (2013). Reinterpreting Core and Periphery in Australia's Mineral and Energy Resources Boom: An Innisian perspective on the Pilbara. *Australian Geographer*, *44*(3), 323–340. doi: 10.1080/00049182.2013.817033

Arteaga Claramunt, A. M., White, N. E., Bunce, M., O'Connell, M., Bullen, R. D., & Mawson, P. R. (2018). Determination of the diet of the ghost bat (Macroderma gigas) in the Pilbara region of Western Australia from dried prey remains and DNA metabarcoding. *Australian Journal of Zoology*, *66*(3), 195. doi: 10.1071/ZO18040

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science and Technology*, *48*(3), 1819–1827. doi: 10.1021/es404734p

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*, 1–12. doi: 10.1038/sdata.2018.214

Blake, M., McKeown, N. J., Bushell, M. L. T., & Shaw, P. W. (2016). DNA extraction from spider webs. *Conservation Genetics Resources*, *8*(3), 219–221. doi: 10.1007/s12686-016-0537-8

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., … de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, *29*(6), 358–367. doi: 10.1016/j.tree.2014.04.003

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*, *16*(1). doi: 10.1111/1755-0998.12428

Burbidge, A. H., Johnstone, R. E., & Pearson, D. J. (2010). A Biodiversity Survey of the Pilbara Region of Western Australia , 2002 – 2007. *Records of the Western Australiam Museum*, *Suppliment*(78), 247–279.

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. doi: 10.1038/nmeth.3869

Calvignac-spencer, A., Merkel, K., & Kutzner, N. (2013). *Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity*. 915–924. doi: 10.1111/mec.12183

Campbell, S. P., Clark, J. A., Crampton, L. H., Guerry, A. D., Hatch, L. T., Hosseini, P. R., … O'Connor, R. (2002). An Assessment of Monitoring Efforts in Endangered Species Recovery Plans. *Ecological Applcations*, *12*(3), 674–681.

Clarke, L. J., Soubrier, J., Weyrich, L. S., & Cooper, A. (2014). Environmental metabarcodes for insects: In silico PCR reveals potential for taxonomic bias. *Molecular Ecology Resources*, *14*(6). doi: 10.1111/1755-0998.12265

Costa, F. O., & Carvalho, G. R. (2017). The Barcode of Life Initiative: synopsis and prospective societal impacts of DNA barcoding of Fish. *Genomics, Society and Policy*, *3*(2), 29–40. doi: 10.1186/1746-5354-3-2-29

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., … Bik, H. M. (2016). The ecologist ' s field guide to sequence-based identification of biodiversity. *Methods in Ecology Adn Evolution*, *7*, 1008–1018. doi: 10.1111/2041-210X.12574

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, *14*(2). doi: 10.1111/1755-0998.12188

De Cáceres, M., & Legendre, P. (2009). Associations between species and groups of sites and statistical inference. *Ecology*, *90*(12), 3566–3574. doi: 10.1890/08-1823.1

de Mattia, F. D. E., Gentili, R., Bruni, I., Galimberti, A., Sgorbati, S., Casiraghi, M., & Labra, M. (2012). A multi-marker DNA barcoding approach to save time and resources in vegetation surveys. *Botanical Journal of Linnean Society*, *169*, 518–529.

Demes, K. W., Graham, M. H., & Suskiewicz, T. S. (2009). Phenotypic plasticity reconciles incongruous molecular and morphological taxonomies: The giant kelp, macrocystis (laminariales, phaeophyceae), is a monospecific genus. *Journal of Phycology*, *45*(6), 1266–1269. doi: 10.1111/j.1529-8817.2009.00752.x

Derocles, S. A. P., Evans, D. M., Nichols, P. C., Evans, S. A., & Lunt, D. H. (2015). Determining Plant – Leaf Miner – Parasitoid Interactions: A DNA Barcoding Approach. *PLOS ONE*, *10*(2), e0117872. doi: 10.1371/journal.pone.0117872

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, *345*(6195), 401–406. doi: 10.1126/science.1251817

Dormontt, E. E., van Dijk, K., 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*(September), 1–12. doi: 10.3389/fevo.2018.00134

Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C. M., … Nelson, N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience*, *4*(46). doi: 10.1186/s13742-015-0086-1

Epanchin-Niell, R. S., Haight, R. G., Berec, L., Kean, J. M., & Liebhold, A. M. (2012). Optimal surveillance and eradication of invasive species in heterogeneous landscapes. *Ecology Letters*, *15*(8), 803–812. doi: 10.1111/j.1461-0248.2012.01800.x

Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-scale monitoring of plants through environmental DNA metabarcoding of soil: Recovery, resolution, and annotation of four DNA markers. *PLoS ONE*, *11*(6), 1–16. doi: 10.1371/journal.pone.0157505

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*, 1–10. doi: 10.1111/rec.12976

Friar, E. A. (2005). Isolation of DNA from Plants with Large Amounts of Secondary Metabolites. *Methods in Enzymology*, *395*, 3–14. doi: 10.1016/S0076-6879(05)95001-5

Galimberti, A., Spinelli, S., Bruno, A., Mezzasalma, V., De Mattia, F., Cortis, P., & Labra, M. (2016). Evaluating the efficacy of restoration plantings through DNA barcoding of frugivorous bird diets. *Conservation Biology*. doi: 10.1111/cobi.12687

Hendrich, L., Morinière, J., Haszprunar, G., Hebert, P. D. N., Hausmann, A., Köhler, F., & Balke, M. (2015). A comprehensive DNA barcode database for Central European beetles with a focus on Germany: adding more than 3500 identified species to BOLD. *Molecular Ecology Resources*, *15*(4), 795–818. doi: 10.1111/1755-0998.12354

Herrick, J. E., Schuman, G. E., & Rango, A. (2006). Monitoring ecological processes for restoration projects. *Journal for Nature Conservation*, *14*(3–4), 161–171. doi: 10.1016/j.jnc.2006.05.001

Huang, J., Yu, H., Guan, X., Wang, G., & Guo, R. (2016). Accelerated dryland expansion under climate change. *Nature Climate Change*, *6*(2), 166–171. doi: 10.1038/nclimate2837

Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Research*, *17*(3), 377–386. doi: 10.1101/gr.5969107

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

Jurado-Rivera, J. A., Vogler, A. P., Reid, C. A. M., Petitpierre, E., & Gómez-Zurita, J. (2009). DNA barcoding insect-host plant associations. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1657), 639–648. doi: 10.1098/rspb.2008.1264

Khanuja, S. P. S., Shansany, A. K., & Kumar, S. (1999). Rapid isolation of DNA from and fresh samples of plants producing large amounts of secondary metabolites and essential Oils. *Plant Molecular Biology Reporter*, *17*, 1–7. doi: 10.1023/A:1007528101452

Koziol, A., Stat, M., Simpson, T., Jarman, S., Harvey, E. S., McDonald, J., … Bunce, M. (2018). Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources*, (September), 1–11. doi: 10.1111/1755-0998.12971

Kraaijeveld, K., de Weger, L. A., Ventayol García, M., Buermans, H., Frank, J., Hiemstra, P. S., & den Dunnen, J. T. (2015). Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Molecular Ecology Resources*, *15*(1), 8–16. doi: 10.1111/1755-0998.12288

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

Murray, D. C., Pearson, S. G., Fullagar, R., Chase, B. M., Houston, J., Atchison, J., … Bunce, M. (2012). High-throughput sequencing of ancient plant and

mammal DNA preserved in herbivore middens. *Quaternary Science Reviews*, *58*, 135–145. doi: 10.1016/J.QUASCIREV.2012.10.021

Nevill, P. G., Zhong, X., Tonti-Filippini, J., Byrne, M., Hislop, M., Thiele, K., van Leeuwen, S., Boykin, L. M., Small, I. (2020). Large scale genome skimming from herbarium material for accurate plant identification and phylogenomics. *Plant Methods, 16*(1), 1-8. doi: 10.1186/s13007-019-0534-5

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., … Maintainer, H. W. (2019). vegan: Community Ecology Package. R package version 2.5-5. https://CRAN.R-project.org/package=vegan. *Community Ecology Package*, *2*(9), 1–297. Retrieved from https://cran.r-project.org/web/packages/vegan/vegan.pdf

Pearson, D. L., Hamilton, A. L., & Erwin, T. L. (2011). Recovery Plan for the Endangered Taxonomy Profession. *BioScience*, *61*(1), 58–63. doi: 10.1525/bio.2011.61.1.11

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., … Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187). doi: 10.1126/science.1246752

Poinar, H. N., Hofreiter, M., Spaulding, W. G., Martin, P. S., Stankiewicz, B. A., Bland, H., … Pääbo, S. (1998). Molecutar Coproscopy : Dung and Diet of the Extinct Ground Stoth. *Science*, *281*(5375), 402–406.

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*(February), 27282. doi: 10.1038/srep27282

Prasifka, J. R., Lopez, M. D., Hellmich, R. L., Lewis, L. C., & Dively, G. P. (2007). Comparison of pitfall traps and litter bags for sampling ground-dwelling arthropods. *Journal of Applied Entomology*, *131*(2), 115–120. doi: 10.1111/j.1439-0418.2006.01141.x

R Core Team. (2018). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. 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), 1–11. doi: 10.1093/nar/gkr732

Richard, B., Decaëns, T., Rougerie, R., James, S. W., Porco, D., & Hebert, P. D. N. (2010). Re-integrating earthworm juveniles into soil biodiversity studies: Species identification through DNA barcoding. *Molecular Ecology Resources*, *10*(4), 606–614. doi: 10.1111/j.1755-0998.2009.02822.x

Rix, M. G., Edwards, D. L., Byrne, M., Harvey, M. S., Joseph, L., & Roberts, J. D. (2015). Biogeography and speciation of terrestrial fauna in the south-western Australian biodiversity hotspot. *Biological Reviews*, *90*(3), 762–793. doi: 10.1111/brv.12132

Santos, S. A. P., Cabanas, J. E., & Pereira, J. A. (2007). Abundance and diversity of soil arthropods in olive grove ecosystem (Portugal): Effect of pitfall trap type. *European Journal of Soil Biology*, *43*(2), 77–83. doi: 10.1016/j.ejsobi.2006.10.001

Sirois, S. H., & Buckley, D. H. (2019). Factors governing extracellular DNA degradation dynamics in soil. *Environmental Microbiology Reports*, *11*(2), 173–184. doi: 10.1111/1758-2229.12725

Sudmeyer, R. (2016). Climate in the Pilbara. *Bulletin 4873*, (Department of Agriculture and Food, Western Australia, Perth).

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., … Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3), e14–e14. doi: 10.1093/nar/gkl938

Taberlet, P., Coissac, E., Pompanon, F. 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

Thompson, K. A., & Newmaster, S. G. (2014). Molecular taxonomic tools provide more accurate estimates of species richness at less cost than traditional

morphology-based taxonomic practices in a vegetation survey. *Biodiversity and Conservation*, *23*(6), 1411–1424. doi: 10.1007/s10531-014-0672-z

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

Titley, M. A., Snaddon, J. L., & Turner, E. C. (2017). Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. *PLoS ONE*, *12*(12), 1–14. doi: 10.1371/journal.pone.0189577

UNEP. (2011). Report of the Tenth Meeting of the Conference of the Parties to the Convention on Biological Diversity. *UNEP/CBD/COP/10/27.*, 18–29 October 2010, Nagoya, Japan.

Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E., Coissac, E., Pompanon, F., … Taberlet, P. (2009). New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: The trnL approach. *Molecular Ecology Resources*, *9*(1), 51–60. doi: 10.1111/j.1755-0998.2008.02352.x

Weigand, A. M., Jochum, A., Pfenninger, M., Steinke, D., & Klussmann-Kolb, A. (2011). A new approach to an old conundrum-DNA barcoding sheds new light on phenotypic plasticity and morphological stasis in microsnails (Gastropoda, Pulmonata, Carychiidae). *Molecular Ecology Resources*, *11*(2), 255–265. doi: 10.1111/j.1755-0998.2010.02937.x

Wiens, J. J. (2016). Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species. *PLoS Biology*, *14*(12), 1–18. doi: 10.1371/journal.pbio.2001104

Williams, A. V., Nevill, P. G., & Krauss, S. L. (2014). Next generation restoration genetics: applications and opportunities. Trends in Plant Science, 19(8), 529-537. doi.org/10.1016/j.tplants.2014.03.011

Yang, C., Wang, X., Miller, J. A., Blécourt, M. De, Ji, Y., Yang, C., … Yu, D. W. (2014). Using metabarcoding to ask if easily collected soil and leaf-litter

samples can be used as a general biodiversity indicator. *Ecological Indicators*, *46*, 379–389. doi: 10.1016/j.ecolind.2014.06.028

Yoccoz, N. G., Brathen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., … Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, *21*, 3647–3655. doi: 10.1111/j.1365-294X.2012.05545.x

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

Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., Lees, D. C., & Jones, G. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, *11*(2), 236–244. doi: 10.1111/j.1755-0998.2010.02920.x

## 2.9 Supplementary Information

**Table S2.9.1** Results of qPCR screening. Higher Cycle threshold (CT) values indicate less DNA available to be amplified.

| Substrate | Assay | % of undetected DNA (CT>37) | avg ct value | sd ct value | sequenced? |
|---|---|---|---|---|---|
| Pitfall | ZBJ-Art | 13.8 | 29.3 | 6.1 | yes |
| | Ant236 | 1.3 | 28.1 | 5.1 | yes |
| | trnL | 11.3 | 29.9 | 5.6 | yes |
| | 12SV5 | 76.3 | 37.8 | 5.1 | no |
| Vane | ZBJ-Art | 11.3 | 27.7 | 5.9 | yes |
| | Ant236 | 7.5 | 30.7 | 4.3 | yes |
| | trnL | 2.5 | 28.9 | 4.6 | yes |
| | 12SV5 | 74.4 | 37.6 | 4.8 | no |
| Scat | ZBJ-Art | 11.1 | 31 | 5.4 | yes |
| | Ant236 | 7.5 | 32.3 | 3.13 | yes |
| | trnL | 1.3 | 24.4 | 3.7 | yes |
| | 12SV5 | 0 | 29.1 | 2.9 | yes |
| Plant | ZBJ-Art | 21.5 | 33.8 | 3.9 | yes |
| | Ant236 | 20 | 34.3 | 2.7 | yes |
| | trnL | 0 | 18.9 | 3.4 | yes |
| | rbcL | 0 | 15.4 | 2.2 | yes |
| | 12SV5 | 100 | 40 | 0 | no |
| Soil | ZBJ-Art | 41.8 | 34.8 | 4.53 | yes |
| | Ant236 | 93 | 39.6 | 1.4 | no |
| | trnL | 61.3 | 36.7 | 4.2 | yes |
| | 12SV5 | 97.6 | 39.8 | 0.9 | no |
| Extracell soil* | trnL | 75 | 37.9 | 4.6 | no |

*The extracellular DNA of 8 soil samples was also extracted using the protocol described by Taberlet et al. (2012) to see if DNA yield improved with this extraction method.

### 2.9.1 Contamination Concerns

Extraction controls are blanks that are run concurrently with the extraction of our samples, and are our primary method of dealing with contamination. Sequence variants found in extraction controls are removed from the dataset before analysis. This worked well to remove contaminants such as human DNA, of which there were 2 ASVs found, both of which were removed because they were also present in the extraction control. Another sequence variant in the 12SV5 data set was also removed as a likely contaminant for a few reasons. First it was a fish family Sigunidae that does not occur in the study areas, second it accounted for only 0.01% of the 12SV5 sequences, and finally because it is a likely lab contaminant, as there was a concurrent project on this family of fish at the time.

**Table S2.9.2** ANOVA and PERMANOVA results for differences in ASV richness between substrates and sites. PERMANOVAs were run using 999 permutations.

**12SV5**

ANOVA, log-transformed richness

| | *Df* | *Sum Sq* | *Mean Sq* | *F value* | *Pr(>F)* | |
|---|---|---|---|---|---|---|
| Site | 1 | 25.2728 | 25.2728 | 2209.1 | < 0.001 | *** |
| Residuals | 38 | 0.4347 | 0.0114 | | | |

**Ant236**

PERMANOVA

| | *Df* | *SumsOfSqs* | *MeanSqs* | *F.Mod* | *R2* | *P* | |
|---|---|---|---|---|---|---|---|
| Site | 1 | 1.7465 | 1.74646 | 27.244 | 0.13361 | 0.001 | *** |
| Substrate | 3 | 3.2397 | 1.07991 | 16.846 | 0.24784 | 0.001 | *** |
| Site:Substrate | 3 | 0.9698 | 0.32328 | 5.0429 | 0.07419 | 0.001 | *** |
| Residuals | 111 | 7.1157 | 0.06411 | | 0.54436 | | |
| Total | 118 | 13.0717 | | | 1 | | |

**ZBJ-Art**

PERMANOVA

| | *Df* | *SumsOfSqs* | *MeanSqs* | *F.Mod* | *R2* | *P* | |
|---|---|---|---|---|---|---|---|
| Site | 1 | 2.4138 | 2.41381 | 30.859 | 0.15081 | 0.001 | *** |
| Substrate | 4 | 1.779 | 0.44474 | 5.6856 | 0.11115 | 0.001 | *** |
| Site:Substrate | 4 | 0.3924 | 0.09811 | 1.2543 | 0.02452 | 0.275 | |
| Residuals | 146 | 11.4204 | 0.07822 | | 0.71352 | | |
| Total | 155 | 16.0056 | | | 1 | | |

**trnL**

| | *Df* | *SumsOfSqs* | *MeanSqs* | *F.Mod* | *R2* | *P* | |
|---|---|---|---|---|---|---|---|
| Site | 1 | 1.7049 | 1.70494 | 48.321 | 0.09129 | 0.001 | *** |
| Substrate | 4 | 9.8189 | 2.45472 | 69.572 | 0.52575 | 0.001 | *** |
| Site:Substrate | 4 | 1.7184 | 0.42961 | 12.176 | 0.09201 | 0.001 | *** |
| Residuals | 154 | 5.4336 | 0.03528 | | 0.29094 | | |
| Total | 163 | 18.6759 | | | 1 | | |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

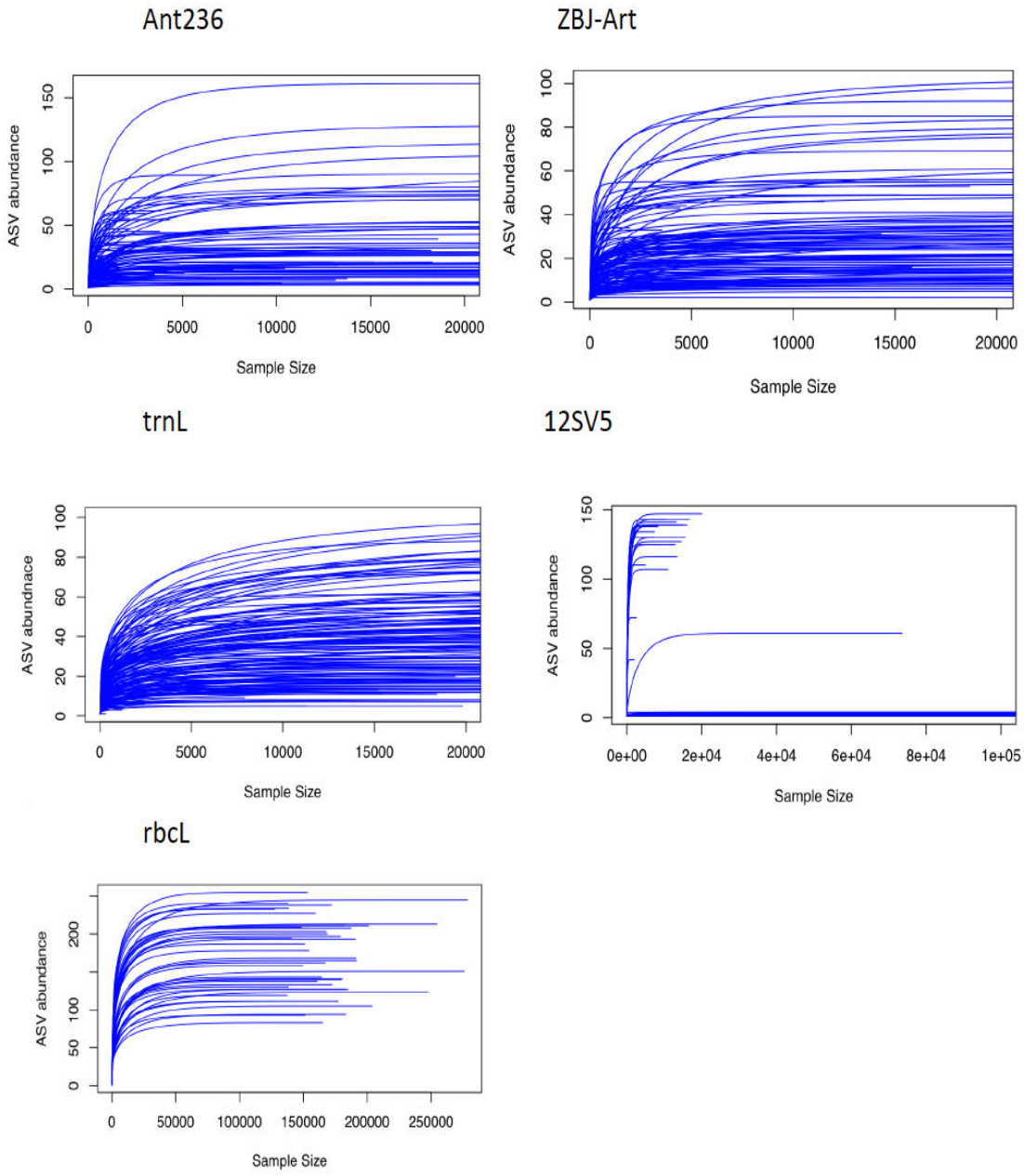


**Figure S2.9.1** Rarefaction curves for each primer set. Created using the 'rarecurve' function from the R package 'vegan' (Oksanen et al. 2018). Rarefaction curves for three assays (trnL, Ant236, MZArt) were cut off at 20 000 sequences to allow lower abundance samples to be seen better.

**Table S2.9.3** PERMANOVA testing the differences in community composition between substrates and sites. Based on family level presence absence table with Bray-curtis similarity applied and 999 permutations.

***PERMANOVA***

| | **Df** | **SumsOfSqs** | **MeanSqs** | **F.Model** | **R2** | **Pr(>F)** |
|---|---|---|---|---|---|---|
| *Site* | 1 | 5.409 | 5.4093 | 38.411 | 0.1013 | 0.001 |
| *Substrate* | 4 | 18.295 | 4.5738 | 32.478 | 0.3427 | 0.001 |
| *Site:Substrate* | 4 | 6.024 | 1.506 | 10.694 | 0.1128 | 0.001 |
| *Residuals* | 168 | 23.659 | 0.1408 | | 0.4432 | |
| *Total* | 177 | 53.388 | | | 1 | |

***Pilbara Pairwise Permanova***

| ***pairs*** | **F.Model** | **R2** | **p.value** | **p.adjusted** |
|---|---|---|---|---|
| *Pitfall vs Plant* | 27.07 | 0.4225085 | 0.001 | 0.01 |
| *Pitfall vs Scat* | 34.462 | 0.4822426 | 0.001 | 0.01 |
| *Pitfall vs Vane* | 11.926 | 0.2437523 | 0.001 | 0.01 |
| *Pitfall vs Soil* | 4.2396 | 0.1824316 | 0.005 | 0.05 |
| *Plant vs Scat* | 42.032 | 0.5251887 | 0.001 | 0.01 |
| *Plant vs Vane* | 13.168 | 0.2573436 | 0.001 | 0.01 |
| *Plant vs Soil* | 14.09 | 0.413323 | 0.006 | 0.06 |
| *Scat vs Vane* | 26.3 | 0.4090243 | 0.001 | 0.01 |
| *Scat vs Soil* | 26.887 | 0.5734444 | 0.005 | 0.05 |
| *Vane vs Soil* | 4.8643 | 0.1956334 | 0.004 | 0.04 |

***SCP pairwise PERMANOVA***

| ***pairs*** | **F.Model** | **R2** | **p.value** | **p.adjusted** |
|---|---|---|---|---|
| *Pitfall vs Plant* | 25.645 | 0.4029362 | 0.001 | 0.01 |
| *Pitfall vs Scat* | 31.128 | 0.4502921 | 0.001 | 0.01 |
| *Pitfall vs Soil* | 30.606 | 0.4665096 | 0.001 | 0.01 |
| *Pitfall vs Vane* | 10.282 | 0.2129645 | 0.001 | 0.01 |
| *Plant vs Scat* | 12.919 | 0.2537189 | 0.001 | 0.01 |
| *Plant vs Soil* | 41.693 | 0.5436369 | 0.001 | 0.01 |
| *Plant vs Vane* | 16.392 | 0.3013712 | 0.001 | 0.01 |
| *Scat vs Soil* | 41.241 | 0.5409324 | 0.001 | 0.01 |
| *Scat vs Vane* | 22.801 | 0.3750107 | 0.001 | 0.01 |
| *Soil vs Vane* | 18.738 | 0.3486921 | 0.001 | 0.01 |

**Table S2.9.4** Results of indicator species analysis for A) Pilbara and B) SCP. These were conducted using the R package 'indicspecies' and show which families are closely associated with the substrates. Only those with p-values <0.1 are shown here.

**A**: Pilbara

| *Families* | *s.Pitfall* | *s.Plant* | *s.Scat* | *s.Soil* | *s.Vane* | *index* | *stat* | *p.value* |
|---|---|---|---|---|---|---|---|---|
| *Macropodidae* | 0 | 0 | 1 | 0 | 0 | 3 | 1 | 0.001 |
| *Formicidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.8607 | 0.001 |
| *Solanaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.8957 | 0.001 |
| *Amaranthaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8176 | 0.001 |
| *Asteraceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8452 | 0.001 |
| *Brassicaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.9271 | 0.001 |
| *Campanulaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8944 | 0.001 |
| *Chenopodiaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8 | 0.001 |
| *Convolvulaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8165 | 0.001 |
| *Cupressaceae* | 0 | 0 | 0 | 0 | 1 | 5 | 0.8581 | 0.001 |
| *Fabaceae* | 0 | 1 | 1 | 0 | 1 | 23 | 0.9244 | 0.001 |
| *Goodeniaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8944 | 0.001 |
| *Malvaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8672 | 0.001 |
| *Rubiaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.8433 | 0.001 |
| *Zygophyllaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.8367 | 0.001 |
| *Poaceae* | 0 | 1 | 1 | 1 | 0 | 22 | 0.8457 | 0.003 |
| *Polygalaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.5916 | 0.005 |
| *Termitidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.5916 | 0.008 |
| *Oxalidaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.7071 | 0.014 |
| *Phyllanthaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.7746 | 0.022 |
| *Dilleniaceae* | 0 | 1 | 1 | 0 | 1 | 23 | 0.7506 | 0.023 |
| *Chromulinaceae* | 0 | 0 | 0 | 1 | 0 | 4 | 0.7071 | 0.025 |
| *Noctuidae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.5916 | 0.027 |
| *Proteaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.6132 | 0.081 |

**B**: SCP

| *Families* | *s.Pitfall* | *s.Plant* | *s.Scat* | *s.Soil* | *s.Vane* | *index* | *stat* | *p.value* |
|---|---|---|---|---|---|---|---|---|
| *Canidae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.59 | 0.001 |
| *Dromaiidae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.55 | 0.001 |
| *Macropodidae* | 0 | 0 | 1 | 0 | 0 | 3 | 1 | 0.001 |
| *Ancylostomatidae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.59 | 0.001 |
| *Asilidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.6 | 0.001 |
| *Bangiaceae* | 0 | 0 | 0 | 1 | 0 | 4 | 0.65 | 0.001 |
| *Cecidomyiidae* | 0 | 1 | 0 | 0 | 0 | 2 | 0.65 | 0.001 |
| *Chloropidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.8 | 0.001 |
| *Entomobryidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.72 | 0.001 |
| *Gracilariaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.77 | 0.001 |
| *Isotomidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.67 | 0.001 |

| *Families* | *s.Pitfall* | *s.Plant* | *s.Scat* | *s.Soil* | *s.Vane* | *index* | *stat* | *p.value* |
|---|---|---|---|---|---|---|---|---|
| *Malawimonadidae* | 0 | 0 | 0 | 1 | 0 | 4 | 0.71 | 0.001 |
| *Noctuidae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.78 | 0.001 |
| *Phoridae* | 0 | 0 | 0 | 0 | 1 | 5 | 0.61 | 0.001 |
| *Pythiaceae* | 0 | 0 | 0 | 1 | 0 | 4 | 0.75 | 0.001 |
| *Scarabaeidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.61 | 0.001 |
| *Sciaridae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.52 | 0.001 |
| *Sphaeroceridae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.59 | 0.001 |
| *Formicidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.82 | 0.001 |
| *Halictidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.9 | 0.001 |
| *Hemiscorpiidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.5 | 0.001 |
| *Leuctridae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.5 | 0.001 |
| *Microbotryaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.52 | 0.001 |
| *Myrtaceae* | 1 | 1 | 1 | 0 | 1 | 27 | 0.95 | 0.001 |
| *Pinaceae* | 0 | 0 | 1 | 0 | 1 | 14 | 0.62 | 0.001 |
| *Staphylinidae* | 0 | 0 | 0 | 0 | 1 | 5 | 0.74 | 0.001 |
| *Aizoaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.82 | 0.001 |
| *Araliaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.55 | 0.001 |
| *Asparagaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.92 | 0.001 |
| *Asteraceae* | 1 | 1 | 1 | 0 | 1 | 27 | 0.94 | 0.001 |
| *Brassicaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.72 | 0.001 |
| *Campanulaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.81 | 0.001 |
| *Chenopodiaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.55 | 0.001 |
| *Convolvulaceae* | 1 | 1 | 1 | 0 | 0 | 16 | 0.78 | 0.001 |
| *Cyperaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.63 | 0.001 |
| *Dilleniaceae* | 1 | 1 | 1 | 0 | 1 | 27 | 0.87 | 0.001 |
| *Ericaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.71 | 0.001 |
| *Fabaceae* | 1 | 1 | 1 | 0 | 1 | 27 | 0.93 | 0.001 |
| *Goodeniaceae* | 1 | 1 | 1 | 0 | 0 | 16 | 0.84 | 0.001 |
| *Haemodoraceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.87 | 0.001 |
| *Iridaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.93 | 0.001 |
| *Macarthuriaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.67 | 0.001 |
| *Menyanthaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.69 | 0.001 |
| *Poaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.89 | 0.001 |
| *Proteaceae* | 1 | 1 | 1 | 0 | 1 | 27 | 0.94 | 0.001 |
| *Restionaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.96 | 0.001 |
| *Rubiaceae* | 1 | 1 | 1 | 0 | 0 | 16 | 0.72 | 0.001 |
| *Rutaceae* | 0 | 1 | 0 | 0 | 0 | 2 | 0.67 | 0.001 |
| *Stylidiaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.86 | 0.001 |
| *Elachistidae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.5 | 0.002 |
| *Sciomyzidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.5 | 0.002 |
| *Aphididae* | 0 | 1 | 0 | 0 | 0 | 2 | 0.45 | 0.003 |
| *Tarsonemidae* | 0 | 1 | 0 | 0 | 0 | 2 | 0.45 | 0.003 |

| *Families* | *s.Pitfall* | *s.Plant* | *s.Scat* | *s.Soil* | *s.Vane* | *index* | *stat* | *p.value* |
|---|---|---|---|---|---|---|---|---|
| *Cupressaceae* | 1 | 0 | 0 | 0 | 1 | 9 | 0.5 | 0.003 |
| *Lamiaceae* | 0 | 1 | 1 | 0 | 0 | 10 | 0.5 | 0.003 |
| *Acanthamoebidae* | 0 | 0 | 1 | 1 | 0 | 13 | 0.43 | 0.004 |
| *Chromulinaceae* | 0 | 0 | 1 | 1 | 0 | 13 | 0.49 | 0.004 |
| *Hypogastruridae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.45 | 0.004 |
| *Casuarinaceae* | 0 | 1 | 1 | 0 | 1 | 23 | 0.57 | 0.004 |
| *Lauraceae* | 1 | 1 | 1 | 0 | 0 | 16 | 0.56 | 0.004 |
| *Crabronidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.45 | 0.005 |
| *Pyralidae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.45 | 0.005 |
| *Carabidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.45 | 0.006 |
| *Chironomidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.46 | 0.006 |
| *Lycosidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.46 | 0.006 |
| *Orchidaceae* | 0 | 1 | 0 | 0 | 0 | 2 | 0.45 | 0.006 |
| *Rosaceae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.45 | 0.006 |
| *Acrididae* | 1 | 0 | 0 | 0 | 1 | 9 | 0.46 | 0.007 |
| *Cicadellidae* | 0 | 0 | 0 | 0 | 1 | 5 | 0.45 | 0.023 |
| *Asphodelaceae* | 0 | 1 | 1 | 0 | 1 | 23 | 0.49 | 0.023 |
| *Clubionidae* | 0 | 1 | 0 | 0 | 0 | 2 | 0.4 | 0.025 |
| *Pygmephoridae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.39 | 0.027 |
| *Malvaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.4 | 0.028 |
| *Mydidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.39 | 0.029 |
| *Primulaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.39 | 0.029 |
| *Amaryllidaceae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.39 | 0.03 |
| *Oxalidaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.4 | 0.03 |
| *Haemonchidae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.39 | 0.031 |
| *Muscidae* | 0 | 0 | 1 | 0 | 1 | 14 | 0.39 | 0.033 |
| *Zygaenidae* | 0 | 1 | 0 | 0 | 0 | 2 | 0.39 | 0.034 |
| *Colletidae* | 0 | 0 | 0 | 0 | 1 | 5 | 0.39 | 0.038 |
| *Euphorbiaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.39 | 0.038 |
| *Galumnidae* | 1 | 1 | 0 | 0 | 0 | 6 | 0.39 | 0.041 |
| *Geometridae* | 0 | 0 | 1 | 0 | 0 | 3 | 0.39 | 0.041 |
| *Vitaceae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.39 | 0.041 |
| *Curculionidae* | 0 | 1 | 1 | 0 | 1 | 23 | 0.43 | 0.042 |
| *Blaberidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0.34 | 0.085 |
| *Hemerobiidae* | 0 | 0 | 0 | 0 | 1 | 5 | 0.34 | 0.092 |
| *Eupodidae* | 1 | 1 | 1 | 0 | 0 | 16 | 0.43 | 0.095 |

**Table S2.9.5** Families detected in each of the eDNA Substrates. Numbers indicate the number of samples in which the family was detected, and they are presented with the most common families on top

| | *Pilbara* | | | | | *SCP* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Family*** | ***Pitfall*** | ***Plant*** | ***Scat*** | ***Soil*** | ***Vane*** | ***Pitfall*** | ***Plant*** | ***Scat*** | ***Soil*** | ***Vane*** |
| *Fabaceae* | 4 | 19 | 20 | 0 | 16 | 18 | 20 | 20 | 1 | 12 |
| *Myrtaceae* | 6 | 6 | 7 | 1 | 14 | 20 | 20 | 16 | 3 | 20 |
| *Proteaceae* | 3 | 12 | 7 | 0 | 2 | 18 | 20 | 20 | 7 | 20 |
| *Dilleniaceae* | 6 | 18 | 8 | 0 | 13 | 13 | 20 | 18 | 1 | 10 |
| *Poaceae* | 9 | 17 | 20 | 2 | 8 | 5 | 19 | 20 | 1 | 3 |
| *Asteraceae* | 5 | 0 | 20 | 0 | 3 | 20 | 19 | 20 | 0 | 11 |
| *Convolvulaceae* | 0 | 6 | 20 | 0 | 4 | 11 | 16 | 14 | 0 | 5 |
| *Goodeniaceae* | 4 | 1 | 20 | 0 | 0 | 17 | 14 | 14 | 0 | 3 |
| *Rubiaceae* | 1 | 18 | 14 | 0 | 3 | 9 | 13 | 11 | 0 | 2 |
| *Formicidae* | 20 | 1 | 0 | 0 | 6 | 19 | 0 | 1 | 0 | 7 |
| *Solanaceae* | 1 | 19 | 19 | 0 | 6 | 4 | 2 | 2 | 0 | 0 |
| *Iridaceae* | 0 | 0 | 1 | 0 | 3 | 3 | 20 | 19 | 0 | 2 |
| *Restionaceae* | 0 | 0 | 2 | 0 | 0 | 1 | 19 | 19 | 0 | 0 |
| *Macropodidae* | 0 | 0 | 20 | 0 | 0 | 0 | 0 | 20 | 0 | 0 |
| *Noctuidae* | 0 | 8 | 6 | 0 | 0 | 0 | 14 | 11 | 0 | 1 |
| *Brassicaceae* | 0 | 0 | 19 | 0 | 2 | 2 | 0 | 14 | 0 | 3 |
| *Campanulaceae* | 0 | 0 | 16 | 0 | 0 | 2 | 2 | 17 | 0 | 1 |
| *Stylidiaceae* | 0 | 0 | 1 | 0 | 0 | 3 | 14 | 19 | 0 | 1 |
| *Aizoaceae* | 0 | 2 | 2 | 0 | 0 | 5 | 2 | 20 | 0 | 3 |
| *Asparagaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 18 | 0 | 0 |
| *Cupressaceae* | 1 | 0 | 3 | 0 | 18 | 5 | 0 | 1 | 0 | 6 |
| *Pinaceae* | 1 | 0 | 0 | 0 | 3 | 4 | 1 | 11 | 1 | 9 |
| *Haemodoraceae* | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 18 | 0 | 0 |
| *Cecidomyiidae* | 0 | 2 | 1 | 0 | 3 | 0 | 14 | 5 | 1 | 3 |
| *Amaranthaceae* | 0 | 6 | 19 | 0 | 2 | 0 | 0 | 2 | 0 | 0 |
| *Casuarinaceae* | 1 | 0 | 0 | 0 | 3 | 3 | 11 | 6 | 0 | 5 |
| *Malvaceae* | 1 | 3 | 19 | 0 | 1 | 0 | 1 | 4 | 0 | 0 |
| *Chenopodiaceae* | 1 | 2 | 16 | 0 | 1 | 0 | 1 | 7 | 0 | 0 |
| *Chloropidae* | 5 | 0 | 0 | 0 | 1 | 16 | 0 | 2 | 0 | 2 |
| *Ericaceae* | 0 | 0 | 0 | 0 | 0 | 3 | 7 | 16 | 0 | 0 |
| *Lauraceae* | 0 | 0 | 0 | 0 | 1 | 4 | 5 | 12 | 2 | 0 |
| *Menyanthaceae* | 1 | 0 | 0 | 0 | 0 | 2 | 7 | 14 | 0 | 0 |
| *Araliaceae* | 1 | 0 | 0 | 0 | 0 | 3 | 8 | 8 | 0 | 2 |
| *Phoridae* | 2 | 0 | 0 | 0 | 3 | 3 | 2 | 0 | 0 | 11 |
| *Pythiaceae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 5 | 14 | 0 |
| *Cyperaceae* | 1 | 0 | 1 | 0 | 0 | 1 | 5 | 12 | 0 | 0 |

| Family | Pilbara Pitfall | Plant | Scat | Soil | Vane | SCP Pitfall | Plant | Scat | Soil | Vane |
|---|---|---|---|---|---|---|---|---|---|---|
| *Halictidae* | 0 | 0 | 0 | 0 | 1 | 17 | 0 | 0 | 0 | 1 |
| *Macarthuriaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 13 | 0 | 1 |
| *Asphodelaceae* | 0 | 0 | 0 | 0 | 0 | 2 | 7 | 5 | 0 | 4 |
| *Entomobryidae* | 2 | 0 | 0 | 0 | 0 | 12 | 2 | 0 | 0 | 0 |
| *Dasypogonaceae* | 0 | 0 | 0 | 0 | 0 | 4 | 3 | 7 | 0 | 2 |
| *Loranthaceae* | 0 | 0 | 2 | 0 | 0 | 2 | 5 | 3 | 1 | 3 |
| *Acrididae* | 2 | 1 | 0 | 0 | 0 | 7 | 1 | 1 | 0 | 3 |
| *Cicadellidae* | 0 | 0 | 0 | 0 | 3 | 2 | 1 | 2 | 0 | 7 |
| *Eupodidae* | 2 | 0 | 0 | 0 | 0 | 5 | 4 | 3 | 1 | 0 |
| *Microbotryaceae* | 0 | 2 | 4 | 0 | 0 | 0 | 2 | 7 | 0 | 0 |
| *Lamiaceae* | 0 | 3 | 1 | 0 | 1 | 0 | 3 | 7 | 0 | 0 |
| *Oxalidaceae* | 0 | 0 | 10 | 0 | 0 | 1 | 0 | 4 | 0 | 0 |
| *Zygophyllaceae* | 0 | 0 | 14 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Staphylinidae* | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 12 |
| *Rutaceae* | 0 | 0 | 2 | 0 | 1 | 0 | 10 | 0 | 0 | 1 |
| *Gelechiidae* | 3 | 7 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 |
| *Carabidae* | 3 | 0 | 0 | 0 | 0 | 6 | 0 | 1 | 0 | 2 |
| *Gracilariaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 0 | 0 |
| *Scarabaeidae* | 0 | 0 | 0 | 0 | 1 | 9 | 0 | 0 | 0 | 2 |
| *Phyllanthaceae* | 0 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Asilidae* | 2 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 1 |
| *Bangiaceae* | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 9 | 0 |
| *Curculionidae* | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 3 | 0 | 4 |
| *Malawimonadidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 10 | 0 |
| *Sciaridae* | 2 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 2 |
| *Ancylostomatidae* | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 7 | 0 | 0 |
| *Chromulinaceae* | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 4 | 5 | 0 |
| *Lycosidae* | 4 | 0 | 0 | 0 | 0 | 5 | 0 | 1 | 0 | 0 |
| *Miridae* | 0 | 3 | 0 | 0 | 4 | 1 | 0 | 0 | 0 | 2 |
| *Psyllidae* | 0 | 4 | 0 | 0 | 4 | 0 | 0 | 1 | 0 | 1 |
| *Acanthamoebidae* | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 4 | 3 | 0 |
| *Isotomidae* | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 0 |
| *Canidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 7 | 0 | 0 |
| *Chironomidae* | 0 | 1 | 0 | 0 | 1 | 5 | 0 | 1 | 0 | 0 |
| *Dolichopodidae* | 0 | 0 | 0 | 0 | 4 | 2 | 0 | 0 | 0 | 1 |
| *Sphaeroceridae* | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 |
| *Thripidae* | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 3 | 0 | 1 |
| *Termitidae* | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Opiliaceae* | 0 | 5 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Polygalaceae* | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Dromaiidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 |
| *Galumnidae* | 0 | 0 | 0 | 0 | 0 | 4 | 2 | 0 | 0 | 0 |

| Family | Pilbara Pitfall | Plant | Scat | Soil | Vane | SCP Pitfall | Plant | Scat | Soil | Vane |
|---|---|---|---|---|---|---|---|---|---|---|
| *Geometridae* | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 3 | 0 | 0 |
| *Muscidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 2 |
| *Chrysomelidae* | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 2 |
| *Sporidiobolaceae* | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 2 | 0 | 0 |
| *Gyrostemonaceae* | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Rosaceae* | 1 | 0 | 1 | 0 | 0 | 4 | 0 | 0 | 0 | 0 |
| *Clubionidae* | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 1 |
| *Crabronidae* | 0 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 |
| *Elachistidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 |
| *Elateridae* | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| *Pyralidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 4 | 0 | 0 |
| *Saprolegniaceae* | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 0 |
| *Sciomyzidae* | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 |
| *Tachinidae* | 0 | 0 | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 1 |
| *Triozidae* | 1 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Hemiscorpiidae* | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 |
| *Leuctridae* | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 |
| *Orchidaceae* | 0 | 1 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 |
| *Acanthaceae* | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Apocynaceae* | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Santalaceae* | 0 | 1 | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| *Aphididae* | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 |
| *Blaberidae* | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 |
| *Blattidae* | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| *Cosmopterigidae* | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 |
| *Crambidae* | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Hemerobiidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 |
| *Tarsonemidae* | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 |
| *Zygaenidae* | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 |
| *Hypogastruridae* | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 |
| *Rhytididae* | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 0 | 0 |
| *Amaryllidaceae* | 0 | 0 | 0 | 0 | 1 | 3 | 0 | 0 | 0 | 0 |
| *Boraginaceae* | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| *Haloragaceae* | 0 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Sapindaceae* | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 |
| *Vitaceae* | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| *Araneidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 |
| *Coniopterygidae* | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| *Culicidae* | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 |
| *Diapriidae* | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| *Gryllidae* | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Limoniidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 |

| Family | Pilbara Pitfall | Plant | Scat | Soil | Vane | SCP Pitfall | Plant | Scat | Soil | Vane |
|---|---|---|---|---|---|---|---|---|---|---|
| *Membracidae* | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| *Nemesiidae* | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| *Nymphalidae* | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 |
| *Oxyopidae* | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Plutellidae* | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| *Salticidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 |
| *Theridiidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 |
| *Apidae* | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| *Colletidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| *Cynipidae* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 |
| *Eulophidae* | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| *Haemonchidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 |
| *Lucanidae* | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| *Mydidae* | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| *Mymaridae* | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Parthenopidae* | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Pygmephoridae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 |
| *Rhiniidae* | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| *Tettigoniidae* | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| *Ustilaginaceae* | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Apiaceae* | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Euphorbiaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 |
| *Meliaceae* | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 |
| *Musaceae* | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| *Oleaceae* | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| *Primulaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 |
| *Typhaceae* | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| *Camelidae* | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Brentidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Cleridae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| *Ephydridae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| *Ichneumonidae* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Lonchaeidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| *Lumbricidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| *Melyridae* | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Pipunculidae* | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| *Pteromalidae* | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| *Scolopendridae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| *Simuliidae* | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| *Tingidae* | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Yponomeutidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| *Baetidae* | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| Family | Pilbara Pitfall | Plant | Scat | Soil | Vane | SCP Pitfall | Plant | Scat | Soil | Vane |
|---|---|---|---|---|---|---|---|---|---|---|
| *Bombyliidae* | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| *Buprestidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| *Dictyonellidae* | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Histeridae* | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| *Proctophyllodidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| *Psychodidae* | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| *Reduviidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| *Vannellidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| *Commelinaceae* | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Geraniaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| *Lythraceae* | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| *Marsileaceae* | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Mazaceae* | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Moraceae* | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Plantaginaceae* | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Zamiaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| *Cacatuidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Falconidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Leporidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Scincidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Varanidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Agromyzidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Anthicidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Bethylidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Braconidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Chrysopidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Corinnidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Cryptocercidae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Cunaxidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| *Cymbaeremaeidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Dryinidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Dytiscidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Ectobiidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Encyrtidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Gnaphosidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Hyalospheniidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Hybotidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Limacodidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Lygaeidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Machilidae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Macrobiotidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Mantidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

| Family | Pilbara Pitfall | Plant | Scat | Soil | Vane | SCP Pitfall | Plant | Scat | Soil | Vane |
|---|---|---|---|---|---|---|---|---|---|---|
| *Nanorchestidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Nepticulidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Nitidulidae* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Peronosporaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Platygastridae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Punctoribatidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Rhyparochromidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Roeslerstammiidae* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Scolopocryptopidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Selenopidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Steinernematidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Syrphidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Tetranychidae* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Thomisidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Tortricidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Trigonidiidae* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Vaejovidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Amaurobiidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Anthocoridae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Anthomyiidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Anystidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Bothriuridae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Bovidae* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Castniidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Ceratopogonidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Chalcididae* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Chondrinidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Cicadidae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Coccinellidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Coleophoridae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Corallinaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Dictynidae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Drosophilidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Erythraeidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Glossosomatidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Haustoriidae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Lauxaniidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Leptohyphidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Lycaenidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Malasseziaceae* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Mithracidae* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Mordellidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

| Family | Pilbara Pitfall | Plant | Scat | Soil | Vane | SCP Pitfall | Plant | Scat | Soil | Vane |
|---|---|---|---|---|---|---|---|---|---|---|
| *Muridae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Neochloridaceae* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Niphargidae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Notodontidae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Oecophoridae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Oribotritiidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Peltoperlidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Pleosporaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *Polyplacidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Pompilidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Potamonautidae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Psephenidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Pseudomonadaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Psocidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Psychidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Sordariaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Sphecidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Tenebrionidae* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Tephritidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Tubificidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Varunidae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Vespidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *Zenarchopteridae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Actinidiaceae* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Celastraceae* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *Crassulaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Cyatheaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| *Fagaceae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Loganiaceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Picrodendraceae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *Pittosporaceae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Podocarpaceae* | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

**Table S2.9.6** Families detected in the morphological plant survey. There were 20 sample points surveyed in the SCP and 8 in the Pilbara. Number indicate the number of surveys in which the family was identified, and the percent of surveys in which the family was identified.

| *Family* | *SCP* | *Pilbara* | *SCP (%)* | *Pilbara (%)* |
|---|---|---|---|---|
| *Fabaceae* | 20 | 8 | 100 | 100 |
| *Poaceae* | 19 | 8 | 95 | 100 |
| *Proteaceae* | 20 | 5 | 100 | 62.5 |
| *Asteraceae* | 20 | 3 | 100 | 37.5 |
| *Goodeniaceae* | 8 | 7 | 40 | 87.5 |
| *Amaranthaceae* | 0 | 8 | 0 | 100 |
| *Iridaceae* | 20 | 0 | 100 | 0 |
| *Myrtaceae* | 20 | 0 | 100 | 0 |
| *Scrophulariaceae* | 0 | 8 | 0 | 100 |
| *Asparagaceae* | 19 | 0 | 95 | 0 |
| *Dilleniaceae* | 19 | 0 | 95 | 0 |
| *Anarthriaceae* | 18 | 0 | 90 | 0 |
| *Ericaceae* | 18 | 0 | 90 | 0 |
| *Solanaceae* | 0 | 7 | 0 | 87.5 |
| *Rubiaceae* | 1 | 6 | 5 | 75 |
| *Stylidaceae* | 16 | 0 | 80 | 0 |
| *Cyperaceae* | 3 | 5 | 15 | 62.5 |
| *Haemodoraceae* | 15 | 0 | 75 | 0 |
| *Hypoxidaceae* | 15 | 0 | 75 | 0 |
| *Rutaceae* | 15 | 0 | 75 | 0 |
| *Pteridaceae* | 0 | 6 | 0 | 75 |
| *Brassicaceae* | 0 | 5 | 0 | 62.5 |
| *Restionaceae* | 12 | 0 | 60 | 0 |
| *Araliaceae* | 11 | 0 | 55 | 0 |
| *Casuarinaceae* | 11 | 0 | 55 | 0 |
| *Convolvulaceae* | 0 | 4 | 0 | 50 |
| *Xanthorrhoeaceae* | 9 | 0 | 45 | 0 |
| *Colchicaceae* | 8 | 0 | 40 | 0 |
| *Loranthaceae* | 8 | 0 | 40 | 0 |
| *Chenopodiaceae* | 0 | 3 | 0 | 37.5 |
| *Santalaceae* | 0 | 3 | 0 | 37.5 |
| *Malvaceae* | 0 | 3 | 0 | 37.5 |
| *Hemerocallidaceae* | 7 | 0 | 35 | 0 |
| *Caryophyllaceae* | 1 | 2 | 5 | 25 |
| *Lauraceae* | 6 | 0 | 30 | 0 |
| *Lamiaceae* | 5 | 0 | 25 | 0 |

| *Family* | *SCP* | *Pilbara* | *SCP (%)* | *Pilbara (%)* |
|---|---|---|---|---|
| *Apiaceae* | 4 | 0 | 20 | 0 |
| *Dasypogonaceae* | 4 | 0 | 20 | 0 |
| *Zamiaceae* | 4 | 0 | 20 | 0 |
| *Montiaceae* | 0 | 1 | 0 | 12.5 |
| *Sapindaceae* | 0 | 1 | 0 | 12.5 |
| *Boraginaceae* | 0 | 1 | 0 | 12.5 |
| *Campanulaceae* | 2 | 0 | 10 | 0 |
| *Aizoaceae* | 1 | 0 | 5 | 0 |
| *Crassulaceae* | 1 | 0 | 5 | 0 |
| *Droseraceae* | 1 | 0 | 5 | 0 |
| *Euphorbiaceae* | 1 | 0 | 5 | 0 |
| *Orchidaceae* | 1 | 0 | 5 | 0 |
| *Orobanchaceae* | 1 | 0 | 5 | 0 |
| *Polygalaceae* | 1 | 0 | 5 | 0 |
| *Violaceae* | 1 | 0 | 5 | 0 |

# CHAPTER 3

## CHANGES IN SOIL MICROBIAL COMMUNITIES IN POST MINE SITE ECOLOGICAL RESTORATION: IMPLICATIONS FOR MONITORING USING HIGH THROUGHPUT DNA SEQUENCING

## 3.1 Preface

*This chapter consists of a published manuscript titled '**Changes in soil microbial communities in post mine ecological restoration: implications for monitoring using high throughput sequencing**' [Science of the Total Environment 2020]. The content in section 3.2 is the same as the published manuscript with only minor changes in formatting to accommodate thesis referencing.*

This chapter examines the potential use of high throughput sequencing (HTS) of soil microbial communities (SMC) to monitor mine site restoration. HTS has been used to characterize SMC for over a decade, and it has become standard technique in many studies looking at the effects of disturbances and other treatments on SMC. SMC form the functional basis of ecosystems, with important roles in plant growth, organic matter decomposition, and nutrient cycling. They also respond quickly to changes in the environment, and they have been suggested as promising targets for restoration monitoring. Studies on SMC changes during restoration are typically in a single location and use either bacteria or fungi. This does not tell us how consistent or inconsistent SMC changes may be across locations, which will affect their usefulness for monitoring. This study characterized SMC changes during restoration across three ecologically different locations in Western Australia, using both bacterial and fungal assays, to examine the implications for restoration monitoring.

