Invertebrate DNA metabarcoding reveals changes in communities across mine site restoration chronosequences

Kristen Fernandes[1,2*], Mieke van der Heyde[1,2], Megan Coghlan[2], Grant Wardell-Johnson[1], Michael Bunce[2], Richard Harris[3], and Paul Nevill[1]

[1]ARC Centre for Mine Site Restoration, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia.

[2]Trace and Environmental DNA Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia.

[3]School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia

*Corresponding Author: Kristen Fernandes, Trace and Environmental DNA Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA 6102, Australia.

Email: Kristen.Fernandes@student.curtin.edu.au

Ph: +61 4 478 801 373

Authors' contributions

KF conducted the study and wrote the paper with the support of; PN, RH, MB with experimental design; PN, RH, MVH with sample collection; GWJ with flora surveys; MC, MB with metabarcoding workflows, including assay design; MVH with statistical analysis. All authors contributed to editing the paper.

Abstract

Invertebrate biomonitoring can reveal crucial information about the status of restoration projects; however, it is routinely underused because of the high level of taxonomic expertise and resources required. Invertebrate DNA metabarcoding has been used to characterise invertebrate biodiversity but its application in restoration remains untested. We use DNA metabarcoding, a new approach for restoration assessment, to explore the invertebrate composition from pitfall traps at two mine site restoration chronosequences in south-western Australia. Invertebrates were profiled using two COI assays to investigate invertebrate biodiversity. The data revealed differences between invertebrate communities at the two mines and between the different age plots of the chronosequences. Several characteristic taxa were identified for each age within the chronosequence, including springtails within the youngest sites (Order: Collembola) and millipedes within the oldest and reference sites (Order: Julida). This study facilitates development of a molecular 'toolkit' for the monitoring of ecological restoration projects. We suggest that a metabarcoding approach shows promise in complementing current monitoring practices that rely on alpha taxonomy.

Implications for Practice

- Invertebrate monitoring is underused in restoration monitoring because it is resource
  intensive and requires taxonomic expertise.

- The DNA metabarcoding approach has potential to add considerable value to fauna
  monitoring programs.

- Using the invertebrate DNA metabarcoding technique it was possible to identify differences
  in invertebrate communities across restoration chronosequences.

- With further research and development of this technique, DNA metabarcoding has the
  potential to become a key component of the 'toolkit' for restoration monitoring.

Introduction

Ecosystem restoration is driven by a desire to address continued global declines in land quality,
biodiversity and ecosystem services (Bullock et al. 2011). The ultimate goals of restoration are to
return a pre-disturbance level of condition and function to degraded systems and to enable the
restored environment to persist without further intervention from land managers (Hobbs & Suding
2009). However, these goals require an understanding of the complex ecological processes
underpinning the environment, as well as the role of biotic and abiotic factors in those processes.
Adequate baseline monitoring is therefore crucial to determine ecosystem functioning and
composition prior to planned environmental degradation (Pereira et al. 2013). In a restoration
context, sites also require consistent monitoring to improve restoration practices through adaptive

management interventions (Palmer & Filoso 2009; Wortley et al. 2013). Furthermore, consistent routine monitoring allows regulators and stakeholders to determine whether restoration goals are being achieved (McDonald et al. 2016; Miller et al. 2017). Monitoring is therefore critical to 'best-practice' ecological restoration.

Historically, monitoring in restoration has focussed on plant species (Young 2000; Ruiz-Jaen & Mitchell Aide 2005). However, plant community monitoring alone cannot ascertain the return of essential ecosystem functions (e.g. nutrient cycling) (Herrick et al. 2006). Invertebrates, for example, are essential to overall ecological function (Rosenberg et al. 1986) and are sensitive to disturbance (McGeoch 1998). Although invertebrates have been used for biomonitoring in aquatic environments, (e.g. Leese et al. 2018) they are infrequently targeted in the assessment of long-term disturbance responses in terrestrial environments (Andersen et al. 2004; Andersen & Majer 2004).