### 3.1.1 Acknowledgements

This work was supported by the Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration (ICI150100041) and the Pawsey Supercomputing Centre, funding from the Australian Government and the Government of Western Australia. We thank the mining companies BHP, Hanson Construction Material, and South32 for facilitating access to sites for sampling. We would also like to thank Sheree Walters for help with sample collection and the members of the Trace and Environmental DNA (TrEnD) Laboratory for support with metabarcoding workflows and bioinformatics.

### 3.1.2 Data Accessibility

All sequencing data and DADA2 scripts can be found online at https://doi.org/10.5061/dryad.4qrfj6q7g

### 3.1.3 Author Contributions

MvH conducted the study and wrote the manuscript. MvH, PN, MB, NW, and GW-J were involved in the experimental design. Samples were collected and processed by MvH; molecular and bioinformatics work was performed by MvH; all data was analyzed and processed by MvH; statistical analysis was done by MvH; the manuscript was edited by all authors.

## CHANGES IN SOIL MICROBIAL COMMUNITIES IN POST MINE ECOLOGICAL RESTORATION: IMPLICATIONS FOR MONITORING USING HIGH THROUGHPUT DNA SEQUENCING

*Science of the Total Environment* 749 (2020), 142262
https://doi.org/10.1016/j.scitotenv.2020.142262

van der Heyde, M.[1,2*], Bunce, M.[2,3], Dixon, K.[1], Wardell-Johnson, G.[1], White, N.E.[2], Nevill, P.[1,2]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, GPO Box U1987, Perth, Western Australia, 6845

[2]Trace and Environmental DNA Laboratory, School of Life and Molecular Sciences, Curtin University, GPO Box U1987, Perth, Western Australia, 6845

[3]Environmental Protection Authority, 215 Lambton Quay, Wellington 6011, New Zealand.

*Corresponding author

## 3.2 Abstract

The ecological restoration of ecosystem services and biodiversity is a key intervention used to reverse the impacts of anthropogenic activities such as mining. Assessment of the performance of restoration against completion criteria relies on biodiversity monitoring. However, monitoring usually overlooks soil microbial communities (SMC), despite increased awareness of their pivotal role in many ecological functions. Recent advances in cost, scalability and technology has led to DNA sequencing being considered as a cost-effective biological monitoring tool, particularly for otherwise difficult to survey groups such as microbes. However, such approaches for monitoring complex restoration sites such as post-mined landscapes have not yet been tested. Here we examine bacterial and fungal communities across chronosequences of mine site restoration at three locations in Western Australia to determine if there are consistent changes in SMC diversity, community composition and functional capacity. Although we detected directional changes in community composition indicative of microbial recovery, these were inconsistent between locations and microbial taxa (bacteria or fungi). Assessing functional diversity provided greater understanding of changes in site conditions and microbial recovery than could be determined through assessment of community composition alone. These results demonstrate that high-throughput amplicon sequencing of environmental DNA (eDNA) is an effective approach for monitoring the complex changes in SMC following restoration. Future monitoring of mine site restoration using eDNA should consider archiving samples to provide improved understanding of changes in communities over time. Expansion to include other biological groups (e.g. soil fauna) and substrates would also provide a more holistic understanding of biodiversity recovery.

## 3.3 Introduction

The loss of biodiversity and ecosystem services due to land degradation is a global crisis that undermines the wellbeing of 3.2 billion people, costing approximately 10% of annual gross domestic product, and prompting the United Nations to declare a UN decade on Ecosystem Restoration (2021-2030) (UN Environment Programme, 2020). Mining represents an extreme form of land degradation, where the ecosystem of a site is removed and is reinstated or rehabilitated often at large scales. For example, the footprint of land degraded by mining in China alone is 3.7 million hectares, an area the size of the Netherlands (Li, 2006). The cost of landforming and restoring such large areas is high (Menz, Dixon, & Hobbs, 2013), with a per hectare cost of up to AU$34,000 in Australia (Gardner & Bell, 2007). Understanding if a restoration target has been reached (see Gann et al. 2019) requires accurate monitoring to ensure the best return on investment is achieved, indicate when restoration has been successful, and apply adaptive management principles to future restoration projects (Herrick, Schuman, & Rango 2006; Miller et al. 2017). There is increasing realization that more nuanced approaches to restoration are needed that take into account the return of ecosystem services and the interactions that occur between lifeforms, from microbes to mammals.

Soil microbial communities (SMC) represent emerging targets for restoration monitoring (Harris, 2003; Nurulita et al., 2016; Gellie, Mills, Breed, & Lowe, 2017; Sun et al., 2017; Yan et al., 2018) as they provide a functional basis for ecosystems and are key agents in the soil-root interface involved in nutrient cycling and decomposition (Meena, Mishra, Bisht, & Pattanayak, 2017), plant performance and community composition (Yang et al., 2018). They also respond rapidly to changes in the environment and are easily affected by soil chemistry (Leff et al., 2015; Šmejkalová, Mikanová, & Borůvka, 2003), physical soil disturbance (Dong et al., 2017; Kabiri, Raiesi, & Ghazavi, 2016), and plant communities (Burns, Anacker, Strauss, & Burke, 2015). As a result, characterizing soil communities could provide indicators of edaphic and biotic capabilities in restoration and act as early indicators of problems or predict restoration trajectory (Muñoz-Rojas, 2018).

Soil bacteria are the most abundant form of soil microbes and have growth rates 10-fold faster than fungi. As a result, they also tend to have higher variation over time (Sun et al., 2017). With slower growth rates, fungi are often more disturbed by soil modifications as they are suppressed by nutrient addition (Rajapaksha, Tobor-Kaplon, & Baath, 2004; Suzuki, Nagaoka, Shimada, & Takenaka 2009) and disruptions to their hyphal networks (Dong et al., 2017; Frey Elliott, & Paustian, 1999). However, patterns of SMC responses to changes are often inconsistent and difficult to predict (Dong et al., 2017; Sipilä, Yrjälä, Alakukku, & Palojärvi, 2012), likely due to the highly diverse and variable nature of these communities.

Several studies have assessed aspects of SMC recovery in a restoration context; most commonly reporting decreased biomass and activity in restoration, and different community composition (Mummey, Stahl, & Buyer, 2002a; Muñoz-Rojas, Erickson, Dixon, & Merritt, 2016; Yan et al., 2018). The increased availability over the last 15 years of high-throughput sequencing has made available an increasingly cost-effective way to monitor SMC community composition (Yan et al., 2018). Several studies have found directional changes in community composition with restoration age (Banning et al. 2011; Gellie et al., 2017; Sun et al., 2017; Yan et al., 2020, 2018). However, these patterns are typically complex. For example, Banning et al. (2011) found that bacteria (and not fungal) communities showed directional changes, with older restoration sites more similar to reference communities. In contrast, Sun et al. (2017) found that fungal communities showed more distinct differences between restoration ages than bacteria. Few if any studies have looked at SMC of restoration at multiple locations. Most studies are limited to either a single restoration and reference site (e.g. Mummey, Stahl, & Buyer, 2002b; Muñoz-Rojas et al., 2016) or one chronosequence of restoration sites (e.g. Gellie et al., 2017; Sun et al., 2017; Yan et al., 2020, 2018). Recent studies using high throughput sequencing of SMC for monitoring highlight the need to firstly test for consistency across locations, and secondly, to define the functional significance of the measured SMC diversity (Gellie et al., 2017; Yan et al., 2018). To accurately assess consistency, it is important to use the same methodological and analytical framework to account for any biases.

By integrating measures of SMC diversity, community composition, and microbial functionality, we test the hypotheses that restoration of mine sites will lead to the recovery of SMC. We used high throughput (amplicon) sequencing of fungi (ITS2) and bacteria (16S) to examine changes in SMC across mine site restoration chronosequences at three locations. We aim to improve the application of high-throughput amplicon sequencing to restoration monitoring by addressing the following questions:

1) Are soil bacterial and fungal communities in older restoration sites more similar to reference communities than those at younger restoration sites?

2) Which functional groups are indicators of the different stages of restoration?

3) Are soil chemical properties (moisture, potassium, carbon, etc.) associated with restoration age, and/or change in bacterial and fungal community composition?

4) Are consistent patterns observed across the three study locations?

The aim of this work is to provide recommendations for future implementation of high throughput sequencing as a more holistic monitoring tool for restoration.

## 3.4 Material and methods

### 3.4.1 Study sites

The term ‘chronosequence’ describes a set of ecological sites that share similar attributes but represent different ages. Traditionally these have been used to describe sites with the same parent material with different periods of soil formation (Stevens & Walker, 1970), but it has also been used to refer to sites with different ages of restoration (Banning et al., 2011; Harris, 2003). Three chronosequences of mine site restoration were studied from three locations in Western Australia; Swan Coastal Plain (SCP), Jarrah Forest (JF) and hot desert Pilbara (PB). Each showed consistency in restoration approaches, soil type, climate and site aspect within the location. All three locations used topsoil in their restoration, and these were stripped to consistent depth within each location and homogenized before application. At each chronosequence, sites of different restoration age were sampled as well as two spatially separated reference sites (see Figure S3.9.1). Reference sites were selected for their proximity to restoration sites and similarity to ecosystems mining companies were attempting to restore. To our knowledge, none of the reference sites were recently impacted by disturbances such as overgrazing or fire. At all three locations, we sampled at least two sites less than 9 years old (Young), and at least two sites older than 9 years (Older) (Figure 3.1).

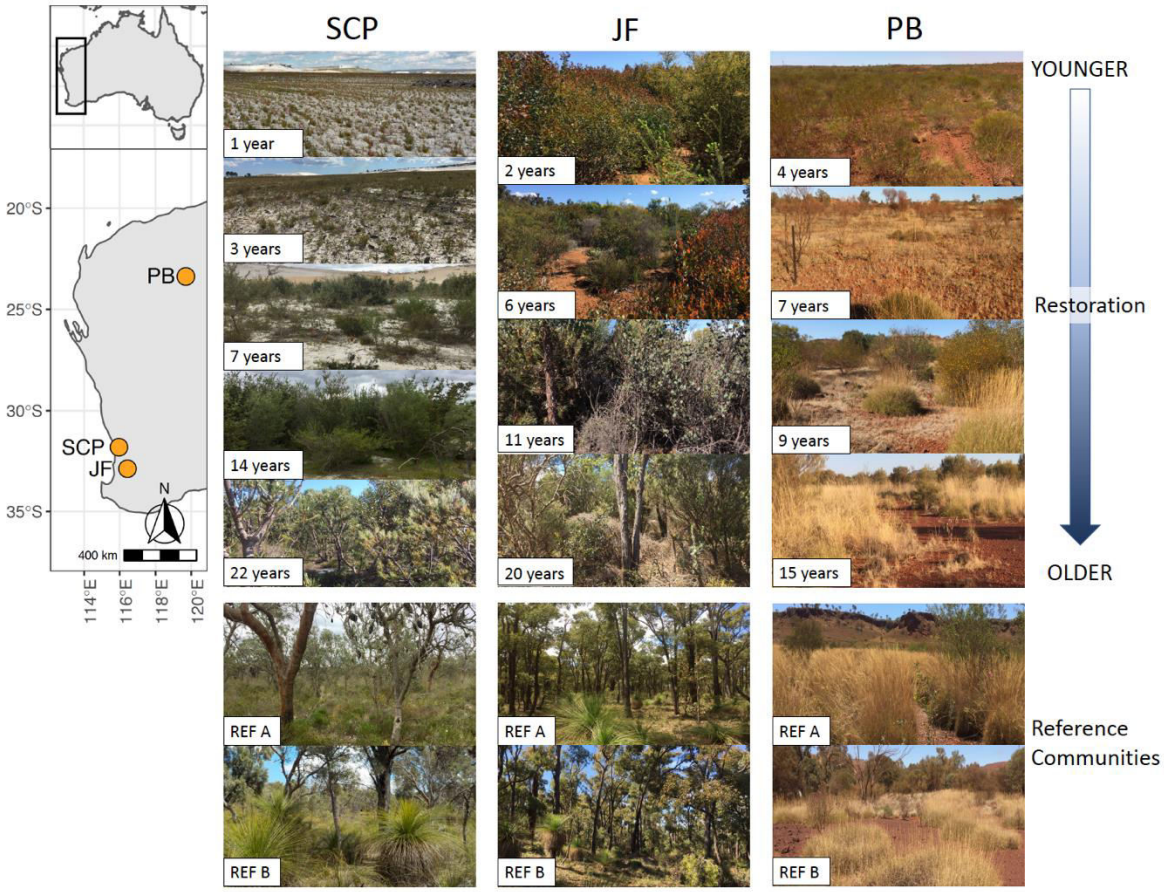


**Figure 3.1** Chronosequences of mining restoration where soil samples were collected. Restoration sites shown with the number of years restoration from 1 to 22 years. Reference sites shown below. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

The Banksia Woodland of the Coastal Plain (SCP) has a warm-summer Mediterranean climate with mild cool wet winters; temperature has a mean minimum of 12.8°C, mean maximum of 24.7°C, with 757 mm mean annual rainfall (Australian Bureau of Meteorology). This location occurs within the Southwest Australian Globally Biodiversity Hotspot (Myers et al., 2007). The mine is located on the silicaceous Bassendean dunes, which are characterized by low nutrient, leached podzols, with high acidity and low water-holding capacity (Dodd & Heddle, 1989; McArthur, 1991). The dominant tree species are *Banksia attenuata* and *B. menziesii*, with less dominant *Eucalyptus todtiana* and *Nuytsia floribunda.* The understory consists of woody species of Myrtaceae, Fabaceae, Proteaceae, and Ericaceae, and non-woody species in Anthericaceae, Stylidiaceae, Cyperaceae, and Haemodoraceae

(Trudgen, 1977). In October 2018, we sampled seven sites at a Hanson Construction Materials sand quarry in Lexia (-31.76°, 115.95°); two reference sites and restoration sites 1, 3, 7, 14, 22 years old. Restoration sites were previously sites of open pit mining. The sites have been restored with the aim of returning mined areas to the surrounding native Banksia woodlands. All restoration was done by Hanson and previous mine owners and included direct transfer of fresh topsoil, ripping, and seeding with native species. A previous study found that species richness and density tended to be higher in restoration than reference sites, and that percent cover increases with restoration age and is highest in reference sites (Benigno, Dixon, & Stevens, 2013). This study also found that restored sites have more basic soils with less organic matter than reference sites (Benigno et al., 2013).

The second chronosequence located in the Jarrah (*E. marginata*) forest is also within the Southwest Australian Biodiversity hotspot (Myers et al., 2007) and has a similar hot-summer Mediterranean climate; temperatures have an mean min. of 8.6°C, mean max. of 23.7°C, and 668.9 mm annual mean rainfall (Australian Bureau of Meteorology). The lateritic soils are nutrient poor and high in gravel with surfaces rich in iron and aluminum (McArthur, 1991). The overstorey vegetation is primarily *E. marginata,* with *E. patens,* and *E. wandoo* also present. The understory is sclerophyllous and dominated by taxa from numerous families, including Fabaceae, Asteraceae, Proteaceae, Dasypogonaceae, and Myrtaceae (Havel, 1975). We sampled six sites from the bauxite mine South32 (-32.96°, 116.48°) in October 2018; two reference sites and restoration sites 2, 6, 11, 20 years old. Restoration sites were previously sites of strip mining. All restoration was undertaken by South32 or the previous mine owners. Post mining the landscape is shaped using waste material and gravel is returned. Topsoil is a homogenized mix of stockpiled topsoil and topsoil that is directly transferred from newly mined areas. The sites are then ripped, seeded with over 100 native species, recalcitrant plants (mostly grasses) are planted, and a one-time treatment of superphosphate is applied. Reference and restoration sites are dominated by Myrtaceae and Fabaceae species. Total cover increases with age of restoration, eventually achieving similar cover percentages to reference sites. Organic carbon increases slowly with age while soil nitrogen increases at a faster rate and soil pH decreases with rehabilitation age (Banning, Grant, Jones, & Murphy, 2008).

The third chronosequence is located in the Pilbara in northwestern Western Australia. The Pilbara has a hot, arid climate with most rainfall occurring in the summer along with cyclonic activity (McKenzie, van Leeuwen, & Pinder 2009). Temperatures have a mean min. of 15°C and mean max of 30.6 °C, with 263.8 mm mean rainfall (Australian Bureau of Meteorology). Soils are acidic stony loams with low fertility, which support open woodlands of snappy gum (*E. leucophloia*) over hummock grasses (*Triodia wiseana, T. basedowii, T. lanigera*) and low Acacia shrubs. (McKenzie et al. 2009). The Pilbara is a significant mining region and accounts for 39% of global iron ore production (Government of Western Australia 2019). We sampled 6 sites at a BHP iron ore mine (-22.84°, 118.95°) in September 2018, 2 reference sites and restoration sites 4, 7, 11, and 15 years old. The restoration sites were primarily borrow pits as these provided the longest, flat chronosequence. Restoration was conducted by the mine owners; landscapes were reformed and stockpiled topsoil (average age 10 years) was applied and then ripped. Restoration areas tended to have higher coverage of woody shrubs (*Acacia*), while reference sites and older restoration areas have more hummock grasses (*Triodia*). Vegetation cover was low in reference sites (~30-40%). Restoration areas also had invasive species such as buffel grass (*Cenchrus ciliaris*) and kapok bush (*Aerva javanica*) which were absent in reference sites (Data from BHP).

### 3.4.2 Sample Collection

Soil samples were collected from 5 points at each restoration/reference site for a total of 95 samples (35 SCP, 30 JF, 30 PB). For each sample, 8 sub-samples were taken randomly in a 10 x10 m plot using a 15 cm soil probe; these were then manually homogenized in a large sample bag and a portion was collected in a 50 mL falcon tube for microbial analyses, while the rest of the sample was kept for soil chemical analyses. The soil probe was cleaned with bleach between each sample and gloves were changed between each sample point. Samples were collected at each location within 2-3 days to minimize variation in environmental conditions caused by weather. Soils were frozen as soon as possible in a mobile freezer and taken to Perth, where they were stored at -20°C until they were processed.

### 3.4.3 Soil chemical properties

Soils to be used for chemical analyses were dried at 50°C for 48 hours and sieved with 2 mm mesh. Soil moisture was determined gravimetrically by measuring a known quantity of soil before and after drying. Further soil chemical analyses were conducted by the CBSP Soil and Plant Analysis Laboratory in Perth. Phosphorus and Potassium were determined using the Colwell method (Colwell, 1965), plant available Sulfur with the Blair/Lefroy Extractable Sulfur method using a 0.25M solution of potassium chloride solution to extract the soil and analyzing it using inductively couple plasma spectroscopy (Blair et al., 1991). Organic carbon was measured using the Walkley Black method (Walkley & Black, 1934). Soil nitrate and ammonium were extracted using a 2M potassium chloride solution and measured colourimetrically after dilution. For pH and conductivity, soils were extracted in deionized water with a 1:5 ratio and then measured with a pH meter and a conductivity electrode. Trace elements (Copper, Zinc, Manganese, Iron) were measured by extracting the soil in a diethylene-triamine-penta-acetic acid (DTPA) solution (ratio of 1:2) measuring with atomic absorption spectroscopy.

### 3.4.4 Soil Microbial Analysis

For DNA extraction, we first used a TissueLyser (Qiagen) to homogenize the soils for 1 min at 30/s in 50 mL falcon tubes. DNA was extracted from 250 mg soil using the DNeasy PowerLyzer PowerSoil kit (Qiagen) on the QiaCube Connect automated platform (Qiagen). The final elution volume was 100 µL, and extraction controls (blanks) were carried out for every set of extractions. Quantitative PCR (qPCR) was run on neat extracts and a 1/10 dilution to see if samples exhibited inhibition, and to determine the optimal DNA input for PCR for each sample to maximise input relative to any inhibitors (Murray, Coghlan, & Bunce, 2015). The qPCR assays were run with two primer assays one targeting the V4 location of the 16S rRNA for Bacteria (16SBact515F -Turner et al., 1999/ 16SBact806R -Caporaso et al., 2011) and the Internal Transcribed Spacer ITS2 for fungi (ITS7F-Ihrmark et al., 2012/ ITS4R-White et al., 1990). These are common regions to target for bacterial and fungal sequencing and are standard for the Earth Microbiome Project (Thompson et al., 2017).

The qPCRs were run on a StepOne Plus (Applied BioSystems) real-time qPCR instrument with the following conditions: 5 min at 95°C, 40 cycles of 95°C for 30s, 30s at the annealing temperature (50°C for Bacteria, 54°C for Fungi) and 45s at 72°C, a melt curve stage of 15s at 95°C 1 min at 60°C and 15s at 95°C, ending with 10 min elongation at 72°C. The PCR mix for quantitation contained: 2.5 mM MgCl2 (Applied Biosystems, USA), 1× PCR Gold buffer (Applied Biosystems), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4 mg/ml bovine serum albumin (Fisher Biotec, Australia), 0.4 µmol/L forward and reverse primer, 1 U AmpliTaq Gold DNA polymerase (Applied Biosystems) and 0.6 µl of a 1:10,000 solution of SYBR Green dye (Life Technologies, USA). Extraction control and non-template controls were included in qPCR assays.

After optimal DNA input was determined by qPCR (most soil extract required a 10x dilution), each sample was assigned a unique combination of multiplex identifier (MID) tags for each primer assay. These MID tags were incorporated into fusion tagged primers, and none of the primer-MID tag combinations had been used previously in the lab to prevent cross contamination. Fusion PCRs were done in duplicate and to minimize PCR stochasticity, the mixes were prepared in a dedicated clean room before DNA was added. The PCRs were carried out under the same conditions as the standard qPCRs described above. Samples were then pooled into approximately equimolar concentrations to produce a PCR amplicon library that was size-selected to remove any primer-dimer that may have accumulated during fusion PCR. Size selection was performed (150-500bp Bacteria, 250-600bp Fungi) using a PippinPrep 2% ethidium bromide cassette (Sage Science, Beverly, MA, U.S.A). Libraries were cleaned using a QIAquick PCR Purification Kit (Qiagen, Germany) and quantified using Qubit Fluorometric Quantitation (Thermo Fisher Scientific). Sequencing was performed on the Illumina MiSeq platform using the 300 cycle V2 (Bacteria), or the 500 cycle V2 (Fungi) as per manufacturer's instructions.

### 3.4.5 Sequencing analysis

Sequences were demultiplexed using OBITools (Boyer et al., 2016) for the Bacterial library and a demultiplex function in the “insect” package (Wilkinson et al., 2018) on the R 3.5.1 platform (R Core Team, 2018). Further sequence processing

was performed in R using the "DADA2" package (Callahan et al., 2016) where sequences were quality filtered, the error rates were estimated, and the sequences were dereplicated. The error rates are then used in the sample inference stage to remove sequences likely to be errors and leave Amplicon Sequence Variants (ASV). These ASVs are equivalent to zero radius operational taxonomic units (ZOTUs) in usearch (Edgar, 2016). The sequences are then merged (Fungi only) and the sequence table is constructed and chimeras removed. Taxonomy was assigned with DADA2 using the naive Bayesian classifier method of Wang et al. (2007). The databases used were Greengenes (DeSantis et al., 2006) for Bacteria and UNITE (Nilsson et al., 2019) for Fungi.

### 3.4.6 Statistics

All statistics were run using R 3.5.1 (R Core Team 2018). Sequence variants that were present in the extraction controls were removed from the dataset, then sequencing depth was rarefied to the minimum in the 'phyloseq' package (McMurdie & Holmes, 2013). Alpha diversity was calculated using the 'phyloseq' package (McMurdie & Holmes 2013) and tested using a two-way analysis of variance (anova) with location and restoration as factors. This was followed by a Tukey HSD test from the 'agricolae' package (deMendiburu, 2019). Community composition was visualized using Non metric multidimensional scaling (NMDS), based on the log transformed ASV table and with Bray-Curtis dissimilarity. Differences between restoration ages were tested using permutational multivariate analysis of variance (PERMANOVA). However, as there was significant spatial autocorrelation between SMC in the Jarrah and Pilbara replicates, we also pooled replicates together and calculated the similarity of each restoration age to reference communities. We also ran regression analyses on these separately for each location. When pooled, there was no spatial autocorrelation, although there was a loss in power. We also looked at the rarefied read abundance of the top 10 most abundant phyla and tested whether that differed across restoration using permuted anovas. We adjusted the *P*-values for multiple tests using the "BH" method (Benjamini & Hochberg, 2007).

Soil chemical variables were tested for homogeneity of variance and log transformed as needed before using one-way anovas to test differences between

restoration age and reference sites within each location. Distance Based Redundancy Analysis (dbRDA) was used to determine the relationship between soil chemical variables and the community composition of the soil. Soil chemical variables were normalized using the *decostand* function in the 'vegan' package (Oksanen et al., 2019) and then used in a dbRDA with a Bray-Curtis dissimilarity matrix of the log-transformed ASV table. Variables with high "vif" or variable inflation factors were removed as they are likely collinear with other variables. All dbRDAs were run separately for each location.

To assess functional differences across restoration ages, we first assigned functionality using FUNGuild (Nguyen et al., 2016) for fungal sequences and METAGENassist (Arndt et al., 2012) for bacterial sequences. FUNGuild uses third party annotation to assign functionality, such as trophic mode, based on taxonomy. METAGENassist uses phenotype information of bacterial species listed on the NCBI database to add information such as metabolism and energy source based on taxonomy (Genus level). For the fungi, we used multipattern analysis from the R package 'indicspecies' (De Cáceres & Legendre, 2009) and then tested the differences in the number of ASVs in each trophic mode across restoration using a chisquare test for associations. For bacterial functionality we looked at the normalized number of reads assigned to each metabolism category, and tested the difference between sites using a two-way PERMANOVA with location and age as the grouping variables. Again, *P*-values were adjusted for multiple comparison using the 'BH' method (Benjamini & Hochberg, 2007).

## 3.5 Results

In total, we generated 4,836,541 quality-filtered bacterial sequences from 93 samples and 4,331,020 quality-filtered fungal sequences from 95 samples. These were rarefied to 23,305 seqs/sample for Bacteria and 11,784 seqs/sample for Fungi

### 3.5.1 SMC diversity, community composition and similarity to reference sites

Bacterial richness (alpha diversity at the local scale) and Shannon diversity responses to restoration were dependent on location (Richness: $F_{3,74}$=4.59, p=0.005; Shannon: $F_{3,74}$=7.19, p<0.001). Tukey HSD results (Figure 3.2) show that reference sites and older restoration sites were not more diverse than younger restoration sites. Similarly, fungal richness responded differently depending on the location ($F_{3,76}$=6.88, p<0.001) but fungal diversity did not ($F_{3,76}$= 2.351, p=0.079). Bacterial diversity varied more in the Pilbara and Coastal Plain sites, while fungal diversity varied more between Jarrah sites (Figure 3.2). In general, there are few clear directional changes in richness and diversity, with the possible exception of fungi in the Jarrah sites where younger sites tend to have the lowest richness and diversity compared to older restoration and reference sites (Figure 3.2).

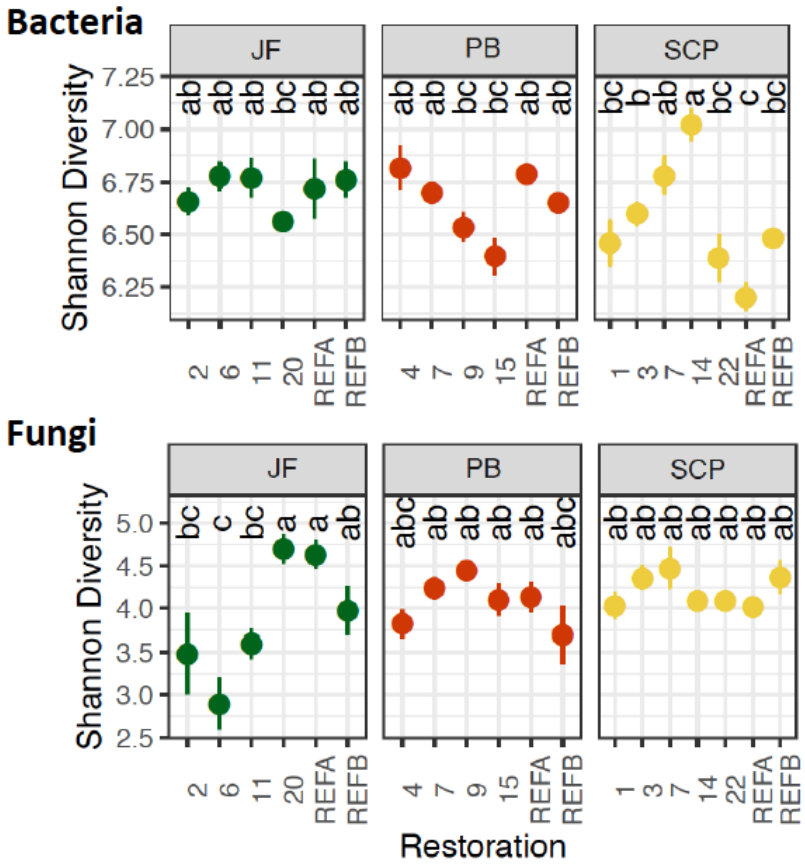


**Figure 3.2** Alpha diversity of bacterial (above) and fungal (below) communities at restoration sites. Letters indicate results of the Tukey HSD test. Richness showed similar patterns to Shannon diversity.

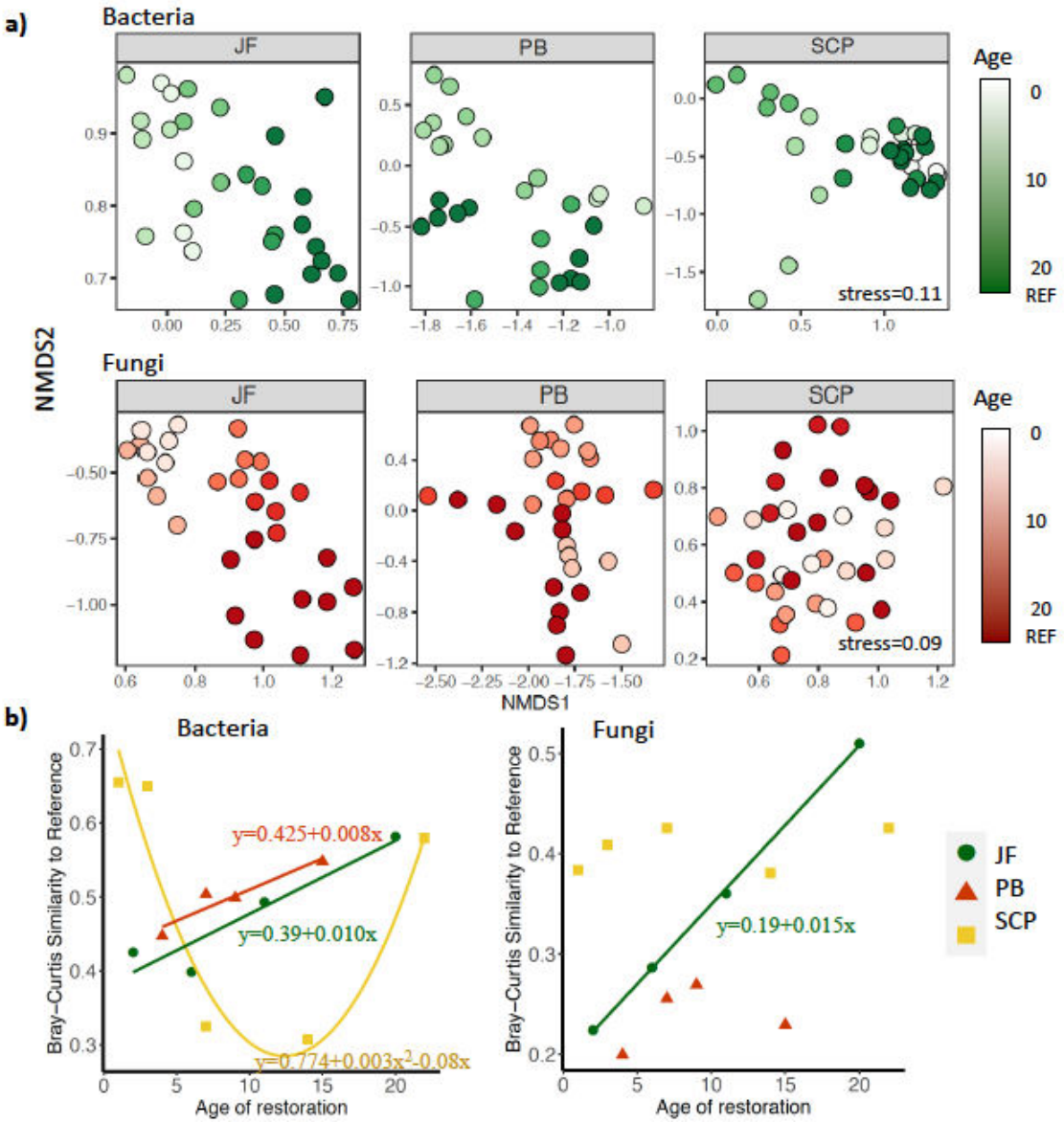


**Figure 3.3** Community composition of soils in restoration and reference sites at three chronosequences of mine site restoration. NMDS Ordinations (a) of bacterial (above) and fungal (below) community composition (similarity=bray curtis). The bray-curtis similarity (b) between each site and the most similar reference site. Lines are included for linear models that were significant (alpha=0.1). JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

Community composition was significantly different across locations for both bacteria and fungi ($p<0.001$, Table S3.9.1). Location and restoration age were also highly significant for both bacteria ($p<0.001$) and fungi ($p<0.001$). However, there was also significant spatial autocorrelation in the Jarrah ($p<0.03$) and Pilbara ($p<0.001$) chronosequences, although the Coastal Plain had no spatial autocorrelation (Table S3.9.2). Because of the loss in power from merging replicates, we use a more

conservative significance level of 0.1 for the linear models predicting similarity to reference sites. For both the Jarrah and the Pilbara chronosequences, Bacterial communities in older restoration are more similar to reference communities than younger restoration sites (Pilbara p=0.05, Adj R2=0.84; Jarrah p=0.06, Adj R2=0.83). In the Coastal Plain, the youngest and oldest sites have the greatest similarity to reference, while the intermediate aged sites are more dissimilar. This relationship follows a quadratic linear model (p=0.09, Adj R2=0.81). For the fungal data, the Jarrah chronosequence was the only one with a significant relationship between restoration age and community similarity to reference ($p<0.001$, Adj R2=0.99). However, for the Coastal Plain and the Pilbara the community similarity between restoration and reference sites approached the community similarity between the two reference sites (PB=0.19, SCP=0.49, JF=0.57)

There were 10 dominant bacterial phyla and 5 dominant fungal phyla (>2% relative abundance), 93% (10 bacteria, 2 fungal) of which showed significant differences in rarefied abundance between restoration ages and reference sites in at least one location (Table 3.1). However, phyla that showed significant differences in rarefied abundance were not significant at all locations. For example, Ascomycota only showed a significant decrease in abundance in the Jarrah location. Gemmatimonadetes was the only phyla that consistently decreased in abundance with restoration age (Table 3.1). Whether rarefied read abundance increased or decreased with restoration age depended on the phyla and the location, and the same phylum (e.g. Chloroflexi, Proteobacteria) could have opposite results in the different locations.

**Table 3.1** Effects of restoration on the rarefied abundance of certain phyla at three Western Australian locations. Only phyla making >2% total abundance were included. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

| | JF | | PB | | SCP | |
|---|---|---|---|---|---|---|
| **Phylum** | **P val** | **direction** | **P val** | **direction** | **P val** | **direction** |
| Acidobacteria | 0.785 | | **0.002** | Decreasing | **0.002** | Variable |
| Actinobacteria | **0.035** | Variable | 0.627 | | **0.002** | Increasing |
| Bacteroidetes | 0.627 | | **0.011** | Variable | 0.627 | |
| Chloroflexi | **0.002** | Decreasing | **0.002** | Increasing | **0.002** | Decreasing |
| Firmicutes | **0.002** | Decreasing | **0.005** | Variable | **0.002** | Variable |
| Gemmatimonadetes | **0.002** | Decreasing | **0.002** | Decreasing | **0.002** | Decreasing |
| Planctomycetes | **0.007** | Decreasing | 0.080 | | **0.002** | Decreasing |
| Proteobacteria | **0.002** | Increasing | **0.002** | Decreasing | **0.002** | Variable |
| Thaumarchaeota | 0.367 | | **0.009** | Decreasing | **0.002** | Decreasing |
| Verrucomicrobia | **0.011** | Variable | **0.026** | Variable | **0.031** | |
| Ascomycota | **0.015** | Decreasing | 0.330 | | 0.4 | |
| Basidiomycota | **0.040** | Increasing | **0.015** | Variable | 0.573 | |
| Glomeromycota | 0.602 | | 0.540 | | NA | |
| Mortierellomycota | 0.625 | | 0.330 | | 0.573 | |

### 3.5.2 Functional groups

Bacterial functional differences in metabolism were heavily driven by location, which accounted for 45.7% of the variation. For example, Dinitrogen fixers were associated with the Coastal Plain sites, sulfide oxidizers with the Pilbara, and bacteria that store polyhydroxybutyrate were characteristic of the Jarrah sites (Figure S3.9.3). Restoration sites tended to have more bacteria that degrade aromatic hydrocarbons (Table 3.2), particularly at the Jarrah location. At both Jarrah and Coastal Plain locations restoration was also associated with Napthalene degrading bacteria and sulfide oxidizers. Reference sites at the Jarrah and Coastal Plain locations were associated with chitin and xylan degradation, dehalogenation, and nitrogen fixation (Table 3.2)

**Table 3.2** Multipattern Analysis showing the bacterial metabolic pathways that are significantly associated with each restoration category at the three locations. Only those with significant (alpha<0.05) were included. Numbers indicate adjusted *P*-values where there were significant associations.

| | JF | | | PB | SCP | |
|---|---|---|---|---|---|---|
| **Metabolism** | **Young** | **Old** | **REF** | **Young** | **Old** | **REF** |
| Ammonia.oxidizer | | | | | | 0.002 |
| Atrazine.metabolism | | | | | 0.002 | |
| Carbon.fixation | | | | | 0.002 | |
| Chitin.degradation | | | 0.011 | | | 0.002 |
| Chlorophenol.degrading | | | | | 0.002 | |
| Degrades.aromatic.hydrocarbons | 0.011 | | | 0.019 | 0.002 | |
| Dehalogenation | | | 0.045 | | | 0.002 |
| Lignin.degrader | | | 0.010 | | | |
| Naphthalene.degrading | 0.010 | | | | 0.004 | |
| Nitrogen.fixation | | | 0.011 | | | 0.002 |
| Streptomycin.producer | | 0.045 | | | | |
| Sulfate.reducer | | | | | | 0.002 |
| Sulfide.oxidizer | 0.013 | | | | 0.002 | |
| Sulfur.metabolizing | | | | | | 0.002 |
| Sulfur.oxidizer | | | | | 0.005 | |
| Xylan.degrader | | | | | | 0.002 |

We were able to assign function to 1209 out of 1678 fungal ASVs, and of these, 492 were identified as having significant (alpha=0.05) associations to one or more groups. The chi-square test for association revealed there were significant differences in the number of indicator ASVs in each trophic mode at the Jarrah ($x^2=51.11$, df=12, $p<0.001$), but not in the Pilbara ($x^2=10.81$ df=10, $p=0.372$), or the Coastal Plain ($x^2=15.42$, df=12, $p=0.219$) (Table 3.2). The reference sites tended to have more symbiotrophic ASVs, the older restoration sites and the reference sites were similar in their levels of saprotrophic ASVs, while the younger restoration sites had less saprotrophic and symbiotrophic ASVs.

**Table 3.3** Number of fungal indicator taxa in each trophic mode at the restoration and reference sites. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

| | JF | | | PB | | | SCP | | |
|---|---|---|---|---|---|---|---|---|---|
| **Trophic Mode** | **Young** | **Old** | **REF** | **Young** | **Old** | **REF** | **Young** | **Old** | **REF** |
| Pathotroph | 6 | 3 | 3 | 8 | 12 | 5 | 3 | 6 | 7 |
| Pathotroph-Saprotroph | 20 | 54 | 42 | 8 | 17 | 5 | 35 | 34 | 27 |
| Pathotroph-Saprotroph-Symbiotroph | 16 | 7 | 2 | 11 | 10 | 9 | 12 | 10 | 9 |
| Pathotroph-Symbiotroph | 2 | 16 | 22 | 0 | 0 | 0 | 2 | 1 | 2 |
| Saprotroph | 20 | 43 | 41 | 22 | 21 | 8 | 38 | 25 | 29 |
| Saprotroph-Symbiotroph | 6 | 8 | 4 | 8 | 5 | 4 | 5 | 3 | 1 |
| Symbiotroph | 5 | 12 | 20 | 13 | 7 | 9 | 3 | 4 | 12 |

### 3.5.3 Soil chemical properties

Responses of soil chemical properties to restoration also varied across the locations. Reference sites tended to be less basic and have higher organic matter (Table 3.4), but this relationship was not significant at all locations. The distance-based Redundancy Analyses show which soil variables were significant in explaining the variations in bacterial and fungal communities (Figure 3.4). Soil pH was one of the few significant variables that was higher in newly restored sites; most significant variables (e.g. organic matter, ammonium, magnesium) were higher in reference soils. In the Pilbara, there were more variables that were higher in restoration, such as calcium, magnesium, and soil moisture. Overall, soil variables explained over 65% of the variation in bacteria (65.7% JF, 78.2% PB, 67.9% SCP) and over 48% of the variation in fungal communities (63.7% JF, 58.7% PB, 48.7% SCP). Soil properties such as Ammonium, pH, Sulfur, and organic carbon are drivers of variation in microbial communities at the Jarrah forest, similar to the Coastal Plain. Ammonium was a significant factor in all three chronosequences for both bacteria and fungi, while other soil properties like pH were significant only in the Jarrah sites. Organic carbon was also identified as common driver of microbial communities, with the exception of bacterial communities in the Coastal Plain.