There are numerous factors underpinning the lack of long-term invertebrate monitoring in terrestrial environments. Morphological comparison has been the method of choice, but is acknowledged to have several limitations. These include the difficulty of measuring species richness and turnover in landscapes with hyper-diverse and cryptic taxa (Smith et al. 2005). Most studies that use invertebrate biodiversity rely on the identification of indicator species that can provide assessment of ecosystem biodiversity dynamics (e.g. Smith et al. 2016; Tizado and Núñez-Pérez 2016). Whilst useful, focal species cannot reflect whole invertebrate diversity and assemblages (Siddig et al. 2016; Silva et al. 2017). From a management perspective, applying a broader approach to invertebrate diversity assessment is imperative to improve monitoring methodologies to achieve best-practice in environmental restoration (Lei et al. 2016).

Invertebrate DNA metabarcoding is a recent approach to monitoring that can effectively characterise invertebrate communities at a faster and cheaper rate than traditional morphological

taxonomic identification (Ji et al. 2013; Yu et al. 2012; Beng et al. 2016). DNA metabarcoding workflows consist of simultaneously amplifying standardised DNA fragments from the total DNA extracted from an environmental sample (Taberlet et al. 2012). These fragments are then sequenced, in parallel, using next-generation sequencing (NGS) before being compared to a reference database of previously generated DNA barcodes for taxonomic assignment. Metabarcoding data has the potential to profile entire invertebrate communities, although, there is possible taxonomic bias based on the gene region examined. This method is relatively fast, inexpensive (Yu et al. 2012) and becoming increasingly portable, so it has potential in biodiversity assessment of terrestrial environments for restoration monitoring (Fernandes et al. 2018; Williams, Nevill, & Krauss 2014).

Here we examine a new tool for restoration monitoring by applying DNA metabarcoding to bulk invertebrate collections from pitfall traps collected across restoration chronosequences. We used high throughput sequencing of invertebrate COI genes to evaluate the effectiveness of DNA metabarcoding methodologies to monitor invertebrate biodiversity in a mine site restoration context. While metabarcoding approaches to restoration monitoring have been recently explored using soil bacterial (16S rRNA) (Gellie et al. 2017) and fungal assays (Yan et al. 2018) there is no study exploring the implementation of these workflows on metazoan biota. This study uses a metabarcoding methodology to determine what invertebrate biodiversity we can reveal from restored mine sites, whether we can detect community composition changes, and identify taxa that characterise different stages of ecological restoration.

Materials and methods

Site description and study design

We sampled two Hanson Construction Materials sand quarries in south-western Australia, located 50 km apart at Gnangara (31.78°S, 115.95°E) and Gingin (31.40°S, 115.93°E) (Fig. 1). The sand

quarries are within the Swan Coastal Plain (SCP) Bioregion, an area that has been the subject of extensive restoration research in response to disturbance (e.g. Benigno, Dixon, & Stevens 2013; Rokich & Dixon 2007; Ritchie et al. 2017). The SCP is characterised by various dune systems and soil types from marine sands to alluvial sediments (Bettenay et al. 1960) and the study mine sites are found on siliceous sands of the Bassendean Dune System. Vegetation at the reference plots is characterised as Banksia woodland. There are three dominant tree species; *Banksia attenuata, B. menziesii*, and *Eucalyptus todtiana*; a diverse midstory featuring a shrub species and an herbaceous understory (Dodd et al. 1984). Native vegetation rehabilitation has been occurring at the Gnangara site since 1991 and at the Gingin site since 2009 (Maher 2009; RPS 2012), with the aim of rehabilitating the mined areas to the surrounding Banksia woodland. All restoration was done by Hanson and previous mine owners. Similar approaches were used at both study locations with the spreading of freshly stripped topsoil and use of local seed mixes. Plots were chosen along a restoration chronosequence spanning 11 years for the Gnangara site (2004, 2007, 2011 and 2015) and 6 years for the Gingin site (2009, 2011, 2013 and 2015) (Fig. 1). At both mine sites, a remnant vegetation plot adjacent to the restored plots was chosen as a reference community. A single reference location was selected at both mine sites due to the difficulty of access. These restored sites were established to address requirements of the mining company for rehabilitation post-mining rather than for our current objectives. As a result, there are limitations in the experimental design including a lack of site replication, and the location of plots generally in a chronosequential line. Despite this, we suggest that our conclusions provide meaningful insight into the return of invertebrate communities following restoration.