**Table 3.4** The effect of restoration on soil chemical variables at three Western Australian locations. Numbers show the mean in each group with the standard error in parantheses. Variables with significant differences (alpha<0.05) are bold, and the letters indicate Tukey HSD test results. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

| | *JF* | | | *PB* | | | *SCP* | | |
|---|---|---|---|---|---|---|---|---|---|
| **Age** | **Young** | **Old** | **REF** | **Young** | **Old** | **REF** | **Young** | **Old** | **REF** |
| Ammonium (mg/kg) | **8.1 (±0.5) b** | **13.6 (±0.7) a** | **15.6 (±1.2) a** | 2.1 (±0.3) | 2.8 (±0.4) | 2.1 (±0.2) | **1.8 (±0.3) b** | **1.8 (±0.2) ab** | **2.7 (±0.2) a** |
| Nitrate (mg/kg) | **1.2 (±0.2)a b** | **1.0 (±0.0) b** | **1.5 (±0.2) a** | 2.0 (±0.4) | 1.4 (±0.3) | 1.1 (±0.4) | 0.5 (±0.0) | 0.5 (±0.0) | 0.5 (±0.0) |
| Phosphorus (mg/kg) | 6.9 (±3.5) | 3.0 (±0.4) | 4.2 (±0.4) | 3.3 (±0.8) | 5.3 (±0.6) | 4.4 (±0.5) | 1.1 (±0.1) | 1.0 (±0.0) | 1.0 (±0.0) |
| Potassium (mg/kg) | 59.2 (±5.0) | 77.5 (±7.8) | 81.7 (±6.8) | 290.4 (±27.3) | 290.8 (±21.6) | 244.5 (±27.5) | 11.3 (±0.7) | 13.5 (±1.5) | 12.8 (±1.0) |
| Sulfur (mg/kg) | 7.3 (±1.1) | 7.1 (±0.2) | 12.1 (±0.9) | 6.6 (±3.9) | 2.5 (±0.3) | 2.3 (±0.4) | 1.5 (±0.1) | 1.9 (±0.2) | 1.4 (±0.1) |
| Organic Carbon (%) | **2.41 (±0.15) c** | **3.41 (±0.09)b** | **4.66 (±0.09) a** | 0.41 (±0.04) | 0.49 (±0.05) | 0.54 (±0.08) | **0.91 (±0.09)a b** | **0.77 (±0.06) b** | **1.14 (±0.07) a** |
| Conductivity (dS/m) | **0.041 b (±0.004)** | **0.063 b (±0.005)** | **0.077 a (±0.006)** | 0.030 (±0.006) | 0.024 (±0.004) | 0.017 (±0.002) | 0.017 (±0.003) | 0.019 (±0.002) | 0.015 (±0.001) |
| pH | **6.4 (±0.0)a** | **6.0 (±0.0)b** | **6.0 (±0.0)b** | **7.1 (±0.2) a** | **6.8 (±0.2)ab** | **6.5 (±0.1) b** | 6.4 (±0.2) | 6.5 (±0.1) | 5.9 (±0.1) |
| Copper (mg/kg) | **1.61 (±0.39) a** | **0.90 (±0.12)a b** | **0.61 (±0.07) b** | 1.01 (±0.13) | 1.18 (±0.08) | 1.38 (±0.14) | 0.32 (±0.03) | 0.28 (±0.07) | 0.25 (±0.04) |
| Iron (mg/kg) | **40.6 (±3.5)b** | **57.9 (±6.3)a** | **59.8 (±5.0)a** | **9.7 (±0.3)b** | **12.4 (±0.5)a** | **12.6 (±0.7)a** | **13.5 (±0.6)ab** | **10.9 (±1.2)b** | **14.2 (±1.0)a** |
| Manganese (mg/kg) | **6.88 (±0.49) c** | **13.54 (±1.30)b** | **30.73 (±1.41) a** | **19.28 (±1.90) b** | **34.31 (±4.11)a** | **33.02 (±2.76) a** | **0.89 (±0.11)b** | **0.79 (±0.09) b** | **1.39 (±0.1)a** |
| Zinc (mg/kg) | 0.82 (±0.27) | 0.53 (±0.06) | 0.30 (±0.03) | 0.29 (±0.01) | 0.38 (±0.08) | 0.32 (±0.04) | 0.31 (±0.02) | 0.37 (±0.05) | 0.40 (±0.13) |
| Aluminium (meq/100g) | **0.05 (±0.01) c** | **0.13 (±0.02)b** | **0.19 (±0.02) a** | 0.13 (±0.02) | 0.14 (±0.01) | 0.11 (±0.02) | **0.04 (±0.00)a b** | **0.03 (±0.00) b** | **0.05 (±0.00) a** |
| Calcium (meq/100g) | **4.72 (±0.44) b** | **5.54 (±0.31)b** | **8.88 (±0.65) a** | **5.81 (±0.90) a** | **4.39 (±0.88)a b** | **2.52 (±0.26) b** | 2.04 (±0.24) | 1.78 (±0.34) | 1.56 (±0.11) |
| Magnesium (meq/100g) | **1.17 (±0.10) b** | **1.42 (±0.06)b** | **2.54 (±0.23) a** | **3.20 (±0.61) a** | **1.25 (±0.18)b** | **1.61 (±0.31) b** | **0.27 (±0.03)a** | **0.16 (±.01)b** | **0.28 (0.02)a** |
| Soil Moisture | **6.9 (±0.3)c** | **8.2 (±0.3)b** | **10.2 (±0.5)a** | **2.9 (±0.1)a** | **2.8 (±0.1)a** | **2.1 (±0.2)b** | 1.5 (±0.1) | 2.0 (±0.1) | 2.0 (±0.3) |

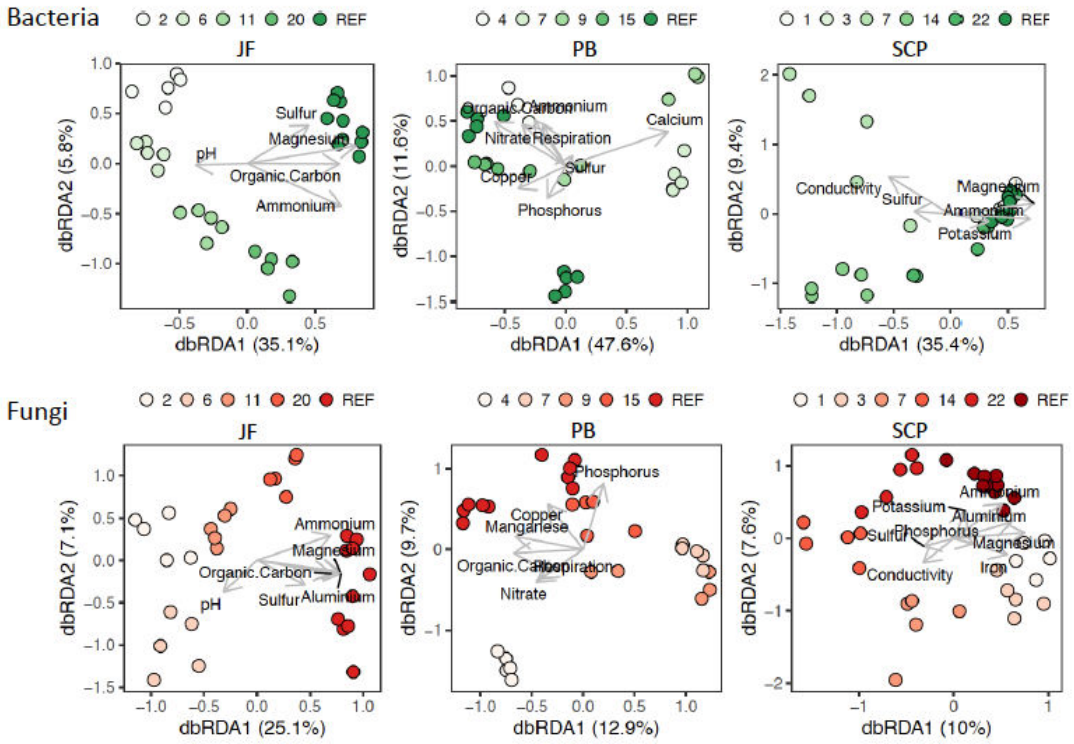


**Figure 3.4** Results of distance-based Redundancy Analyses for bacterial communities above (green) and fungal communities below (red). Significant soil terms (alpha=0.05) are shown using arrows and labels. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

## 3.6 Discussion

In this study, we assessed SMC across three restoration chronosequences using a high-throughput amplicon sequencing approach. We demonstrated changes in SMC at restored sites, but found that patterns were complex. Understanding the responses of soil microbes to restoration is important as they are increasingly popular targets for monitoring biodiversity recovery.

### *3.6.1* Are SMC in older restoration sites more similar to reference communities than those at younger restoration sites?

Overall, there were no consistent changes in microbial richness or diversity across the three restoration chronosequences (Figure 3.2). These trends support previous work where the greatest changes identified were in the shifts in community composition, rather than diversity (Banning et al. 2011; Sun et al. 2017; Yan et al. 2018). Our results indicated strong compositional differences within each of the chronosequences. However, these should be interpreted with caution, as there was also significant spatial autocorrelation between replicates at two out of three locations (Jarrah and Pilbara). Spatial autocorrelation is an important consideration when using SMC for monitoring (Yan et al. 2020), especially in mine site restoration where site locations are dependent on presence of resources, rather than the ecology of the surrounding environment. While spatial scale is a strong driver of microbial diversity (Nunan et al., 2002; O'Brien et al., 2016), especially in reconstructed soils (Mummey et al. 2002b), soil chemical factors and plant associations account for more variation in SMC (Burns et al., 2015; Nunan et al., 2002). Accordingly, spatial autocorrelation does not mean the data are not informative for restoration. Instead, it indicates the importance of multiple reference sites for comparison as done in this study where we were able to include spatially separated reference sites located near the restoration sites. The collection and archiving of soil samples throughout restoration might help generate time-stamped data that is less impacted by autocorrelation as sampling sites can be more closely controlled.

Several studies have found a trend of increasing similarity to reference bacterial communities with restoration age (Banning et al. 2011, Yan et al 2020, Sun et al. 2017). We found similar directional changes in bacterial communities at the

Jarrah and Pilbara locations. In contrast, at the Coastal Plain, the youngest and oldest sites were most similar to the reference soils. The use of stored topsoil for restoration of study sites presents a potential confounding factor as stockpiling topsoil under different conditions may cause variation in SMC; the Coastal Plain was the only location in our study where direct transfer of topsoil during the restoration process was conducted. During direct transfer, soil is stripped from an area to be mined and transferred immediately to a site to be restored. This approach to restoration preserves the integrity of the soil seed banks (Rokich, Dixon, Sivasithamparam, & Meney, 2000) and may also allow the bacterial communities to be maintained in the short term. However, following this initial phase the dynamic interactions of edaphic and vegetation factors (e.g. absence of mature trees) present in the re-growing restoration may cause shifts in the bacterial communities in the intermediate aged sites. In contrast, the Jarrah and Pilbara sites stockpile and store topsoil until required (from months to years) during which time there may be reduction of microbial communities to those capable of surviving the biologically hostile conditions within a stockpile (Birnbaum, Bradshaw, Ruthrof, & Fontaine, 2017).

Higher orders of bacteria such as phyla are considered to share some general life history strategies as a result of shared evolutionary pathways (Fierer, Bradford, & Jackson 2007; Philippot et al., 2010). Therefore, despite the diversity within the phyla, they can be an indicator of successional trajectories (Banning et al. 2011, Yan et al. 2020). We found that the response of bacterial rarefied phyla abundance was very different between locations. For example, with increasing restoration age, Proteobacteria were more abundant at the Jarrah location, less abundant in the Pilbara, and their response was variable in the Coastal Plain. Other studies have also found increases of Proteobacteria with restoration age (Yan et al. 2020, Banning et al. 2011, Gellie et al. 2017), similar to our results in the Jarrah forest. However, none of these studies were located in a hot arid climate, and the decrease of Proteobacteria in the Pilbara may be due to the extreme climatic and soil conditions of the arid zone. Proteobacteria abundance is often related to carbon availability (Fierer et al. 2007) and there were no significant changes in organic carbon at the Pilbara chronosequence sites. However, there are also examples such as the phylum Chloroflexi, which showed a pattern of decrease in the Jarrah site, yet another study

in the same ecosystem identified it as one of the phyla that increases with restoration (Banning et al. 2011).

Bacteria are highly variable, and because of their high growth rates communities can change rapidly in composition within a year (Lauber et al., 2013; Sun et al., 2017). It is common in microbial studies to find distinct communities between disturbed and undisturbed sites, but the taxa driving those differences are often inconsistent between studies (Lauber et al. 2013). Study-specific soil and site conditions likely drive this variability in taxa, which is why using only certain taxa as indicators is questionable. It also emphasizes the importance of reference sites near restoration sites sampled concurrently, to account for variability over time and space to provide an indication of general trajectory due to climatic variables.

Unlike bacteria, fungal communities showed progressive recovery towards the reference communities only at the Jarrah location, with no clear trends at the other two sites. For both the Pilbara and Coastal Plain locations, there are two possible explanations for why fungal communities are not becoming more similar to reference communities with increasing age. Firstly, fungal communities may resist restoration as they show no trajectory towards reference communities with increasing restoration age. This was a key finding in a previous study at the Coastal Plain location (Hart et al. 2019). Secondly, the similarity between restoration and reference sites was approximately the same as the similarity between the two reference sites, implying that the communities may have achieved maximum similarity to reference communities early in the restoration process. The use of topsoil in restoration may have adequately preserved the fungal communities in these systems. The latter is supported by the fact that fungal phyla also showed few significant differences in rarefied abundance in the Pilbara or Coastal Plain, whereas phyla in the Jarrah chronosequence showed clear differences in phyla (Ascomycota and Basidiomycota) that agreed with previous studies (Yan et al. 2018). As has been previously emphasized (Lauber et al. 2013) community composition alone may not be as important as the presence of particular functional groups which may also vary as site conditions mature from a disturbance event.

### 3.6.2 Which functional groups are indicators of the different stages of restoration?

The functional capacity of SMC is an appealing target for monitoring restoration because there is considerable functional redundancy and it is less variable than community composition over small spatial and time scales (Kumarasan et al. 2017). Despite this, there are few examples where high-throughput sequencing has been used to explore change in functional groups across a restoration chronosequence (Yan et al. 2020). This information is important because SMC underpin many ecological and physiological functions (e.g., organic matter decomposition, regulation of mineral nutrient availability) (Meena et al., 2017; Yang et al., 2018) that are essential to building ecological resilience. We observed a higher incidence of bacteria involved in organic matter decomposition (chitin degradation, lignin degradation, xylan degradation) in reference sites at the Jarrah and Coastal Plain locations. Chitin is the structural element of organisms such as fungi and insects (Merzendorfer, 2006; Roncero, 2002), while lignin and xylan are biopolymers in plant cell walls (Ochoa-Villarreal et al. 2012). All of these are more abundant in reference ecosystems, resulting in a higher prevalence of organic matter degrading bacteria in those sites. Reference sites were also associated with nitrogen fixing bacteria, providing plants with an important limiting nutrient (Vitousek et al., 2002). Conversely, restoration sites were associated with bacteria that degrade aromatic hydrocarbons and naphthalene at Jarrah and Coastal Plain (Table 4). These organisms are likely responding to a major shift in chemical composition of the soil as a result of topsoil stripping and storage. This may shift microbe abundance to reflect the disequilibrium of very altered substrates that are not present in the reference sites.

Including the fungal functionality analysis improves the interpretation of the community composition results by showing there were also no differences in fungal trophic modes between the sites at those locations. However, in the Jarrah forest, where there was a trajectory in fungal community composition, we also found significant differences in trophic modes between younger restoration and reference sites. Reference sites tended to have more saprotrophs, necessary for decomposing accumulated leaf litter. They also had more symbiotrophs, which are fungi that

exchange nutrients with host cells (Nguyen et al. 2016) such as mycorrhizae, providing nitrogen and phosphorus to their plant partners in exchange for carbohydrates (Glen et al., 2008). These mycorrhizal networks underpin forest growth and health as found in the jarrah forest ecosystem (Glen et al., 2008). In contrast, Yan et al. (2018) working in a coastal revegetation system found little difference between the number of indicator taxa in each trophic mode. The differences between the Jarrah forest location and the Coastal Plain and Pilbara locations reflect the higher biomass in Jarrah, and higher proportion of root biomass attributed to mycorrhizal species.

### 3.6.3 Are soil chemical properties associated with restoration age, and/or change in SMC composition?

The effect of restoration age on some soil chemical variables was consistent across sites and the patterns similar to those found in previous studies on restoration chronosequences (Banning et al. 2008; Munoz-Rojas 2016; Yan et al. 2018). For example, there was an increase in organic carbon and decrease in pH at older sites, and the direction of this change was towards the values found at reference sites. However, we found trends in other soil chemical variables (e.g. calcium, magnesium) that tended to be different in the Pilbara compared to the Jarrah and Coastal Plain, likely reflecting differences in climate and vegetation between these locations. Similarly, the soil variables that are significant in explaining the variation in SMC changed between locations although there were some common trends (e.g. organic carbon driving communities closer to reference). Soil abiotic variables are known drivers of microbial community composition (Burns et al. 2015; Yan et al. 2018), although the mechanisms behind many of these relationships is not fully understood.

## 3.7 Conclusion

Our findings show that ecological restoration of mine sites can lead to the development of soil microbial communities, which over time become increasingly similar in composition to those of natural reference sites. However, the trajectory response of SMC was location and organism (fungal vs bacterial) specific and affected by topsoil application. Thus, high throughput monitoring of SMC changes should be treated with caution and applied to appropriate ecosystems (i.e. monitoring fungi in ecosystems more reliant on fungal symbioses). Further studies are needed that include sites located in different climate zones, on different soil types or with different plant communities. Also needed are studies of older restoration sites, studies with multiple time points and across different seasons, to understand of background levels of variability. Archiving of samples is suggested, to enable better understanding of how SMC communities change over time. Our results also emphasize the importance of multiple reference sites to account for the variability over space that is common in soil microbial communities.

Including functional analyses of microbial data improved our understanding of the microbial responses to restoration. Currently, the tools to examine functionality from high-throughput sequencing data are available, and will continue to develop in the future especially as microbial analysis is increasingly employing metagenomic (i.e. shotgun) approaches (Kumaresan et al. 2017). We advocate that restoration studies involving SMC should explore functionality as well as composition, but that measurements of richness are less informative. In addition, assessing functionality using non sequencing based methods such as microbial respiration (Haney et al. 2008; Munoz-Rojas et al. 2016) and plant bioassays will be important in validating high-throughput sequencing results.

eDNA studies could also be extended to include other biological groups such as soil fauna (Eaton et al. 2017) or to other sources of DNA (Fernandes et al. 2019; Heyde et al. 2020), enabling a more holistic understanding of biodiversity recovery. Many companies in the resources sector strive towards ‘best practice’ restoration, although what constitutes best practice is not always clear. The approaches herein and in other published studies show great promise in our capacity to incorporate a wider microbial lens on the issue (Gellie et al. 2017; Yan et al. 2018). With further

refinement to experimental design and better ways to study microbial function, these approaches may help guide future restoration efforts and interventions (i.e. microbial inoculation) and expand past mining to agricultural land and contaminated sites.

## 3.8 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.


Arndt, D., Xia, J., Liu, Y., Zhou, Y., Guo, A.C., Cruz, J.A., Sinelnikov, I., Budwill, K., Nesbo, C.L., Wishart, D.S. (2012). METAGENassist: a comprehensive web server for comparative metagenomics. *Nucleic Acids Research 40*, W88–W95. doi:10.1093/nar/gks497.

Australian Bureau of Meteorology. Climate data online, www.bom.gov.au, 2019

Banning, N.C., Grant, C.D., Jones, D.L., Murphy, D.V. (2008). Recovery of soil organic matter, organic matter turnover and nitrogen cycling in a post-mining forest rehabilitation chronosequence. *Soil Biology Biochemistry 40*, 2021–2031. doi:10.1016/j.soilbio.2008.04.010.

Banning, N.C., Gleeson, D.B., Grigg, A.H., Grant, C.D., Andersen, G.L., Brodie, E.L., Murphy, D.V. (2011). Soil microbial community successional patterns during forest ecosystem restoration. *Applied Environmental Microbiology 77*, 6158–6164. doi:10.1128/AEM.00764-11.

Benigno, S.M., Dixon, K.W., Stevens, J.C., (2013). Increasing soil water retention with native-sourced mulch improves seedling establishment in postmine Mediterranean sandy soils. *Restoration Ecology 21*, 617–626.

Benjamini, Y., and Hochberg, Y. (2007). Controlling the false discovery rate : a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B 57*, 289–300.

Birnbaum, C., Bradshaw, L.E., Ruthrof, K.X., Fontaine, J.B. (2017). Topsoil stockpiling in restoration: impact of storage time on plant growth and symbiotic soil biota. *Ecological Restoration 35* (3), 237–245. doi:10.3368/er.35.3.237.

Blair, G., Chinoim, N., Lefroy, R., Anderson, G., Crocker, G. (1991). A soil sulfur test for pastures and crops. *Soil Research 29*, 619–626.

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 16*(1), 176-182. doi:10.1111/1755-0998.12428.

Burns, J.H., Anacker, B.L., Strauss, S.Y., Burke, D.J. (2015). Soil microbial community variation correlates most strongly with plant species identity, followed by soil chemistry, spatial location and plant genus. *AoB Plants 7*, 1–10. doi:10.1093/aobpla/plv030.

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*, 581–583. doi:10.1038/nmeth.3869.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences U. S. A. 108*, 4516–4522. doi:10.1073/pnas.1000080107.

Colwell, J.D. (1965). An automatic procedure for the determination of phosphorus in sodium hydrogen bicarbonate extract of soil. *Chemistry and Industry* 749–895.

De Cáceres, M. and Legendre, P. (2009). Associations between species and groups of sites and statistical inference. *Ecology 90*, 3566–3574. doi:10.1890/08-1823.1.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., …Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied Environmental Microbiology 72*, 5069–5072. doi:10.1128/AEM.03006-05.

Dodd, J. and Heddle, E.M. (1989). Water relations of Banksia wood-lands. *Journal of the Royal Society of Western Australia 71*, 91–92.

Dong, W., Liu, E., Yan, C., Tian, J., Zhang, H., Zhang, Y. (2017). Impact of no tillage vs. conventional tillage on the soil bacterial community structure in a winter wheat cropping succession in northern China. *European Journal of Soil Biology 80*, 35–42. doi:10.1016/j.ejsobi.2017.03.001.

Eaton, W.D., Shokralla, S., McGee, K.M., Hajibabaei, M. (2017). Using

metagenomics to show the efficacy of forest restoration in the New Jersey pine barrens. Edited by F.J.J. chain. *Genome 60* (10), 825–836. doi:10.1139/gen-2015-0199.

Edgar, R.C. (2016). UNOISE2: Improved Error-Correction for Illumina 16S and ITS Amplicon Sequencing. bioRxiv 081257. doi:10.1101/081257.

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. doi:10.1111/rec.12976.

Fierer, N., A. Bradford, M., B. Jackson, R. (2007). Toward an ecological classification of soil bacteria. *Ecology 88*, 1354–1364. doi:10.1890/05-1839.

Frey, S.D., Elliott, E.T., Paustian, K. (1999). Bacterial and fungal abundance and biomass in conventional and no-tillage agroecosystems along two climatic gradients. *Soil Biology and Biochemistry 31*, 573–585. doi:10.1016/S0038-0717(98)00161-8.

Gann, G.D., McDonald, T., Walder, B., Aronson, J., Nelson, C.R., Jonson, …Dixon, K.W. (2019). International principles and standards for the practice of ecological restoration, 2nd edn. *Restoration Ecology 27*(S1), S1-S46. doi:10.1111/rec.13035.

Gardner, J.H. and Bell, D.T. (2007). Bauxite mining restoration by Alcoa world alumina Australia in Western Australia: social, political, historical, and environmental contexts. *Restoration Ecology 15* (Suppl. 4), 3–10. doi:10.1111/j.1526-100X.2007.00287.x.

Gellie, N.J.C., Mills, J.G., Breed, M.F., Lowe, A.J. (2017). Revegetation rewilds the soil bacterial microbiome of an old field. *Molecular Ecology 26*, 2895–2904. doi:10.1111/mec.14081.

Glen, M., Bougher, N.L., Colquhoun, I.J., Vlahos, S., Loneragan, W.A., O'Brien, P.A., Hardy, G.E.S.J. (2008). Ectomycorrhizal fungal communities of rehabilitated bauxite mines and adjacent, natural jarrah forest in Western Australia. *Forest Ecology Management 255*, 214–225. doi:10.1016/j.foreco.2007.09.007.

Government of Western Australia, Department of Jobs, Tourism, Science and Innovation. (2019). Western Australia Iron ore profile, www.jtsi.wa.gov.au.

Haney, R., Brinton, W., Evans, E. (2008). Soil CO2 respiration: comparison of chemical titration, IRGA and Solvita gel system. *Renewable Agriculture and Food Systems 23*, 1–6.

Harris, J.A., (2003). Measurements of the soil microbial community for estimating the success of restoration. *European of Journal Soil Sciences 54*, 801–808. doi:10.1046/j.1351-0754.2003.0559.x.

Hart, M., Gorzelak, M.A., McAmmond, B.M., Van Hamme, J.D., Stevens, J., Abbott, L., Whiteley, A.S., Nevill, P. (2019). Fungal communities resist recovery in sand mine restoration. *Frontiers in Forests and Global Change 2*, 1–9. doi:10.3389/ffgc.2019.00078.

Havel, J.J. (1975). Site-Vegetation Mapping in the Northern Jarrah Forest. I. Definition of Site-Vegetation Types. *Forests Department of WA* Bulletin N.

Herrick, J.E., Schuman, G.E., Rango, A. (2006). Monitoring ecological processes for restoration projects. *Journal for Nature Conservation 14*, 161–171. doi:10.1016/j.jnc.2006.05.001.

van der Heyde, M., Bunce, M.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 00*, 1–14. doi:10.1111/1755-0998.13148.

Ihrmark, K., Bödeker, I.T.M., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., … Lindahl, B.D. (2012). New primers to amplify the fungal ITS2 region - evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiology Ecology 82*, 666–677. doi:10.1111/j.1574-6941.2012.01437.x.

Kabiri, V., Raiesi, F., Ghazavi, M.A. (2016). Tillage effects on soil microbial biomass, SOM mineralization and enzyme activity in a semi-arid Calcixerepts. *Agriculture, Ecosystems and Environment 232*, 73–84. doi:10.1016/j.agee.2016.07.022.

Kumaresan, D., Cross, A.T., Moreira-Grez, B., Kariman, K., Nevill, P., Stevens, J.,

…Whiteley, A.S. (2017). Microbial functional capacity is preserved within engineered soil formulations used in mine site restoration. *Scientific Reports 7* (1), 1–9. doi:10.1038/s41598-017-00650-6.

Lauber, C.L., Ramirez, K.S., Aanderud, Z., Lennon, J., Fierer, N. (2013). Temporal variability in soil microbial communities across land-use types. *ISME* J 7, 1641–1650. doi:10.1038/ismej.2013.50.

Leff, J.W., Jones, S.E., Prober, S.M., Barberán, A., Borer, E.T., Firn, J.L., … Fierer, N. (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academdy of Sciences 112*, 10967–10972. doi:10.1073/pnas.1508382112.

Li, M.S. (2006). Ecological restoration of Mineland with particular reference to the metalliferous mine wasteland in China: a review of research and practice. *Science of the Total Environment 357* (1–3), 38–53. doi:10.1016/ j.scitotenv.2005.05.003.

McArthur, W.M. (1991). Reference Soils of South-Western Australia. Department of Agriculture, Perth.

McKenzie, N.L., van Leeuwen, S., Pinder, A.M. (2009). Introduction to the Pilbara biodiversity survey, 2002–2007. *Records of the Western Australian Museum*, Suppl. 78, 3. doi:10.18195/issn.0313-122x.78(1).2009.003-089.

McMurdie, P.J. and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One 8*. doi:10.1371/journal.pone.0061217.

Meena, V.S., Mishra, P.K., Bisht, J.K., Pattanayak, A. (2017). Strength of microbes in nutrient cycling: a key to soil health. In: *Agriculturally Important Microbes for Sustainable Agriculture*. pp. 1–356. doi:10.1007/978-981-10-5589-8.

deMendiburu, F. (2019). Agricolae: statistical procedures for agricultural research. R package version 1.3-1https://CRAN.R-project.org/ package=agricolae2019https://CRAN.R-project.org/package=agricolae

Menz, M.H.M., Dixon, K.W., Hobbs, R.J. (2013). Hurdles and opportunities for landscape-scale restoration. *Science 339* (6119), 526–527. doi:10.1126/ science.1228334.

Merzendorfer, H., (2006). Insect chitin synthases: A review. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology 176*, 1–15. doi:10.1007/s00360-005-0005-3.

Miller, B.P., Sinclair, E.A., Menz, M.H.M., Elliott, C.P., Bunn, E., Commander, L.E., Dalziell, E., et al. (2017). A framework for the practical science necessary to restore sustainable, resilient, and biodiverse ecosystems. *Restoration Ecology 25* (4), 605–617. doi:10.1111/rec.12475.

Mummey, D.L., Stahl, P.D., Buyer, J.S. (2002a). Microbial biomarkers as an indicator of ecosystem recovery following surface mine reclamation. *Applied Soil Ecology 21*, 251–259. doi:10.1016/S0929-1393(02)00090-2.

Mummey, D.L., Stahl, P.D., Buyer, J.S. (2002b). Soil microbiological properties 20 years after surface mine reclamation: spatial analysis of reclaimed and undisturbed sites. *Soil Biology and Biochemistry 34*, 1717–1725. doi:10.1016/S0038-0717(02)00158-X.

Muñoz-Rojas, M. (2018). Soil quality indicators: critical tools in ecosystem restoration. *Current Opinion in Environment Science and Health 5*, 47–52. doi:10.1016/j.coesh.2018.04.007.

Muñoz-Rojas, M., Erickson, T.E., Dixon, K.W., Merritt, D.J. (2016). Soil quality indicators to assess functionality of restored soils in degraded semiarid ecosystems. *Restoration Ecology 24*, S43–S52. doi:10.1111/rec.12368.

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

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2007. Biodiversity hotspots for conservation priorities. *Biodiversity and Conservation 16*, 853–858. doi:10.1080/21564574.1998.9650003.

Nguyen, N., Song, Z., Bates, S., Branco, S., Tedersoo, L., Menke, J., Schilling, J., Kennedy, P. (2016). FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology 20*, 241–248.

Nilsson, R.H., Larsson, K.H., Taylor, A.F.S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., … Abarenkov, K. (2019). The UNITE database for molecular

identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research 47*, D259–D264. doi:10.1093/nar/gky1022.

Nunan, N., Wu, K., Young, I.M., Crawford, J.W., Ritz, K. (2002). In situ spatial patterns of soil bacterial populations, mapped at multiple scales, in an arable soil. *Microbial Ecology 44*, 296–305. doi:10.1007/s00248-002-2021-0.

Nurulita, Y., Adetutu, E.M., Gunawan, H., Zul, D., Ball, A.S. (2016). Restoration of tropical peat soils: the application of soil microbiology for monitoring the success of the restoration process. *Agriculture Ecosystems and Environment 216*, 293–303. doi:10.1016/j.agee.2015.09.031.

O'Brien, S.L., Gibbons, S.M., Owens, S.M., Hampton-Marcell, J., Johnston, E.R., Jastrow, J.D., … Antonopoulos, D.A. (2016). Spatial scale drives patterns in soil bacterial diversity. *Environmental Microbiology 18*, 2039–2051. doi:10.1111/1462-2920.13231.

Ochoa-Villarreal, M., Aispuro-Hernndez, E., Vargas-Arispuro, I., Ngel, M. (2012). Plant cell wall polymers: function, structure and biological activity of their derivatives. In: *Polymerization*. InTech. doi:10.5772/46094.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., … Maintainer, H. W. (2019). vegan: Community Ecology Package. R package version 2.5-5. https://CRAN.R-project.org/package=vegan. *Community Ecology Package*, *2*(9), 1–297.

Philippot, L., Andersson, S.G.E., Battin, T.J., Prosser, J.I., Schimel, J.P., Whitman, W.B., Hallin, S. (2010). The ecological coherence of high bacterial taxonomic ranks. *Nature Reviews Microbiology 8*, 523–529. doi:10.1038/nrmicro2367.

R Core Team. (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org.

Rajapaksha, R.M.C.P., Tobor-Kaplon, M.A., Ba˚a˚th, E. (2004). Metal toxicity affects fungal and bacterial activities in soil differently. *Applied Environmental Microbiology 70*, 2966–2973. doi:10.1128/AEM.70.5.2966.

Rokich, D.P., Dixon, K.W., Sivasithamparam, K., Meney, K.A. (2000). Topsoil

handling and storage effects on woodland restoration in Western Australia. *Restoration Ecology 8* (2), 196–208. doi:10.1046/j.1526-100X.2000.80027.x.

Roncero, C. (2002). The genetic complexity of chitin synthesis in fungi. *Current Genetics 41*, 367–378. doi:10.1007/s00294-002-0318-7.

Sipilä, T.P., Yrjälä, K., Alakukku, L., Palojärvi, A. (2012). Cross-site soil microbial communities under tillage regimes: fungistasis and microbial biomarkers. *Applied Environmental Microbiology 78*, 8191–8201. doi:10.1128/ aem.02005-12.

Šmejkalová, M., Mikanová, O., Borůvka, L. (2003). Effects of heavy metal concentrations on biological activity of soil micro-organisms. *Plant, Soil and Environment 49*, 321–326.

Stevens, P.R. and Walker, T.W. (1970). The chronosequence concept and soil formation. *The Quaterly Review of Biology 45*, 333–350.

Sun, S., Li, S., Avera, B.N., Strahm, B.D., Badgley, B.D. (2017). Soil bacterial and fungal communities show distinct recovery patterns during Forest ecosystem restoration. *Applied Environmental Microbiology 83*, 1–14. doi:10.1128/ AEM.00966-17.

Suzuki, C., Nagaoka, K., Shimada, A., Takenaka, M. (2009). Bacterial communities are more dependent on soil type than fertilizer type, but the reverse is true for fungal communities. *Soil Science and Plant Nutrition 55*, 80–90. doi:10.1111/j.1747-0765.2008.00344.x.

Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., … Knight, R. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature 551*, 457–463. doi:10.1038/nature24621.

Trudgen, M. (1977). A Report on the Flora and Floristic Community Types Found on Parts of Sand Mining Leases Held by Rocla in the Gnangara Area. Report Prepared for Rocla Quarry Products, Perth, Western Australia.

Turner, S., Pryer, K.M., Miao, V.P.W., Palmer, J.D. (1999). Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *The Journal of Eukaryotic Microbiology 46*, 327–338. doi:10.1111/j.1550-7408.1999.tb04612.x.

UN Environment Programme. (2020). Decade on ecosystem restoration 2021-2030. https://www.decadeonrestoration.org/

Vitousek, P.M., Cassman, K., Cleveland, C., Crews, T., Field, C.B., Grimm, N.B., … Sprent, J.I. (2002). Towards an ecological understanding of biological nitrogen fixation. *Biogeochemistry 57/58*, 1–45. doi:10.1023/A:1015798428743.

Walkley, A., and Black, I.A. (1934). An examination of the Degtjareff method for determining soil organic matter, a proposed modification of the chromic acid titration method. *Soil Science 37*, 29–39.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied Environmental Microbiology 73*, 5261–5267. doi:10.1128/AEM.00062-07.

White, T.J., Bruns, T.D., Lee, S.B., Taylor, J.W. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols: A Guide to Methods and Applications. pp. 315–323.

Wilkinson, S., Davy, S., Bunce, M., Stat, M. (2018). Taxonomic identification of environmental DNA with informatic sequence classification trees. *PeerJ Preprints 1–2*. doi:10.7287/peerj.preprints.26812.

Yan, D.F., Mills, J.G., Gellie, N.J.C., Bissett, A., Lowe, A.J., Breed, M.F. (2018). High-throughput eDNA monitoring of fungi to track functional recovery in ecological restoration. *Biological Conservation 217*, 113–120. doi:10.1016/j.biocon.2017.10.035.

Yan, D., Bissett, A., Gellie, N., Mills, J.G., Lowe, A.J., Breed, M.F. (2020). Soil bacterial community differences along a coastal restoration chronosequence. *Plant Ecology 221*(9), 795-811. doi:10.1007/s11258-019-00979-0.

Yang, G., Wagg, C., Veresoglou, S.D., Hempel, S., Rillig, M.C. (2018). How soil biota drive ecosystem stability. *Trends in Plant Science 23*, 1057–1067. doi:10.1016/j.tplants.2018.09.007.

## 3.9 Supplementary Information

**Table S3.9.1** Results for the two-way ANOVAs run on alpha diversity of bacteria and fungi in restoration and reference sites

| ***Grouping factor*** | ***ANOVA reslts*** | | |
|---|---|---|---|
| Bacteria: Observed Richness | *Df* | *F* | *P value* |
| *Restoration* | 13 | 4.3055 | <0.001 |
| *Location* | 2 | 1.6513 | 0.198 |
| *Restoration: Location* | 3 | 4.5874 | 0.005 |
| *Residuals* | 74 | | |
| Bacteria: Shannon Diversity | *Df* | *F* | *P value* |
| *Restoration* | 13 | 6.8157 | <0.001 |
| *Location* | 2 | 6.4058 | 0.003 |
| *Restoration: Location* | 3 | 7.189 | <0.001 |
| *Residuals* | 74 | | |
| Fungi: Observed Richness | *Df* | *F* | *P value* |
| *Restoration* | 13 | 4.494 | <0.001 |
| *Location* | 2 | 27.794 | <0.001 |
| *Restoration: Location* | 3 | 6.884 | <0.001 |
| *Residuals* | 76 | | |
| Fungi: Shannon Diversity | *Df* | *F* | *P value* |
| *Restoration* | 13 | 5.615 | <0.001 |
| *Location* | 2 | 2.359 | 0.101 |
| *Restoration: Location* | 3 | 2.351 | 0.079 |
| *Residuals* | 76 | | |

**Table S3.9.2** Mantel test results for spatial autocorrelation. r indicates the correlation between the community similarity and the spatial distribution. p indicates the significance of the correlation (alpha=0.05). JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

| | *Bacteria* | | *Fungi* | |
|---|---|---|---|---|
| | r | p | r | p |
| *Overall* | 0.673 | 0.001 | 0.446 | 0.001 |
| *JF* | 0.31 | 0.03 | 0.347 | 0.001 |
| *PB* | 0.478 | 0.001 | 0.318 | 0.001 |
| *SCP* | 0.027 | 0.317 | 0.05 | 0.311 |
| *Pooled replicates* | r | p | r | p |
| *JF* | -0.035 | 0.383 | -0.086 | 0.536 |
| *PB* | 0.32 | 0.117 | 0.197 | 0.187 |
| *SCP* | -0.192 | 0.613 | 0.196 | 0.264 |

**Table S3.9.3** The effect of restoration age on abundance of fungal sequences in each trophic mode

| | *Jarrah* | | | *Pilbara* | | | *SCP* | | |
|---|---|---|---|---|---|---|---|---|---|
| ***Trophic Mode*** | **F** | **p.adj** | **Direction** | **F** | **p.adj** | **Direction** | **F** | **p.adj** | **Direction** |
| *Pathotroph* | 2.53 | 0.017 | Decreasing | 2.67 | 0.019 | Variable | 1.14 | 0.389 | |
| *Pathotroph-Saprotroph* | 4.67 | 0.008 | Increasing | 0.79 | 0.594 | | 5.20 | 0.009 | Variable |
| *Pathotroph-Symbiotroph* | 4.82 | 0.004 | Decreasing | 2.78 | 0.027 | Decreasing | 4.33 | 0.004 | Variable |
| *Pathotroph-Saprotroph-Symbiotroph* | 8.77 | 0.004 | Increasing | NA | NA | NA | 2.64 | 0.015 | |
| *Saprotroph* | 1.38 | 0.289 | | 1.64 | 0.185 | | 1.64 | 0.181 | |
| *Saprotroph-Symbiotroph* | 1.12 | 0.389 | | 3.56 | 0.006 | Decreasing | 3.56 | 0.006 | Decreasing |
| *Symbiotroph* | 6.23 | 0.004 | Increasing | 2.17 | 0.033 | Variable | 4.64 | 0.004 | Increasing |

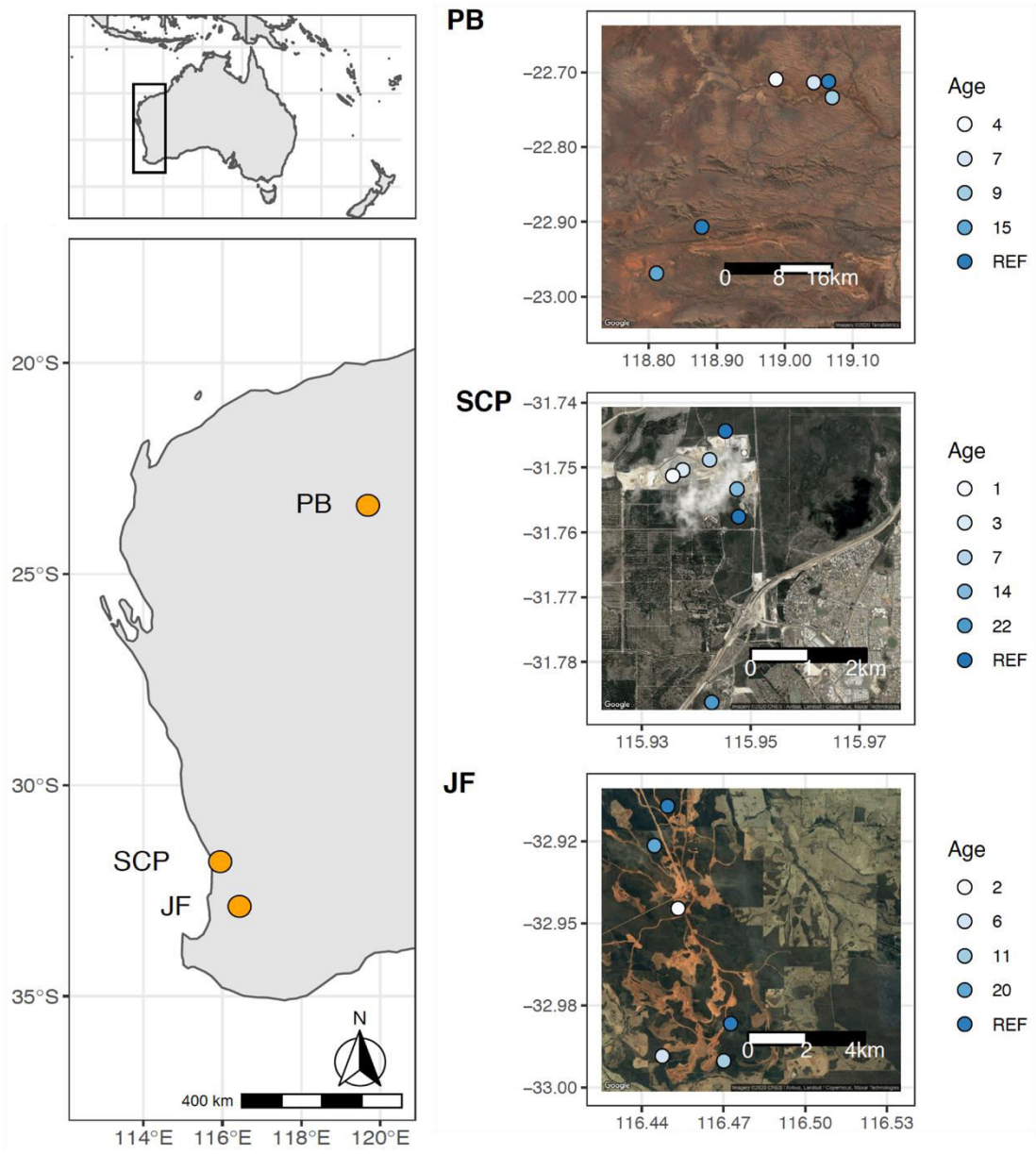


**Figure S3.9.1** Map of site locations. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

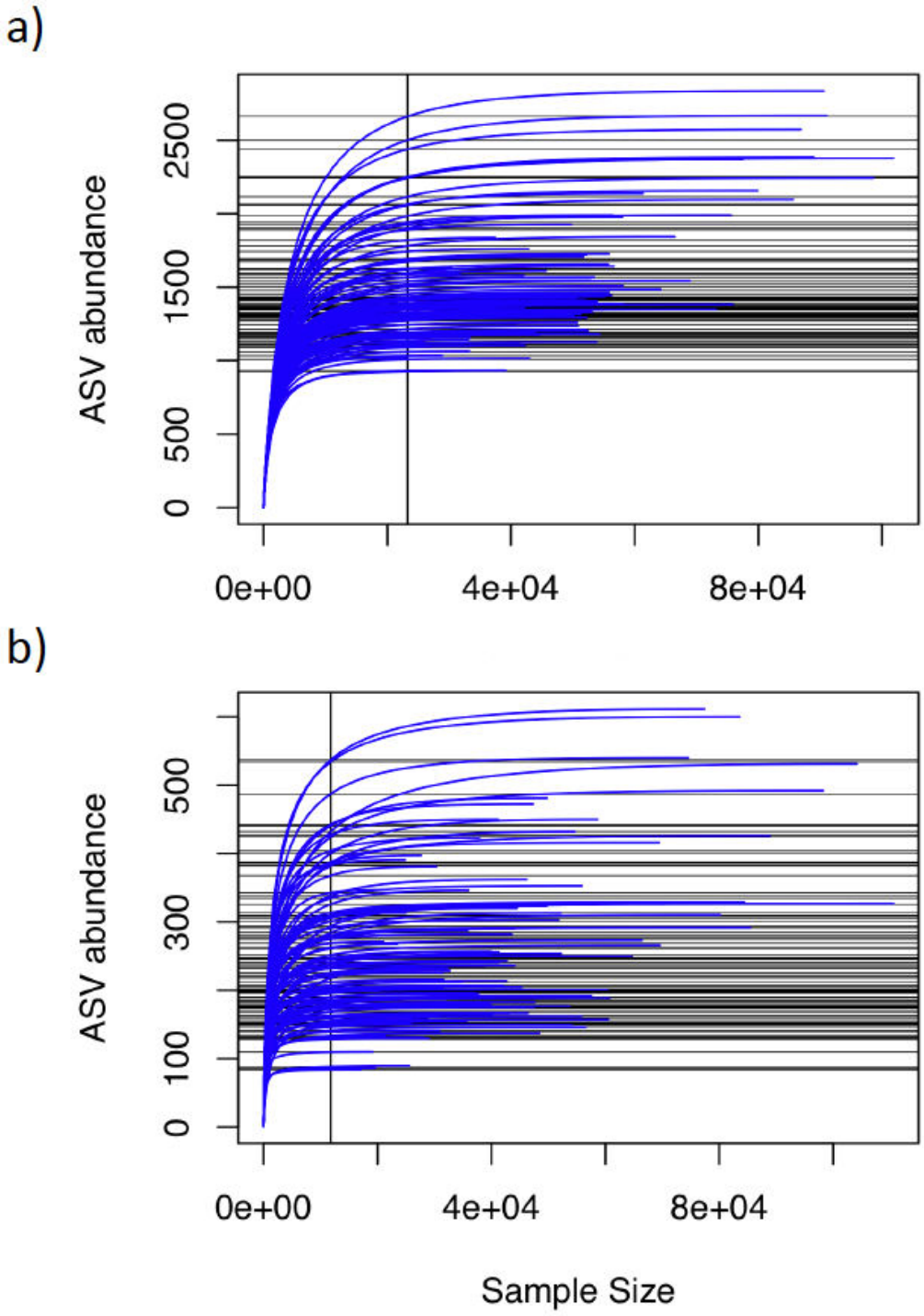


**Figure S3.9.2** Rarefaction curves for a) 16S Bacteria and b) ITS2 fungal sequences. Vertical lines indicate level to which samples were rarefied.