Sample collection and processing

Samples were collected in April 2017. Five pitfall traps were set within three 10 m x 10 m quadrats per restoration age and left for 7 days. Sampling methodology followed that used by Ward, New, and Yen (2001). Five pitfall traps were pooled to form one sample, giving three samples per plot with

30 in total across the two sites. Pitfall traps consisted of 120 mL sampling vials of 4 cm diameter

filled with ~40 mL of 60% ethylene glycol. Samples were washed in ethanol and ultrapure water

before being stored in 70% ethanol at -18˚C. Environmental factors such as vegetation height, slope,

and bare ground cover were recorded at each vegetation quadrat (Table S3).

Preliminary tests indicated that the ZBJ-Art primers (Zeale et al. 2011) did not amplify sequences

from the order Hymenoptera, so samples of mixed invertebrates were sorted and ants (Family:

Formicidae) separated so that an ant specific assay could be developed. The ant primers were

designed using Primer 3 (v2.3.7 in Geneious v. R10.2.3) and an alignment of 1,020 COI sequences

downloaded from the National Centre for Biotechnology Information (NCBI) GenBank and 81

Western Australian Ant species (Majer et al. 2018). In silico analysis was performed in Geneious v.

R10.2.3 to confirm the new assay would amplify ant sequences. The assay was then optimised for

PCR on positive and negative controls, and the size of the amplicon generated was confirmed by

electrophoresis and visualisation under a UV light using a 2% agarose gel stained with GelRed (Fisher

Biotec, Australia). Ant samples from our study were halved and representative species selected and

morphologically identified. These samples were sequenced using the developed COI ant assay, with

the same methodology as the rest of the samples. These sequences, the 1,020 COI sequences

downloaded from GenBank and sequences from 81 Western Australian Ant species were then used

as an ant reference database for the study. The other half of the ant samples were then put through

the metabarcoding workflow.

Before the samples of mixed invertebrates were homogenised, invertebrates were size selected to

combat potential bias of large organism size by using two legs of all individuals with a body size

equal or larger than the size of a honeybee, and whole specimens of any organisms smaller (Ji et al.

2013; Meier et al. 2016). Samples were then homogenised using a hand-held homogeniser (Omni TH

Homogeniser Drill). Individual reference ant samples, after being identified, were homogenised

differently because of their small size, using a PreCellLys24 2.8mm Ceramic Bead Kit and a Minilys Personal Homogeniser for 1 minute at 5000 rpm (Bertin Instruments [Bertin], France). Homogenized samples were frozen and stored at -18˚C until extraction.

DNA Metabarcoding

Up to 290 mg of each homogenate was used for extraction, and the remaining homogenate was stored at -18°C. The subsampled material was digested overnight in ATL buffer and proteinase K (Qiagen, Netherlands) at 56°C. DNA extraction was done using a Qiagen QIAamp DNA Mini kit modified with a starting volume of 400 μL of digest fluid and a 100 μL elution on an automated Qiacube (Qiagen, Netherlands). DNA extraction controls were carried out for every set of extractions.

Extracts were diluted (1/10 and 1/100) to assess the amplification efficiency and presence of PCR inhibitors by running quantitative PCR (qPCR) on the undiluted and diluted DNA extracts. The qPCR assays were run using two invertebrate mitochondrial cytochrome oxidase subunit 1 (COI) primer sets - one specifically looking at ant taxa (Ant236/361 from this paper) and the other a more general invertebrate assay (ZBJ-ArtF1c/ZBJ-ArtR2c from Zeale et al. (2011)) (Supplementary Table 1), and all samples were amplified with both assays. PCR reactions were carried out in 25 μL with final concentrations of the following reagents: 1 x PCR Gold buffer(Applied Biosystems Inc. [ABI], USA), 2 mmol/L MgCl2 (ABI), 0.4 mg/ml Bovine serum albumin ([BSA] Fisher Biotec, Australia), 0.25 mmol/L dNTPs (Astral Scientific, Australia), 0.4 μmol/L of each forward and reverse primer (Integrated DNA Technologies [IDT], Australia), 0.6 μL of 1/10000 SYBR Green (Invitrogen, USA), 1 U of AmpliTaq Gold (ABI), 2 μL of template DNA, and made to volume with ultrapure water. All qPCR reactions were run on a Step-One Plus qPCR thermocycler (ABI) under the following conditions: initial denaturation at 95˚C for 5 min, followed by 50 cycles of 95˚C for 30s, 50˚C or 54˚C for 30s, the ant and the

invertebrate assay, respectively (Supplementary Table 1), 72˚C for 30s, and a final extension of 72˚C for 10 min. Extraction controls and non-template PCR controls were also included in the qPCR reaction. DNA dilutions with the highest relative proportion of starting template DNA that showed uninhibited amplification (determined using the qPCR cycle threshold (CT) values between 15 and 38) were selected for metabarcoding using assay-specific fusion tagged primers.