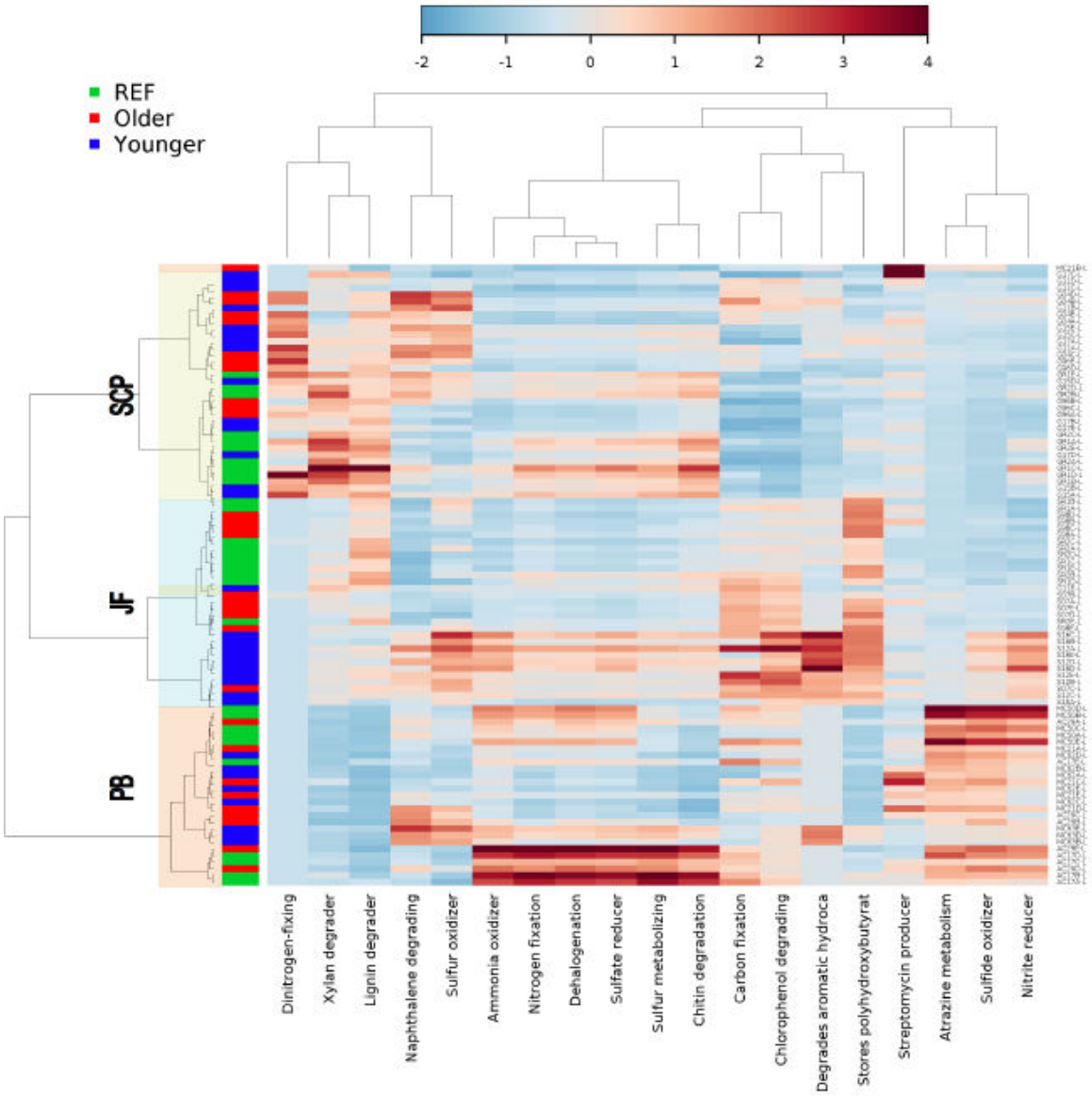


**Figure S3.9.3** Heatmap of bacterial metabolism in restoration and reference sites at the three locations. There was a significant interaction between location and restoration ($F_{4,92} = 4.072$, $p < 0.001$), and restoration ($F_{2,92} = 5.22$, $p < 0.001$), and location ($F_{2,92} = 46.704$, $p < 0.001$) were both highly significant.

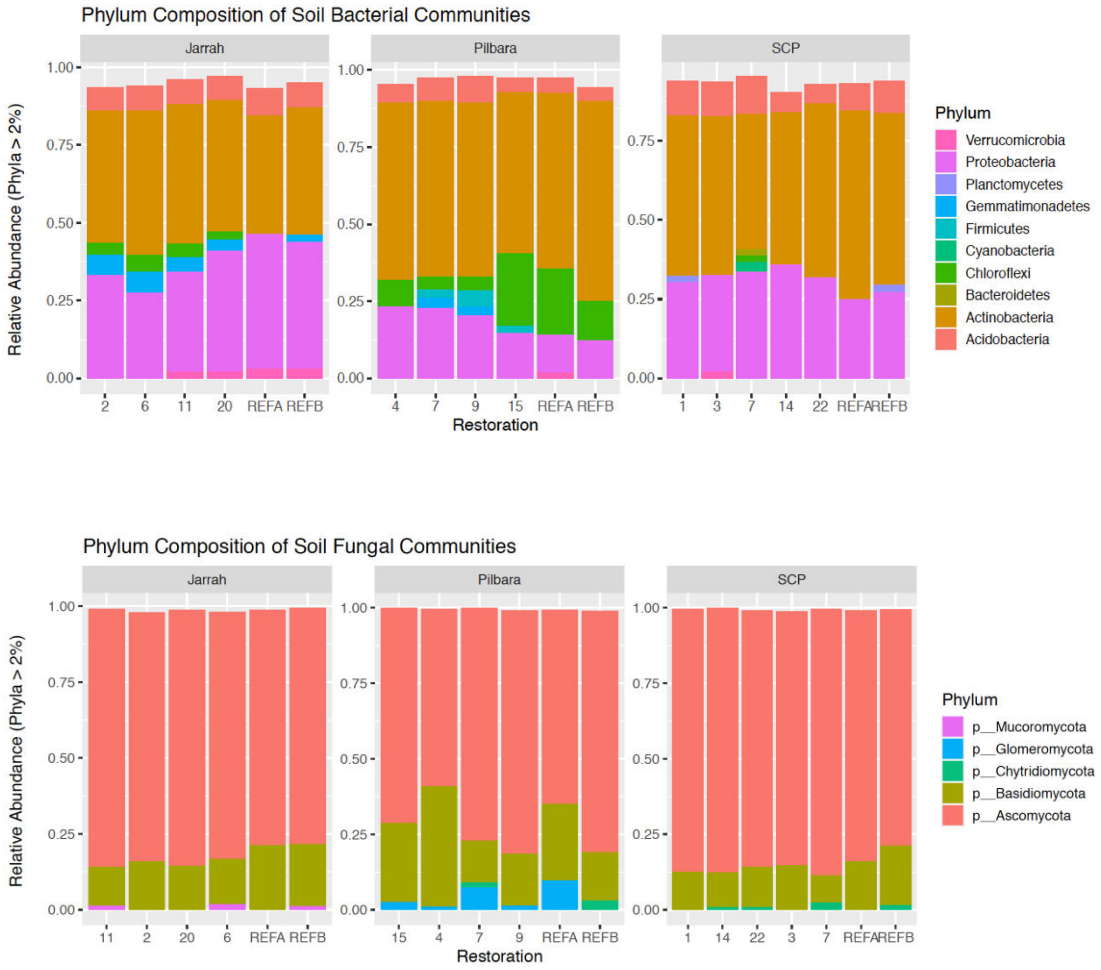


**Figure S3.9.4** Phylum composition of bacteria and fungi at restoration and reference sites.

# CHAPTER 4

## SCAT DNA PROVIDES IMPORTANT DATA FOR EFFECTIVE MONITORING OF MAMMAL AND BIRD BIODIVERSITY





## 4.1 Preface

*This chapter consists of a manuscript under review titled '**Scat DNA provides important data for effective monitoring of bird and mammal biodiversity**' [Biodiversity and Conservation 2020]. The content in section 4.2 is the same as the submitted manuscript with only minor changes in formatting to accommodate thesis referencing.*

This chapter evaluates the potential of pooled scat collections, a novel eDNA substrate, to detect vertebrate diversity. Fauna recovery during restoration is often overlooked, though they can provide indication of ecosystem functionality such as seed dispersal, forage availability, habitat suitability. Here, this new substrate is tested across three different ecosystems to determine suitability in a range of habitats. We examine the differences in bird and mammal diversity detected from the three different ecosystem, and whether this method can differentiate between restored and reference sites within each ecosystem.

### 4.1.1 Acknowledgements

This work was supported by the Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration (ICI150100041) and the Pawsey Supercomputing Centre, funding from the Australian Government and the Government of Western Australia. We thank the mining companies BHP, Hanson Construction Material, and South32 for facilitating access to sites for sampling. We would also like to thank Sheree Walters for help with sample collection and the members of the Trace and Environmental DNA (TrEnD) Laboratory for support with metabarcoding workflows and bioinformatics.

### 4.1.2 Data Accessibility

Sequencing and sample data and is available at the Dryad Digital Repository: https://doi.org/10.5061/dryad.p2ngf1vnt

### 4.1.3 Author Contributions

MvH conducted the study and wrote the manuscript. MvH, PN, MB, NW, and GW-J were involved in the experimental design. Samples were collected and

processed by MvH; molecular and bioinformatics work was performed by MvH; all data was analyzed and processed by MvH; statistical analysis was done by MvH; the manuscript was edited by all authors.

## SCAT DNA PROVIDES IMPORTANT DATA FOR EFFECTIVE MONITORING OF MAMMAL AND BIRD BIODIVERSITY

van der Heyde, M.[1,3*], Bateman, P.W[2], Bunce, M. [3,4], Wardell-Johnson, G.[1], White, N.E. [3], Nevill, P.[1, 3]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, GPP Box U1987, Perth, Western Australia, 6845

[2]Behavioural Ecology Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, GPP Box U1987, Perth, Western Australia, 6845

[3]Trace and Environmental DNA Laboratory, School of Life and Molecular Sciences, Curtin University, GPP Box U1987, Perth, Western Australia, 6845

[4]Environmental Protection Authority, 215 Lambton Quay, Wellington 6011, New Zealand.

*Corresponding author

## 4.2 Abstract

Fauna has long been neglected in the monitoring of ecological restoration, despite the key role they play in ecosystem function. Vertebrate surveys can be time consuming and costly, often requiring multiple methodologies and taxonomic expertise, making comprehensive monitoring cost prohibitive. Here we evaluate a new method of assessing mammal and bird diversity through the genetic identification of scat collections. Using DNA metabarcoding of scat collections from three bioregions we generated bird and mammalian assemblage data and distinguished between sites with different restoration histories. However, scat detectability was affected by environmental conditions (e.g. rainfall and soil), suggesting that our approach is most applicable at certain times of year or in arid (or semi-arid) environments with rocky soils, where conditions are favorable for scat preservation. Taken together these data provide a pathway to: plan, monitor and establish best-practice when restoring landscapes and add to the growing body of literature on the value of DNA metabarcoding in biomonitoring applications.

## 4.3 Introduction

The United Nations Strategic plan for Biodiversity 2011-2020 has at its core, the goal to reduce species decline. Efforts to prevent losses in biodiversity depend on monitoring in order to (i) identify and preserve high biodiversity areas (Sarkar et al. 2006), (ii) detect and manage invasive species (Clavero et al. 2009; Doherty et al. 2016), and (iii) assess and improve ecological restoration outcomes (Mcdonald et al. 2016; Miller et al. 2017). Ecological restoration is a globally vital tool for combating the loss of biodiversity and ecosystem services (Benayas et al. 2009; UN Environment Programme 2020). Assessing the success of restoration programs is crucial to improve outcomes and ensure that goals are being met (S. A. Thompson and Thompson 2004). However, there is a strong bias towards vegetation measures in restoration monitoring and a lack of focus on faunal communities (Cristescu, Frère, and Banks 2012; Cross, Tomlinson, Craig, Dixon, et al. 2019). It is often assumed that fauna return to disturbed areas following the return of vegetation (Palmer, Ambrose, and Poff 1997; Cristescu, Frère, and Banks 2012), but this is not always the case (Cristescu et al. 2013). Fauna are responsible for ecosystem functions such as pollination, seed dispersal, and nutrient cycling (Herrera 1995; Fleming and Muchhala 2008), and can act as facilitators of restoration (Catterall 2018). As such, they should be included in assessments of restoration success (Cross, Bateman, and Cross 2020).

Faunal recolonization in restoration areas is variable, and affected by factors such as time since restoration work was completed, and the taxa in question (Cristescu et al. 2012). Birds, for example, are generally successful colonizers that soon reach levels of richness and diversity equivalent to undisturbed areas, while mammal and reptile richness often remains lower in restoration sites (Cristescu et al. 2012). Community composition usually differs because sites undergoing restoration host more opportunistic and generalist species than do undisturbed areas (Nichols and Nichols 2003; Cristescu, Frère, and Banks 2012). Older restoration sites are more likely to have equivalent levels of richness and community composition to undisturbed sites (Nichols and Nichols 2003; Cristescu, Frère, and Banks 2012). Unfortunately, less than half of the literature on the role of fauna in ecological restoration represents vertebrates, resulting in a poor understanding of vertebrate recolonization (Cross, Tomlinson, Craig, Dixon, et al. 2019).

The costs of comprehensive ecological monitoring of vertebrates are often prohibitive. This is because vertebrate surveys require a combination of different methodologies (e.g. bird surveys, camera traps, pitfall traps, and trace identifications) necessitating teams of experts in the field for long periods, often in remote locations (Environmental Protection Authority 2010). This has spurred the development of new technologies for biodiversity monitoring, including environmental DNA (eDNA) or DNA metabarcoding (Pierre Taberlet, Coissac, et al. 2012). DNA metabarcoding refers to a process wherein specific 'barcoding' regions of DNA are sequenced in parallel from environmental samples such as soil or water (Taberlet et al. 2012). It is relatively fast and inexpensive (Yu et al. 2012), and able to identify multiple species in a single sample. While shown to be successful in detecting vertebrate diversity in aquatic systems (Goldberg et al. 2011; Jerde et al. 2011; West et al. 2020), there are fewer studies testing eDNA methods for vertebrate diversity assessment in terrestrial systems (Fernandes et al. 2018; Mieke van der Heyde et al. 2020). Some studies have had success using soil to detect vertebrate diversity (K. Andersen et al. 2012; Drummond et al. 2015). However, these were conducted in temperate climates which were conducive for DNA preservation. The method may not necessarily be as appropriate in environments that have high temperatures and UV radiation, such as many environments in Australia. Previous work in Australia testing multiple terrestrial substrates suggests bulk scat samples are the most reliable method of detecting vertebrate DNA (van der Heyde et al., 2020). Scats are among the most commonly found signs of animal presence, and often the only sign of animal species that may be crepuscular, nocturnal, or occur in low densities (Wilson and Delahay 2001). With genetic analyses, scats may also be used to determine diet (Berry et al. 2017; Arteaga Claramunt et al. 2018), although this would require targeting scat collection rather than the broad collection of 'pooled' scats to determine vertebrate assemblages.

Here we test a new tool for monitoring vertebrate communities in mine site restoration by applying DNA metabarcoding to bulk scat samples collected in restoration sites and compare the data to reference sites. This study was conducted in three different bioregions of Western Australia to determine if results were consistent across different climates and ecosystems. Specifically, we use metabarcoding to compare bird and mammal assemblages and better understand the strengths and

limitations of this new approach to biomonitoring. The results from this study will be used to assess the feasibility of scat metabarcoding as a practical tool to deploy in restoration monitoring. Given the Bonn Challenge goal to restore 350 million $km^2$ of degraded terrestrial ecosystems by 2030 (K. Suding et al. 2015) improved monitoring of restoration outcomes will ensure we get the best value from the considerable financial investment required to meet this ambitious target.

## 4.4 Methods

### 4.4.1 Study Sites

Samples were collected at three locations. The Swan Coastal Plain (SCP) and Jarrah Forest (*Eucalyptus marginata*) (JF) Bioregions are both within the Mediterranean-climate Southwest Australian Global Biodiversity Hotspot (Myers et al. 2007), while the third is in the semi-arid Pilbara Bioregion (PB). At each location, we sampled two restoration sites under nine years old, two restoration sites over 9 years old, and two reference sites (Figure 4.1). Therefore, a total of 18 sites were targeted for scat collection.

The Banksia Woodland of the Coastal Plain is a biodiverse ecosystem with a Mediterranean climate with cool, wet winters and hot dry summer. The area has a mean annual minimum temperature of 12.8°C, a mean annual maximum of 24.7°C, and a mean annual rainfall of 757 mm (Australian Bureau of Meteorology). The dominant tree species are *Banksia attenuata* and *B. menziesii*, with less dominant *Eucalyptus todtiana* and *Nuytsia floribunda.* The understory consists of woody species of Myrtaceae, Fabaceae, Proteaceae, and Ericaceae, and non-woody species in Anthericaceae, Stylidaceae, Cyperaceae, and Haemodoraceae (Trudgen 1977). In October 2018, we sampled six sites at a Hanson Construction Materials sand quarry in Lexia (-31.76°, 115.95°); two reference sites and restoration sites 1, 3, 14, 22 years old. The four restoration sites were established to restore mined areas to surrounding native Banksia woodlands. All restoration was done by Hanson and previous mine owners and included seeding with native species and planting of dominant tree species (*Banksia attenuata*, *B. menziesii*, and *Eucalyptus todtiana*). Plant species richness and density tends to be higher in restoration sites, and percent cover increases with restoration age and is highest in reference sites (Benigno, Dixon, and Stevens 2013). According to a previous study looking at banksia woodland restoration, birds colonize rapidly and there is little difference in richness or assemblages between restoration and reference (Comer and Wooller 2002). There is less data available of mammalian recovery; some of the more common animals include the Western Grey Kangaroo (*Macropus fugilinosus*) and the Australian Raven (*Corvus coronoides*) (www.ala.org.au).

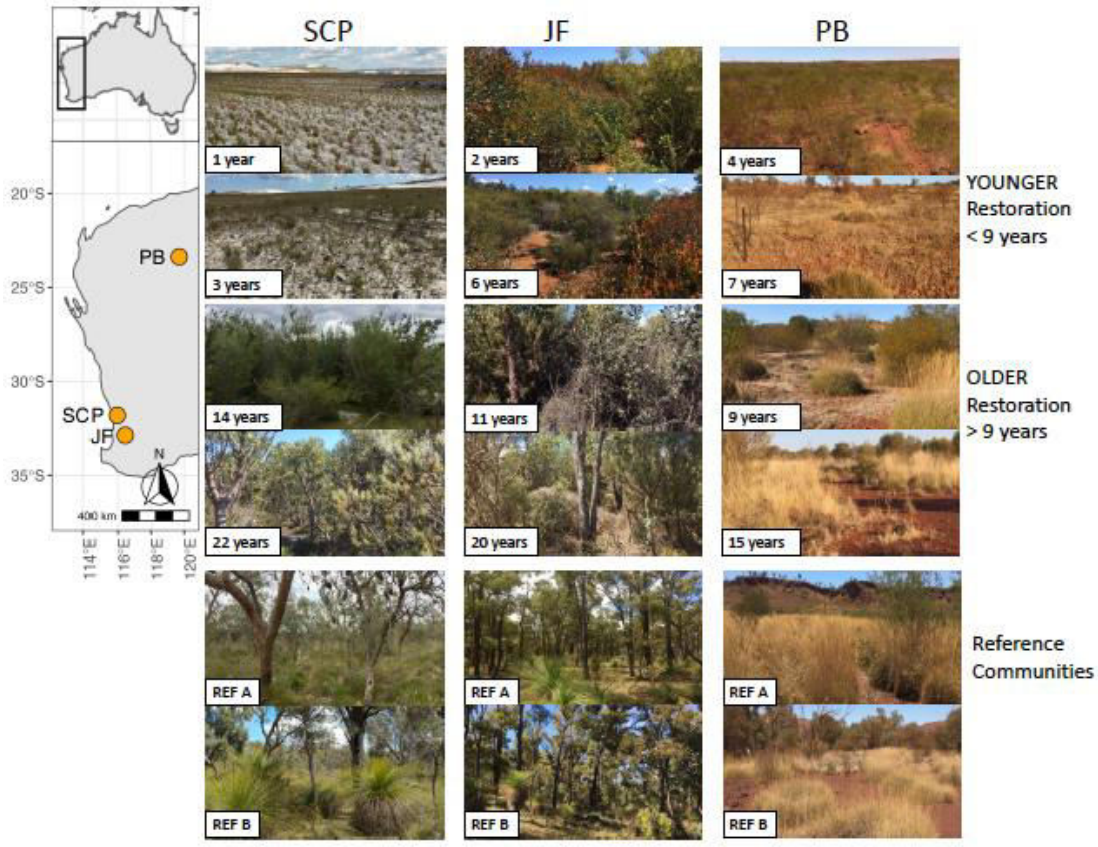


**Figure 4.1** Chronosequences of mining restoration where scat samples were collected. Restoration sites shown with the number of years restoration from 1 to 22 years. Reference sites shown below. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

The Jarrah location has a similar Mediterranean climate to the Coastal Plain site. The area has a mean annual minimum temperatures of 8.6°C, a mean annual maximum of 23.7°C, and mean annual rainfall of 668.9 mm (Australian Bureau of Meteorology). The overstory vegetation is dominated by *E. marginata,* with *E. patens,* and *E. wandoo* also common. The understory is sclerophyllous and floristically diverse, dominated by several families (i.e. Fabaceae, Asteraceae, Proteaceae, Dasypogonaceae, Myrtaceae) (Havel 1975). We sampled six sites from the bauxite mine South32 (-32.96°, 116.48°) in October 2018; two reference sites and restoration sites 2, 6, 11, 20 years old (Figure 4.1). All restoration was undertaken by South32 or the previous mine owners and included seeding with over 100 native species. Reference and restoration sites are dominated by species of Myrtaceae and Fabaceae. Total cover increases with age of restoration and reaches

similar cover percentages to reference sites. There are uncontrolled feral animals including *Sus scrofa* (pig), and *Vulpes vulpes* (fox) (Williams and Mitchell 2002). Fauna surveys indicate bird diversity increases with age of restoration and a greater abundance of Carnaby's black-cockatoo (*Calyptorhynchus latirostris)* and new-holland honeyeater (*Phylidonyris novaehollandiae*) in restoration sites (Data from South32). Mammalian richness remains similar throughout restoration, although the restoration sites have a higher abundance of house mouse (*mus musculus*) and Western Grey Kangaroo (*M. fugilinosus*) (Data from South32).

The open woodlands of the Pilbara occur in a hot, arid location of north-western Australia with most rainfall occurring in the summer, associated with cyclonic activity (McKenzie, van Leeuwen, and Pinder 2009). Temperatures have a mean annual minimum of 15°C and mean annual maximum of 30.6 °C, with a mean annual rainfall of 263.8 mm (Australian Bureau of Meteorology). The open snappy gum (*E. leucophloia*) woodland occurs over hummock grasses (*Triodia wiseana, T. basedowii, T. lanigera*) and low acacia shrubs (McKenzie et al. 2009). This site is located in a region that includes globally significant iron ore resources and accounts for up to 39% of global iron ore production (Government of Western Australia 2019). We sampled 6 sites at a BHP iron ore mine (-22.84°, 118.95°) in September 2018, two reference sites and restoration sites 4, 7, 11, and 15 years old (Figure 4.1). These sites are within 1 km of several ephemeral creeks. Restoration was conducted by the mine owners and restoration areas tended to have higher coverage of woody shrubs (*Acacia*), while reference sites and older restoration areas has more hummock grasses (*Triodia*) and mature gum (*E. leucophloia*) (Data from BHP). There is limited data in the Pilbara, none of the 31 studies examining vertebrate recovery in Australasia used sites restored post iron-ore extraction (Cross, Tomlinson, Craig, Dixon, et al. 2019). Some of the more common animals in the area, according to the Atlas of Living Australia (www.ala.org.au) include the Pebble-mound mouse (*Pseudomys chapmani*) and the Galah (*Eolophus raseicapilla*)

### 4.4.2 Sample Collection

Scat samples were collected from 4 randomly chosen points at each restoration/reference site. At each point, scat was collected over an 8-minute period along a 200 m transect, at least 5 m from the edge of the site. A portion of all scats

seen during this transect were collected; herbivore, carnivore, and bird, no matter how degraded. All scats collected on a transect were pooled to form one sample. Latex gloves were used in collection to minimize contamination and new gloves were used for each sample. A total of 72 pooled scat samples were collected, 24 from each location. Scat samples were frozen as soon as possible in a mobile freezer and transported to the Trace and Environmental DNA lab where they were stored at -20°C until processed.

### 4.4.3 Sample processing

A portion of each scat in the sample was put in a 50 mL falcon tube. Scats that could be easily identified was noted for each sample, to provide a morphological comparison to sequencing results. These included identifying Emu scat because it is very distinctive, bird scat (not including Emu), and kangaroo and wallaby scat. Scat was homogenized in 50mL falcon tubes using a TissueLyser (Qiagen) with 4 steel balls (5mm) for 3 minutes in 30 second intervals. DNA was extracted from 250uL scat homogenate, using the DNeasy PowerLyzer PowerSoil kit (Qiagen) on the QiaCube Connect automated platform (Qiagen). A final elution volume of 100 µL (Tris buffer) was used, and extraction controls (blanks) were included in every set of extractions. Quantitative PCR (qPCR) was run on neat extracts and a 1/10 dilution to see if samples exhibited inhibition, and to determine optimal DNA input for each sample (Dáithí C. Murray, Coghlan, and Bunce 2015). The qPCR assays were run with primers that target the 12S gene *12SV5*-F/R (Riaz et al. 2011). This assay is both short enough (98 bp), to pick up degraded DNA and broad enough to amplify mammalian, avian, and reptilian DNA.

The PCR mix for quantitation contained: 2.5 mM MgCl2 (Applied Biosystems, USA), 1× PCR Gold buffer (Applied Biosystems), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4 mg/ml bovine serum albumin (Fisher Biotec, Australia), 0.4 µmol/L forward and reverse primer, 1 U AmpliTaq Gold DNA polymerase (Applied Biosystems) and 0.6 µl of a 1:10,000 solution of SYBR Green dye (Life Technologies, USA). The qPCRs were run on a StepOne Plus (Applied BioSystems) real-time qPCR instrument with the following conditions: 5 min at 95°C, 40 cycles of 95°C for 30s, 30s at the 60°C and 45s at 72°C, a melt curve stage of 15s at 95°C, 1 min at 60°C and 15s at 95°C, ending with 10 min elongation at

72°C. Extraction controls were included in qPCR assays. DNA dilutions that showed uninhibited amplification (Murray et al. 2015) were selected for further analysis.

The DNA extracts that were deemed of sufficient template number and free of inhibition (qPCR described above) were assigned a unique combination of multiplex identifier (MID) tags. These MID tags were incorporated into fusion tagged primers, and, to prevent cross contamination, none of the primer-MID tag combinations had been used previously in the lab. Fusion PCRs for scat samples were done in duplicate and the PCR master mixes were prepared in a dedicated ultra-clean lab before DNA was added to minimize the potential for contamination. The PCR cycling conditions were as stated above for the quantification qPCRs. Samples that amplified successfully were then pooled into approximately equimolar concentrations to produce a final metabarcoding library that was size-selected (150-450bp) using a PippinPrep 2% ethidium bromide cassette (Sage Science, Beverly, MA, U.S.A). Libraries were cleaned using a QIAquick PCR Purification Kit (Qiagen, Germany) and final library quantification with a Qubit Fluorometric Quantitation (Thermo Fisher Scientific). Sequencing was performed on the Illumina MiSeq platform using the 300 cycle V2 kit following the manufacturer's protocol.

### 4.4.4 Sequencing analysis

Bioinformatics and taxonomic assignments were performed on the Pawsey Supercomputer. Sequences were demultiplexed using OBITools (Boyer et al. 2016b). Further sequence processing carried out in R using the "DADA2" package (Callahan et al. 2016) where sequences were quality filtered, the error rates were estimated, and the sequences were dereplicated. The error rates are then used in the sample inference stage to remove sequences likely to be errors and leave Amplicon Sequence Variants (ASV). These ASVs are equivalent to zero radius operational taxonomic units (ZOTUs) in usearch (Edgar 2016). The sequence table was constructed, chimeras removed, and curated using LULU (Frøslev et al. 2017). Taxonomy was assigned using BLASTN to search against the GenBank reference database (https://www.ncbi.nlm.nih.gov/genba nk/). Minimum percent coverage was set at 97%, minimum percent identification at 95%, the top 10 hits were considered and taxonomic identification was dropped to the lowest common ancestor if sequences matched with more than one taxa. Any identifications with less than 97%

identity were dropped to family level. Taxonomic identifications were validated against the Atlas of Living Australia database (www.ala.org.au) to verify the organisms occurred in the area. If an identification did not occur in the area, the identification was dropped to the lowest common ancestor that occurred in the area or removed from the dataset (See Table S4.9.2 for details). For example, the African bird *Ardeotis kori* was assigned to an ASV, and this identification was dropped to genus level as there is another *Ardeotis* species that occurs in the area.

### 4.4.5 Statistics

All statistics were performed using R 3.5.3 (R Core Team 2018). Sequence variants that were present in the extraction controls were removed from the dataset, then all human sequences were removed using the 'phyloseq' package (McMurdie and Holmes 2013). Taxa within scat samples were filtered by relative sequence abundance varying from 0.01% to 0.5% relative abundance. The results of the filtered data sets were then compared to morphological identification of scats to see what level of filtering produced the fewest false positives and negatives in the sequencing data. Alpha diversity was calculated using the 'phyloseq' package (McMurdie & Holmes 2013) and tested using a one way analysis of variance (ANOVA). If the ANOVA indicated there were significant differences between sites a Tukey HSD test was performed from the 'agricolae' package (Mendiburu 2019) to show which sites were significantly different from the others. A similarity profile (SIMPROF) analysis was performed using the R package 'clustsig' (Whitaker and Christman 2014) to determine if restoration and reference sites are clustered together. To see which species were associated with restoration or reference sites at the three locations we performed a multipattern analysis using the 'indicspecies' package in R (De Caceres and Legendre 2009). P-values for the multipattern analysis were adjusted using the 'BH' method (Benjamini and Hochberg 2007) to account for multiple testing.

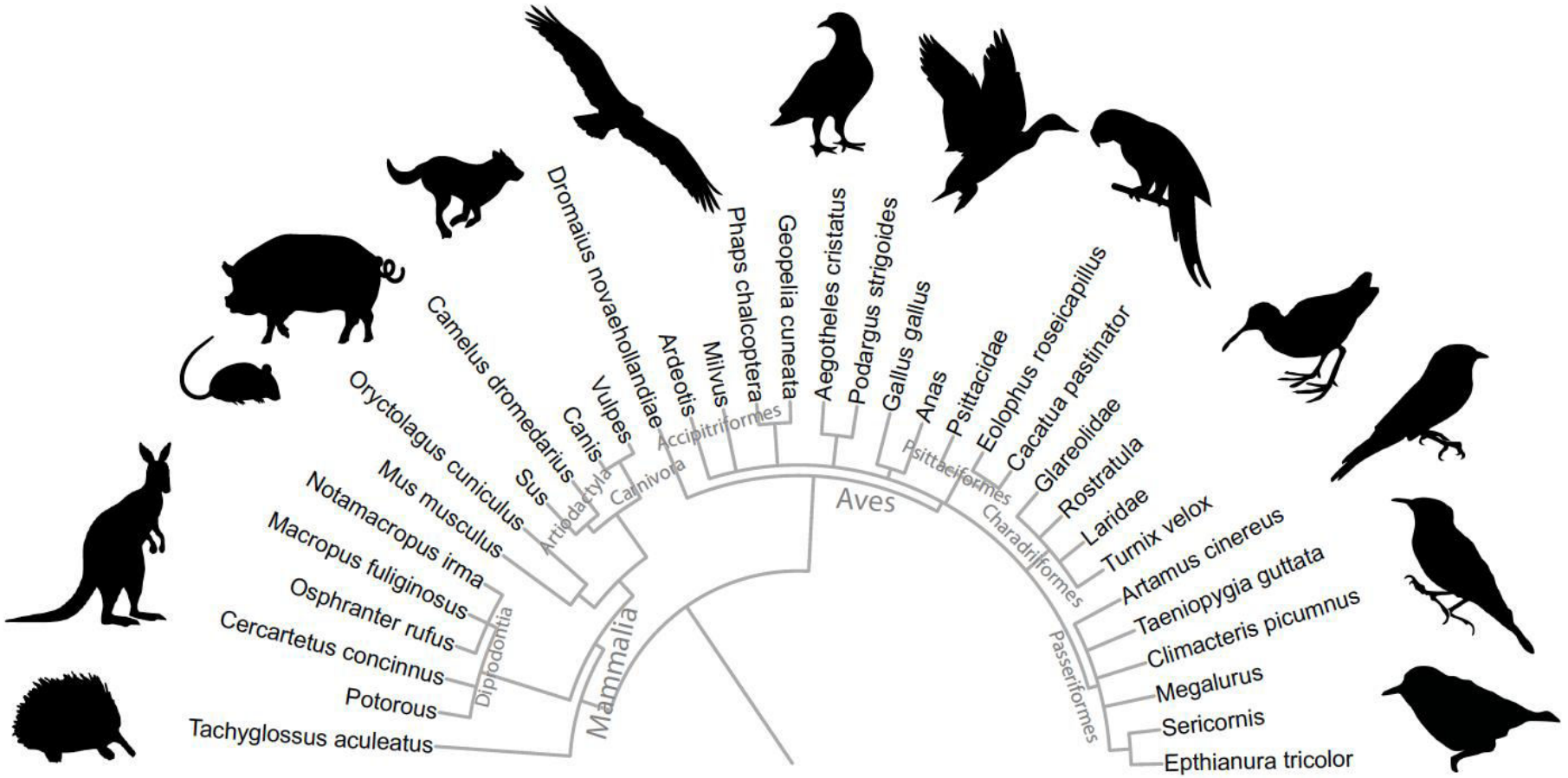


**Figure 4.2** Phylogenetic tree of vertebrate diversity detected from scat samples (pooled across sites)

## 4.5 Results

### 4.5.1 DNA Metabarcoding results

We generated 2,597,756 raw sequence reads from the 74 samples including 2 extraction controls. There were 66 samples with a minimum depth of 3500 sequences/sample. Post quality filtering the average sequencing depth was 35,558 ± 3491 (mean ± SE). There were 85 curated ASVs, 81 of which were assigned taxonomy, and 61 of which remained following all filtering steps (see section 4.4.5). Filtering reads within samples at a minimum 0.1% relative abundance resulted in the least false identifications while keeping the maximum read depth (Table 4.1). Occasionally samples were incorrectly identified in the field, for example degraded pig scat was misidentified as Macropodidae. The majority (71.6%) of reads were mammalian (17 ASV): including 8 Dipodonitia (kangaroos, wallabies etc.), 2 Carnivora (foxes, dogs), 2 Artiodactyla (pig, camel), 2 Lagomorpha (rabbit), 1 Monotremata (Echidna) and 1 Rodentia (mouse) ASV.The remaining 28.4% of reads (47 ASV) were bird: including 19 Passeriformes (songbirds), 5 Columbiformes (pigeons and doves), 4 Psittaciformes (parrots), 4 Charadriiformes (shore birds), 3 Anseriformes (waterbirds), 2 Accipitriformes (raptors) ASVs, and more (Figure 4.2, Figure 4.3). No reptilian DNA was identified from the samples analysed (Figure 4.2). Carnivora sequences were detected in 4 samples, 1 in the Coastal Plain and 3 in Jarrah. Organisms detected in these samples included the Common Bronzewing (*Phaps chalcoptera*), Western Pygmy-possum (*Cercartetus concinnus*), House mouse (*Mus musculus*), large macropods and pig (*Sus scrofa*).

### 4.5.2 Location differences

The richness of vertebrate taxa was highest in the Pilbara (5.38 ± 0.83 SE) followed by the Jarrah forest (2.91 ± 0.31 SE) and finally the Coastal Plain (1.76 ± 0.15 SE). Differences in richness among locations can be largely attributed to the bird taxa, 39 of which were detected in the Pilbara while less than 5 were detected in the other locations. Approximately the same number of mammalian taxa were detected at the three location (9 JF, 7 PB, 7 SCP)

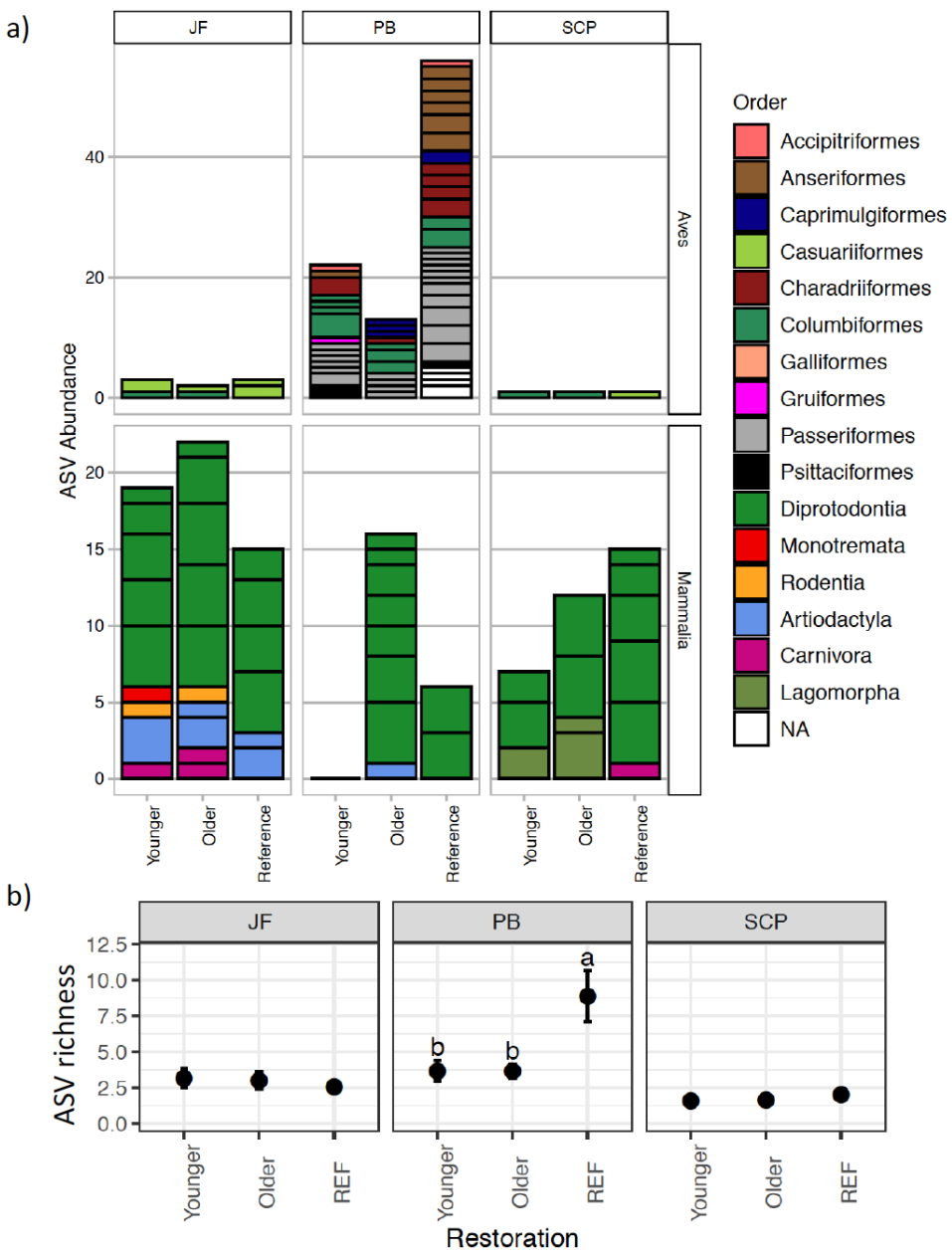


**Figure 4.3** Differences in Vertebrate diversity between sites of different restoration history. a) Vertebrate taxa detected from scat samples collected at restoration and reference sites. Younger indicates restoration sites of less than 9 years, older sites older than 9 years old (n=8). Each segment indicates different ASV, and the size of the segment shows the number of samples in which the ASV was detected. b) Taxa richness of vertebrates sequences collected from restoration. and reference sites grouped into Younger (<9 years old) Older (>9 years old) and reference. Letters indicate significant difference between groups (alpha=0.05). ANOVA results show significant differences in Observed richness in the PB ($F_{2, 18}$ = 6.064, p= 0.01), but not in the JF ($F_{2, 19}$ = 0.278, p=0.76) or SCP location ($F_{2, 18}$ = 0.728, p= 0.496). Richness values in the PB were square root transformed for statistical tests but are shown untransformed in the plots. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

**Table 4.1** False positive and false negatives in sequence results at different levels of filtering compared to morphological identification of scat. Filtering level indicates the minimum within sample relative abundance to keep the taxa in the sample. (-) False negative, (+) false positive

| | ***Unfiltered*** | | ***0.01%*** | | ***0.10%*** | | ***0.50%*** | |
|---|---|---|---|---|---|---|---|---|
| ***Taxa*** | (-) | (+) | (-) | (+) | (-) | (+) | (-) | (+) |
| *Diprodontia (Kangaroos and Wallabies)* | 0 | 13 | 0 | 7 | 1 | 2 | 1 | 2 |
| *Casuariformes (Emu)* | 0 | 2 | 0 | 2 | 0 | 1 | 0 | 0 |
| *Bird* | 1 | 7 | 2 | 4 | 3 | 0 | 4 | 0 |
| *% false ID* | *9.83* | | *6.41* | | *2.99* | | *2.99* | |

### 4.5.3 Vertebrate diversity and community composition in restoration sites

Richness was higher in reference communities but only in the Pilbara (Figure 3), where the mean richness was more than twice as high in reference sites (8.86 ± 1.79 SE) as it was in the younger (3.67 ± 0.71 SE) or older (3.63 ± 0.48 SE) restoration sites. The SIMPROF analysis also showed that the restoration and reference sites clustered separately for the Pilbara location (Figure 4.4). There were no significant differences in richness within the Coastal Plain and Jarrah location; some sites clustered together, despite being separated by over 100km.

In the Pilbara, several taxa were significantly associated with reference samples (9 taxa) and less so with restoration samples (0 taxa with younger, 2 with older) (Figure 4.5). Most (9/10) of the taxa associated with Pilbara reference samples were birds such as *Rostrulata* species (*Rostrulata australis*; Painted Snipe) and Estrilididae (estrilid finch). Taxa associated with the Jarrah forest samples were associated with all three site histories (young, old, reference) and comprised a feral mammal *Sus* (pig), and two macropod species that were also associated with Coastal Plain samples: *Notamacropus Irma* (western brush wallaby) and *Macropus fuliginosus* (western grey kangaroo).

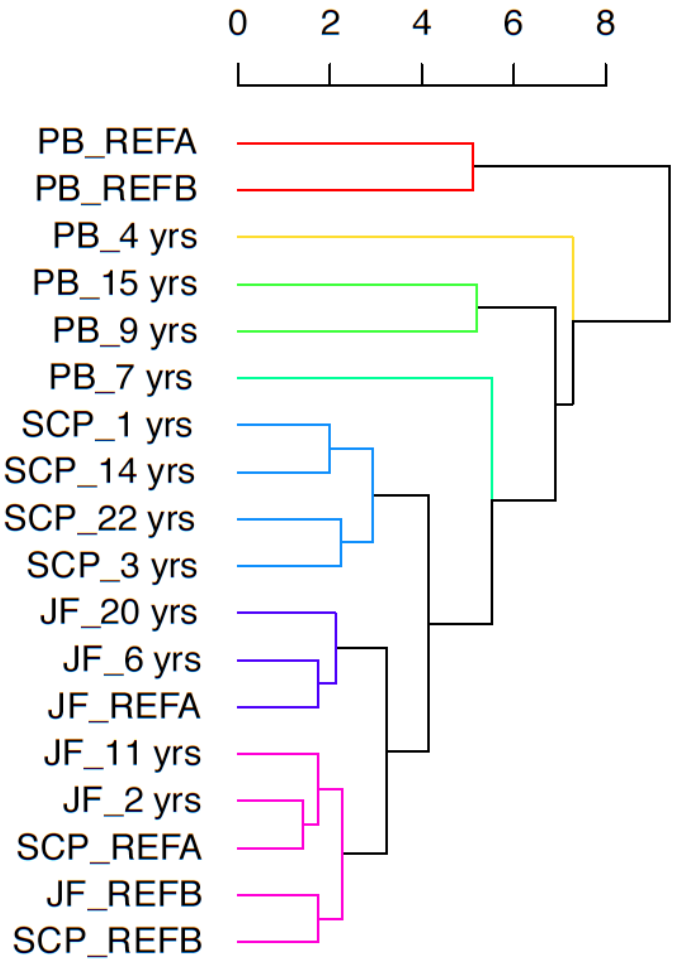


**Figure 4.4** SIMPROF plot showing the clustering of sites based on Euclidean distance. Colours represent clusters that are significantly different (alpha=0.05) from each other. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

## 4.6 Discussion

Using adventitiously collected scat samples we were able to generate bird and mammalian diversity data, but no reptile DNA was detected. Bird and mammalian communities were different between locations, with the greatest diversity detected in the Pilbara. The ability to discriminate between site histories is due in large part to the increased avifaunal diversity detected in the Pilbara (Figure 4.3). Previous studies (reviewed in Cristescu, et al. 2012) have suggested that richness is higher in reference sites, and that avian richness recovers rapidly (within a few years). The results of decreased richness in restoration sites is consistent with previous studies. However, this study shows richness (mainly avian) remained low in restoration up to 14 years old (Figure 4.2, Figure 4.3). This apparent contradiction may be explained by the geographic bias in previous studies towards sites with higher rainfall (Cristescu et al. 2012), compared to the arid Pilbara environment. Taxa associated with reference sites included finches (Estrilididae), who feed primarily on seeds of the *Triodia* grasses (Zann et al. 1995), which are more abundant in reference sites. Reference sites in the Pilbara were also strongly associated with various water birds (e.g. *Anas* - duck); of note was the detection of the rare and endangered wader *Rostratula australis* (Painted Snipe) in several reference sites. Considered in the context of foodwebs, the presence of birds at these sites may, in-turn, be related to the reestablishment of the insect diversity. Many wetland systems in the Pilbara are ephemeral and dependent on the flash flooding caused by the monsoon season and impervious geology (Pinder et al. 2010). The presence of these waterbirds was detected from scat, despite the nearby ephemeral creek being dry at the time of sampling, indicating the longevity of the scat samples and the potential to detect year-round vertebrate diversity. Conducting surveys at certain times of years (or set times following rainfall) may further enhance detection rates.

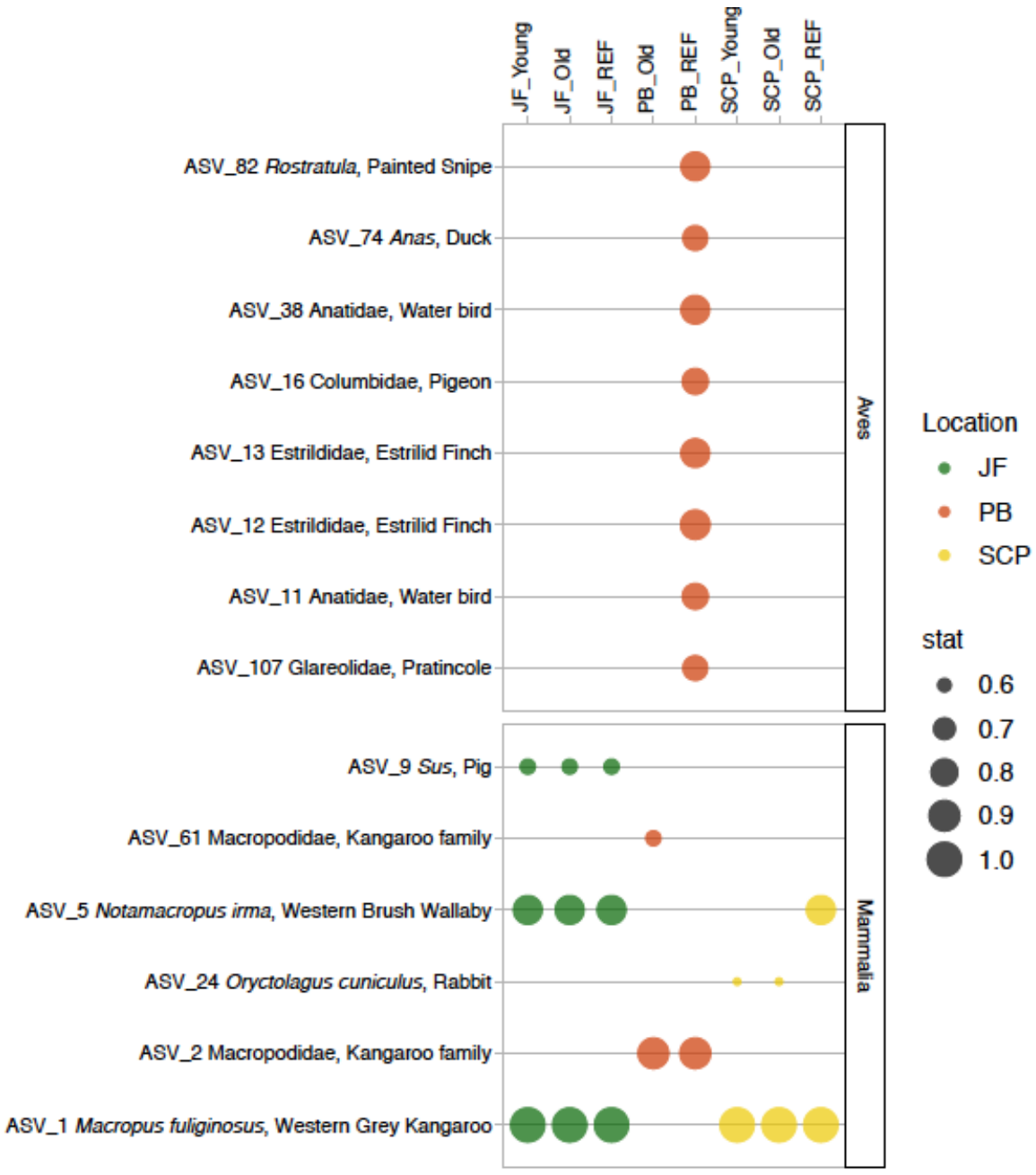


**Figure 4.5** Multipattern analysis of vertebrate taxa in restoration and reference sites in three locations. Only ASVs with significant associations (adjusted p value<0.05) are shown, size of the circles indicates the strength of the association. PB-Pilbara, SCP-Swan Coastal Plain, JF-Jarrah forest

An important sample bias is illustrated by the differences in taxa detection between the hot, arid Pilbara location and the more densely vegetated Coastal Plain and Jarrah locations. There were far more bird taxa detected in the Pilbara (Figure 4.2), despite similar levels of bird diversity in the Pilbara and the Coastal Plain (Gole and Project 2006; Burbidge, Johnstone, and Pearson 2010). Environmental

conditions are known to affect both the detection and decay rates of scats (Rhodes et al. 2011; Brown, Ramsey, and Gaffney 2014; Poggenburg et al. 2018). The Pilbara has the highest proportion of bare ground, and a hard, rocky, red surface, where white bird scats are relatively easy to see and collect. This is especially so, compared to the Coastal Plain and Jarrah where scats are difficult to detect in the sand or leaf litter. Telfer et al. (2006) also found a greater loss of scats in areas with litter over sand compared to rocky surfaces. The higher rainfall in the Coastal Plain and Jarrah may also contribute to higher degradation rates (Poggenburg et al. 2018), limiting detectability and should be taken into account in any comparison between locations.

The lack of diversity detected in the Coastal Plain and Jarrah results in a lack of discrimination between the locations (Figure 4). The vertebrate communities are different between these bioregions (Williams and Mitchell 2002; Department of the Environment and Energy 2016), but the data generated from pooled scat samples is insufficient to distinguish them. Indeed, Jarrah scat samples produced far less diversity than a recent fauna survey using a combination of traditional methods such as pitfall traps, box traps, funnel traps, camera traps, bird surveys, and ultrasonic recorders (Data from South32). However, all methods of biodiversity assessment come with their own intrinsic biases (La and Nudds 2016; Kolowski and Forrester 2017). Understanding those biases is important for selection of appropriate techniques to suit the purpose of the survey, and to prevent misinterpretation of data. While we acknowledge this method will not detect all vertebrate diversity at a site, it can still provide a useful rapid assessment of birds and mammals, particularly in some (less vegetated) locations. A salient point is that, even in its current state of development, data may be more useful (or cost effective) than current survey methods.

Despite Australia being home to 14% of the world's reptile species (Australian Museum 2019) we detected no reptile DNA. Reptile DNA is proving to be notoriously difficult to detect from environmental samples (Kucherenko et al. 2018; Adams et al. 2019; Baker et al. 2020). This has been attributed to the 'Shedding Hypothesis' which proposes that because of the keratinized skin, reptiles shed DNA at a lower rate than other organisms and are therefore less detectable using eDNA methods from soil or water (Adams et al. 2019). Collecting scat

samples was thought to be a possible solution, as they would contain more DNA. However, we found that reptile scats were hard to find given our experimental design (i.e. transects). Reptilian metabolism is typically many times lower than mammals (Bartholomew and Tucker 1963; Else and Hulbert 1981), requiring less fuel and consequently producing less waste products in the form of scats. Collecting reptile scat may require more extensive sample collection (i.e. longer transects, more time) as well as more targeted sampling.

The most commonly detected species were large mammals (Macropodidae) with larger scats, making them more visible and likely to be collected. This ascertainment bias will need to be considered when comparing faunal assemblages. The smallest organisms, the house mouse (*M. musculus*) and the Western pygmy-possum (*C. concinnus*), were only detected in samples that contained Carnivore sequences (dog and fox). Rather than collecting their scat, these smaller creatures were likely eaten and detected as part of the carnivore diet. Another potential prey item is the Common Bronze-wing (*P. chalcoptera*) but this is also detected in samples without carnivore scat and it is difficult to determine whether it was a prey item or scat of this bird was collected. These 'russian doll' effects also needs to be factored into comparative analyses. Most mammalian scat collected came from herbivores, which feed continuously and produce more scats than carnivores (Munn et al. 2012; De Cuyper et al. 2020). As DNA metabarcoding scat collections can detect large mammals and birds, it is complementary with pitfall and funnel traps, which are the primary survey tools to detect reptiles and small mammals (Environmental Protection Authority 2010).

One advantage of morphological scat surveys is the relative abundance data they can provide, although factors of scat degradation and detection also need to be taken into account (Brodie 2006; Lonsinger et al. 2016). Morphological identification of scats is commonly used in fauna surveys, but these identifications are often inaccurate (Harrington et al. 2010; Monterroso et al. 2013; Spitzer et al. 2019), limiting their usefulness of these surveys. This is a significant issue where species in the same family (e.g. Macropodidae) (Wadley, Austin, and Fordham 2013; Spitzer et al. 2019) co-occur. We limited our identification to only the most easily identified scat, which allowed us to determine optimum filtering for our dataset

(Table 1). However, comparing the presence of those taxa in the DNA and morphological sample data also showed examples where a scat was misidentified. Additionally, bird scats are almost impossible to identify morphologically (with a few notable exceptions-Emu), limiting the use of bird scat surveys. DNA methods, such as those demonstrated in this study, can improve the range of taxa that can be surveyed using scat and improve identifications.

We demonstrate the potential of pooled scat collections to provide assemblage data for birds and mammals; however, in some cases it may be preferable to use individual scats. Pooled scat collections results in lower costs overall, as multiple scats can be sequenced per sample. Targeted scat collection would allow dietary analyses to be used to evaluate not only the presence of certain animals, but also their interactions with other organisms in the area (Galimberti et al. 2016; Arteaga Claramunt et al. 2018). For example, scat from a generalist frugivorous bird can be used to evaluate plant diversity in restoration and how it changes over time (Galimberti et al. 2016). However, collecting scats for dietary analyses has similar constraints in that it is best to collect from hard surfaces (rock and ice) and difficult to find in vegetated areas (McInnes et al. 2017). Galimberti et al. (2016) overcame these limitations by collecting fresh scat from birds caught using mist nets.

## 4.7 Conclusion

We examined whether opportunistic and rapid collection of scats could be used for eDNA based biodiversity assessment of birds and mammals. This study has shown that eDNA analyses of scat samples can generate valuable diversity data that enable differentiation between sites with different restoration histories. Our results are consistent with other eDNA based studies that have demonstrated various substrates can be used for terrestrial biomonitoring (Schnell et al., 2015; van der Heyde et al. 2020). However, the detectability of scats was affected by environmental conditions (e.g. soil colour and site rainfall) and this approach may be most suitable in locations favourable to the collection of scats from many species, such as the Pilbara. Therefore, it may be beneficial to explore other sources of vertebrate DNA for some locations such as invertebrate derived DNA (Calvignac-spencer, Merkel, and Kutzner 2013; Schnell et al. 2015), or targeted samples from high traffic areas such as log piles or tree hollows (Mazurek and Zielinski 2004).

While this approach is applicable to restoration monitoring (in certain locations), DNA does not give good indications of bird or mammal behaviour or abundance at a site. Presence of an organism does not necessarily indicate persistence (Cross et al. 2020) or that the site is providing the same level of resources as a reference site (Cross, Tomlinson, Craig, and Bateman 2019). Future research should include a comparison of molecular and other methods to better understand the advantages and limitations of each. A comprehensive monitoring program should include multiple measures of vertebrate recovery (Cross, Tomlinson, Craig, Dixon, et al. 2019; Cross, Bateman, and Cross 2020) and this study demonstrates that eDNA is one tool that can be used to quickly and easily survey bird and mammal diversity. Arguably one of the most pressing needs in seeking to operationalise this tool is to develop better decision trees on what method(s) and/or substrate to apply at a given site - practical guidance in this area will help shape best-practice restorations.

## 4.8 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.


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). doi: 10.3390/d11040050

Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjær, K. H., … Willerslev, E. (2012). Meta-barcoding of "dirt" DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, *21*(8), 1966–1979. doi: 10.1111/j.1365-294X.2011.05261.x

Arteaga Claramunt, A. M., White, N. E., Bunce, M., O'Connell, M., Bullen, R. D., & Mawson, P. R. (2018). Determination of the diet of the ghost bat (Macroderma gigas) in the Pilbara region of Western Australia from dried prey remains and DNA metabarcoding. *Australian Journal of Zoology*, *66*(3), 195. doi: 10.1071/ZO18040

Australian Museum. (2019). Reptiles: A diverse group of animals including turtles, lizards, snakes and crocodiles.

Baker, S. J., Niemiller, M. L., Stites, A. J., Ash, K. T., Davis, M. A., Dreslik, M. J., & Phillips, C. A. (2020). Evaluation of environmental DNA to detect Sistrurus catenatus and Ophidiomyces ophiodiicola in crayfish burrows. *Conservation Genetics Resources*, *12*(1), 13–15. doi: 10.1007/s12686-018-1053-9

Bartholomew, G. A., & Tucker, V. A. (1963). Control of changes in body temperature, metabolism and circulation by the agamid lizard, Amphibolurus barbatus. *Physiol. Zool.*, *36*, 199–218.

Benayas, J. M. R., Newton, A. C., Diaz, A., & Bullock, J. M. (2009). Enhancement of Biodiversity and Ecosystem Services by Ecological Restoration: A Meta-Analysis. *Science*, *325*(5944), 1121–1124. doi: 10.1126/science.1172460

Benigno, S. M., Dixon, K. W., & Stevens, J. C. (2013). Increasing soil water

retention with native-sourced mulch improves seedling establishment in postmine Mediterranean sandy soils. *Restoration Ecology*, *21*(5), 617–626. doi: 10.1111/j.1526-100X.2012.00926.x

Benjamini, Y., & Hochberg, Y. (2007). Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing pp . *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Berry, T. E., Osterrieder, S. K., Murray, D. C., Coghlan, M. L., Richardson, A. J., Grealy, A. K., … Bunce, M. (2017). DNA metabarcoding for diet analysis and biodiversity: A case study using the endangered Australian sea lion (Neophoca cinerea). *Ecology and Evolution*, *7*(14). doi: 10.1002/ece3.3123

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*, *16*(1), 176–182. doi: 10.1111/1755-0998.12428

Brodie, J. F. (2006). An Experimentally Determined Persistence-Rate Correction Factor for Scat-Based Abundance Indices. *Wildlife Society Bulletin*, *34*(4), 1216–1219. doi: 10.2193/0091-7648(2006)34[1216:aedpcf]2.0.co;2

Brown, W. E., Ramsey, D. S. L., & Gaffney, R. (2014). Degradation and detection of fox (Vulpes vulpes) scats in Tasmania: Evidence from field trials. *Wildlife Research*, *41*(8), 681–690. doi: 10.1071/WR14152

Burbidge, A. H., Johnstone, R. E., & Pearson, D. J. (2010). A Biodiversity Survey of the Pilbara Region of Western Australia , 2002 – 2007. *Records of the Western Australiam Museum*, *Suppliment*(78), 247–279.

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. doi: 10.1038/nmeth.3869

Calvignac-spencer, A., Merkel, K., & Kutzner, N. (2013). *Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity*. 915–924. doi: 10.1111/mec.12183

Catterall, C. P. (2018). Fauna as passengers and drivers in vegetation restoration: A

synthesis of processes and evidence. *Ecological Management and Restoration*, *19*(August), 54–62. doi: 10.1111/emr.12306

Clavero, M., Brotons, L., Pons, P., & Sol, D. (2009). Prominent role of invasive species in avian biodiversity loss. *Biological Conservation*, *142*(10), 2043–2049. doi: 10.1016/j.biocon.2009.03.034

Comer, S. J., & Wooller, R. D. (2002). A comparison of the passerine avifaunas of a rehabilitated minesite and a nearby reserve in south-western Australia. *Emu*, *102*(3), 305–311. doi: 10.1071/MU00042

Cristescu, R. H., Frère, C., & Banks, P. B. (2012). A review of fauna in mine rehabilitation in Australia: Current state and future directions. *Biological Conservation*, *149*(1), 60–72. doi: 10.1016/j.biocon.2012.02.003

Cristescu, R. H., Rhodes, J., Frére, C., & Banks, P. B. (2013). Is restoring flora the same as restoring fauna? Lessons learned from koalas and mining rehabilitation. *Journal of Applied Ecology*, *50*(2), 423–431. doi: 10.1111/1365-2664.12046

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 and Restoration*, *21*(1), 4–8. doi: 10.1111/emr.12393

Cross, S. L., Craig, M. D., Tomlinson, S., Dixon, K. W., & Bateman, P. W. (2020). Using monitors to monitor ecological restoration: Presence may not indicate persistence. *Austral Ecology*, 1–12. doi: 10.1111/aec.12905

Cross, S. L., Tomlinson, S., Craig, M. D., & Bateman, P. W. (2019). The Time Local Convex Hull method as a tool for assessing responses of fauna to habitat restoration: A case study using the perentie (Varanus giganteus: Reptilia: Varanidae). *Australian Journal of Zoology*, *67*(1), 27–37. doi: 10.1071/ZO19040

Cross, S. L., Tomlinson, S., Craig, M. D., Dixon, K. W., & Bateman, P. W. (2019). Overlooked and undervalued: The neglected role of fauna and a global bias in ecological restoration assessments. *Pacific Conservation Biology*, *25*(4), 331–341. doi: 10.1071/PC18079

De Caceres, M., & Legendre, P. (2009). Associations between species and groups of sites : indices and statistical inference. *Ecology*, *90*(12), 3566–3574.

De Cuyper, A., Meloro, C., Abraham, A. J., Müller, D. W. H., Codron, D., Janssens, G. P. J., & Clauss, M. (2020). The uneven weight distribution between predators and prey: Comparing gut fill between terrestrial herbivores and carnivores. *Comparative Biochemistry and Physiology -Part A : Molecular and Integrative Physiology*, *243*(January), 110683. doi: 10.1016/j.cbpa.2020.110683

Department of the Environment and Energy. (2016). Banksia woodlands of the Swan Coastal Plain ecological community. *Commonwealth of Australia*, 117.

Doherty, T. S., Glen, A. S., Nimmo, D. G., Ritchie, E. G., & Dickman, C. R. (2016). Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(40), 11261–11265. doi: 10.1073/pnas.1602480113

Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C. M., … Nelson, N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience*, *4*(46). doi: 10.1186/s13742-015-0086-1

Edgar, R. C. (2016). UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, 081257. doi: 10.1101/081257

Else, P. L., & Hulbert, A. J. (1981). Comparison of the “mammal machine” and the “reptile machine”: energy production. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, *9*(1), 3–9. doi: 10.1152/ajpregu.1981.240.1.r3

Environmental Protection Authority and Department of Environment and Conservation. (2010). Technical Guidance - Sampling Methods for Terrestrial Vertebrate Fauna (eds B.M. Hyder, J. Dell and M.A Cowan). *Perth, Western Australia*.

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

Fleming, T. H., & Muchhala, N. (2008). Nectar-feeding bird and bat niches in two worlds: pantropical comparisons of vertebrate pollination systems. *Journal of Biogeography*, *35*(5), 764–780. doi: 10.1111/j.1365-2699.2007.01833.x

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). doi: 10.1038/s41467-017-01312-x

Galimberti, A., Spinelli, S., Bruno, A., Mezzasalma, V., De Mattia, F., Cortis, P., & Labra, M. (2016). Evaluating the efficacy of restoration plantings through DNA barcoding of frugivorous bird diets. *Conservation Biology*. doi: 10.1111/cobi.12687

Goldberg, C. S., Pilliod, D. S., Arkle, R. S., & Waits, L. P. (2011). Molecular detection of vertebrates in stream water: A demonstration using rocky mountain tailed frogs and Idaho giant salamanders. *PLoS ONE*, *6*(7). doi: 10.1371/journal.pone.0022746

Gole, C. A., & Project, P. B. (2006). *Bird surveys in selected Perth metropolitan reserves : Round 3 survey report Bird surveys in selected Perth metropolitan reserves : Round 3 survey report*.

Harrington, L. A., Harrington, A. L., Hughes, J., Stirling, D., & Macdonald, D. W. (2010). The accuracy of scat identification in distribution surveys: American mink, Neovison vison, in the northern highlands of Scotland. *European Journal of Wildlife Research*, *56*(3), 377–384. doi: 10.1007/s10344-009-0328-6

Havel, J. J. (1975). Site-vegetation mapping in the northern jarrah forest. I. Definition of site-vegetation types. *Forests Department WA*, *Bulletin N*.

Herrera, C. M. (1995). Plant-vertebrate seed dispersal systems in the Mediterranean: Ecological, evolutionary, and historical determinants. *Annual Review of Ecology and Systematics*, *26*(1), 705–727. doi: 10.1146/annurev.es.26.110195.003421

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). “Sight-

unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, *4*(2), 150–157. doi: 10.1111/j.1755-263X.2010.00158.x

Kolowski, J. M., & Forrester, T. D. (2017). Camera trap placement and the potential for bias due to trails and other features. *PLoS ONE*, *12*(10), 1–20. doi: 10.1371/journal.pone.0186679

Kucherenko, A., Herman, J. E., III, E. M. E., & Urakawa, H. (2018). Terrestrial snake environmental DNA accumulation and degradation dynamics and its environmental application. *Herpetologica*, *74*, 38–49.

La, V. T., & Nudds, T. D. (2016). Estimation of avian species richness: Biases in morning surveys and efficient sampling from acoustic recordings. *Ecosphere*, *7*(4), 1–13. doi: 10.1002/ecs2.1294

Lonsinger, R. C., Gese, E. M., Knight, R. N., Johnson, T. R., & Waits, L. P. (2016). Quantifying and correcting for scat removal in noninvasive carnivore scat surveys. *Wildlife Biology*, *22*(2), 45–54. doi: 10.2981/wlb.00179

Mazurek, M. J., & Zielinski, W. J. (2004). Individual legacy trees influence vertebrate wildlife diversity in commercial forests. *Forest Ecology and Management*, *193*(3), 321–334. doi: 10.1016/j.foreco.2004.01.013

Mcdonald, T., Gann, G. D., Jonson, J., & Dixon, K. W. (2016). International standards for the practice of ecological restoration - including principles and key concepts. *Communications Standards*, (December), 427–437. doi: 10.1016/b978-0-08-034092-0.50030-2

McInnes, J. C., Alderman, R., Deagle, B. E., Lea, M. A., Raymond, B., & Jarman, S. N. (2017). Optimised scat collection protocols for dietary DNA metabarcoding in vertebrates. *Methods in Ecology and Evolution*, *8*(2), 192–202. doi: 10.1111/2041-210X.12677

McKenzie, N. L., van Leeuwen, S., & Pinder, A. M. (2009). Introduction to the Pilbara Biodiversity Survey, 2002–2007. *Records of the Western Australian Museum, Supplement*, *78*(1), 3. doi: 10.18195/issn.0313-122x.78(1).2009.003-089

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible

Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4). doi: 10.1371/journal.pone.0061217

Mendiburu, F. de. (2019). agricolae: Statistical Procedures for Agricultural Research. R package version 1.3-1. https://CRAN.R-project.org/package=agricolae. *Https://CRAN.R-Project.Org/Package=agricolae*.

Miller, B. P., Sinclair, E. A., Menz, M. H. M., Elliott, C. P., Bunn, E., Commander, L. E., … Stevens, J. C. (2017). A framework for the practical science necessary to restore sustainable, resilient, and biodiverse ecosystems. *Restoration Ecology*, *25*(4), 605–617. doi: 10.1111/rec.12475

Monterroso, P., Castro, D., Silva, T. L., Ferreras, P., Godinho, R., & Alves, P. C. (2013). Factors affecting the (in)accuracy of mammalian mesocarnivore scat identification in South-western Europe. *Journal of Zoology*, *289*(4), 243–250. doi: 10.1111/jzo.12000

Munn, A. J., Tomlinson, S., Savage, T., & Clauss, M. (2012). Retention of different-sized particles and derived gut fill estimate in tammar wallabies (Macropus eugenii): Physiological and methodological considerations. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, *161*(2), 243–249. doi: 10.1016/j.cbpa.2011.11.003

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

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2007). Biodiversity hotspots for conservation priorities. *Biodiversity and Conservation*, *16*(4), 853–858. doi: 10.1080/21564574.1998.9650003

Nichols, O. G., & Nichols, F. M. (2003). Long-term trends in faunal recolonization after bauxite mining in the Jarrah forest of Southwestern Australia. *Restoration Ecology*, *11*(3), 261–272. doi: 10.1046/j.1526-100X.2003.00190.x

Palmer, M. A., Ambrose, R. F., & Poff, N. L. (1997). Ecological Theory and Community Restoration Ecology. *Restoration Ecology*, *5*(4), 291–300. doi: 10.1046/j.1526-100X.1997.00543.x

Pinder, A. M., Halse, S. A., Shiel, R. J., & McRae, J. M. (2010). An arid zone awash with diversity: patterns in the distribution of aquatic invertebrates in the Pilbara region of Western Australia. *Records of the Western Australian Museum, Supplement*, *78*(1), 205. doi: 10.18195/issn.0313-122x.78(1).2010.205-246

Poggenburg, C., Nopp-Mayr, U., Coppes, J., & Sachser, F. (2018). Shit happens … and persists: decay dynamics of capercaillie (Tetrao urogallus L.) droppings under natural and artificial conditions. *European Journal of Wildlife Research*, *64*(3). doi: 10.1007/s10344-018-1187-9

R Core Team. (2018). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org*.

Rhodes, J. R., Lunney, D., Moon, C., Matthews, A., & Mcalpine, C. A. (2011). The consequences of using indirect signs that decay to determine species' occupancy. *Ecography*, *34*(1), 141–150. doi: 10.1111/j.1600-0587.2010.05908.x

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), 1–11. doi: 10.1093/nar/gkr732

Sarkar, S., Pressey, R. L., Faith, D. P., Margules, C. R., Fuller, T., Stoms, D. M., … Andelman, S. (2006). Biodiversity Conservation Planning Tools: Present Status and Challenges for the Future. *Annual Review of Environment and Resources*, *31*(1), 123–159. doi: 10.1146/annurev.energy.31.042606.085844

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*(1), 1–14. doi: 10.1186/s12983-015-0115-z

Spitzer, R., Churski, M., Felton, A., Heurich, M., Kuijper, D. P. J., Landman, M., … Cromsigt, J. P. G. M. (2019). Doubting dung: eDNA reveals high rates of misidentification in diverse European ungulate communities. *European Journal of Wildlife Research*, *65*(2). doi: 10.1007/s10344-019-1264-8

Suding, K., Higgs, E., Palmer, M., Callicott, J. B., Anderson, C. B., Baker, M., … Schwartz, K. Z. S. (2015). Committing to ecological restoration. *Science*, *348*(6235), 638–640. doi: 10.1126/science.aaa4216

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8, SI), 1789–1793. doi: 10.1111/j.1365-294X.2012.05542.x

Telfer, W. R., Griffiths, A. D., & Bowman, D. M. J. S. (2006). Scats can reveal the presence and habitat use of cryptic rock-dwelling macropods. *Australian Journal of Zoology*, *54*(5), 325–334. doi: 10.1071/ZO05074

Thompson, S. A., & Thompson, G. G. (2004). Adequacy of rehabilitation monitoring practices in the Western Australian mining industry. *Ecological Management and Restoration*, *5*(1), 30–33. doi: 10.1111/j.1442-8903.2004.00172.x

Trudgen, M. (1977). *A report on the flora and floristic communitytypes found on parts of sand mining leases held by Rocla in the Gnangara area. Report prepared for Rocla Quarry Prod-ucts, Perth, Western Australia*.

UN Environment Programme. (2020). Decade on Ecosystem Restoration 2021-2030.

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*, (00), 1–14. doi: 10.1111/1755-0998.13148

Wadley, J. J., Austin, J. J., & Fordham, D. A. (2013). Rapid species identification of eight sympatric northern Australian macropods from faecal-pellet DNA. *Wildlife Research*, *40*(3), 241. doi: 10.1071/WR13005

West, K. M., Stat, M., Harvey, E. S., Skepper, C. L., DiBattista, J. D., Richards, Z. T., … Bunce, M. (2020). eDNA metabarcoding survey reveals fine-scale coral reef community variation across a remote, tropical island ecosystem. *Molecular Ecology*, *29*(6), 1069–1086. doi: 10.1111/mec.15382

Whitaker, D., & Christman, M. (2014). clustsig: Significant Cluster Analysis. R package version 1.1. *Https://CRAN.R-Project.Org/Package=clustsig*. Retrieved from http://www.douglaswhitaker.com

Williams, K., & Mitchell, D. (2002). Jarrah Forest 1 (JF1 - Northern Jarrah Forest subregion). *A Biodiversity Audit of Western Australia's 53 Biogeographical Subregions in 2002*, *1*(September 2001), 369–381. Retrieved from http://www.naturebase.net/content/view/960/1397/

Wilson, G. J., & Delahay, R. J. (2001). A review of methods to estimate the abundance of terrestrial carnivores using field signs and observation. *Wildlife Research*, *28*(1), 151–164.

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

Zann, R. A., Morton, S. R., Jones, K. R., & Burley, N. T. (1995). The Timing of Breeding by Zebra Finches in Relation to Rainfall in Central Australia. *Emu - Austral Ornithology*, *95*(3), 208–222. doi: 10.1071/MU9950208

## 4.9 Supplementary Information

### 4.9.1 Supplementary Methods

In addition to scat samples, we also collected soil samples to test whether vertebrate diversity could be sequenced from this substrate. Although a previous study had shown that soil was not an ideal substrate for vertebrate DNA (Mieke van der Heyde et al. 2020), we attempted to optimize sample collection and processing by collecting deeper soil not as exposed to UV radiation and more subsamples per sample.

#### *4.9.1.1 Sample Collection*

Soil samples were collected from 4 points at each restoration/reference site for a total of 74 samples. For each soil sample, 8 subsamples were taken randomly in a 10x10m plot to a depth of 15cm; these were then manually homogenized in a large sample bag and a portion was collected in a 50mL falcon tube for microbial analyses, while the rest of the sample was kept for soil chemical analyses. The soil probe was cleaned with bleach between each sample and gloves were changed between each sample point. Soil samples were frozen as soon as possible in a mobile freezer and taken to Perth, where they were stored at -20°C until they were processed.

#### *4.9.1.2 Sample processing*

Soil samples were homogenized in 50mL tubes for 30 seconds. DNA was extracted and 250mg soil using the DNeasy PowerLyzer PowerSoil kit (Qiagen) on the QiaCube Connect automated platform (Qiagen). The final elution volume was 100 µL, and extraction controls (blanks) were carried out for every set of extractions. Quantitative PCR (qPCR) was run on neat extracts and a 1/10 dilution to see if samples exhibited inhibition, and to determine optimal dilution level for each sample (Dáithí C. Murray, Coghlan, and Bunce 2015). The qPCR assays were run with primers that target the 12S gene *12SV5*-F/R (98 bp, Riaz et al., 2011)

The qPCRs were run on a StepOne Plus (Applied BioSystems) real-time qPCR instrument with the following conditions: 5 min at 95°C, 40 cycles of 95°C for 30s, 30s at the annealing temperature and 45s at 72°C, a melt curve stage of 15s at 95°C 1 min at 60°C and 15s at 95°C, ending with 10 min elongation at 72°C. In an

effort to amplify as much as possible from the soil samples we lowered the annealing temperature to 52°C, added twice as much template DNA from the DNA extractions, and ran the PCR for 50 cycles instead of 40. The PCR mix for quantitation were the same as used for the scat samples.We followed the same sequencing procedure as the scat samples, except for soil samples, we first sequenced a subset of the samples, 4 samples from a reference site in each location. The Fusion PCR for soil was performed in quadruplet to maximize to probability of generating vertebrate sequences.

### 4.9.2 Supplementary Results

We generated 1,482,322 reads from the 12 soil samples that were sequenced. Post quality filtering the average sequencing depth was 113,867 ± 26599 (mean ± SE). There were 10 curated ASVs generated from the soil samples, all of which were assigned taxonomy (Table S4.9.1). Results were over 90% human DNA with some chicken and pig sequences. This was deemed likely contamination and the remaining soil samples were not sequenced.

**Table S4.9.1** Results of soil sample sequencing a reference site at each of the three locations. EC-Extraction control, NTC – Non template control.

| | ***Percentage of reads in each taxa*** | | | | |
|---|---|---|---|---|---|
| ***Taxa*** | SCP | Pilbara | Jarrah | EC | NTC |
| *Human* | 100.0 | 94.0 | 92.3 | 100.0 | 100.0 |
| *Pig* | 0.0 | 0.0 | 7.7 | 0.0 | 0.0 |
| *Chicken* | 0.0 | 6.0 | 0.0 | 0.0 | 0.0 |
| *Total reads* | 435768 | 497994 | 398876 | 67696 | 2 |

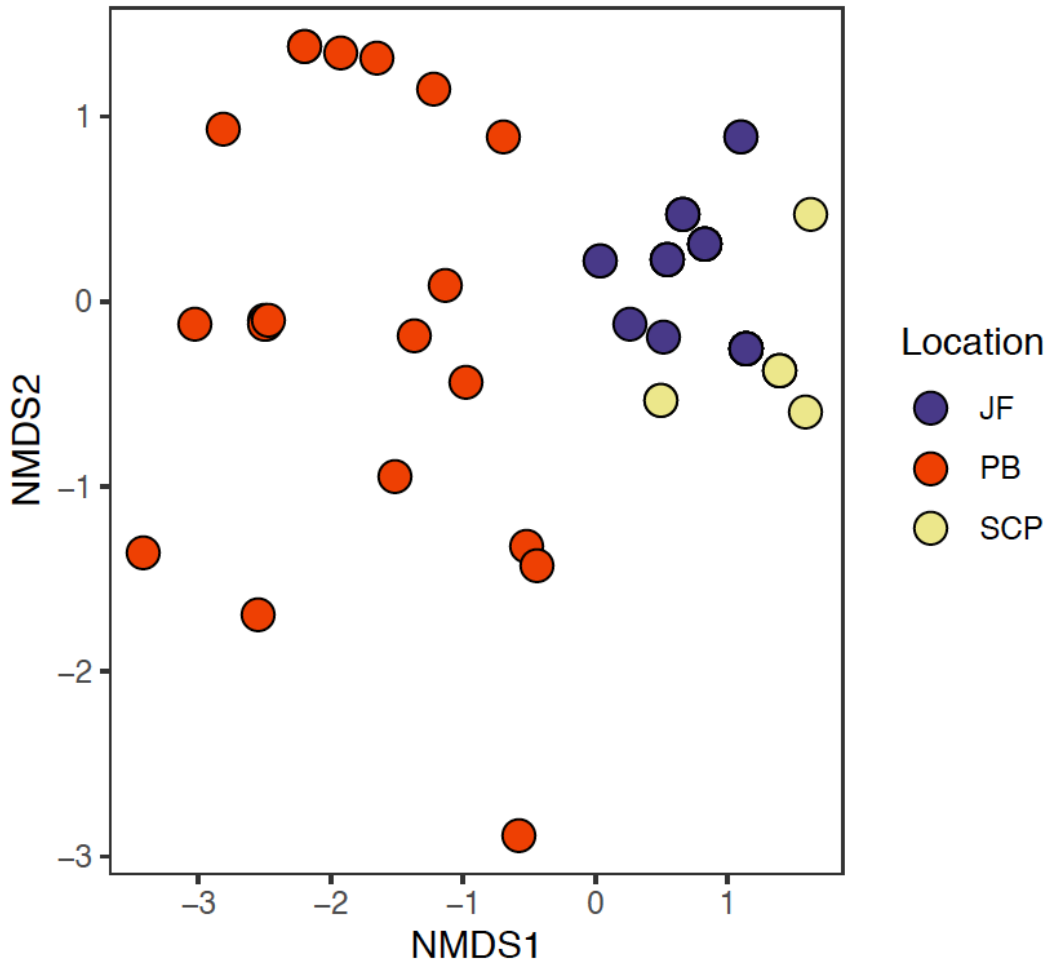


**Figure S 4.9.1** NMDS ordination of samples (stress =0.03) showing that Pilbara samples are different from the Jarrah and SCP samples, Jarrah and SCP seem to cluster together, and this is likely because there was so little detected in the Jarrah and SCP that it wasn't able to distinguish between them.

**Table S4.9.2** Taxonomic assignments of all ASVs checked against occurrence records of the study locations.

| OTU | Class | Order | Family | Genus | Species | Notes* |
|---|---|---|---|---|---|---|
| **ASV_85** | Aves | Accipitriformes | Accipitridae | dropped | dropped | Y |
| **ASV_8** | Aves | Accipitriformes | Accipitridae | *Milvus* | dropped | Y |
| **ASV_74** | Aves | Anseriformes | Anatidae | *Anas* | dropped | Y |
| **ASV_11** | Aves | Anseriformes | Anatidae | dropped | dropped | <97% ID |
| **ASV_38** | Aves | Anseriformes | Anatidae | dropped | dropped | Y |
| **ASV_23** | Aves | Caprimulgiformes | Aegothelidae | *Aegotheles* | *Aegotheles cristatus* | Y |
| **ASV_66** | Aves | Caprimulgiformes | Aegothelidae | *Aegotheles* | *Aegotheles cristatus* | Y |
| **ASV_160** | Aves | Caprimulgiformes | Podargidae | *Podargus* | *Podargus strigoides* | Y |
| **ASV_21** | Aves | Casuariiformes | Dromaiidae | *Dromaius* | *Dromaius novaehollandiae* | Y |
| **ASV_107** | Aves | Charadriiformes | Glareolidae | *Pluvianus* | *Pluvianus aegyptius* | Family |
| **ASV_47** | Aves | Charadriiformes | Laridae | *Saundersilarus* | *Saundersilarus saundersi* | Family |
| **ASV_82** | Aves | Charadriiformes | Rostratulidae | *Rostratula* | *Rostratula benghalensis* | Genus |
| **ASV_37** | Aves | Charadriiformes | Turnicidae | *Turnix* | *Turnix velox* | Y |
| **ASV_16** | Aves | Columbiformes | Columbidae | *Ducula* | *Ducula zoeae* | Family |
| **ASV_30** | Aves | Columbiformes | Columbidae | *Geopelia* | *Geopelia cuneata* | Y |
| **ASV_3** | Aves | Columbiformes | Columbidae | *Phaps* | *Phaps chalcoptera* | Y |
| **ASV_29** | Aves | Columbiformes | Columbidae | *Phaps* | *Phaps chalcoptera* | Y |
| **ASV_135** | Aves | Columbiformes | Columbidae | *Uropelia* | *Uropelia campestris* | Family |
| **ASV_43** | Aves | dropped | dropped | dropped | dropped | Y |
| **ASV_125** | Aves | dropped | dropped | dropped | dropped | Y |
| **ASV_174** | Aves | dropped | dropped | dropped | dropped | Y |
| **ASV_227** | Aves | dropped | dropped | dropped | dropped | Y |
| **ASV_72** | Aves | Galliformes | Phasianidae | *Gallus* | *Gallus gallus* | Y |
| **ASV_35** | Aves | Gaviiformes | Gaviidae | *Gavia* | dropped | remove |
| **ASV_111** | Aves | Gaviiformes | Gaviidae | *Gavia* | dropped | remove |
| **ASV_19** | Aves | Gruiformes | Otididae | *Ardeotis* | *Ardeotis kori* | Genus |
| **ASV_197** | Aves | Passeriformes | Acanthizidae | *Sericornis* | dropped | Y |
| **ASV_211** | Aves | Passeriformes | Climacteridae | *Climacteris* | *Climacteris picumnus* | Y |
| **ASV_236** | Aves | Passeriformes | Corvidae | *Artamus* | *Artamus cinereus* | Y |
| **ASV_80** | Aves | Passeriformes | dropped | dropped | dropped | Y |
| **ASV_109** | Aves | Passeriformes | dropped | dropped | dropped | Y |
| **ASV_156** | Aves | Passeriformes | dropped | dropped | dropped | Y |
| **ASV_194** | Aves | Passeriformes | dropped | dropped | dropped | Y |
| **ASV_15** | Aves | Passeriformes | Estrildidae | dropped | dropped | Y |
| **ASV_172** | Aves | Passeriformes | Estrildidae | dropped | dropped | Y |
| **ASV_12** | Aves | Passeriformes | Estrildidae | *Spermestes* | *Spermestes cucullata* | Family |
| **ASV_13** | Aves | Passeriformes | Estrildidae | *Spermestes* | *Spermestes cucullata* | <97% ID |

| OTU | Class | Order | Family | Genus | Species | Notes* |
|---|---|---|---|---|---|---|
| **ASV_31** | Aves | Passeriformes | Estrildidae | *Spermestes* | *Spermestes cucullata* | <97% ID |
| **ASV_95** | Aves | Passeriformes | Estrildidae | *Spermestes* | *Spermestes cucullata* | <97% ID |
| **ASV_181** | Aves | Passeriformes | Estrildidae | *Spermestes* | *Spermestes cucullata* | <97% ID |
| **ASV_27** | Aves | Passeriformes | Estrildidae | *Taeniopygia* | *Taeniopygia guttata* | Y |
| **ASV_18** | Aves | Passeriformes | Locustellidae | *Poodytes* | *Poodytes punctatus* | Genus |
| **ASV_71** | Aves | Passeriformes | Meliphagidae | dropped | dropped | Y |
| **ASV_32** | Aves | Passeriformes | Meliphagidae | *Epthianura* | *Epthianura tricolor* | Y |
| **ASV_28** | Aves | Passeriformes | Pachycephalidae | *Colluricincla* | *Colluricincla harmonica* | Y |
| **ASV_62** | Aves | Psittaciformes | Cacatuidae | *Cacatua* | *Cacatua pastinator* | Y |
| **ASV_6** | Aves | Psittaciformes | Cacatuidae | *Eolophus* | *Eolophus roseicapillus* | Y |
| **ASV_146** | Aves | Psittaciformes | Cacatuidae | *Eolophus* | *Eolophus roseicapillus* | Y |
| **ASV_53** | Aves | Psittaciformes | Psittacidae | dropped | dropped | Y |
| **ASV_243** | Betaproteobacteria | Burkholderiales | Burkholderiaceae | *Ralstonia* | *Ralstonia solanacearum* | remove |
| **ASV_243** | Betaproteobacteria | dropped | dropped | dropped | dropped | remove |
| **ASV_17** | dropped | dropped | dropped | dropped | dropped | remove |
| **ASV_25** | dropped | dropped | dropped | dropped | dropped | remove |
| **ASV_93** | dropped | dropped | dropped | dropped | dropped | remove |
| **ASV_238** | dropped | dropped | dropped | dropped | dropped | remove |
| **ASV_253** | dropped | dropped | dropped | dropped | dropped | remove |
| **ASV_255** | dropped | dropped | dropped | dropped | dropped | remove |
| **ASV_4** | Mammalia | Artiodactyla | Camelidae | *Camelus* | *Camelus dromedarius* | Y |
| **ASV_9** | Mammalia | Artiodactyla | Suidae | *Sus* | dropped | Y |
| **ASV_92** | Mammalia | Carnivora | Canidae | *Canis* | dropped | Y |
| **ASV_10** | Mammalia | Carnivora | Canidae | *Vulpes* | dropped | Y |
| **ASV_79** | Mammalia | Carnivora | Felidae | *Felis* | dropped | Y |
| **ASV_14** | Mammalia | Diprotodontia | Burramyidae | *Cercartetus* | *Cercartetus concinnus* | Y |
| **ASV_2** | Mammalia | Diprotodontia | Macropodidae | dropped | dropped | Y |
| **ASV_61** | Mammalia | Diprotodontia | Macropodidae | dropped | dropped | Y |
| **ASV_1** | Mammalia | Diprotodontia | Macropodidae | *Macropus* | *Macropus fuliginosus* | Y |
| **ASV_5** | Mammalia | Diprotodontia | Macropodidae | *Notamacropus* | *Notamacropus irma* | Y |
| **ASV_7** | Mammalia | Diprotodontia | Macropodidae | *Osphranter* | *Osphranter rufus* | Y |
| **ASV_240** | Mammalia | Diprotodontia | Macropodidae | *Thylogale* | dropped | Family |
| **ASV_63** | Mammalia | Diprotodontia | Potoroidae | *Potorous* | *Potorous platyops* | Genus |
| **ASV_24** | Mammalia | Lagomorpha | Leporidae | *Oryctolagus* | *Oryctolagus cuniculus* | Y |
| **ASV_168** | Mammalia | Lagomorpha | Leporidae | *Oryctolagus* | *Oryctolagus cuniculus* | Y |
| **ASV_148** | Mammalia | Monotremata | Tachyglossidae | *Tachyglossus* | *Tachyglossus aculeatus* | Y |
| **ASV_20** | Mammalia | Primates | Hominidae | *Homo* | *Homo sapiens* | Y |
| **ASV_39** | Mammalia | Primates | Hominidae | *Homo* | *Homo sapiens* | Y |
| **ASV_68** | Mammalia | Primates | Hominidae | *Homo* | *Homo sapiens* | Y |
| **ASV_193** | Mammalia | Primates | Hominidae | *Homo* | *Homo sapiens* | Y |

| OTU | Class | Order | Family | Genus | Species | Notes* |
|---|---|---|---|---|---|---|
| **ASV_254** | Mammalia | Primates | Hominidae | *Homo* | *Homo sapiens* | Y |
| **ASV_26** | Mammalia | Rodentia | Muridae | *Mus* | *Mus musculus* | Y |
| **ASV_251** | | | | | uncultured bacterium | remove |
| **ASV_251** | | | | | uncultured Gemmatimonadetes bacterium | remove |

* Y-taxa occurs, Family-family occurs in the area, dropped to family level, Genus-genus occurs in the area, dropped to genus level, <97% ID-dropped to Family level.

**Table S4.9.3** Comparison between morphological and DNA identification of scats. Numbers indicate the number of samples at each site where the taxa was identified. JF-Jarrah Forest, PB-Pilbara, SCP-Swan Coastal Plain

| | *Macropod* | | *Bird* | | *Emu* | |
|---|---|---|---|---|---|---|
| *Site* | Morph | DNA | Morph | DNA | Morph | DNA |
| *JF_2* | 4 | 4 | 2 | 1 | 0 | 0 |
| *JF_6* | 4 | 4 | 1 | 1 | 2 | 2 |
| *JF_11* | 4 | 4 | 1 | 0 | 0 | 0 |
| *JF_20* | 4 | 4 | 1 | 1 | 0 | 1 |
| *JF_REFA* | 4 | 4 | 0 | 0 | 2 | 2 |
| *JF_REFB* | 4 | 4 | 0 | 0 | 1 | 1 |
| *PB_4* | 0 | 0 | 4 | 4 | 0 | 0 |
| *PB_7* | 0 | 1 | 4 | 3 | 0 | 0 |
| *PB_9* | 4 | 4 | 4 | 4 | 0 | 0 |
| *PB_15* | 4 | 3 | 3 | 3 | 0 | 0 |
| *PB_REFA* | 3 | 3 | 3 | 3 | 0 | 0 |
| *PB_REFB* | 3 | 3 | 4 | 4 | 0 | 0 |
| *PB_REFC* | 0 | 1 | 4 | 4 | 0 | 0 |
| *SCP_1* | 4 | 4 | 0 | 0 | 0 | 0 |
| *SCP_3* | 3 | 3 | 1 | 1 | 0 | 0 |
| *SCP_11* | 4 | 4 | 0 | 0 | 0 | 0 |
| *SCP_14* | 4 | 4 | 1 | 1 | 0 | 0 |
| *SCP_22* | 4 | 4 | 0 | 0 | 0 | 0 |
| *SCP_REFA* | 4 | 4 | 0 | 0 | 0 | 0 |
| *SCP_REFB* | 4 | 4 | 0 | 0 | 1 | 1 |

### 4.9.3 Supplementary References

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

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), 1–11. doi: 10.1093/nar/gkr732

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*, (00), 1–14. doi: 10.1111/1755-0998.13148

# CHAPTER 5

## EVALUATING RESTORATION TRAJECTORIES USING DNA METABARCODING OF INVERTEBRATES AND THEIR ASSOCIATED PLANT COMMUNITIES

## 5.1 Preface

*This chapter consists of a manuscript in preparation titled* **'Evaluating restoration trajectories using DNA metabarcoding of invertebrates and their associated plant communities'**

This chapter investigates the use of eDNA metabarcoding to monitor the recovery of invertebrate communities, well established indicators of restoration. Invertebrates are good indicators of ecosystem condition because they play a vital role in many ecosystem functions such as pollination, seed dispersal, nutrient cycling, and they are also numerous and easy to capture. However, invertebrate surveys are challenging because they require many person hours of specialist taxonomists to identify specimens, and many species remain undescribed. DNA metabarcoding offer a potential rapid survey method that can assess invertebrate diversity as well as the diversity of plant the invertebrates are interacting with. However, many species will not be identifiable using eDNA methods because they are not present in the reference databases. This chapter examines the recovery of ground dwelling and airborne invertebrates and their associated plant communities across three different ecosystem in Western Australia. Here we assess if DNA metabarcoding of invertebrates can evaluate restoration trajectories despite limited taxonomic identification of sequence variants.

### 5.1.1 Acknowledgements

This work was supported by the Australian Research Council Industrial Transformation Training Centre for Mine Site Restoration (ICI150100041) and the Pawsey Supercomputing Centre, with funding from the Australian Government and the Government of Western Australia. We thank the mining companies BHP, Hanson Construction Material, and South32 for facilitating access to sites for sampling. We would also like to thank Sheree Walters for help with sample collection and the members of the Trace and Environmental DNA (TrEnD) Laboratory for support with metabarcoding workflows and bioinformatics.

### 5.1.2 Data Accessibility

Sequencing and sample data and is available at the Dryad Digital Repository: https://doi.org/10.5061/dryad.q573n5tgw

### 5.1.3 Author Contributions

MvH conducted the study and wrote the manuscript. MvH, PN, MB, NW, and GW-J were involved in the experimental design. Samples were collected and processed by MvH; molecular and bioinformatics work was performed by MvH; all data was analyzed and processed by MvH; statistical analysis was done by MvH; the manuscript was edited by all authors.

## EVALUATING RESTORATION TRAJECTORIES USING DNA METABARCODING OF INVERTEBRATES AND THEIR ASSOCIATED PLANT COMMUNITIES

van der Heyde, M.[1,2*], Bunce, M. [2,3], Dixon, K.W.[2], Majer, J. D.[1], Wardell-Johnson, G.[1], White, N.E. [2], Nevill, P.[1, 2]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, GPP Box U1987, Perth, Western Australia, 6845

[2]Trace and Environmental DNA Laboratory, School of Life and Molecular Sciences, Curtin University, GPP Box U1987, Perth, Western Australia, 6845

[3]Environmental Protection Authority, 215 Lambton Quay, Wellington 6011, New Zealand.