Fusion tagged primers are gene specific primers that incorporate multiplex identifier (MID) tags of six-eight bases and appropriate Illumina clustering and sequencing adaptors. Unique combinations of MID tags were assigned to individual DNA extracts to allow for the pooled sequences to be assigned to samples post sequencing. No primer-MID combinations had been previously used for any samples processed within the same laboratory to avoid cross contamination between runs and samples. Fusion PCR reactions were done in duplicate on DNA extracts using the same thermocycling conditions as the standard qPCR reactions. Duplicates were pooled if they amplified successfully on the fusion-tagged qPCR. Samples were then pooled into approximately equimolar concentrations to produce a DNA library of all extracts for sequencing. The library was size-selected using a Pippin Prep 2% agarose with ethidium bromide cassette (Sage Sciences, USA) for fragments between 160-450 bp, and then purified using a QIAquick PCR purification kit (Qiagen) as per the manufacturer's instructions with the addition of a 5 min incubation prior to elution at room temperature. The purified library was eluted in 40 µL of EB buffer and quantified on a QuBit (Invitrogen, CA, USA) using dsDNA high sensitivity reagents to determine the optimal volume of library required for sequencing. The amplicon library was sequenced in a single direction on an Illumina MiSeq (Illumina, USA) using a 300 cycle V2 reagent kit with a V2 Nano flow cell, as per manufacturer's protocol.

Data filtering and Operational Taxonomic Unit Analysis

Sequences were assigned to respective samples based on MID tags using Geneious v. R10.2.3. Only amplicons that were a 100% match to the MID, gene specific primer, and sequencing adaptor regions were kept for analysis. Adaptor and primer regions were removed, and remaining amplicon sequences were filtered using USEARCH (Edgar et al. 2011) v8's fastq filter with a maximum error of 0.5. We used R v 3.4. (R Core Team 2014) to run a Pearson correlation between sequence number and the numbers of OTUs found per sample to determine if it was necessary to subsample sequences to standardize sequence depth among samples. Sequences within each sample were randomly subsampled to the lowest number of sequences found for each respective assay using USEARCH (Edgar et al. 2011) v8's fastq subsample function. Sequences were then grouped into Operational Taxonomic Units (OTUs) using a 97% similarity threshold point (Edgar 2013, Ji et al. 2013; Yu et al. 2012). Chimeras and low abundance OTU clusters (<5 sequences) were removed using USEARCH v8's quality control filters (Edgar et al. 2011).

OTUs were compared to the GenBank nucleotide reference sequence database (Clark et al. 2015) and the custom reference database created from the reference ant specimens and Australian ant species sequences using the Basic Local Alignment Search Tool (BLASTn) with the default parameters (Altschul et al. 1990). BLAST output files for invertebrate OTUs were then analysed using MEtaGenome ANalyzer (MEGAN) (Huson et al. 2011) and visualised using the NCBI taxonomic framework with the lowest common ancestry (LCA) parameters: reporting all reads, a minimum bit score 50.0, and reports limited to the top 10% of matches. Assignment of sequences to taxa was only considered where a sequence match was made along 100% of the length of the query, with species level matches only considered at a 98-100% match to the database. The Atlas of Living Australia (www.ala.org.au) was considered in interpreting matches made by the BLAST search to determine the probability that the identification made was accurate to known ranges and habitat types of

assigned species. Once OTUs were assigned to taxa, both assays were combined and statistical

analysis was conducted on the combined dataset.

Statistical Analysis

Differences in OTU richness between age plots was tested separately for each site using a one way

analysis of variance (ANOVA) in R v 3.4 (R Core Team 2014).

We created a Bray-Curtis dissimilarity matrix using the OTU presence/absence data and visualized it

using non-metric multidimensional scaling (NMDS) from the R package 'vegan' (Oksanen et al 2016).