*Corresponding author

## 5.2 Abstract

Invertebrate communities provide many critical ecosystem functions, including pollination, decomposition, herbivory and soil creation. They have also been identified as indicators of ecological restoration and therefore make excellent monitoring targets. Unfortunately, invertebrates are often overlooked in restoration monitoring because they are time consuming to survey, require often rare taxonomic expertise to identify, and there are many undescribed species. DNA metabarcoding is a tool to rapidly survey invertebrate communities, which can also provide additional information about the plants those invertebrates are interacting with. Here we evaluate how invertebrate communities may be used to determine ecosystem trajectories during restoration. We collected invertebrates from vane and pitfall traps across chronosequences of mine site restoration in three ecologically different locations in Western Australia and used DNA metabarcoding to identify the invertebrate community and plants they have interacted with. Ground-dwelling invertebrates showed the clearest restoration signals, with communities becoming more similar to reference communities over time. These patterns of community recovery were weaker in airborne invertebrates, which have higher dispersal abilities and therefore less local fidelity to environmental conditions. Invertebrate community recovery was most evident in ecosystems with relatively stable climax communities, while the trajectory in the Pilbara, with its unpredictable monsoonal flooding, was unclear. Results from the plant assay indicate that invertebrates are foraging locally, and provides additional functional data about the interaction between invertebrates and their environment. Thus, we show how DNA metabarcoding of invertebrate communities can be used to evaluate likely trajectories for restoration, which enables the definition of success criteria and informs on the required time scale for restoration monitoring. Testing and incorporating new monitoring techniques like DNA metabarcoding is critical to improving restoration outcomes, and is particularly salient now given the ambitious global restoration targets associated with the recently announced UN decade on Ecosystem Restoration.

## 5.3 Introduction

Fauna are often overlooked in restoration monitoring in favor of vegetation (Cross, Tomlinson, Craig, Dixon, et al. 2019; Ruiz-jaen and Aide 2005), with the general assumption that they will naturally recolonize an area with the return of plant communities (Palmer, Ambrose, and Poff 1997). However, this is not always the case (Cristescu et al. 2013), and understanding the recovery of fauna is important because they play a vital role in many ecosystem function including pedogenesis, seed dispersal, pollination and nutrient cycling (Catterall 2018; Bronstein, Alarcón, and Geber 2006; Ness et al. 2004; Hunter 2001). Recently, greater attention has been paid to fauna to both assess and facilitate ecological restoration (Catterall 2018; Cross, Bateman, and Cross 2020; Majer 2009).

Invertebrates are of particular interest as they have long been used as indicators of ecosystem recovery in both aquatic and terrestrial systems (Andersen et al. 2002; Andersen and Sparling 1997; Folgarait 1998; Majer 2009). They are sensitive to disturbances and essential for ecosystem function (Rosenberg, Danks, and Lehmkuhl 1986; Folgarait 1998), not to mention being numerous, easy to capture, and incredibly diverse (Gaston 1991). Because studies tend to target particular groups of arthropods, responses to restoration are mixed, depending on the target taxa (Cristescu, Frère, and Banks 2012). Some of the variation in responses to restoration among different arthropod classes may be attributed to dispersal ability. For example, beetles with high dispersal abilities are able to recolonize more quickly than millipedes in a regenerating forest (Magura et al. 2015). Along with dispersal ability, changes in community composition during restoration (Andersen et al. 2002; Majer 2009) have been attributed to a shift from generalist r-strategist species, which thrive in disturbed and unpredictable environments, to K-selected species, which require predictable, and favorable environments (Majer 1989;). As such, invertebrate communities may be used to evaluate restoration trajectories of recovery or convergence, where the objective and expectation is *directional* change in composition towards a reference community (Mcdonald et al. 2016; Suding and Gross 2006). However, in harsher ecosystems that are often naturally unpredictable, the lower diversity and selection of A (Adversity)-strategists ( Southwood 1977; Dunlop et al. 1985; Majer 1989) may make directional changes during restoration to be less likely.

Despite being excellent indicators of ecosystem change, the high diversity within invertebrate communities makes it particularly difficult to identify invertebrate specimens, often requiring many expert person-hours from multiple taxonomists specializing in different invertebrate taxa (Majer 1983). This process is costly and time consuming, and dependent on taxonomic expertise that is dwindling worldwide (Pearson, Hamilton, and Erwin 2011; Majer et al. 2013). Additionally, many invertebrate taxa are cryptic (Smith, Fisher, and Hebert 2005) or have yet to be identified, especially in Australia with its high degree of endemism (Austin et al. 2004; Rix et al. 2015) and as much as 75% of the Australian arthropod diversity being undescribed (Austin et al. 2004; Yeates, Harvey, and Austin 2003). Consequently, most studies looking at invertebrate responses to restoration have targeted certain taxa either because they have previously shown to be good bioindicators (Andersen et al. 2002), or they are threatened and of legal and conservation value (i.e. Lepidoptera) (Majer, 2009).

Some of the difficulties associated with invertebrate monitoring can be reduced using DNA metabarcoding to provide taxonomic assignments. This process uses high-throughput sequencing to determine invertebrate diversity from small barcoding regions of the genome (Beng et al. 2016; Ji et al. 2013; Yu et al. 2012). Compared to morphological identification where each specimen has to be identified individually, DNA metabarcoding has been shown to be accurate, reliable, and faster than conventional morphological methods (Beng et al. 2016; Ji et al. 2013). As an added benefit, the sequencing data can be readily stored and analyzed by a third party, such as regulators (Fernandes et al. 2018). Although abundance estimates using DNA metabarcoding are often skewed by primer bias (Elbrecht and Leese 2015) or DNA extraction method (Majaneva et al. 2018), presence/absence data has been used to demonstrate arthropod responses to post mine site restoration (Fernandes, van der Heyde, et al. 2019) and land use change (Beng et al. 2016).

One of the advantages of DNA metabarcoding is its ability to detect not only invertebrate diversity and composition but also provide functional data by identifying the organisms they have been interacting with (Jurado-Rivera et al. 2009; Pornon et al. 2016). In the case of arthropods, previous studies suggest that DNA from arthropod samples should be able to identify which plant species pollinators have

visited (Pornon et al. 2016) and which plant species they have consumed (Jurado-Rivera et al. 2009). However, these studies have hitherto not been undertaken in a restoration context, so the utility of such approaches for restoration monitoring is unknown. Presumably, assessing these communities can illustrate the interaction between invertebrates and plants during restoration. However, since the invertebrates may carry plant DNA from outside the restoration area (van der Heyde et al., 2020a), they may not necessarily have high fidelity to local conditions.

Our earlier work has explored the use of DNA metabarcoding of ground-dwelling invertebrates to monitor mine site restoration (Fernandes, van der Heyde, et al. 2019); however, this study used a single reference site per mine and the results were spatially auto-correlated as older sites were closest to the reference sites. Here we use two spatially separated reference sites per mine, two trap types that capture ground dwelling and airborne invertebrates, and study sites in multiple locations with different climates and ecosystems. This study evaluates whether we can use DNA metabarcoding of invertebrates to evaluate restoration trajectories (convergence to reference communities) in restored sites. We have four hypotheses:

i) Ground dwelling invertebrates will show recovery trajectories better than airborne invertebrates because with lower dispersal abilities they better reflect local environmental conditions.

ii) Ecosystems with stable climax communities demonstrate trajectories of recovery more clearly than less diverse, climatically harsher unpredictable ecosystems.

iii) The plants associated with invertebrates will not show trajectories of recovery as well as invertebrates because plant DNA may be sourced from outside the site area.

iv) Metabarcoding provides functional information by indicating how invertebrate communities are interacting with the plants in and around restoration sites

## 5.4 Material and Methods

### 5.4.1 Study Sites

Restoration and reference sites were sampled from three locations up to 1000km apart in Western Australia, namely: the Swan Coastal Plain (SCP), Jarrah Forest (JF) and Pilbara (PB). There was consistency in restoration approaches, soil type, climate and site aspect of the sites within each location. At each location, sites of different restoration age were sampled along with two spatially separated reference sites (Figure 1). At all three locations, we sampled at least two sites less than 9 years old (Young), and at least two sites older than 9 years (Older). These sites are previously described in van der Heyde et al. (2020b) (Chapter 2), and briefly below. At all locations two reference sites were selected on the basis of the following criteria: similarity to ecosystems that are the target of restoration efforts, proximity to restoration sites, similarity in slope and aspect, and spatially separate from each other to account for variation in reference communities.

The Coastal Plain has a warm-summer Mediterranean climate with mild cool wet winters; temperature has a mean minimum of 12.8°C, mean maximum of 24.7°C, with 757 mm mean annual rainfall (Australian Bureau of Meteorology). This location is part of the broader region of South-Western Australia, a globally recognized biodiversity hotspot (Myers et al. 2007). The mine is located on the silicaceous Bassendean dunes, with high acidity and low water-holding capacity (Dodd and Heddle 1989; McArthur 1991). The ecosystem is referred to as Banksia Woodland after the dominant tree species, *Banksia attenuata* and *B. menziesii*. Other trees include less dominant *Eucalyptus todtiana* and *Nuytsia floribunda.* The understory consists of woody species of Myrtaceae, Ericaceae, Proteaceae, and Epacridaceae, and non-woody species in Anthericaceae, Stylidiaceae, Cyperaceae, and Haemodoraceae (Trudgen 1977). In October 2018, we sampled eight sites at a Hanson Construction Materials sand quarry in Lexia (31.76 °S, 115.95 °E), with two reference sites and restoration sites 1, 3, 7, 11,14, 22 years old. The sites have been restored with the aim of returning mined areas to the surrounding native Banksia woodlands. All restoration was done by Hanson and previous mine owners and included direct transfer of fresh topsoil, ripping, and seeding with native plant species. Plant species richness and density tended to be higher in restoration than

reference sites, and percent cover has increased with restoration age and is highest in reference sites (Benigno, Dixon, and Stevens 2013).

The second location in the Jarrah (*E. marginata*) forest is also part of the Southwest Australia Biodiversity hotspot (Myers et al. 2007) and has a similar hot-summer Mediterranean climate; temperatures have an mean minimum of 8.6°C, mean maximum of 23.7°C, and 668.9 mm annual mean rainfall (Australian Bureau of Meteorology). The lateritic soils are nutrient poor and high in gravel with surfaces rich in iron and aluminum (McArthur 1991). The vegetation is dominated by *E. marginata;* other common trees are *E. patens,* and *E. wandoo*. The understory consists of sclerophyllous shrubs from several families including Anthericaceae, Fabaceae, Asteraceae, Proteaceae, Dasypogonaceae, and Myrtaceae (Havel 1975). We sampled six sites from the bauxite mine which is now run by South32 (32.96°S, 116.48°E) in October 2018; two reference sites and restoration sites 2, 6, 11, and 20 years old. All restoration was undertaken by South32 or the previous mine owners. Post mining the landscape was shaped using waste material and gravel. Fresh topsoil was directly transferred from newly mined areas to the restoration area and supplemented with stockpiled topsoil as needed. The sites were then ripped, seeded with over 100 native species, recalcitrant plants (mostly grasses) were planted, and a one-time treatment of superphosphate is applied (Data from South32). Reference and restoration sites are dominated by Myrtaceae and Fabaceae species. Total cover has increased with age of restoration to similar cover percentages of reference sites (Data from South32).

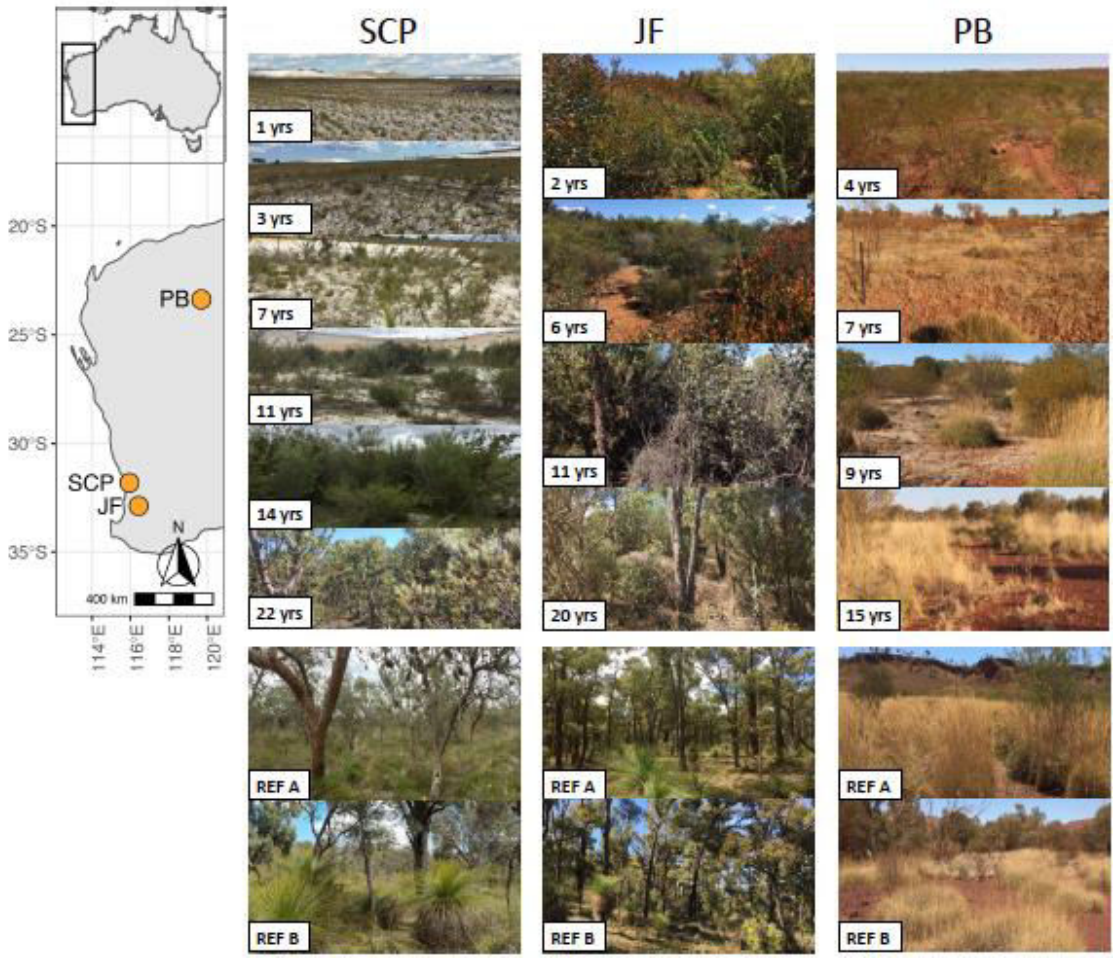


**Figure 5.1** Chronosequences of mining restoration where invertebrate samples were collected. Restoration sites shown above, with the number of years of restoration from 1 to 22 years. Reference sites shown below. SCP-Swan Coastal Plain, JF-Jarrah Forest, PB-Pilbara

The third location, the Pilbara, is in northwest of Western Australia. The Pilbara has a hot, arid climate with most rainfall occurring in the summer, and associated with cyclonic activity (McKenzie, van Leeuwen, and Pinder 2009) causing unpredictable flooding in areas. Temperatures have a mean minimum of 15°C and mean maximum of 30.6 °C, with 263.8 mm mean rainfall (Australian Bureau of Meteorology). The unfavourable conditions and large variation in yearly rainfall are thought to select for a wide range of r-and A-strategist invertebrates (Majer, 1989). Soils are acidic stony loams with low fertility, which support open woodlands of snappy gum (*E. racemosa*) over hummock grasses (*Triodia wiseana, T. basedowii, T. lanigera*) and low Acacia shrubs (McKenzie, van Leeuwen, and Pinder 2009). The Pilbara is a significant mining region and accounts for 39% of

global iron ore production (Government of Western Australia, 2019). We sampled six sites at a BHP iron ore mine (22.84 °S, 118.95 °E) in September 2018, with two reference sites and restoration sites 4, 7, 11, and 15 years old. Restoration was conducted by the mine owners; landscapes were reformed and stockpiled topsoil (average age 10 years) was applied and then ripped. Restoration areas tended to have higher coverage of woody shrubs (*Acacia*), while reference sites and older restoration areas have more hummock grasses (*Triodia*). Restoration areas also had invasive species such as buffel grass (*Cenchrus ciliaris*) and kapok bush (*Aerva javanica*), which were absent in reference sites (Data from BHP).

### 5.4.2 Sample Collection

At each site we collected 10 invertebrate samples, five from vane traps and five from pitfall traps (n=200). Each vane trap sample included the contents of a yellow and blue vane trap with 150 mL of ethylene glycol and was left on the site for 7 days. Each pitfall trap sample included the contents of four pitfall traps (4 cm diameter, 12 cm deep with ethylene glycol as a capture fluid), and was also left in the field for 7 days Pitfall traps were spaced 10 m apart in a square around the vane traps in the center for each sample point.

### 5.4.3 Sample Processing

For DNA extraction, we first rinsed off the ethylene glycol with de-ionized water using 20-µm sieves that were sterilized in bleach and under UV light between every sample. Samples were then homogenized using a TissueLyser (Qiagen) for 2 min in 30 sec increments at 30/s in 50mL falcon tubes with 4 steel balls (4mm diameter). 400µL of the homogenate was digested overnight and the DNA extracted using the DNeasy Blood and Tissue kit (Qiagen) on the QiaCube Connect automated platform (Qiagen). The final elution volume was 200 µL, and extraction controls (blanks) were carried out for every set of extractions. Quantitative PCR (qPCR) was run on neat extracts and a 1/10 dilution to see if samples exhibited inhibition, and to determine optimal DNA input for PCR for each sample to maximize input relative to any inhibitors (Dáithí C. Murray, Coghlan, and Bunce 2015). Two assays were used in this study to target invertebrate and plant diversity. The invertebrate assay used the primers fwhF2/fwhR2n (Vamos, Elbrecht, and Leese 2017) to amplify a 205bp

section of the cytochrome c oxidase I (COI) region. For plants we used the trnlc/h primers (Pierre Taberlet et al. 2007a) which targets the chloroplast trnL (UAA) intron

The qPCRs were run on a StepOne Plus (Applied BioSystems) real-time qPCR instrument with the following conditions: 5 min at 95°C, 40 cycles of 95°C for 30s, 30s at the annealing temperature (50°C for invertebrates, 52°C for plants) and 45s at 72°C, a melt curve stage of 15s at 95°C 1 min at 60°C and 15s at 95°C, ending with 10 min elongation at 72°C. The PCR mix for quantitation contained: 2.5 mM MgCl2 (Applied Biosystems, USA), 1× PCR Gold buffer (Applied Biosystems), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4 mg/ml bovine serum albumin (Fisher Biotec, Australia), 0.4 μmol/L forward and reverse primer, 1 U AmpliTaq Gold DNA polymerase (Applied Biosystems) and 0.6 μl of a 1:10,000 solution of SYBR Green dye (Life Technologies, USA). Extraction control and non-template controls were included in qPCR assays.

After optimal DNA input was determined by qPCR, each sample was assigned a unique combination of multiplex identifier (MID) tags for each primer assay. These MID tags were incorporated into fusion tagged primers, and none of the primer-MID tag combinations had been used previously in the lab to prevent cross contamination. Fusion PCRs were done in duplicate and to minimize PCR stochasticity, the mixes were prepared in a dedicated clean room before DNA was added. The PCRs were done with the same conditions as the standard qPCRs described above. Samples were then pooled into approximately equimolar concentrations to produce a PCR amplicon library that was size-selected to remove any primer-dimer that may have accumulated during fusion PCR. Size selection was performed (150-450bp) using a PippinPrep 2% ethidium bromide cassette (Sage Science, Beverly, MA, U.S.A). Libraries were cleaned using a QIAquick PCR Purification Kit (Qiagen, Germany) and quantified using Qubit Fluorometric Quantitation (Thermo Fisher Scientific). Sequencing was performed on the Illumina MiSeq platform using the 300 cycle V2 as per manufacturer's instructions.

### 5.4.4 Sequencing analysis

Sequences were demultiplexed using a demultiplex function in the "insect" package (Wilkinson et al. 2018) on the R 3.5.3 platform (R Core Team, 2018). Further sequence processing was performed in R using the "DADA2" package (Callahan et al. 2016) where sequences were quality filtered, the error rates were estimated, and the sequences were dereplicated. The error rates were then used in the sample inference stage to remove sequences likely to be errors and leave Amplicon Sequence Variants (ASV). These ASVs are equivalent to zero radius operational taxonomic units (ZOTUs) in usearch (Edgar 2016). The sequence table was then constructed and chimeras removed. Taxonomy was determined using the Basic Local Alignment Search Tool (blastn) on a high-performance cluster computer (Pawsey Supercomputing Centre) to search against the online reference database GenBank (https://www.ncbi.nlm.nih.gov/genbank/). Invertebrate sequences were also searched against and arthropod COI reference sequences extracted from the Barcode of Life Database (BOLD: https://www.barcodeoflife.org), because there are reference sequences that are found uniquely on one of the two databases. We used MEGAN (Huson et al. 2007) to assign taxonomy with a minimum support of 205 considering the top 50 blast hits.

### 5.4.5 Statistics

All statistics were run using R 3.5.3 (R Core Team 2018). Samples with low sequencing depth were removed and ASVs that were present in the extraction controls were removed from the dataset. We selected ASVs in the phylum 'Arthropoda' for the invertebrate assay and 'Plantae' for the plant assay. Copy numbers in each sample were filtered to a minimum of 0.05% within sample abundance. We verified there was no correlation between sequencing depth and ASV richness before continuing. Read counts were transformed to presence absence to reduce the effects of biases (Elbrecht and Leese 2015; Majaneva et al. 2018). Spatial autocorrelation was tested using the Mantel test in the 'ade4' package in R (Mantel 1967). Three criteria were examined to determine if communities showed a trajectory of recovery or convergence to the reference community. First, community composition should be different between younger restoration, older restoration, and reference sites. This was visualized using Non metric multidimensional scaling

(NMDS), based on presence/absence ASV table and with Bray-Curtis dissimilarity. The 'ordiellipse' function from the 'vegan' R package was used to draw ellipses showing the 95% confidence interval of the group (Oksanen et al. 2018). Differences between restoration ages were tested using permutational multivariate analysis of variance (PERMANOVA). Second, establishing a restoration trajectory requires *directional* change; we expect that restored communities become more similar to reference communities over time. Replicates at each site were pooled and the similarity between each site and the reference sites was calculated. This relationship was tested using linear models separately for each assay and location. Third, we expect that the proportion of 'reference' ASVs, that is, ASVs that were found in reference sites, would increase over time. This relationship was tested using a simple linear model. For all three, we tested the SCP data with and without the extra two sites (7 years and 11 years) to ensure that any comparisons of trajectory between the locations was fair. This analysis is based on the prediction of changing composition from r- or A- to K-strategists and provides additional information about whether the patterns in community similarity to reference communities is driven by compositional changes, or richness. Finally, to understand the taxa associated with restoration and reference sites, we ran a multipattern analysis for each site using the R package 'indicspecies' (De Caceres and Legendre 2009).

## 5.5 Results

In total, we generated 14,780,759 quality-filtered invertebrate sequences from 196 samples with a minimum of 3,000 reads/sample. Out of 5862 initial ASVs, 2635 belonged to the phylum Arthropoda. The remaining ASVs were either unidentified or fungi, and only made up 23.7% of the read count. In the plant assay, we generated 13,441,527 filtered plant sequences from 197 samples with a minimum of 5600 sequences/sample. From the initial 511 plant ASVs, 381 remained post filtering and these accounted for 87.8% of the sequences. Overall, there were fewer ASVs in the Pilbara compared to the Coastal Plain or Jarrah, especially in the pitfall traps where the Pilbara had 17-32% fewer invertebrate ASVs

### 5.5.1 Community Composition

Invertebrate diversity in the vane traps was dominated by Hymenoptera, Coleoptera, Diptera, Hemiptera, and Lepidoptera. Some of these (Hymenoptera, Coleoptera, and Hemiptera) also made up most of the diversity in the pitfall traps, along with Collembola and Aranae. Collembola were largely absent from the Pilbara, which had more Orthoptera ASVs. The majority (67%) of invertebrate ASVs could not be identified beyond order level. However, 99% of plant ASVs could be identified to family level. Plant diversity in the SCP and Jarrah was dominated by Myrtaceae, Fabaceae, Dilleniaceae, and Proteaceae, while in the Pilbara the richest families were Fabaceae, Poaceae and Malvaceae (Figure 5.2). Because of the poor taxonomic assignments, we confined our considerations to ASVs for our subsequent analyses.

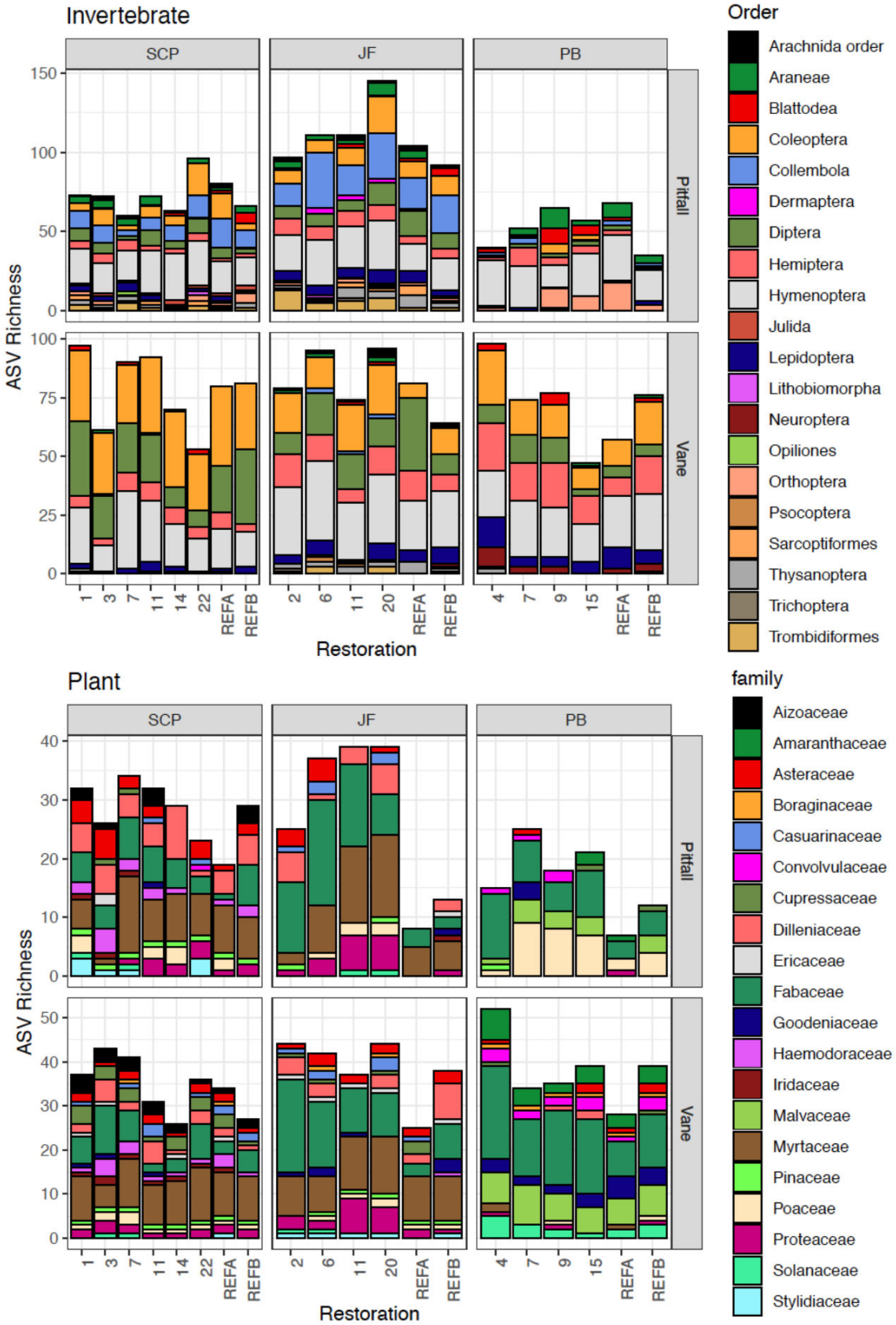


**Figure 5.2** Composition of invertebrates (above) and plant (below) communities detected from pitfall and vane traps. Shows the number of ASVs in each order (invertebrates) or family (plants) at all restoration and reference sites.

There were significant differences in community composition between younger restoration, older restoration and reference sites in all locations for both trap types and assays (Figure 5.3, PERMANOVA, alpha=0.05). Similarly, the pairwise analysis showed all restoration ages were significantly different from each other (alpha=0.05), with the exception of the plant community in the Pilbara samples, where the reference samples were not significantly different from the younger (vane) or the older (pitfall) restoration samples (Table S5.9.2). The Mantel tests showed significant spatial autocorrelation in the invertebrate communities from pitfall traps but not the vane traps (alpha=0.05, Table 5.1). Similarly, the spatial correlation with community dissimilarity was lower in the plant sequences compared to the invertebrate assay (Table 5.1).

**Table 5.1** Results of the Mantel test showing the correlation between spatial distances and community dissimilarity. Results for the samples separately, and pooled (sites) are shown. Inverterbate assay-fwh, plant assay-trnl

| | | | **Samples** | | **Sites** | |
|---|---|---|---|---|---|---|
| **Trap** | **Assay** | **Location** | r | p | r | p |
| **Pitfall** | fwh | JF | 0.161 | 0.001 | 0.189 | 0.205 |
| | | PB | 0.308 | 0.002 | 0.566 | 0.018 |
| | | SCP | 0.375 | 0.001 | 0.349 | 0.116 |
| | trnl | JF | 0.069 | 0.064 | -0.1858 | 0.706 |
| | | PB | -0.0118 | 0.508 | -0.203 | 0.838 |
| | | SCP | -0.077 | 0.785 | -0.2365 | 0.789 |
| **Vane** | fwh | JF | 0.131 | 0.011 | 0.237 | 0.104 |
| | | PB | 0.114 | 0.082 | -0.084 | 0.548 |
| | | SCP | -0.117 | 0.851 | -0.074 | 0.555 |
| | trnl | JF | 0.0466 | 0.147 | -0.314 | 0.916 |
| | | PB | 0.233 | 0.007 | 0.079 | 0.319 |
| | | SCP | 0.115 | 0.125 | -0.359 | 0.874 |

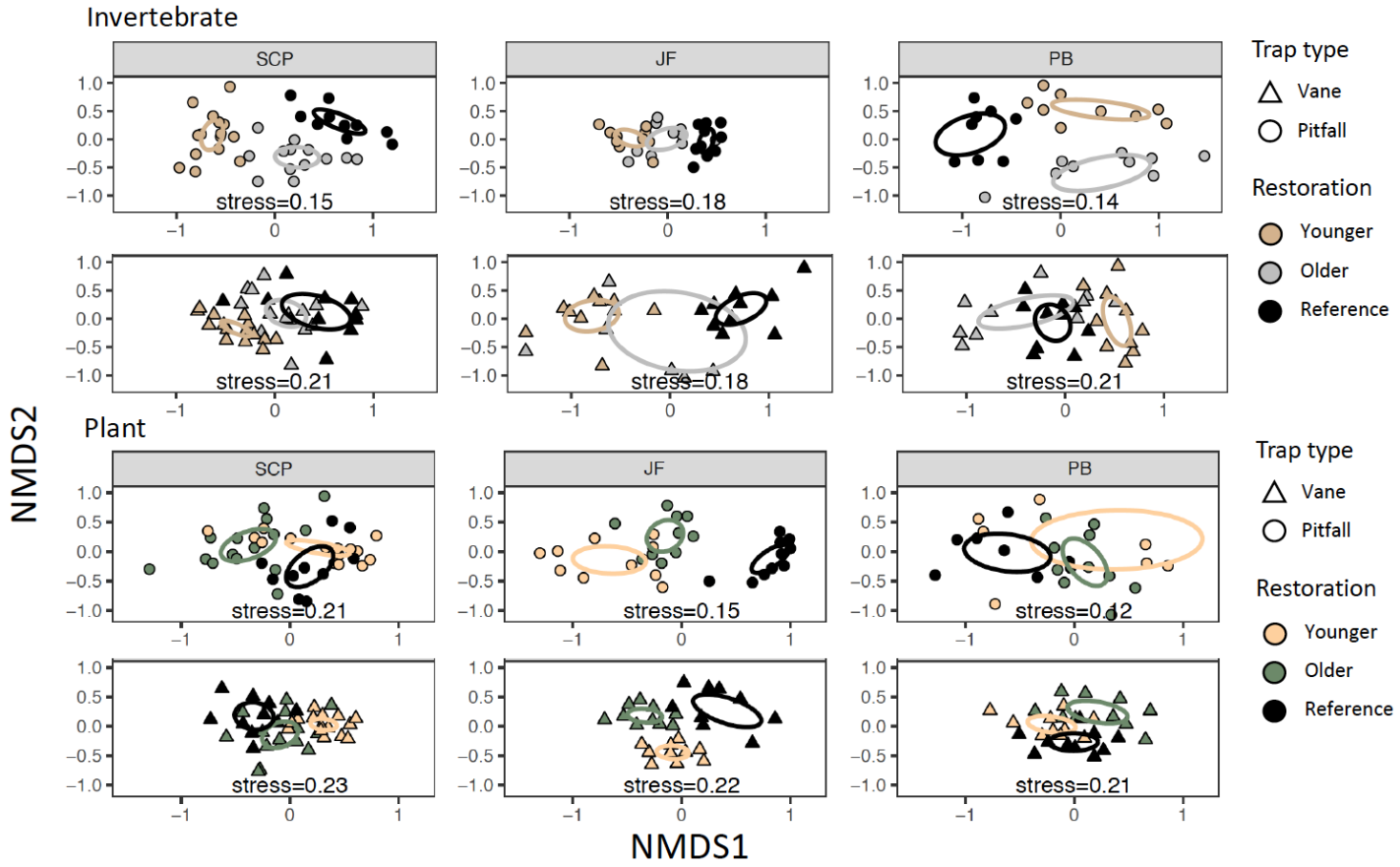


**Figure 5.3** NMDS ordinations or Invertebrate and plant communities in restoration and reference sites. Ellipses were drawn using 'Ordiellipse' in the vegan R package and indicate 95% confidence interval of the group. PERMANOVAs were significant for all facets (alpha=0.05).

### 5.5.2 Similarity to reference communities

The invertebrate communities showed clear directional changes (increasing similarity to reference over time) in the pitfall traps from the Coastal Plain and the Jarrah (Figure 5.4). This trajectory is less evident (SCP) or entirely absent in the vane traps (JF, PB). There were no directional changes in invertebrate community composition observed in the Pilbara. The results from the plant communities were different. In the Coastal Plain, there was a significant relationship between similarity to reference communities and age of restoration in the vane traps, but not the pitfall traps. The directional change in plant communities occurred in both the pitfall and vane trap samples in the Jarrah, but was significant in the pitfall traps only. Similarly, the plant communities became more similar to reference communities in the Pilbara pitfall traps while there was not a relationship in the vane traps (Figure 5.4).

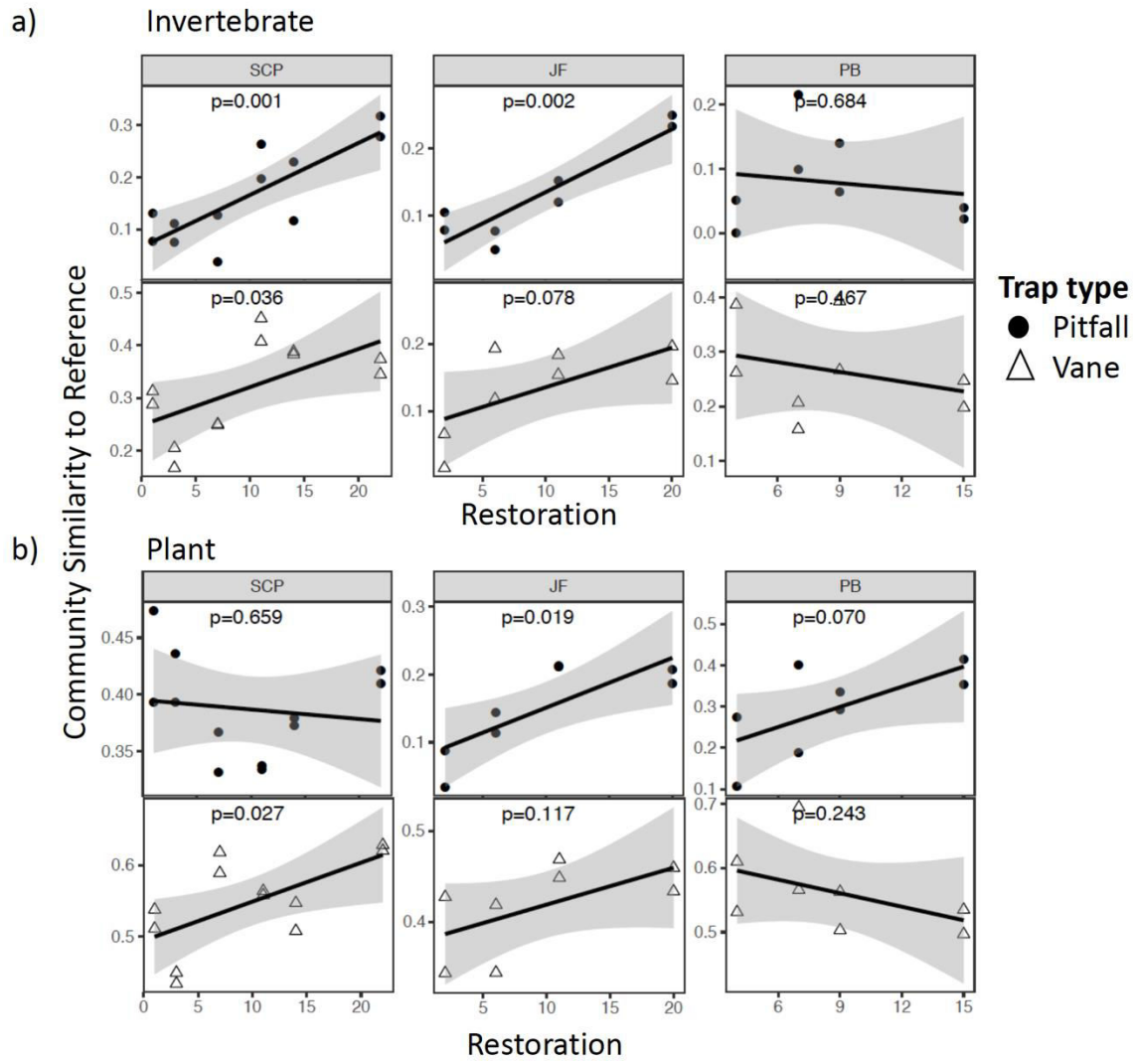


**Figure 5.4** Similarity (bray-curtis) of restoration sites of different ages (years) to communities in reference sites. Lines indicate linear models with 95% confidence interval shown with shading. P-values for the linear models shown for each plot. Removing the two extra sites in the SCP (7,11 years) did not change the relationships or the significance of the models, with the exception of invertebrate communities from vane traps (from p=0.036 to p=0.053)

### 5.5.3 Proportion of "reference" associated ASVs

Only the invertebrate communities from the pitfall trap samples from the Coastal Plain and the Jarrah forest showed significant increases in the proportion of 'reference' ASVs over time. For plant sequences, only pitfall traps in the Pilbara that showed increasing 'reference' ASVs over time (Figure 5.5). Overall the vane traps had a higher proportion of ASVs that were shared with reference samples than pitfall traps. This was true for both the invertebrate assay (49.4% vs 22.2% 'reference'

ASVs) and the plant assay (78.6% vs 59.7% 'reference' ASVs). There was also a higher proportion of shared ASVs in the plant assay compared to the invertebrate assay (Figure 5.5). Between the two reference sites, there was variation in the number of ASVs shared with each other. The pitfall traps in the Pilbara only had 3 ASVs shared between the two reference sites (average of 1.2 ±0.4ASVs per sample). The amount of shared ASVs was higher between the Coastal Plain and Jarrah pitfall traps (10 and 8 respectively).

### 5.5.4 Multipattern Analysis

Across the three locations, there were 82 invertebrate ASVs with significant association (alpha =0.05) with younger restoration (<9 years), older (>9 years), reference sites, or a combination (Table S5.9.3). Of these, 44 were assigned to family, 16 to genus, and only 3 to species level. This includes *Iridomyrmex sanguineus*, which was associated with younger restoration in the Pilbara and *Monomorium rothsteini*, associated with reference sites in the Pilbara. Most Coleoptera (12/16) were associated with older restoration or reference sites and 13 of those were from vane trap samples. Apidae ASVs were found mainly in the younger restoration vane traps. For the plant assay, there were 59 ASVs with significant association (Table S5.9.4), 52 of which were assigned to family, 21 to genus, and 8 to species level. Some of the species include *Petrophile squamata,* found in older restoration in the Jarrah, and *Porana commixta* found in vane traps of reference sites in the Pilbara (Table 5.2). Most Fabaceae ASVs (13/14) were associated with younger restoration in the Coastal Plain and Jarrah.

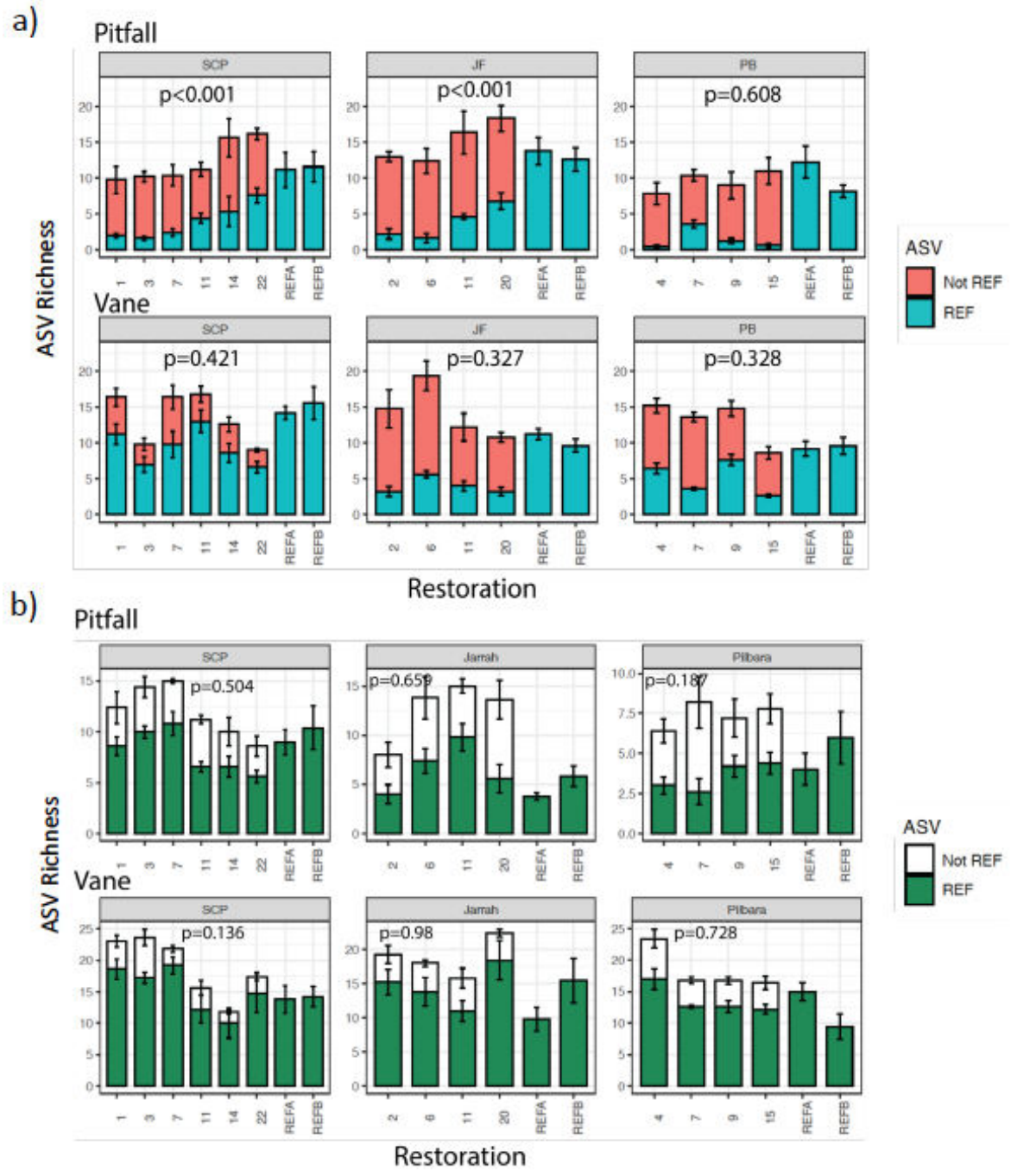


**Figure 5.5** ASV richness in the different sites, separated into ASVs that were present in reference sites and those not found in reference sites. Separated into a) invertebrate and b) plant communities. P-values indicate the significance of the relationship between the proportion of reference ASVs and age of restoration (years).

**Table 5.2** Taxa of interest, based on general observations of the data and indicator species analysis

| Taxa of interest | Name | Reason |
|---|---|---|
| ***Melophorus*** | Australian genus of ant | Associated with younger restoration sites in both the SCP and PB. Species in this genus are known as 'sun-loving' (Andersen et al. 2002) and are often found in restoration sites (Andersen, Hoffmann, and Somes 2003). |
| ***Iridomyrmex sanguineus*** | Northern meat ant | Associated with younger restoration sites in PB. *Iridomyrmex* species are among the first to colonize revegetated sites (Andersen 1993). |
| **Julida - *Ommatoiulus*** | Portugese millipede | Invasive detritivore species found in great abundance in the SCP, particularly in older restoration and reference sites. Feeds on litter, which is more available in those sites. |
| **Fabaceae** | Legume family | ASVs in this family are strongly associated with younger restoration in JF and PB. Acacia shrubs tend to establish rapidly at restored sites in these locations (Data from BHP, Data from South32). |
| ***Goodenia microptera*** | *Goodenia microptera* | An insect pollinated species found predominantly in vane traps of PB reference sites. |
| ***Anigozanthos*** | Kangaroo paw | Associated with younger restoration in SCP pitfall traps. These grow quickly (within a year) in SCP restoration. |

## 5.6 Discussion

Terrestrial invertebrate fauna are key indicators of ecosystem change (Andersen et al., 2002; Majer, 2009; Majer, Brennan, & Moir, 2007), and in this study, we show that even without taxonomic identification, DNA metabarcoding of invertebrate samples can be used to rapidly assess complex biological interactions and evaluate restoration trajectories. These trajectories of community recovery were more evident in stable climax ecosystems and in ground-dwelling invertebrates with lower dispersal ability than airborne invertebrates. Examining plant diversity associated with invertebrate samples also showed some indications of directional changes in community composition and demonstrates that invertebrates are likely foraging locally.

### 5.6.1 Ground-dwelling vs airborne invertebrate

Vane traps do not show the same local fidelity as pitfall traps; and, as expected, tend to have weaker indications of community recovery (Figure 5.3, Figure 5.4). Vane traps capture airborne invertebrates, typically pollinators (Hall 2018), and can trap organisms that may come from more than 1.8 km away (Jha and Dick 2010) while species caught by pitfall traps have more limited dispersal (Majer, 1980; Ness et al., 2004; Ward, New, & Yen, 2001). This would also explain the greater proportion of shared taxa in the vane traps compared to the pitfall traps (Figure 5.5), and the greater spatial correlation in pitfall trap samples (Table 5.1). Beyond the differences in attraction distance of the traps, our results also suggest quicker recolonization of airborne invertebrates as the number of 'reference' associated taxa is similar to reference sites within a few years (Figure 5.5, SCP, PB). Variation in dispersal abilities is important as those with more mobility are able to recolonize areas more quickly (Magura et al. 2015) and from greater distance (Knop, Herzog, and Schmid 2011). Fortunately, there is no sign of thermophilic or other barriers (Cranmer, McCollin, and Ollerton 2012; Tomlinson et al. 2018) preventing invertebrates from accessing and using restoration sites. Because of their more sedentary nature, ground-dwelling invertebrates are good indicators of organisms that are likely reproducing in situ, while airborne invertebrates can indicate the forage support and attractiveness of a site. Our findings indicate that invertebrate communities are demonstrating a reasonable ability to recover without intervention

following the establishment of plant communities, conforming with the 'Field of Dreams' Hypothesis which posits that if suitable habitat can be created, species will colonize it and function will be restored (Palmer, Ambrose, and Poff 1997). However, this is dependent on the presence of source populations. In this study, all sites were near remnant vegetation that could act as a taxa pool; in cases of isolated restoration sites it may be more difficult to evaluate restoration trajectories using invertebrate communities.