Environmental variables collected from the plots (bare ground cover, maximum height of the

vegetation and slope) were tested for significant correlations with the NMDS using the envfit

function of the 'vegan' package (Oksanen et al 2016, permutations=9999). To determine if the

invertebrate communities i) differed between sites and ii) differed between age plots within sites,

we ran a permutational multivariate analysis of variance (PERMANOVA, permutations = 9999) using

PRIMER v.7 (Clarke & Gorley 2015) with age of plot nested within site. The dispersion between the

centroids were also calculated with PRIMER v.7 (Clarke & Gorley 2015).

In R, we ran a regression analysis based on Bray-Curtis similarity of restoration plots to reference

plots to evaluate whether invertebrate communities became more similar to reference communities

over time. . Mantel tests were also conducted using the mantel.rtest function of the 'ade4' package

(Dray and Dufour 2007, permutations=9999) to determine if the Bray-Curtis similarity between plots

was spatially autocorrelated. Finally, indicator species analysis was done using the R package

'indicspecies' (De Caceras and Legendre 2009) implemented with 9999 permutations.

Results

The two DNA metabarcoding assays yielded 1,846,984 quality filtered sequences across the two assays with an average of 10,261 (SD: ± 5,843) reads per sample. We subsampled to the lowest amount of reads: 3,328 sequences for the ant assay and 1,439 sequences for the invertebrate assay. Quality-filtered and demultiplexed sequence data is available from the Open Scientific Framework Repository: https://osf.io/fpwr3/. We identified 76 invertebrate OTUs; 43 from the Ant COI assay and 33 from the Invertebrate COI assay (Fig. 2). Using databases to match the OTUs to the highest level of classification, 2 were matched at only class level, 14 could only be matched to order, 2 at suborder, 26 to family level, 7 at subfamily, 8 to genus and 13 could be matched to species level. Only 4 OTUs could not be matched to any taxonomic level (Supplementary Table 2). Of the 24 ant taxa described from samples using a morphological approach, 11 of those species were found within the total data set. However, DNA sequences of the remaining species were too similar to be able to confidently be assigned to a higher taxonomic resolution and remained at genus or order level.

There were no significant differences in OTU richness between restoration ages at Gingin ($F_{4,10}$ =0.74, p=0.59) or Gnangara ($F_{4,10}$=1.94, p=0.18) (Table 1). However, the NMDS showed distinct clustering of samples from the same age plot (Fig. 3b), and the PERMANOVA with the restoration ages nested within the two sites, showed significant differences between the sites ($F_{1,8}$ =6.546, p < 0.001) and between the restoration ages within sites ($F_{8,20}$=2.369, p < 0.0001). The centroid dispersion of samples in reference plots tended to be higher than restoration plots, particularly the younger restoration plots. Larger dispersion of centroids indicates greater variation between samples within older communities. Environmental values such as maximum vegetation height, slope, or bare ground cover were not significantly related with the invertebrate community composition (R^2<0.19, p>0.10) (Supplementary Table 3).

A simple linear regression (Fig. 3a) indicated that older restoration plots were more similar to reference communities than younger plots ($F_{1,6}$=10.61, p=0.016, R^2= 0.645). Mantel tests indicated that there was significant spatial autocorrelation in the data at both the Gingin (p<0.0001, R= 0.389) and Gnangara sites (p=0.004, R= 0.398) (Fig. 4).

An indicator species analysis identified several invertebrates that characterised the restored plots and particular age plots at the two sites (Table 2). Two Collembola (springtail) OTUs were identified as indicators of young restoration plots. There were several OTUs characteristic of restoration plots (See Table 2 for summary). Of note, is *Rhytidoponera metallica* (Family: Formicidae), which occurred at all restoration plots, but was not detected in the reference plots.

Discussion

We investigated the recovery of invertebrate communities post sand mining in former Banksia woodlands in SW Australia. Metabarcoding of invertebrate DNA showed clear differentiation between the mine sites and age plots within those sites, with restoration communities becoming increasingly similar to reference communities over time.

Were older sites most similar to reference sites?

Understanding whether a restored site is approaching a reference community is a key challenge for mine site restoration given the need to demonstrate achievement of biodiversity goals as part of regulatory processes (Bullock et al. 2011; Cross et al. 2018). We found that invertebrate OTU richness was similar between the sites and age plots. Studies have shown that richness for ground

dwelling invertebrates was approximately the same between sites exhibiting different habitat composition (Lengyel et al. 2016); instead, it was the species composition that had changed.