Stable vs unpredictable ecosystems

The r/K selection theory is a predictive model for life history strategies that vary from r selected (high fecundity, short lifespan, small bodies, opportunistic, high dispersal) in unpredictable environments to K-selected (low fecundity, long lifespan, large bodies, low dispersal) in predictable environments (Pianka 1970). In ecological succession and restoration, it is expected that systems are dominated by r-selected species initially as they take advantage of the disturbance, followed by a shift to K-selected species as the system develops towards a stable climax community (Majer, 1989). This concept is developed further by Southwood (1977) and Greenslade and Greenslade (1983) as a 'habitat template', which condenses the variety of habits onto two axes equivalent to their favourableness and predictability. As well as explaining the conditions for r- and K-strategists, this template introduces a third adversity or A-selection strategy, which is selected for in environments that are very unfavourable and not always predicable. Such environments, including the Pilbara, support lower diversities of organisms with lower interaction between species (Greenslade and Greenslade 1983). In this study, we classified taxa based on whether they were found in reference sites as a proxy for selection strategy, since there was inadequate information to classify them based on taxonomic identification. As expected, in older restored sites we saw significant increases in the proportion of 'reference' taxa with time in both the Jarrah and Coastal Plain (Figure 5.5), which shows a directional change in community composition toward that of the reference community.

The Pilbara location, which has a more unpredictable and harsher climate, did not show a similar trajectory of community recovery. Dunlop et al. (1985), and to a lesser extent, Fletcher (1990), observed ant richness fully recovered in Pilbara rehabilitation, but, similar to our results, the species composition remained different

between natural and restored sites. In the Pilbara, the main factors driving compositional turnover in terrestrial fauna are regolith/soil and landform/hydrogeologic, as well as climate (Gibson et al. 2015), all were factors that were shared between Pilbara restored and reference sites. Here, the structure of the revegetation rapidly came to resemble the structure of the original predominantly grassland habitat (see Figure 5.1), which is in marked contrast to the situation at the other two locations. In that regard, the reference areas may provide conditions that are as unpredictable and unfavourable as the areas under restoration; compared with the other two regions, they are also less rich in species. Thus, recolonization of Pilbara sites may be more stochastic and less influenced by selection pressures than in the Coastal Plain and Jarrah. However, there was a particularly low proportion of shared 'reference' taxa overall in the Pilbara pitfall traps (Figure 5.5), so ecosystem recovery is far from complete.

### 5.6.2 Plants associated with invertebrate samples

Generally, directional changes in community composition were less evident in the plant diversity associated with invertebrate samples (Figure 5.3, Figure 5.4). This was expected, as we hypothesized that the signal would be diluted because invertebrates can carry plant DNA from outside the study area (van der Heyde et al., 2020a). The lack of abundance or behavior data is a commonly acknowledged limitation of DNA metabarcoding (Elbrecht and Leese 2015; Elbrecht, Peinert, and Leese 2017; Fernandes et al. 2018; Lim et al. 2016). However, the clear difference in plant assay community composition between restoration sites (Figure 5.3) indicates the invertebrates are interacting with plants locally on the restoration site, rather than only passing through. Plant sequences reflected some site characteristics, generating a greater richness of Fabaceae ASVs in younger restoration sites observed to have high cover of acacia shrubs (Data from BHP, South32). While some plant DNA may originate from debris falling in to traps, there is also evidence that these are plants that were ingested or otherwise visited by invertebrates. For example, plants in the family Goodeniaceae require insect pollination (Jabaily et al. 2012; Keighery 1980); there are virtually no Goodeniaceae ASVs in the pitfall traps (PB and JF), but they are present in most sites in vane traps (PB and JF, Figure 5.2). Unfortunately, we cannot identify which invertebrates are interacting with certain plants. This would

require isolating invertebrates and extracting DNA from each species separately, for example by extracting DNA from the pollen loads (Bell et al. 2017; Pornon et al. 2016). Alternatively, DNA from flowers has also been used to identify probable pollinators (Thomsen and Sigsgaard 2019), however, these methods require species-specific sampling and therefore far more samples and greater costs. We argue that using bulk arthropod samples is a cost, time and resource efficient method that allows researchers to gain an informative snapshot of the invertebrate community and the plants they are interacting with.

Importantly, as this study was conducted in the spring/early summer, we cannot confirm whether the same patterns would exist throughout the year. Seasonality affects invertebrate communities (Santorufo, Van Gestel, and Maisto 2014; Shimazaki and Miyashita 2005), plant communities, and especially the interaction between the two (CaraDonna et al. 2017; Rico-Gray et al. 1998). A previous study conducted during autumn (April) in the Coastal Plain sites using pitfall traps also detected directional changes in invertebrate communities (Fernandes, van der Heyde et al. 2019), but no differences in plant communities generated from pitfall traps (unpublished). In the spring there is more new plant growth and flowering resulting in more invertebrates that use those resources (Clark and Dallwitz 1974; Herrera 1988). This study offers preliminary testing of consistency in restoration patterns across space, but not temporally within or between years.

## 5.7 Conclusion

We have demonstrated the use of high throughput sequencing of invertebrate samples to evaluate restoration trajectories. Defining the likely trajectory of a restored site is important as it enables the definition of success criteria, and the required time scales for restoration monitoring. We show that trajectories towards reference ecosystems were more evident in ground dwelling invertebrates in stable climax ecosystems. Despite the lack of abundance data, metabarcoding can indicate functional ecosystem recovery by showing how the invertebrates are interacting with the plant community. Understanding restoration trajectories using DNA metabarcoding will require additional research to determine the effects of seasonal variation, and consistency of patterns across multiple years and different ecosystems. It is important to remember that ecosystems are dynamic, determining whether sites have been fully restored depends heavily on the selection of appropriate reference sites to capture the natural variation in the reference ecosystem. The Bonn Challenge goal to restore 350 million $km^2$ of degraded terrestrial ecosystems by 2030 (Suding et al., 2015) means we must ensure we get the best value from the considerable financial investment required to meet ambitious global restoration targets. Testing new monitoring techniques like DNA metabarcoding and evaluating where they prove beneficial is critical to potentially incorporating them in restoration projects and improving restoration outcomes.

## 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.


Andersen, A. N., Hoffmann, B. D., Müller, W. J., & Griffiths, A. D. (2002). Using ants as bioindicators in land management: Simplifying assessment of ant community responses. *Journal of Applied Ecology*, *39*(1), 8–17. doi: 10.1046/j.1365-2664.2002.00704.x

Andersen, A. N., & Sparling, G. P. (1997). Ants as indicators of restoration success : Relationship with soil microbial biomass in the Australian seasonal tropics. *Society for Ecological Restoration*, *5*(2), 109–114.

Austin, A. D., Yeates, D. K., Cassis, G., Fletcher, M. J., La Salle, J., Lawrence, J. F., … Taylor, G. S. (2004). Insects “Down Under” - Diversity, endemism and evolution of the Australian insect fauna: Examples from select orders. *Australian Journal of Entomology*, *43*(3), 216–234. doi: 10.1111/j.1326-6756.2004.00448.x

Bell, K. L., Fowler, J., Burgess, K. S., Dobbs, E. K., Gruenewald, D., Lawley, B., … Brosi, B. J. (2017). Applying pollen DNA metabarcoding to the study of plant–pollinator interactions. *Applications in Plant Sciences*, *5*(6), 1600124. doi: 10.3732/apps.1600124

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

Benigno, S. M., Dixon, K. W., & Stevens, J. C. (2013). Increasing soil water retention with native-sourced mulch improves seedling establishment in postmine Mediterranean sandy soils. *Restoration Ecology*, *21*(5), 617–626. doi: 10.1111/j.1526-100X.2012.00926.x

Bronstein, J. L., Alarcón, R., & Geber, M. (2006). The evolution of plant-insect mutualisms. *New Phytologist*, *172*(3), 412–428. doi: 10.1111/j.1469-8137.2006.01864.x

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. doi: 10.1038/nmeth.3869

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. doi: 10.1111/ele.12740

Catterall, C. P. (2018). Fauna as passengers and drivers in vegetation restoration: A synthesis of processes and evidence. *Ecological Management and Restoration*, *19*(August), 54–62. doi: 10.1111/emr.12306

Clark, L. R., & Dallwitz, M. J. (1974). On the relative abundance of some Australian psyllidae that coexist on Eucalyptus blakelyi. *Australian Journal of Zoology*, *22*(3), 263–276. doi: 10.1071/ZO9740387

Cranmer, L., McCollin, D., & Ollerton, J. (2012). Landscape structure influences pollinator movements and directly affects plant reproductive success. *Oikos*, *121*(4), 562–568. doi: 10.1111/j.1600-0706.2011.19704.x

Cristescu, R. H., Frère, C., & Banks, P. B. (2012). A review of fauna in mine rehabilitation in Australia: Current state and future directions. *Biological Conservation*, *149*(1), 60–72. doi: 10.1016/j.biocon.2012.02.003

Cristescu, R. H., Rhodes, J., Frére, C., & Banks, P. B. (2013). Is restoring flora the same as restoring fauna? Lessons learned from koalas and mining rehabilitation. *Journal of Applied Ecology*, *50*(2), 423–431. doi: 10.1111/1365-2664.12046

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 and Restoration*, *21*(1), 4–8. doi: 10.1111/emr.12393

Cross, S. L., Tomlinson, S., Craig, M. D., Dixon, K. W., & Bateman, P. W. (2019). Overlooked and undervalued: The neglected role of fauna and a global bias in ecological restoration assessments. *Pacific Conservation Biology*, *25*(4), 331–

341. doi: 10.1071/PC18079

De Caceres, M., & Legendre, P. (2009). Associations between species and groups of sites : indices and statistical inference. *Ecology*, *90*(12), 3566–3574.

Dodd, J., & Heddle, E. M. (1989). Water relations of Banksia wood-lands. *Journal of the Royal Society of Western Australia*, *71*, 91–92.

Dunlop, J. N., Majer, J. D., Morris, C. J., & Walker, K. J. (1985). A preliminary assessment of minesite rehabilitation in the pilbara iron ore province using ant communities as ecological indicators. *Mulga Res. Cen. J.*, *8*, 25–31. doi: 10.13140/RG.2.1.1778.8240

Edgar, R. C. (2016). UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, 081257. doi: 10.1101/081257

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. doi: 10.1371/journal.pone.0130324

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

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. doi: 10.1111/rec.12976

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

Fletcher, D. (1990). Ant succession in Pilbara borrow pits. *Mulga Res. Cen. J.*, *10*, 25–31

Folgarait, P. J. (1998). Ant biodiversity and its relationship to ecosystem functioning : a review. *Biodiversity and Conservation, 7,*1221-1244

Gaston, K. J. (1991). The Magnitude of global insect species richness. *Conservation Biology*, *5*(3), 283–296. doi: 10.1111/j.1523-1739.1991.tb00140.x

Gibson, L. A., Williams, K. J., Pinder, A. M., Harwood, T. D., McKenzie, N. L., Ferrier, S., … Manion, G. (2015). Compositional patterns in terrestrial fauna and wetland flora and fauna across the Pilbara biogeographic region of Western Australia and the representativeness of its conservation reserve system. *Records of the Western Australian Museum, Supplement*, *78*(2), 515. doi: 10.18195/issn.0313-122x.78(2).2015.515-545

Greenslade, P. J. M. & Greenslade, P. (1983). Ecology of soil invertebrates. IN (ed. K. Lee) *Soils, an Austrlian viewpoint.* CSIRO, Melbourne, pp 645-649.

Government of Western Australia Department of Jobs Tourism Science and Innovation. (2019). *Western Australia Iron Ore*. Retrieved from www.jtsi.wa.gov.au.

Hall, M. (2018). Blue and yellow vane traps differ in their sampling effectiveness for wild bees in both open and wooded habitats. *Agricultural and Forest Entomology*, *20*(4), 487–495. doi: 10.1111/afe.12281

Havel, J. J. (1975). Site-vegetation mapping in the northern jarrah forest. I. Definition of site-vegetation types. *Forests Department WA*, *Bulletin N*.

Herrera, C. M. (1988). Variation in mutualisms: the spatiotemporal mosaic of a pollinator assemblage. *Biological Journal of the Linnean Society*, *35*(2), 95–125. doi: 10.1111/j.1095-8312.1988.tb00461.x

Hunter, M. D. (2001). Insect population dynamics meets ecosystem ecology: Effects of herbivory on soil nutrient dynamics. *Agricultural and Forest Entomology*, *3*(2), 77–84. doi: 10.1046/j.1461-9563.2001.00100.x

Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Research*, *17*(3), 377–386. doi: 10.1101/gr.5969107

Jabaily, R. S., Shepherd, K. A., Gustafsson, M. H. G., Sage, L. W., Kraus, S. L., Howarth, D. G., & Motley, T. J. (2012). Systematics of the Austral-Pacific family Goodeniaceae: Establishing a taxonomic and evolutionary framework. *Taxon*, *61*(2), 419–436. doi: 10.1002/tax.612012

Jha, S., & Dick, C. W. (2010). Native bees mediate long-distance pollen dispersal in a shade coffee landscape mosaic. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(31), 13760–13764. doi: 10.1073/pnas.1002490107

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

Jurado-Rivera, J. A., Vogler, A. P., Reid, C. A. M., Petitpierre, E., & Gómez-Zurita, J. (2009). DNA barcoding insect-host plant associations. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1657), 639–648. doi: 10.1098/rspb.2008.1264

Keighery, G. J. (1980). Bird pollination in South Western Australia: A checklist. *Plant Systematics and Evolution*, *135*(3–4), 171–176. doi: 10.1007/BF00983185

Knop, E., Herzog, F., & Schmid, B. (2011). Effect of connectivity between restoration meadows on invertebrates with contrasting dispersal abilities. *Restoration Ecology*, *19*(201), 151–159. doi: 10.1111/j.1526-100X.2010.00737.x

Lim, N. K. M., Tay, Y. C., Tan, J. W. T., Kwik, J. T. B., Baloğlu, B., Meier, R., & Yeo, D. C. J. (2016). Next-generation freshwater bioassessment : eDNA metabarcoding with a conserved metazoan primer reveals species-rich and communities. *Royal Society Open Science*, *3*, 160635. doi: 10.1098/rsos.160635

Magura, T., Bogyó, D., Mizser, S., Nagy, D. D., & Tóthmérész, B. (2015). Recovery of ground-dwelling assemblages during reforestation with native oak depends on the mobility and feeding habits of the species. *Forest Ecology and Management*, *339*, 117–126. doi: 10.1016/j.foreco.2014.12.015

Majaneva, M., Diserud, O. H., Eagle, S. H. C., Hajibabaei, M., & Ekrem, T. (2018). Choice of DNA extraction method affects DNA metabarcoding of unsorted invertebrate bulk samples. *Metabarcoding and Metagenomics*, *2*, 1–12. doi: 10.3897/mbmg.2.26664

Majer, J. D. (1980). The influence of ants on broadcast and naturally spread seed in Rehabilitated Bauxite Mined Areas. *Reclamation Review*, *3*, 3–9.

Majer, J. D. (1983). Ants: bioindicators of minesite rehabilitation, land-use and conservation. *Environmental Managment*, *7*(4), 375–383.

Majer, J. D. (1989). Long-term colonization of fauna in reclaimed land. In (ed. J. D. Majer). In *Animals in Primary Succession. The Role of Fauna in Reclaimed Land. Cambridge University Press, Cambridge.* (pp. 143-174.).

Majer, J. D. (2009). Animals in the restoration process — Progressing the trends. *Restoration Ecology*, *17*(4), 315–319. doi: 10.1111/j.1526-100X.2009.00528.x

Majer, J. D., Brennan, K. E. C., & Moir, M. L. (2007). Invertebrates and the restoration of a forest ecosystem: 30 years of research following bauxite mining in Western Australia. *Restoration Ecology*, *15*(SUPPL. 4), 104–115. doi: 10.1111/j.1526-100X.2007.00298.x

Majer, J. D., Gunawardene, N. R., Taylor, C. K. & Harvey, M. S. (2013). A last word. *Records of the Western Australian Museum, Supplement,* 83, 405-406.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(1), 209–220. doi: 10.1136/bmj.2.5525.1328-b

McArthur, W. M. (1991). Reference soils of South-Western Australia. In *Department of Agriculture, Perth.*

Mcdonald, T., Gann, G. D., Jonson, J., & Dixon, K. W. (2016). International standards for the practice of ecological restoration - including principles and key concepts. *Communications Standards*, (December), 427–437. doi: 10.1016/b978-0-08-034092-0.50030-2

McKenzie, N. L., van Leeuwen, S., & Pinder, A. M. (2009). Introduction to the Pilbara Biodiversity Survey, 2002–2007. *Records of the Western Australian Museum, Supplement*, *78*(1), 3. doi: 10.18195/issn.0313-122x.78(1).2009.003-089

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

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2007). Biodiversity hotspots for conservation priorities. *Biodiversity and Conservation*, *16*(4), 853–858. doi: 10.1080/21564574.1998.9650003

Ness, J. H., Bronstein, J. L., Andersen, A. N., & Holland, J. N. (2004). Ant body size predicts dispersal distance of ant-adapted seeds: Implications of small-ant invasions. *Ecology*, *85*(5), 1244–1250. doi: 10.1890/03-0364

Oksanen, A. J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., … Szoecs, E. (2018). vegan: Community Ecology Package. R package version 2.5-5. *Https://CRAN.R-Project.Org/Package=vegan*.

Palmer, M. A., Ambrose, R. F., & Poff, N. L. (1997). Ecological theory and community restoration ecology. *Restoration Ecology*, *5*(4), 291–300. doi: 10.1046/j.1526-100X.1997.00543.x

Pearson, D. L., Hamilton, A. L., & Erwin, T. L. (2011). Recovery plan for the endangered taxonomy profession. *BioScience*, *61*(1), 58–63. doi: 10.1525/bio.2011.61.1.11

Pianka, E. R. (1970). On r- and K-Selection. *The American Naturalist*, *104*(940), 592–597.

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*(February), 27282. doi: 10.1038/srep27282

R Core Team. (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna. https://www.R-project.org*.

Rico-Gray, V., García-Franco, J. G., Palacios-Rios, M., Díaz-Castelazo, C., Parra-Tabla, V., & Navarro, J. A. (1998). Geographical and seasonal variation in the richness of ant-plant interactions in Mexico. *Biotropica*, *30*(2), 190–200. doi: 10.1111/j.1744-7429.1998.tb00054.x

Rix, M. G., Edwards, D. L., Byrne, M., Harvey, M. S., Joseph, L., & Roberts, J. D. (2015). Biogeography and speciation of terrestrial fauna in the south-western Australian biodiversity hotspot. *Biological Reviews*, *90*(3), 762–793. doi: 10.1111/brv.12132

Rosenberg, D. M., Danks, H. V., & Lehmkuhl, D. M. (1986). Importance of insects in environmental impact assessment. *Environmental Management*, *10*(6), 773–783. doi: 10.1007/BF01867730

Ruiz-jaen, M. C., & Aide, T. M. (2005). Restoration Success : How Is It Being Measured ? *Restoration Ecology*, *13*(3), 569–577.

Santorufo, L., Van Gestel, C. A. M., & Maisto, G. (2014). Sampling season affects conclusions on soil arthropod community structure responses to metal pollution in Mediterranean urban soils. *Geoderma*, *226–227*(1), 47–53. doi: 10.1016/j.geoderma.2014.02.001

Shimazaki, A., & Miyashita, T. (2005). Variable dependence on detrital and grazing food webs by generalist predators: Aerial insects and web spiders. *Ecography*, *28*(4), 485–494. doi: 10.1111/j.0906-7590.2005.04105.x

Smith, M. A., Fisher, B. L., & Hebert, P. D. N. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: The ants of Madagascar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1825–1834. doi: 10.1098/rstb.2005.1714

Southwood, T. . R. . E. . (1977). Habitat, the Templet for Ecological Strategies ? *Journal of Animal Ecology*, *46*(2), 336–365.

Suding, K., Higgs, E., Palmer, M., Callicott, J. B., Anderson, C. B., Baker, M., … Schwartz, K. Z. S. (2015). Committing to ecological restoration. *Science*, *348*(6235), 638–640. doi: 10.1126/science.aaa4216

Suding, K. N., & Gross, K. L. (2006). The dynamic nature of ecological systems: Multiple states and restoration trajectories. In *Falk DA, Palmer MA, Zedler JB (eds) Foundations of restoration ecology. Island Press, Washington, D.C.* (pp. 190–209).

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., … Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3), e14–e14. doi: 10.1093/nar/gkl938

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

Tomlinson, S., Webber, B. L., Bradshaw, S. D., Dixon, K. W., & Renton, M. (2018). Incorporating biophysical ecology into high-resolution restoration targets: insect pollinator habitat suitability models. *Restoration Ecology*, *26*(2), 338–347. doi: 10.1111/rec.12561

Trudgen, M. (1977). *A report on the flora and floristic communitytypes found on parts of sand mining leases held by Rocla in the Gnangara area. Report prepared for Rocla Quarry Prod-ucts, Perth, Western Australia*.

Vamos, E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, *1*, e14625. doi: 10.3897/mbmg.1.14625

van der Heyde, M, Bunce, M., Dixon, K., Wardell-johnson, G., White, N. E., & Nevill, P. (2020b). Changes in soil microbial communities in post mine ecological restoration : Implications for monitoring using high throughput DNA sequencing. *Science of the Total Environment*, *749*, 142262. doi: 10.1016/j.scitotenv.2020.142262

van der Heyde, Mieke, Bunce, M., Wardell-Johnson, G., Fernandes, K., White, N. E., & Nevill, P. (2020a). Testing multiple substrates for terrestrial biodiversity monitoring using environmental DNA metabarcoding. *Molecular Ecology Resources*, (00), 1–14. doi: 10.1111/1755-0998.13148

Ward, D. F., New, T. R., & Yen, A. L. (2001). Effects of pitfall trap spacing on the abundance, richness and composition of invertebrate catches. *Journal of Insect Conservation*, *5*(1), 47–53. doi: 10.1023/A:1011317423622

Wilkinson, S., Davy, S., Bunce, M., & Stat, M. (2018). Taxonomic identification of environmental DNA with informatic sequence classification trees. *PeerJ Preprints*, (June), 1–2. doi: 10.7287/peerj.preprints.26812

Yeates, D. K., Harvey, M. S., & Austin, A. D. (2003). New estimates for terrestrial arthropod species-richness in Australia. *Records of the South Australian Museum Monograph Series*, *7*(Harvey 2002), 231–241.

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

## 5.9 Supplementary Information

### Table S5.9.1 Summary of ASVs

| Location | Trap | fwh ASVs | trnl ASVs |
|---|---|---|---|
| **JF** | Pitfall | 236 | 92 |
| | Vane | 181 | 94 |
| **PB** | Pitfall | 160 | 71 |
| | Vane | 162 | 90 |
| **SCP** | Pitfall | 193 | 116 |
| | Vane | 168 | 85 |

### Table S5.9.2 PERMANOVA results showing differences in community composition based on the age of the restoration

| Assay | Substrate | Location | Df | F | R2 | P |
|---|---|---|---|---|---|---|
| **fwh** | Vane | JF | 2/29 | 3.060 | 0.185 | 0.001 |
| | | SCP | 2/39 | 3.821 | 0.171 | 0.001 |
| | | PB | 2/29 | 2.146 | 0.137 | 0.001 |
| | Pitfall | JF | 2/29 | 3.522 | 0.207 | 0.001 |
| | | SCP | 2/39 | 6.423 | 0.274 | 0.001 |
| | | PB | 2/29 | 3.141 | 0.195 | 0.001 |
| **trnl** | Vane | JF | 2/29 | 5.423 | 0.287 | 0.001 |
| | | SCP | 2/39 | 5.336 | 0.224 | 0.001 |
| | | PB | 2/29 | 2.338 | 0.148 | 0.001 |
| | Pitfall | JF | 2/29 | 7.553 | 0.359 | 0.001 |
| | | SCP | 2/39 | 4.492 | 0.195 | 0.001 |
| | | PB | 2/29 | 2.203 | 0.155 | 0.006 |

**Table S5.9.3** Multipattern analysis of invertebrate assay showing the ASVs that were significantly (alpha=0.05) associated with younger restoration, older restoration, and/or reference sites

| Location-Trap | Restoration | taxa |
|---|---|---|
| **JF_Pitfall** | Reference | ASV_112: Hymenoptera |
| **JF_Pitfall** | Reference | ASV_165: Arthropoda order |
| **JF_Pitfall** | Younger | ASV_221: Hemiptera Cydnidae |
| **JF_Pitfall** | Older+Reference | ASV_240: Collembola |
| **JF_Pitfall** | Younger | ASV_292: Collembola Hypogastruridae *Ceratophysella Ceratophysella gibbosa* |
| **JF_Pitfall** | Younger+Older | ASV_34: Hymenoptera Formicidae |
| **JF_Pitfall** | Younger+Older | ASV_45: Hymenoptera Formicidae *Iridomyrmex* |
| **JF_Pitfall** | Younger | ASV_531: Hymenoptera Formicidae *Cardiocondyla* |
| **JF_Pitfall** | Younger+Older | ASV_55: Hymenoptera Formicidae *Camponotus* |
| **JF_Pitfall** | Reference | ASV_84: Hymenoptera Formicidae |
| **JF_Pitfall** | Younger | ASV_850: Collembola |
| **JF_Vane** | Younger | ASV_102: Hymenoptera Apidae |
| **JF_Vane** | Younger | ASV_123: Hymenoptera |
| **JF_Vane** | Reference | ASV_14: Coleoptera Carabidae |
| **JF_Vane** | Reference | ASV_176: Hymenoptera Halictidae |
| **JF_Vane** | Younger | ASV_188: Hymenoptera Apidae |
| **JF_Vane** | Reference | ASV_189: Diptera |
| **JF_Vane** | Younger | ASV_191: Hymenoptera Apidae |
| **JF_Vane** | Younger | ASV_197: Hymenoptera Apidae |
| **JF_Vane** | Younger | ASV_199: Hymenoptera Apidae *Apis* |
| **JF_Vane** | Younger | ASV_234: Hymenoptera Apidae |
| **JF_Vane** | Younger | ASV_241: Hymenoptera Apidae |
| **JF_Vane** | Reference | ASV_3: Diptera |
| **JF_Vane** | Younger | ASV_315: Hemiptera Aphididae |
| **JF_Vane** | Younger | ASV_374: Hymenoptera Apidae |
| **PB_Pitfall** | Reference | ASV_121: Orthoptera |
| **PB_Pitfall** | Older | ASV_124: Blattodea Blattidae |
| **PB_Pitfall** | Older+Reference | ASV_167: Orthoptera |
| **PB_Pitfall** | Younger | ASV_29: Hymenoptera Formicidae *Iridomyrmex Iridomyrmex sanguineus* |
| **PB_Pitfall** | Older | ASV_294: Hymenoptera |
| **PB_Pitfall** | Younger | ASV_38: Hymenoptera Formicidae *Melophorus* |
| **PB_Pitfall** | Younger | ASV_458: Hymenoptera |
| **PB_Pitfall** | Reference | ASV_464: Hymenoptera Formicidae |
| **PB_Pitfall** | Younger | ASV_476: Hymenoptera Formicidae *Camponotus* |

| Location-Trap | Restoration | taxa |
|---|---|---|
| **PB_Pitfall** | Younger | ASV_61: Hymenoptera Formicidae |
| **PB_Pitfall** | Older | ASV_7: Hymenoptera |
| **PB_Pitfall** | Reference | ASV_86: Hymenoptera Formicidae *Monomorium Monomorium rothsteini* |
| **PB_Pitfall** | Younger+Older | ASV_9: Hymenoptera |
| **PB_Pitfall** | Older | ASV_97: Hymenoptera Formicidae |
| **PB_Vane** | Younger | ASV_1186: Lepidoptera Noctuidae *Helicoverpa* |
| **PB_Vane** | Younger | ASV_1439: Hemiptera |
| **PB_Vane** | Older | ASV_217: Hemiptera Cixiidae |
| **PB_Vane** | Older | ASV_362: Hemiptera Membracidae *Rigula Rigula* sp. BOLD:AAG8547 |
| **PB_Vane** | Older | ASV_41: Hemiptera |
| **PB_Vane** | Younger | ASV_424: Hemiptera |
| **PB_Vane** | Younger | ASV_513: Coleoptera Chrysomelidae |
| **PB_Vane** | Younger | ASV_586: Coleoptera |
| **PB_Vane** | Older | ASV_651: Hymenoptera Halictidae *Lipotriches* |
| **SCP_Pitfall** | Younger+Older | ASV_103: Hymenoptera Formicidae |
| **SCP_Pitfall** | Younger | ASV_114: Coleoptera |
| **SCP_Pitfall** | Older+Reference | ASV_12: Arthropoda order |
| **SCP_Pitfall** | Younger | ASV_126: Hymenoptera Colletidae |
| **SCP_Pitfall** | Older+Reference | ASV_146: Coleoptera Staphylinidae |
| **SCP_Pitfall** | Younger | ASV_251: Hemiptera Aphididae *Aphis* |
| **SCP_Pitfall** | Younger | ASV_271: Hymenoptera Formicidae *Melophorus* |
| **SCP_Pitfall** | Older | ASV_30: Hymenoptera Formicidae |
| **SCP_Pitfall** | Older+Reference | ASV_309: Julida Julidae Ommatoiulus |
| **SCP_Pitfall** | Younger | ASV_349: Orthoptera Acrididae |
| **SCP_Pitfall** | Older | ASV_355: Lithobiomorpha Henicopidae |
| **SCP_Pitfall** | Older+Reference | ASV_43: Arthropoda order |
| **SCP_Pitfall** | Older | ASV_473: Hymenoptera |
| **SCP_Pitfall** | Older | ASV_49: Diptera |
| **SCP_Pitfall** | Reference | ASV_540: Diptera Sciaridae |
| **SCP_Pitfall** | Older | ASV_70: Hymenoptera |
| **SCP_Pitfall** | Reference | ASV_707: Coleoptera |
| **SCP_Pitfall** | Older+Reference | ASV_79: Arthropoda order |
| **SCP_Pitfall** | Older | ASV_981: Hymenoptera |
| **SCP_Pitfall** | Younger+Older | ASV_99: Hymenoptera Formicidae *Rhytidoponera* |
| **SCP_Vane** | Reference | ASV_1049: Diptera |
| **SCP_Vane** | Reference | ASV_1169: Arthropoda order |
| **SCP_Vane** | Older | ASV_120: Coleoptera |
| **SCP_Vane** | Older+Reference | ASV_13: Coleoptera |
| **SCP_Vane** | Older+Reference | ASV_18: Coleoptera |
| **SCP_Vane** | Younger+Older | ASV_2: Coleoptera Scarabaeidae |

| Location-Trap | Restoration | taxa |
|---|---|---|
| **SCP_Vane** | Older+Reference | ASV_20: Coleoptera |
| **SCP_Vane** | Reference | ASV_334: Arthropoda order |
| **SCP_Vane** | Younger | ASV_4: Coleoptera Carabidae |
| **SCP_Vane** | Reference | ASV_475: Coleoptera |
| **SCP_Vane** | Reference | ASV_534: Arthropoda order |
| **SCP_Vane** | Reference | ASV_556: Coleoptera |
| **SCP_Vane** | Older+Reference | ASV_63: Coleoptera |
| **SCP_Vane** | Older+Reference | ASV_76: Coleoptera |

**Table S5.9.4 Multipattern analysis of plant assay showing the ASVs that were significantly (alpha=0.05) associated with younger restoration, older restoration, and/or reference sites**

| Location-Trap | Restoration | Taxa |
|---|---|---|
| **JF_Pitfall** | Reference | ASV_195: Poales Cyperaceae Tetraria *Tetraria capillaris* |
| **JF_Pitfall** | Reference | ASV_82: Fabales Fabaceae *Bossiaea* |
| **JF_Pitfall** | Younger | ASV_13: Fabales Fabaceae |
| **JF_Pitfall** | Younger+Older | ASV_1: Myrtales Myrtaceae |
| **JF_Pitfall** | Older | ASV_19: Myrtales Myrtaceae |
| **JF_Pitfall** | Younger+Older | ASV_28: Myrtales Myrtaceae |
| **JF_Pitfall** | Reference | ASV_347: Myrtales Myrtaceae |
| **JF_Pitfall** | Older | ASV_95: Myrtales Myrtaceae |
| **JF_Pitfall** | Younger+Older | ASV_101: Plant order |
| **JF_Pitfall** | Older | ASV_18: Proteales Proteaceae Petrophile *Petrophile squamata* |
| **JF_Pitfall** | Older | ASV_214: Proteales Proteaceae Petrophile *Petrophile squamata* |
| **JF_Vane** | Younger | ASV_102: Asterales Asteraceae |
| **JF_Vane** | Older | ASV_191: Dilleniales Dilleniaceae |
| **JF_Vane** | Younger | ASV_125: Fabales Fabaceae |
| **JF_Vane** | Younger | ASV_186: Fabales Fabaceae |
| **JF_Vane** | Younger | ASV_187: Fabales Fabaceae |
| **JF_Vane** | Younger | ASV_306: Fabales Fabaceae |
| **JF_Vane** | Younger+Older | ASV_39: Fabales Fabaceae |
| **JF_Vane** | Younger+Older | ASV_5: Fabales Fabaceae |
| **JF_Vane** | Reference | ASV_114: Myrtales Myrtaceae |
| **JF_Vane** | Older | ASV_164: Myrtales Myrtaceae |
| **JF_Vane** | Older+Reference | ASV_71: Myrtales Myrtaceae |
| **JF_Vane** | Reference | ASV_72: Myrtales Myrtaceae |
| **JF_Vane** | Older | ASV_181: Plant order |
| **JF_Vane** | Older | ASV_2141: Plant order |
| **JF_Vane** | Older | ASV_224: Proteales Proteaceae Petrophile *Petrophile squamata* |
| **JF_Vane** | Older | ASV_466: Proteales Proteaceae Petrophile *Petrophile squamata* |
| **PB_Pitfall** | Younger | ASV_27: Fabales Fabaceae *Indigofera* |
| **PB_Pitfall** | Younger | ASV_14: Fabales Fabaceae |
| **PB_Pitfall** | Older | ASV_103: Plant order |
| **PB_Vane** | Older | ASV_199: Solanales Convolvulaceae Porana *Porana commixta* |
| **PB_Vane** | Younger | ASV_203: Solanales Convolvulaceae |
| **PB_Vane** | Younger+Older | ASV_105: Fabales Fabaceae *Indigofera* |
| **PB_Vane** | Reference | ASV_240: Asterales Goodeniaceae Goodenia *Goodenia microptera* |
| **SCP_Pitfall** | Reference | ASV_53: Apiales Araliaceae *Trachymene* |
| **SCP_Pitfal** | Younger | ASV_35: Asterales Asteraceae |

| Location-Trap | Restoration | Taxa |
|---|---|---|
| **SCP_Pitfall** | Younger+Reference | ASV_37: Asterales Asteraceae |
| **SCP_Pitfall** | Younger | ASV_120: Dilleniales Dilleniaceae *Dillenia* |
| **SCP_Pitfall** | Younger | ASV_157: Dilleniales Dilleniaceae *Dillenia* |
| **SCP_Pitfall** | Younger+Reference | ASV_94: Dilleniales Dilleniaceae |
| **SCP_Pitfall** | Younger | ASV_65: Commelinales Haemodoraceae Anigozanthos *Anigozanthos flavidus* |
| **SCP_Pitfall** | Younger+Reference | ASV_2: Myrtales Myrtaceae |
| **SCP_Pitfall** | Younger+Older | ASV_3: Myrtales Myrtaceae |
| **SCP_Pitfall** | Older+Reference | ASV_26: Proteales Proteaceae |
| **SCP_Pitfall** | Older | ASV_61: Proteales Proteaceae |
| **SCP_Vane** | Younger | ASV_104: Caryophyllales Aizoaceae *Lampranthus* |
| **SCP_Vane** | Younger | ASV_309: Caryophyllales Aizoaceae *Lampranthus* |
| **SCP_Vane** | Younger | ASV_185: Asterales Campanulaceae |
| **SCP_Vane** | Reference | ASV_261: Dilleniales Dilleniaceae Dillenia |
| **SCP_Vane** | Younger | ASV_73: Dilleniales Dilleniaceae |
| **SCP_Vane** | Younger | ASV_253: Fabales Fabaceae *Gompholobium* |
| **SCP_Vane** | Younger | ASV_49: Fabales Fabaceae *Gompholobium* |
| **SCP_Vane** | Younger | ASV_10: Fabales Fabaceae |
| **SCP_Vane** | Younger | ASV_205: Asterales Goodeniaceae *Dampiera* |
| **SCP_Vane** | Younger | ASV_24: Asparagales Iridaceae |
| **SCP_Vane** | Younger+Older | ASV_74: Myrtales Myrtaceae |
| **SCP_Vane** | Younger | ASV_651: Plant order |
| **SCP_Vane** | Younger+Older | ASV_31: Poales Poaceae |
| **SCP_Vane** | Younger | ASV_216: Asparagales |

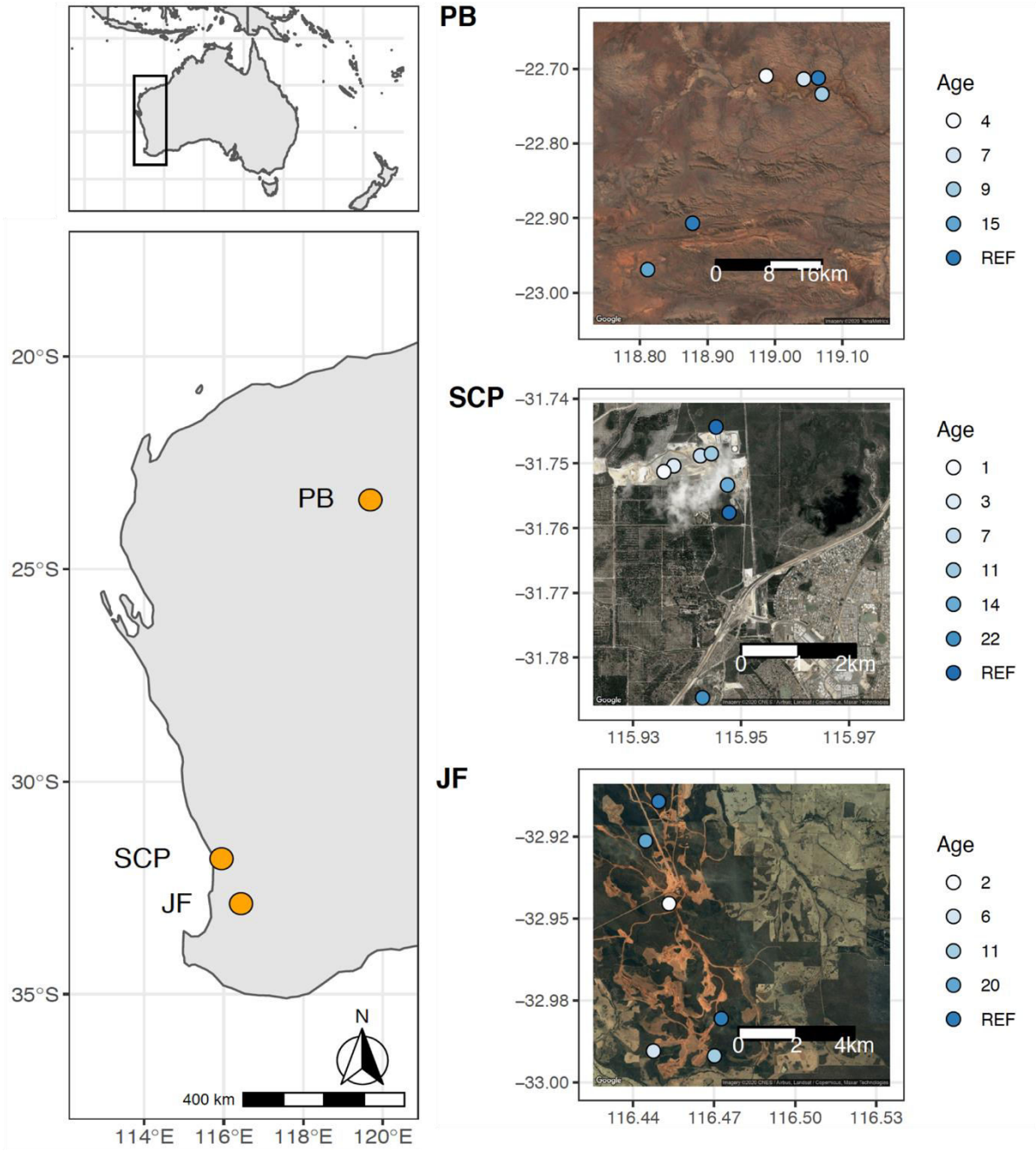


**Figure S5.9.1** Map of study sites

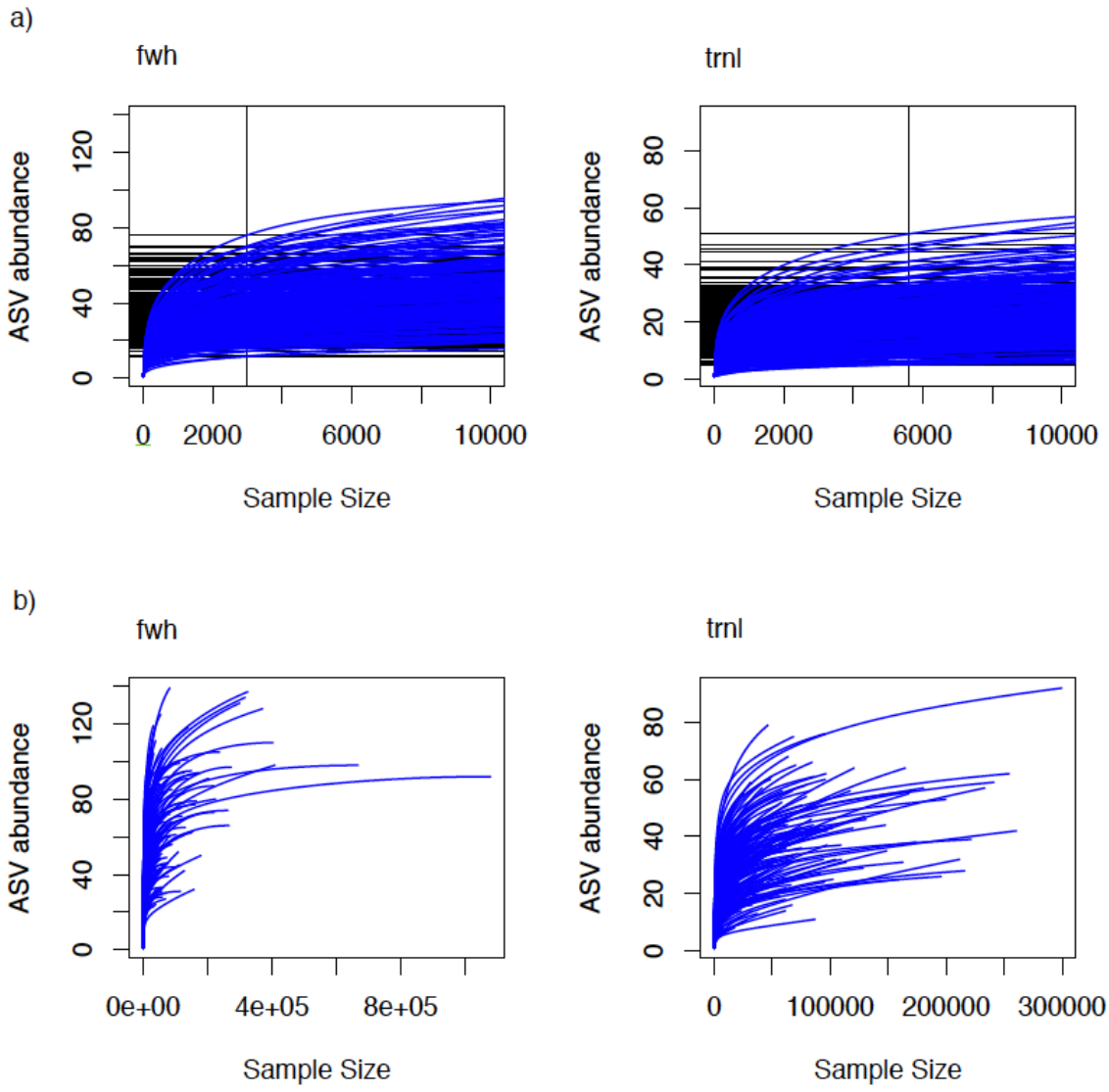


**Figure S5.9.2** Rarefaction curves for both assays showing the ASV accumulation with sample size. a) rarefaction curves up to 10, 000 reads per sample, the line indicates the minimum cut off point to keep the sample, b) total rarefaction curves for all samples.

# CHAPTER 6

## GENERAL DISCUSSION

## 6.1 Summary of findings

Terrestrial eDNA is an emerging field with great potential for biodiversity monitoring. This thesis explores various DNA metabarcoding substrates and their application to mine site restoration monitoring. This final discussion chapter synthesizes the main findings of the previous chapters and examines important considerations to develop DNA metabarcoding for terrestrial surveys. Finally, this chapter discusses future directions for research in this area that will further develop metabarcoding for terrestrial biodiversity monitoring. The overall goal of this thesis is to develop and assess a tool that can be incorporated in restoration monitoring and ultimately improve restoration outcomes. The main findings of the thesis are summarized below.

***Substrate selection critically affects eDNA metabarcoding studies***

Chapter 2 empirically tested the diversity that could be detected in various terrestrial substrates and confirmed that eDNA metabarcoding can provide a broad survey of terrestrial biodiversity. However, the choice of substrate heavily influenced the taxa that were detected, as was previously confirmed in aquatic substrates (Koziol et al., 2018). Multiple substrates detected greater biodiversity and the substrate that generated the greatest diversity was scat followed by bulk invertebrates. Vertebrate DNA was detected in scat samples only, while invertebrates and plants were detected in all substrates (bulk invertebrates, plant material, and soil). Scat samples generated the most diversity followed by bulk invertebrates. This published chapter also illustrated an important limitation of eDNA surveys; which is that DNA may originate from organisms outside of a particular study area. Substrate selection should be tailored to the purpose and limitations of the survey.

***Soil microbial communities can show directional change with restoration, but are highly affected by environmental conditions and topsoil addition***

Chapter 3 tested whether soil microbial communities (SMC) could be used to assess restoration, and if microbial changes were consistent across locations. This

chapter demonstrated that microbial communities in restoration sites *can* become more similar to reference communities with time (Gellie, Mills, Breed, & Lowe, 2017; Yan et al., 2019), but these directional changes are dependent on the location and organisms (fungi or bacteria) examined. Additionally, topsoil application during restoration can act as a confounding factor, especially with bacteria. Fungi only showed patterns of recovery in the highly mycorrhizal forest study location, indicating that ecosystem properties must be taken into account when selecting monitoring targets. The inclusion of functional analyses in this chapter improved the understanding of the changes occurring over time (Kumaresan et al., 2017). Overall, high throughput DNA sequencing for monitoring of SMC changes should be treated with caution and applied to appropriate ecosystems.

***Bulk scat samples can be used to assess bird and mammal diversity, but suitability of this substrate depends on scat detectability***

The methods used in Chapter 4 represent an attempt to generate vertebrate diversity data from environmental samples. A previous chapter (chapter 2) identified scat samples as the best source of vertebrate DNA, therefore Chapter 4 interrogated the potential of this substrate was assessed by collecting bulk scat samples from restored and reference sites across three WA locations, and sequencing these samples using a vertebrate assay. This chapter demonstrated that the bird and mammal diversity generated from these samples was capable of distinguishing between restoration and reference sites, but only in the arid/semi-arid locations where detectability was high because of the open areas and low vegetation cover (McInnes et al., 2017; Poggenburg, Nopp-Mayr, Coppes, & Sachser, 2018). This chapter illustrated some important biases to consider, such as the tendency to pick up more large mammals because larger scats are more visible, as well as detecting more herbivores because they produce more scats than carnivores (De Cuyper et al., 2020; Munn, Tomlinson, Savage, & Clauss, 2012). Bird diversity included several water birds, despite the dryness of nearby ephemeral creeks. This indicated that scat samples were likely preserved and scat samples may be capable of some past diversity in suitable environments. Fauna are often overlooked in restoration

monitoring (Cross, Tomlinson, Craig, Dixon, & Bateman, 2019); despite limitations, bulk scat collections could be used to rapidly assess bird and mammal diversity.

***Soil samples in Australia are not suitable for vertebrate monitoring using eDNA***

In both Chapter 2 and Chapter 4 (Supplementary information) soil samples were collected and tested using vertebrate assays. For chapter 2, no samples contained vertebrate DNA, but these samples were collected from the surface 5cm, with only 5 subsamples per sample, and during the summer with high temperatures and UV radiation. A second attempt was made in Chapter 4 by increasing the depth of sampling to 15 cm, collecting more subsamples per sample (8), and sampling earlier in the year in the spring. Additionally, increased effort for PCR optimization was undertaken, for example input template DNA from the extracts were increased in the PCR reactions, the annealing temperature was dropped, and fusion tagging was performed in quadruplicate instead of duplicate. Despite these attempts, the vertebrate diversity generated from soil consisted of almost entirely contamination (human sequences). Soil samples assessed in this thesis are not suited to vertebrate surveys in Australia, the heat and UV radiation in Australia likely degrades vertebrate DNA rapidly (Barnes et al., 2014; Levy-booth et al., 2007)

***Invertebrate samples have great potential for evaluating restoration trajectories***

In Chapter 5 eDNA metabarcoding was used to assess the change in ground-dwelling and airborne invertebrate communities during restoration across three WA locations. The invertebrate samples were also sequenced using a plant assay to examine the interaction between invertebrate and plant communities. The patterns of invertebrate recovery, communities becoming more similar to reference communities over time, were more evident in ground-dwelling than airborne invertebrates. Ground dwelling invertebrates have lower dispersal abilities and are more sedentary (Magura, Bogyó, Mizser, Nagy, & Tóthmérész, 2015), and therefore have greater local fidelity. On the other hand airborne invertebrates indicate the forage capabilities and attractiveness of a site. Patterns of recovery are also more evident in

ecosystems with relatively stable climax ecosystems, while in arid locations with unpredictable monsoonal flooding, recolonization appeared more stochastic (Majer, 1989). Plant sequences showed that invertebrates were interacting locally with the plant community (Pornon et al., 2016). This indicates that restoration sites are providing adequate resources (i.e. forage) to support diverse invertebrate communities, and that the invertebrates captured are using the restoration sites rather than simply passing through. Overall, eDNA metabarcoding of invertebrate samples can provide indications of restoration trajectory toward reference communities, and indications of ecosystem functions through interactions between different taxonomic groups.

***Restoration trajectories can be complex***

In this thesis, I used eDNA metabarcoding to assess restoration trajectories by examining if communities in restored sites were becoming more similar to reference communities over time. Chapters 3 (SMC) and 5 (inverterbates) illustrate how trajectories are not always in the direction of reference communities, and can be highly site and taxa specific (Suding & Gross, 2006; Wallace, Laughlin, & Clarkson, 2017). To assess ecosystem recovery, the monitoring target must be appropriate for the ecosystem in question. This is discussed further in the next section (Section 6.2)

## 6.2 Methodological considerations for the application of eDNA metabarcoding to restoration monitoring

The purpose of this thesis was to assess eDNA metabarcoding as a tool for monitoring mine site restoration. The potential of this technique has been demonstrated in multiple taxonomic groups (See section 6.1) and across multiple, diverse ecosystems. Accordingly, this thesis identified many factors that need to be considered when designing an eDNA survey. This section synthesizes the lessons learned, the limitations, and the future research needed to further develop this tool for monitoring. I use these considerations to develop a framework for eDNA study design (Figure 6.1)

### 6.2.1 Sampling design

The size of an area to be surveyed influences the appropriate survey targets and substrates because organisms operate on different spatial scales (Wiens, 1989). For example, to survey a small 20 m x 20 m plot, it may be best to survey sedentary organisms such as plants or soil microbes, rather than more mobile organisms like airborne invertebrates. Conversely, to survey a large area, perhaps hundreds of hectares, using vertebrates as environmental samplers may provide an effective survey with the fewer samples, as scats tend to detect the greatest diversity (Chapter 2). This will affect the choice of substrate and appropriate metabarcoding assay in survey design.

In additional to scale, the spatial variation in ecosystems can have implications for monitoring (Larsen, Kincaid, Jacobs, & Urquhart, 2001). Spatial autocorrelation is defined as the "*property of random variables taking values, at pairs of locations a certain distance apart, that are more similar (positive autocorrelation) or less similar (negative auto correlation) than expected for randomly associated pairs of observations*" (Legendre, 1993). It can affect the significance of classical statistical tests and should be accounted for in sample design (Legendre et al., 2002). For restoration monitoring, it can be problematic when variation that may be a result of spatial autocorrelation, is attributed to other treatments factors. For example, it can make it difficult to determine whether the recovery of invertebrate communities is a result of community recovery over time, or

because the oldest sites are closest to the reference site (Fernandes et al., 2019). A solution is to have multiple, spatially separated, reference sites to determine the magnitude of the variation. Gann et al. (2019) discuss the principles of choosing appropriate reference systems, and the need to account for natural variation (i.e. ecological mosaics). This thesis has shown that when it comes to eDNA surveys, distance between reference sites and restoration sites should also be considered because the potential for spatial autocorrelation to influence the results. This may not be feasible in many places without available intact, native ecosystems to use as references (Gann et al., 2019).

### 6.2.2 Substrate selection

This thesis examined a variety of substrates for eDNA surveys, but there are many other potential substrates that could have been used (e.g. flowers, sediment in logs, animal tracks) and more are being tested every day (Franklin et al., 2019; Schnell et al., 2015; Thomsen & Sigsgaard, 2019). There are several criteria that can help narrow down the most appropriate substrate/s for a survey. First, substrates are suited to certain environments. For example, scat collections for vertebrate diversity are appropriate in environments with less vegetative cover, because this increases their persistence and detectability (Chapter 4) (McInnes et al., 2017; Sanchez, Krausman, Livingston, & Gipson, 2004). Second, the target of the survey, which also depends on the environment, will determine which substrates can be used. This thesis has shown that the organisms that can demonstrate community recovery during restoration depend on the environmental conditions of the sites. For example, patterns of fungal recovery were only seen in a highly mycorrhizal ecosystem with high fungal biomass, and not in coastal plain or arid grassland (chapter 3). Third, validation or pilot studies are required in the development of each novel substrate. This includes the application of a substrate to a novel environment, to calibrate the substrate for the site conditions.

### 6.2.3 Assay choice and development

The choice of assay is a crucial decision for any eDNA survey. One limitation of this thesis is that it did not develop any new assays, nor can it provide a comparison or recommendation of any particular assay. Instead, I relied on

validations and tests performed by other researchers (Fahner, Shokralla, Baird, & Hajibabaei, 2016; Taberlet et al., 2007a; Vamos, Elbrecht, & Leese, 2017). This thesis focused mainly on broad assays capable of detecting large taxonomic groups, and as a result, the taxonomic resolution was to the family or genus level, with some species level assignments (Chapter 2, Chapter 4). However, there are also species specific primers that be used to detect the presence or absence of particular species. This is most often used for invasive species management (Biggs et al., 2015) or the detection of conservation priority species (Currier, Morris, Wilson, & Freeland, 2018). Primers need to be validated *in silico* (on the computer), to verify there are no taxonomic biases in the organisms that get amplified (Clarke, Soubrier, Weyrich, & Cooper, 2014; Riaz et al., 2011). Primer validation *in vitro* (in the lab) using DNA from known tissue samples ensures the assay amplifies the target DNA. For species specific probes, *in vitro* validation verifies that the probe does not amplify any closely related species (Biggs et al., 2015; Currier et al., 2018). Finally, *in situ* (in the field) validation is necessary to demonstrate that the assay will detect target DNA in a less controlled environment (Foote et al., 2012).

### 6.2.4 Abundance

Another concern in restoration monitoring is the inherent unreliability of eDNA metabarcoding to assess abundance of different organisms (Elbrecht & Leese, 2015). First, the amplification step in metabarcoding skews sequence abundances such that relative sequence abundance of a species is an unreliable predictor the abundance in the sample material (Clarke et al., 2014; Elbrecht & Leese, 2015). In the future, this may not be an issue if sequencing without PCR becomes a viable possibility (Taberlet et al., 2012). Second, the amount of DNA in the soil is affected by the biomass of the organism more than the population density (Andersen et al. 2012; Elbrecht and Leese 2015), and this would not be resolved by eliminating the dependency on PCR. In addition, other variables such as seasonal spawning (de Souza, Godwin, Renshaw, & Larson, 2016) or even different DNA shedding rates between organisms can reduce the ability to determine population abundance from DNA. While Murray et al. (Daithi C Murray et al., 2011) were able to use relative sequence abundance for dietary analysis, this does not necessarily reflect abundances of prey species. Elbrecht and Leese (2015) strongly recommend using

presence/absence data rather than relative sequence abundance. We may be able to use the number of positive samples as a coarse proxy of species abundance, but any estimations of population densities will likely be taxa dependent and require calibration. If community recovery cannot be detected from a particular substrate using presence absence data alone, the applications for restoration monitoring are limited.

### 6.2.5 Validation of eDNA methods

Developing eDNA as a monitoring tool will rely on determining how closely eDNA estimates of diversity reflect actual diversity. This thesis used morphological comparisons and benchmarking in some chapters (Chapter 2, Chapter 4), while relied on previous validation work (Ji et al., 2013; Yu et al., 2012) in one other chapter (Chapter 5). Comparing eDNA results to other monitoring methods is important to confirm taxonomic identifications and delineate the limitations (Leempoel, Hebert, & Hadly, 2020; Yoccoz et al., 2012). Primer validation should ensure there are no taxonomic biases in the DNA that is amplified, but there may be other biases introduced by the type of substrate, the environment, or if certain taxa are absent from reference databases (Chapter 2, Chapter 4). Each new combination of assay, substrate, and ecosystem should have some level of *in situ* validation. Understanding the relationship between eDNA surveys and actual diversity also requires further work on the DNA deposition and degradation rates of different organisms in different environments (Klymus, Richter, Chapman, & Paukert, 2015; Levy-booth et al., 2007). Quantifying those rates can help determine how many samples needed to detect an organism and the biomass of an organism required to be detectable (Furlan, Gleeson, Wisniewski, Yick, & Duncan, 2019).

### 6.2.6 Reference databases

While some community analyses can be accomplished without taxonomic identification (Chapter 5), for comprehensive biodiversity assessments, species level identifications are important. For example, without accurate taxonomic identification, it can be difficult to determine which species are failing to recolonize restored areas, and predict their ecological niches. Improving reference libraries is key to accurate taxonomic identification (Stoeckle, Das Mishu, & Charlop-Powers,

2020). The international barcode of life consortium (iBOL) launched a program in 2019 to generate barcode coverage for 2 million species by 2026 (iBOL, 2020), which will greatly improve our ability to generate biodiversity data from DNA, if the metabarcoding assays targets the barcode regions of iBOL. However, there are an estimated 8.7 million species on earth (1.2 million described) and species are being described and revised everyday (Mora, Tittensor, Adl, Simpson, & Worm, 2011). Additionally, many species may have significant intra-specific variation (Oberprieler, Andersen, & Moritz, 2018) such that even with representatives in the database taxonomic identification is challenging. While species coverage in genetic databases improves over time as more barcodes are uploaded, these massive databases also require curation to deal with name changes and other taxonomic revisions.

### 6.2.7 Defining appropriate thresholds

This thesis demonstrates how changes in community composition can show directional changes over time towards a reference community; however, defining thresholds of 'successful' or 'stalled' restoration is more challenging. Thresholds should be defined based on current reference ecosystems rather than historical baseline surveys that may have been conducted decades ago. This is because environmental changes in climate and other factors such as invasive species introductions can make historical ecosystems unviable in the current environment and into the future (Gann et al., 2019). The international standards for ecological restoration (Gann et al., 2019) proposed evaluating restoration based on a number of aspects that may be assessed with DNA metabarcoding (species composition, ecosystem function, absence of threats). However, these metrics need to be defined with measurable attributes (e.g. community composition, richness, % reference taxa) with specific targets to function as acceptable completion criteria (Young et al., 2019). Setting specific targets is difficult because ecosystems are dynamic by nature (Suding & Gross, 2006) with temporal variation caused by succession dynamics, seasonal, interannual, and decadal variation (Kirkman et al., 2017). Also, the spatial variation (see Section 6.2.1) in both restoration and reference communities makes it difficult to evaluate what level of an attribute would indicate a community that was effectively 'restored'. Defining thresholds for monitoring depends on understanding

both sources of variation in a particular reference ecosystem so that acceptable levels of certain attributes can be quantified (Young et al., 2019).

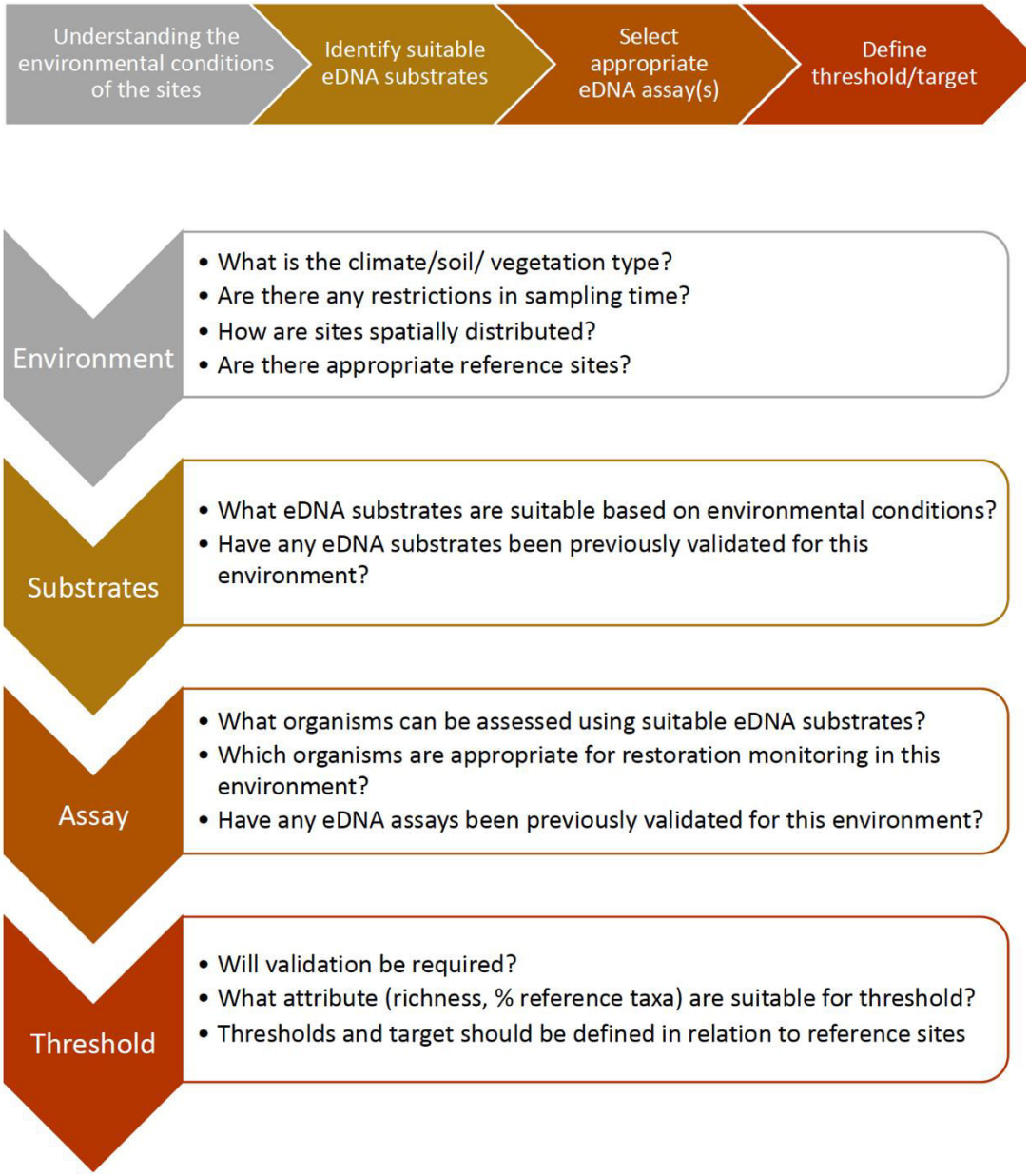


**Figure 6.1** Critical considerations for the design of environmental DNA surveys in restoration monitoring.

### 6.2.8 Using emerging monitoring technologies

For all the benefits of DNA metabarcoding, we must recognize that there are possible disadvantages to using new technology in a monitoring program. In a well

designed monitoring program, the aim would be to use consistent methods over time to examine recovery during restoration. The DNA metabarcoding field is changing rapidly as new substrates are trialed (Calvignac-spencer, Merkel, & Kutzner, 2013; Schnell et al., 2015; Thomsen & Sigsgaard, 2019), new assays are developed (Taberlet et al., 2007b; Vamos et al., 2017), and we develop a better understanding of how DNA metabarcoding and other monitoring methods complement each other (Leempoel et al., 2020; West et al., 2020). As a result, the 'best practice' for eDNA surveys may change between the initial design of a monitoring program and the implementation years later. This is not an issue unique to DNA metabarcoding, even plant survey methods that have been in use for decades are still compared and evaluated for different purposes (Rochefort, Isselin-Nondedeu, Boudreau, & Poulin, 2013). However, as a rapidly evolving field, it may be more difficult to maintain backward compatibility in DNA metabarcoding than more established methodologies. Although, one advantage of using eDNA is that should a new assay be preferred for monitoring, archived DNA samples can be re-sequenced if necessary to provide continuity with future surveys. Similarly, sequence data can be re-analyzed using new bioinformatics pipelines and searched against reference databases that will continue to be populated.

### 6.2.9 Reporting guidelines

Creating comprehensive reporting guidelines that take into account the biases that can be generated through substrate choice, sample design, primers, and bioinformatics is important for accuracy and transparency. For example, minimum number of reads to count a taxa as present in the sample differs between studies. It can be set at a particular cutoff, such as having a minimum copy number of 2 (Ji et al., 2013) or 10 (Fahner et al., 2016) or is based on relative abundance of the taxa (Leempoel et al., 2020; Vamos et al., 2017) or a combination of the two (Murray, Coghlan, & Bunce, 2015). This is one consideration out of many (Table 6.1) that may affect results. Ideally, the minimum reporting information should include all potential sources of variation that may affect the diversity detected through DNA from sample collection to bioinformatic filtering (Goldberg et al., 2016). All

sequencing data should also be made available so it can be audited if necessary, or reanalyzed using different bioinformatics pipelines.

**Table 6.1** Minimum information for reporting. Adapted from a similar table in Goldburg et al. (2016) "Critical considerations for the application of environmental DNA methods to detect aquatic species" *Methods in Ecology and Evolution.*

| Stage | Information |
|---|---|
| **Sample Collection** | Environment and potential limitations on collection: climate, rainfall, vegetation, spatial factors |
| | Contamination precautions |
| | Collection method: substrate, tools, replicates, volume/size |
| **Sample preservation** | Method, duration, temperature |
| **DNA extraction** | Methods including kit and any adjustments to kit protocol, sample volume, replication |
| | Contamination precautions and extraction controls |
| **Sequencing** | Library type, shotgun or amplicon sequencing, and any enrichment strategy |
| | Screening of DNA extracts for quality and inhibitors |
| | Amplification: primers, amplicon length, target taxa, specificity and biases, validations, DNA dilution |
| | Technical replicates and their treatment |
| | Sequencing adaptors, sample index tags |
| | Library preparation protocol and kit, |
| | Sequencing platform, read length, expected fragment size |
| **Post sequencing** | Bioinformatics: the tools used, read trimming and length filtering, quality filtering thresholds, OTU clustering or amplicon sequence variants (ASV) |
| | Taxonomic assignment method, reference database, parameters, presence of relevant taxa on reference databases |
| | Controls (positive and negative) and their interpretation |
| | Number of raw reads and final reads, sequencing depth per sample |

## 6.3 Significance of the thesis

Mining impacts disproportionately affect pristine environments (Bridge, 2004; Miranda et al., 2003), and lands managed by Traditional Owners (First Nations of Australia). To maintain a social license to operate, mining companies must demonstrate that they can meet restoration targets (Burton, Jasmine Zahedi, & White, 2012). The research in this thesis evaluates eDNA metabarcoding as a potential monitoring tool that can assess present diversity and indicate whether ecosystems are on a trajectory indicating recovery. This was the first study to systematically test multiple terrestrial substrates, and my findings validate previous research in aquatic substrates (Koziol et al., 2018) confirming that substrate selection is critical in eDNA surveys.

Within the last year there have been several publications using eDNA metabarcoding of various substrates for terrestrial biodiversity assessment in a variety of contexts (e.g. Leempoel et al., 2020; Liu, Baker, Burridge, Jordan, & Clarke, 2020; Sales et al., 2020; Valentin et al., 2020) and at times, it seems a field of infinite possibilities. However, in addition to the established limitations of eDNA metabarcoding (lack of abundance data, reference databases, etc., see Section 6.2), my research shows that restoration trajectories revealed by eDNA metabarcoding were site- and taxa-dependent. The management implications of this finding are that eDNA surveys may require site-calibration before being applied to restoration monitoring. As a result, it may be challenging to standardize terrestrial eDNA surveys for multiple ecosystems, unlike in aquatic systems. Additionally, the calibration required to prove the efficacy in an environment may increase costs, reducing the potential cost-effectiveness of eDNA for restoration monitoring. Despite the limitations, this thesis demonstrates the potential of this technique to rapidly assess complex, biodiverse systems; and evaluate restoration trajectories that illustrate ecosystem recovery towards reference conditions.

## 6.4 Looking forward

During the completion of this thesis the fields of DNA sequencing and bioinformatics have evolved rapidly and continue to do so. Here I discuss potential future directions that will no doubt advance the utility of eDNA based monitoring.

### 6.4.1 Bioinformatics

Bioinformatics describes that data filtering and analyses tools applied to sequencing data to generate biodiversity information. Differences in bioinformatic filtering can affect the resulting biodiversity data (Evans et al., 2017). For example, all sequencing platforms generate certain amount of sequencing error (Loman et al., 2012), and bioinformatics involves removing error while keeping as much of the genuine sequences as possible (Callahan et al., 2016; Edgar, 2016; Edgar & Flyvbjerg, 2015). If filtering is too stringent you risk missing organisms that could otherwise be detected, and if too relaxed, there may be false positives because of sequencing errors (Evans et al., 2017; Ficetola, Taberlet, & Coissac, 2016). This is an evolving field and researchers are continually refining bioinformatics tools to improve biodiversity assessments. Currently, there is a need to learn and use new bioinformatics tools as they develop to increase the quality of the data (Coissac, Riaz, & Puillandre, 2012; Wilkinson, Davy, Bunce, & Stat, 2018). Fortunately, old projects can benefit from innovations as the sequences can be reanalysed with new methods and improved reference databases. In the future, creating user-friendly platforms for data exploration will be important to increasing accessibility and greater adoption of DNA metabarcoding, as users will be able to see the effects of various bioinformatics parameters on the data produced (Bohmann et al., 2014; Deiner et al., 2017; Ruppert, Kline, & Rahman, 2019).

### 6.4.2 PCR-free sequencing

PCR-free sequencing could theoretically address some of the biases introduced during the amplification step. Shotgun sequencing uses similar next generation sequencing technology as metabarcoding and can be applied to environmental samples (Bista et al., 2018; Cowart, Murphy, & Cheng, 2018; Deiner et al., 2017; Ruppert et al., 2019). This process involves breaking down genome into fragments and using next generation sequencing to sequence the fragments (Clark &

Pazdernik, 2013). However, because all the DNA is being sequenced, a majority of the sequences are bacterial (Cowart et al., 2018), or part of the organisms genome without a reference for identification purposes or mapping of raw data to complete genomes. In order to detect certain taxa high sequencing depth is needed, which increases sequencing costs and the bioinformatics resources needed to process the data (Cowart et al., 2018; Ruppert et al., 2019). Further development in sequencing technologies and reduction in costs is important to making shotgun sequencing accessible for biodiversity assessment.

Some of the new sequencing technologies under development include Third Generation Sequencing (TGS) which can sequence DNA molecules without fragmenting. There are two currently available platforms for TGS: single molecule real time sequencing (SMRT) by PacBio and Oxford Nanopore technology (Logsdon, Vollger, & Eichler 2020) . The advantages of TGS include the ability to sequence longer reads than previous sequencing platforms, and increased portability and speed because TGS requires minimal preprocessing (Lu, Giordano, & Ning, 2016; Roberts, Carneiro, & Schatz, 2013, Latorre-Perez, Pascual, Porcar, & Vilanova 2020). Nanopore has recently developed MinION sequencer which is the size of a USB which can be used to sequence samples in the field (Pomerantz et al., 2018). However, this is a developing field and these technologies are also known to have high error rates (10-15%) (Logsdon, Vollger, & Eichler 2020; Lu et al., 2016; Roberts et al., 2013) that can limit their usefulness in biodiversity assessment.

### 6.4.3 Functional metagenomics

This thesis analyzed functionality by using taxonomic identification to infer function (Douglas et al., 2014; Nguyen et al., 2016). However, HTS can also enable analysis of functional loci (reviewed in Barnes and Turner, 2016). Practical applications include identifying functional gene diversity, potential inbreeding depression, and genomic functional traits associated with survival (Aalismail et al. 2019; Liang et al., 2011; Paige, 2010; Zhen et al. 2019). This approach has been applied to microbial metagenomics using shotgun sequencing (Enagbonma, Aremu, & Babalola 2019; Kumaresan et al., 2017; Liang et al., 2011; Mendoza, Sicheritz-Pontén, & Thomas Gilbert, 2015). Moving from microbial to macrobial functional genomics is challenging, as the concentration of mitochondrial DNA in

environmental samples can be very small (Turner et al., 2014). Applying functional analyses of metagenomics to bulk samples may improve the detection of genes from eukaryotes as there is a greater concentration of eukaryote DNA, but further development in this area is required.

## 6.5 Thesis conclusion

In many ways, the development of eDNA metabarcoding for terrestrial survey is as much about understanding exactly what the constraints and caveats are as it is about improving methodologies. Without a thorough understanding of both the strengths and the weaknesses of the technique, we cannot make correct inferences from the data we generate. Restoration monitoring needs effective tools to assess ecosystem recovery in order to improve restoration methods and evaluate trajectories. The overarching questions of this thesis was "Is eDNA metabarcoding a viable tool to monitor mine site restoration". This thesis demonstrates the viability of eDNA metabarcoding as a tool for restoration monitoring. It shows the influence of substrate selection on biodiversity detection, the increased diversity generated from multiple substrates, and the unsuitability for some substrates (i.e. no vertebrate DNA in soil from hot environments). It highlights that changes in soil microbial communities are not consistent between locations or taxonomic group (fungi or bacteria), but they can show patterns of recovery in certain systems. It also demonstrated the use of a novel substrate (pooled scat) to detect bird and mammal diversity, and showed the usefulness of this method was dependent on scat detectability. The wide-scale testing across different ecosystems revealed the importance of environment as a filter for suitable eDNA substrates. DNA metabarcoding enabled rapid survey of invertebrate communities which, even with poor taxonomic identification, demonstrated strong signal of community recovery over time. Lastly, this thesis demonstrated the ability of DNA metabarcoding to evaluate functional aspects of ecosystem recovery by detecting not only invertebrate communities, but also the interactions between groups of organisms such as invertebrates and plants. Whole ecosystem evaluation of restoration can improve restoration methods in the future, and demonstrate whether ecosystems are recovering. With further development, this technique will be a useful addition to the restoration monitoring 'toolkit', and other applications of terrestrial biodiversity assessment

## 6.6 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.


Aalismail, N. A., Ngugi, D. K., Díaz-Rúa, R., Alam, I., Cusack, M., & Duarte, C. M. (2019). Functional metagenomic analysis of dust-associated microbiomes above the Red Sea. *Scientific Reports*, *9*(1), 1–12. doi: 10.1038/s41598-019-50194-0

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*(1), 1–17. doi: 10.1007/s10592-015-0775-4

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science and Technology*, *48*(3), 1819–1827. doi: 10.1021/es404734p

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., … Dunn, F. (2015). Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt ( Triturus cristatus ). *Biological Conservation*, *183*, 19–28. doi: 10.1016/j.biocon.2014.11.029

Bista, I., Carvalho, G. R., Tang, M., Walsh, K., Zhou, X., Hajibabaei, M., … Creer, S. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.12888

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., … de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, *29*(6), 358–367. doi: 10.1016/j.tree.2014.04.003

Bridge, G. (2004). Contested terrain: mining and the environment. *Annual Review of Environment and Resources*, *29*(1), 205–259. doi: 10.1146/annurev.energy.28.011503.163434

Burton, M., Jasmine Zahedi, S., & White, B. (2012). Public preferences for

timeliness and quality of mine site rehabilitation. The case of bauxite mining in Western Australia. *Resources Policy*. doi: 10.1016/j.resourpol.2012.01.004

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. doi: 10.1038/nmeth.3869

Calvignac-spencer, A., Merkel, K., & Kutzner, N. (2013). *Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity*. 915–924. doi: 10.1111/mec.12183

Clark, D. P., & Pazdernik, N. J. (2013). Genomics and Systems Biology. In *Molecular Biology* (pp. e110–e117). doi: 10.1016/B978-0-12-378594-7.00043-3

Clarke, L. J., Soubrier, J., Weyrich, L. S., & Cooper, A. (2014). Environmental metabarcodes for insects: In silico PCR reveals potential for taxonomic bias. *Molecular Ecology Resources*, *14*(6). doi: 10.1111/1755-0998.12265

Coissac, E., Riaz, T., & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology*. doi: 10.1111/j.1365-294X.2012.05550.x

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*(October 2017), 148–160. doi: 10.1016/j.margen.2017.11.003

Cross, S. L., Tomlinson, S., Craig, M. D., Dixon, K. W., & Bateman, P. W. (2019). Overlooked and undervalued: The neglected role of fauna and a global bias in ecological restoration assessments. *Pacific Conservation Biology*, *25*(4), 331–341. doi: 10.1071/PC18079

Currier, C. A., Morris, T. J., Wilson, C. C., & Freeland, J. R. (2018). Validation of environmental DNA (eDNA) as a detection tool for at-risk freshwater pearly mussel species (Bivalvia: Unionidae). *Aquatic Conservation: Marine and Freshwater Ecosystems*, *28*(3), 545–558. doi: 10.1002/aqc.2869

De Cuyper, A., Meloro, C., Abraham, A. J., Müller, D. W. H., Codron, D., Janssens,

G. P. J., & Clauss, M. (2020). The uneven weight distribution between predators and prey: Comparing gut fill between terrestrial herbivores and carnivores. *Comparative Biochemistry and Physiology -Part A : Molecular and Integrative Physiology*, *243*(January), 110683. doi: 10.1016/j.cbpa.2020.110683

Deiner, K., Bik, H. M., Elvira, M., Seymour, M., Creer, S., Bista, I., … Vere, N. De. (2017). Environmental DNA metabarcoding : Transforming how we survey animal and plant communities. *Molecular Ecology*, (September), 1–24. doi: 10.1111/mec.14350

Douglas, G. M., Maffei, V. J., Zaneveld, J., Yurgel, S. N., Brown, J. R., Taylor, C. M., … Langille, M. G. I. (2014). *PICRUSt2: An improved and extensible approach for metagenome inference*. 2011–2012. doi: doi.org/10.1101/672295

Edgar, R. C. (2016). UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, 081257. doi: 10.1101/081257

Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, *31*(21), 3476–3482. doi: 10.1093/bioinformatics/btv401

Aalismail, N. A., Ngugi, D. K., Díaz-Rúa, R., Alam, I., Cusack, M., & Duarte, C. M. (2019). Functional metagenomic analysis of dust-associated microbiomes above the Red Sea. *Scientific Reports*, *9*(1), 1–12. doi: 10.1038/s41598-019-50194-0

Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., … Pfrender, M. E. (2017). Fish community assessment with eDNA metabarcoding: Effects of sampling design and bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(9), 1362–1374. doi: 10.1139/cjfas-2016-0306

Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-scale monitoring of plants through environmental DNA metabarcoding of soil: Recovery, resolution, and annotation of four DNA markers. *PLoS ONE*, *11*(6), 1–16. doi: 10.1371/journal.pone.0157505

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

Ficetola, G. F., Taberlet, P., & Coissac, E. (2016). How to limit false positives in environmental DNA and metabarcoding? *Molecular Ecology Resources*, *16*(3), 604–607. doi: 10.1111/1755-0998.12508

Foote, A. D., Thomsen, P. F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L. A., … Gilbert, M. T. P. (2012). Investigating the Potential Use of Environmental DNA (eDNA) for Genetic Monitoring of Marine Mammals. *PLoS ONE*, *7*(8), 2–7. doi: 10.1371/journal.pone.0041781

Franklin, T. W., Mckelvey, K. S., Golding, J. D., Mason, D. H., Dysthe, J. C., Pilgrim, K. L., … Ave, E. B. (2019). Using environmental DNA methods to improve winter surveys for rare carnivores : DNA from snow and improved noninvasive. *Biological Conservation*, *229*(November 2018), 50–58. doi: 10.1016/j.biocon.2018.11.006

Furlan, E. M., Gleeson, D., Wisniewski, C., Yick, J., & Duncan, R. P. (2019). eDNA surveys to detect species at very low densities: A case study of European carp eradication in Tasmania, Australia. *Journal of Applied Ecology*, *56*(11), 2505–2517. doi: 10.1111/1365-2664.13485

Gann, G. D., McDonald, T., Walder, B., Aronson, J., Nelson, C. R., Jonson, J., … Dixon, K. W. (2019). International principles and standards for the practice of ecological restoration. Second edition. *Restoration Ecology*, *27*(S1), S1–S46. doi: 10.1111/rec.13035

Gellie, N. J. C., Mills, J. G., Breed, M. F., & Lowe, A. J. (2017). Revegetation rewilds the soil bacterial microbiome of an old field. *Molecular Ecology*, *26*(11), 2895–2904. doi: 10.1111/mec.14081

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., … Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, *7*(11), 1299–1307. doi: 10.1111/2041-210X.12595

iBOL. (2020). BIOSCAN 2019-2026. Revealing species, their interactions, and dynamics. Retrieved September 22, 2020, from https://ibol.org/programs/bioscan/

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

Kirkman, L. K., Barnett, A., Williams, B. W., Hiers, J. K., Pokswinski, M., Mitchell, R. J., … Mitchell, R. J. (2017). A dynamic reference model : a framework for assessing biodiversity restoration goals in a fire-dependent ecosystem. *Ecological Applications*, *23*(7), 1574–1587.

Klymus, K. E., Richter, C. A., Chapman, D. C., & Paukert, C. (2015). Quantification of eDNA shedding rates from invasive bighead carp Hypophthalmichthys nobilis and silver carp Hypophthalmichthys molitrix. *Biological Conservation*, *183*, 77–84. doi: 10.1016/j.biocon.2014.11.020

Koziol, A., Stat, M., Simpson, T., Jarman, S., Harvey, E. S., McDonald, J., … Bunce, M. (2018). Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources*, (September), 1–11. doi: 10.1111/1755-0998.12971

Kumaresan, D., Cross, A. T., Moreira-Grez, B., Kariman, K., Nevill, P., Stevens, J., … Whiteley, A. S. (2017). Microbial Functional Capacity Is Preserved Within Engineered Soil Formulations Used in Mine Site Restoration. *Scientific Reports*, *7*(1), 1–9. doi: 10.1038/s41598-017-00650-6

Larsen, D. P., Kincaid, T. M., Jacobs, S. E., & Urquhart, N. S. (2001). Designs for evaluating local and regional scale trends. *BioScience*, *51*(12), 1069–1078. doi: 10.1641/0006-3568(2001)051[1069:DFELAR]2.0.CO;2

Latorre-Pérez, A., Pascual, J., Porcar, M., & Vilanova, C. (2020). A lab in the field: applications of real-time, in situ metagenomic sequencing. *Biology Methods and Protocols*. doi: 10.1093/biomethods/bpaa016

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*, *287*(1918). doi:

10.1098/rspb.2019.2353

Legendre, P. (1993). Spatial Autocorrelation : Trouble or New Paradigm ? *Ecological Society of America*, *74*(6), 1659–1673.

Legendre, P., Dale, M. R. T., Fortin, M. J., Gurevitch, J., Hohn, M., & Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, *25*(5), 601–615. doi: 10.1034/j.1600-0587.2002.250508.x

Levy-booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, R., Klironomos, J. N., … Dunfield, K. E. (2007). Cycling of extracellular DNA in the soil environment. *Soil Biology & Biochemistry*, *39*, 2977–2991. doi: 10.1016/j.soilbio.2007.06.020

Liang, Y., Van Nostrand, J. D., Deng, Y., He, Z., Wu, L., Zhang, X., … Zhou, J. (2011). Functional gene diversity of soil microbial communities from five oil-contaminated fields in China. *ISME Journal*, *5*(3), 403–413. doi: 10.1038/ismej.2010.142

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*, rec.13236. doi: 10.1111/rec.13236

Logsdon, G. A., Vollger, M. R., & Eichler, E. E. (2020). Long-read human genome sequencing and its applications. *Nature Reviews Genetics*, *21*(10), 597–614. doi: 10.1038/s41576-020-0236-x

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., & Pallen, M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology*, *30*(5), 434–439. doi: 10.1038/nbt.2198

Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics and Bioinformatics*, *14*(5), 265–279. doi: 10.1016/j.gpb.2016.05.004

Magura, T., Bogyó, D., Mizser, S., Nagy, D. D., & Tóthmérész, B. (2015). Recovery of ground-dwelling assemblages during reforestation with native oak depends

on the mobility and feeding habits of the species. *Forest Ecology and Management*, *339*, 117–126. doi: 10.1016/j.foreco.2014.12.015

Majer, J. D. (1989). Long-term colonization of fauna in reclaimed land. In (ed. J. D. Majer). In *Animals in Primary Succession. The Role of Fauna in Reclaimed Land. Cambridge University Press, Cambridge.* (pp. 143-174.).

McInnes, J. C., Alderman, R., Deagle, B. E., Lea, M. A., Raymond, B., & Jarman, S. N. (2017). Optimised scat collection protocols for dietary DNA metabarcoding in vertebrates. *Methods in Ecology and Evolution*, *8*(2), 192–202. doi: 10.1111/2041-210X.12677

Mendoza, M. L. Z., Sicheritz-Pontén, T., & Thomas Gilbert, M. P. (2015). Environmental genes and genomes: Understanding the differences and challenges in the approaches and software for their analyses. *Briefings in Bioinformatics*, *16*(5), 745–758. doi: 10.1093/bib/bbv001

Miranda, M., Burris, P., Froy Bincang, J., Shearman, P., Oliver Briones, J., La Viña, A., & Menard, S. (2003). *Mining and Critical Ecosystems: Mapping the Risks*.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How many species are there on earth and in the ocean? *PLoS Biology*, *9*(8), 1–8. doi: 10.1371/journal.pbio.1001127

Munn, A. J., Tomlinson, S., Savage, T., & Clauss, M. (2012). Retention of different-sized particles and derived gut fill estimate in tammar wallabies (Macropus eugenii): Physiological and methodological considerations. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, *161*(2), 243–249. doi: 10.1016/j.cbpa.2011.11.003

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

Nguyen, N., Song, Z., Bates, S., Branco, S., Tedersoo, L., Menke, J., … Kennedy, P. (2016). FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology*, *20*, 241–248.

Oberprieler, S. K., Andersen, A. N., & Moritz, C. C. (2018). Ants in Australia's

monsoonal tropics: CO1 barcoding reveals extensive unrecognised diversity. *Diversity*, *10*(2), 1–23. doi: 10.3390/D10020036

Paige, K. N. (2010). The Functional Genomics of Inbreeding Depression: A New Approach to an Old Problem. *BioScience*, *60*(4), 267–277. doi: 10.1525/bio.2010.60.4.5

Poggenburg, C., Nopp-Mayr, U., Coppes, J., & Sachser, F. (2018). Shit happens … and persists: decay dynamics of capercaillie (Tetrao urogallus L.) droppings under natural and artificial conditions. *European Journal of Wildlife Research*, *64*(3). doi: 10.1007/s10344-018-1187-9

Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., … Prost, S. (2018). Real-time DNA barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building. *GigaScience*, *7*(4), 1–14. doi: 10.1093/gigascience/giy033

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*(February), 27282. doi: 10.1038/srep27282

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), 1–11. doi: 10.1093/nar/gkr732

Roberts, R. J., Carneiro, M. O., & Schatz, M. C. (2013). The advantages of SMRT sequencing. *Genome Biology*, *14*(6), 2–5. doi: 10.1186/gb-2013-14-6-405

Rochefort, L., Isselin-Nondedeu, F., Boudreau, S., & Poulin, M. (2013). Comparing survey methods for monitoring vegetation change through time in a restored peatland. *Wetlands Ecology and Management*, *21*(1), 71–85. doi: 10.1007/s11273-012-9280-4

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. doi: 10.1016/j.gecco.2019.e00547

Sales, N. G., Kaizer, M. da C., Coscia, I., Perkins, J. C., Highlands, A., Boubli, J. P., … Mcdevitt, A. D. (2020). Assessing the potential of environmental DNA metabarcoding for monitoring Neotropical mammals: a case study in the Amazon and Atlantic Forest, Brazil. *Mammal Review*, *50*(3), 221–225. doi: 10.1111/mam.12183

Sanchez, D. M., Krausman, P. R., Livingston, T. R., & Gipson, P. S. (2004). Persistence of carnivore scat in the Sonoran Desert. *Wildlife Society Bulletin*, *32*(2), 366–372. doi: 10.2193/0091-7648(2004)32[366:pocsit]2.0.co;2

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*(1), 1–14. doi: 10.1186/s12983-015-0115-z

Stoeckle, M. Y., Das Mishu, M., & Charlop-Powers, Z. (2020). Improved Environmental DNA Reference Library Detects Overlooked Marine Fishes in New Jersey, United States. *Frontiers in Marine Science*, *7*(May). doi: 10.3389/fmars.2020.00226

Suding, K. N., & Gross, K. L. (2006). The dynamic nature of ecological systems: Multiple states and restoration trajectories. In *Falk DA, Palmer MA, Zedler JB (eds) Foundations of restoration ecology. Island Press, Washington, D.C.* (pp. 190–209).

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., … Willerslev, E. (2007a). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3), e14–e14. doi: 10.1093/nar/gkl938

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., … Willerslev, E. (2007b). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3). doi: 10.1093/nar/gkl938

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

Turner, C. R., Barnes, M. A., Xu, C. C. Y., Jones, S. E., Jerde, C. L., & Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution*, *5*(7), 676–684. doi: 10.1111/2041-210X.12206

Valentin, R. E., Fonseca, D. M., Gable, S., Kyle, K. E., Hamilton, G. C., Nielsen, A. L., & Lockwood, J. L. (2020). Moving eDNA surveys onto land: Strategies for active eDNA aggregation to detect invasive forest insects. *Molecular Ecology Resources*, *20*(3), 746–755. doi: 10.1111/1755-0998.13151

Vamos, E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, *1*, e14625. doi: 10.3897/mbmg.1.14625

Wallace, K. J., Laughlin, D. C., & Clarkson, B. D. (2017). Exotic weeds and fluctuating microclimate can constrain native plant regeneration in urban forest restoration. *Ecological Applications*, *27*(4), 1268–1279. doi: 10.1002/eap.1520

West, K. M., Stat, M., Harvey, E. S., Skepper, C. L., DiBattista, J. D., Richards, Z. T., … Bunce, M. (2020). eDNA metabarcoding survey reveals fine-scale coral reef community variation across a remote, tropical island ecosystem. *Molecular Ecology*, *29*(6), 1069–1086. doi: 10.1111/mec.15382

Wiens, J. A. (1989). Spatial Scaling in Ecology Published by : British Ecological Society Stable URL : https://www.jstor.org/stable/2389612 Spatial scaling in ecology1. *Functional Ecology*, *3*(4), 385–397.

Wilkinson, S., Davy, S., Bunce, M., & Stat, M. (2018). Taxonomic identification of environmental DNA with informatic sequence classification trees. *PeerJ Preprints*, (June), 1–2. doi: 10.7287/peerj.preprints.26812

Yan, D., Bissett, A., Gellie, N., Mills, J. G., Lowe, A. J., & Breed, M. F. (2019). Soil bacterial community differences along a coastal restoration chronosequence. *Plant Ecology*, *221*(9), 795–811. doi: 10.1007/s11258-019-00979-0

Yoccoz, N. G., Brathen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., …

Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, *21*, 3647–3655. doi: 10.1111/j.1365-294X.2012.05545.x

Young, R., Manero, A., Miller, B., Kragt, M. E., Standish, R. J., Jasper, D., & Boggs, G. (2019). *A framework for developing mine-site completion criteria in Western Australia.* Retrieved from https://research-repository.uwa.edu.au/en/publications/a-framework-for-developing-minesite-completion-criteria-in-western-australia(a578f8b3-4b21-479e-92f2-c628a54a0de9).html

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

Aalismail, N. A., Ngugi, D. K., Díaz-Rúa, R., Alam, I., Cusack, M., & Duarte, C. M. (2019). Functional metagenomic analysis of dust-associated microbiomes above the Red Sea. *Scientific Reports*, *9*(1), 1–12. doi: 10.1038/s41598-019-50194-0

# Appendix I. Copyright statements

Chapter 2: Published in the journal Molecular Ecology Resources. As an author, permission is automatically granted to reproduce this copyright material to include the article in my thesis.

To Whom It May Concern, I, Mieke van der Heyde, contributed to the conceptualization of the study, collected and processed the samples, analyzed the data, and wrote and edited the resulting manuscript:

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 Ecolology Resources* 1–14. doi: 10.1111/1755-0998.13148

---

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Michael Bunce _____

Grant Wardell-Johnson _____

Kristen Fernandes _____

Nicole E. White _____

Paul Nevill _____

Chapter 3: Published in the journal Science of the Total Environment. As an author, permission is automatically granted to reproduce this copyright material to include the article in my thesis.

To Whom It May Concern, I, Mieke van der Heyde, contributed to the conceptualisation of the study, collected and processed the samples, analysed the data, and wrote and edited the resulting manuscript:

van der Heyde M., Bunce M., Dixon, K., Wardell-Johnson G., White, N. E., Nevill, P. (2020) Changes in soil microbial communities in post mine ecological restoration: Implications for monitoring using high throughput DNA sequencing. *Science of the Total Environment* https://doi.org/10.1016/j.scitotenv.2020.142262

---

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Michael Bunce _____

Kingsley Dixon _____

Grant Wardell-Johnson _____

Nicole E. White _____

Paul Nevill _____

Chapter 4: Under review at Biodiversity and Conservation. As an author, permission is automatically granted to reproduce this copyright material to include the article in my thesis.

To Whom It May Concern, I, Mieke van der Heyde, contributed to the conceptualisation of the study, collected and processed the samples, analysed the data, and wrote and edited the resulting manuscript:

van der Heyde M., Bateman, P.W., Bunce M., Wardell-Johnson G., White, N. E., Nevill, P. (2020) Scat DNA provides important data for effective monitoring of bird and mammal diversity. *Biodiversity and Conservation*

_____

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Michael Bunce _____

Bill Bateman _____

Grant Wardell-Johnson _____

Nicole E. White _____

Paul Nevill _____

Chapter 5: In preparation for publication

To Whom It May Concern, I, Mieke van der Heyde, contributed to the conceptualisation of the study, collected and processed the samples, analysed the data, and wrote and edited the resulting manuscript:

van der Heyde M., Bunce M., Dixon, K., Wardell-Johnson G., White, N. E., Nevill, P. (2020) Evaluating restoration trajectories using DNA metabarcoding of invertebrates and their associated plant communities. *In preparation*

_____

I, as a co-author, endorse that this level of contribution by the candidate indicated above is appropriate.

Michael Bunce _____

Kingsley Dixon _____

Grant Wardell-Johnson _____

Nicole E. White _____

Paul Nevill _____

Jonathan Majer _____

https://doi.org/10.1111/1755-0998.13148

https://doi.org/10.1016/j.scitotenv.2020.142262

# Appendix II: Additional Publications

There were two additional articles published during PhD that were not included in the thesis. As an author, permission is automatically granted to reproduce this copyright material to include the article in my thesis. These articles are listed below.

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*, Vol. 26, pp. 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*, 1–10. doi: 10.1111/rec.12976

https://doi.org/10.1111/rec.12868

https://doi.org/10.1111/rec.12976