In our study, invertebrate communities at older restored plots were approaching a reference state after 13 years. Invertebrate communities gradually return after restoration of ecosystems (Watts & Mason 2015) and are influenced by structural elements including vegetation and other invertebrates within the system (Hawkins & Porter 2003; Moir et al. 2005; Yeeles et al. 2017). Although we found older restored plots were more similar to reference plots, the environmental variables we measured (vegetation height, slope, bare ground cover) were not statistically significant in explaining invertebrate community variation, despite increases in vegetation height and vegetation cover in older plots. The choice of variable and the small sample size may be partially responsible for the lack of significance. Future studies determining which environmental variables are associated with invertebrate recovery may help land managers facilitate restoration by simulating the relevant conditions.

Whilst a significant relationship between restoration age and similarity to reference community was found, it is possible that other factors contributed to this pattern. Restoration ages could not be replicated at mines, and construction limited access to reference remnant vegetation. Also, plots were generally in a chronosequential line and a significant spatial autocorrelation was detected in the data. These are all common issues in retrospective studies of chronosequences. Distance to the reference plot may at least partially explain the patterns detected. It could be assumed that invertebrate communities disperse more readily from remnant vegetation to nearby sites, regardless of their age. Although, at the Gingin site, all plots were a similar distance from remnant vegetation that was comparable in species composition and quality to that of the reference plot. However, we cannot rule out that distance to remnant vegetation may be an important factor in the restoration of

invertebrate communities. Accordingly, retention of strategically located remnants may be a critical

factor in the successful restoration of invertebrate communities (Sundermann et al. 2011).

Invertebrate taxa characterising restoration ages

Whilst DNA assays can provide an understanding of total invertebrate communities, the

identification of indicator species remains worthwhile as indicators drive many policy decisions and

are important to investigate from a land management perspective (Ji et al. 2013). Ant fauna are

often a focal group for morphologically based biomonitoring in Australia (Andersen et al., 2002), and

we found several indicator species at the two sites. For example, at Gnangara, *Iridomyrmex*

splendens was found in 87% of the reference plots but not in restored plots. A generalist species, this

ant is found in forested environments and is generally not tolerant of ecosystem degradation or

urbanisation (Heterick & Shattuck, 2011). *Iridomyrmex splendens'* presence is therefore indicative of

an undisturbed environment, and is a species of interest when examining restored sites to

demonstrate return of a pre-disturbance community. Another species, the Green-head ant

(*Rhytidoponera metallica)*, was found at 95% of all the restored plots at both Gingin and Gnangara

but not in the reference plots. *Rhytidoponera metallica* is a known disturbance indicator (Hoffmann

& Andersen 2003), which thrives in disturbed environments that may not be suitable for colonisation

by other ant species (Heterick, 2009).

Younger plots at both mines were characterised by the presence of an *Entomobrya* OTU, a genus of

springtails (Order: Collembola). Springtails are one of the most abundant soil dwelling invertebrates,

with densities of up to 60,000 individuals/m^2 (Eisenhauer, Sabais, & Scheu, 2011). They play a critical

role in mobilising nutrients stored in microbial biomass through feeding on fungi and microbial

matter within soils (Bardgett et al. 1993; Hopkin 1997). The presence of springtails and other

decomposer groups in a system may impact plant performance, with springtail species composition a good predictor for ecosystem functioning; influencing root depth, root biomass, and root litter decomposition (Eisenhauer et al. 2011, 2010). In oldest restored plots and reference plots, millipedes (Order: Julida) were more frequently detected, especially at Gingin where they were found in 90% of the samples. Out of the 9 orders of millipedes in Australia, Julida is the only introduced order. The ubiquity of Julida at Gingin indicates that this alien species has become a key community component, complicating the return to a historic invertebrate assemblage.

Will a DNA approach improve restoration monitoring?

Current monitoring for invertebrates uses morphospecies and taxonomic minimalism to address constraints associated with morphology-based identification, such as time and expertise required (Pik et al. 1999; Beattle & Oliver 1994; Cole et al. 2016). However, morphospecies may not be the most ideal solution especially in cases where organisms have polymorphic life stages or variable phenotypes (for example the polymorphic workers in the ant genus *Melophorus* (Meier et al. 2016). Metabarcoding shows potential for dealing with these issues (Taberlet et al. 2012), although, we did encounter some problems that may limit its effectiveness in restoration monitoring at the present time. Firstly, primer amplification bias can cause considerable issues with accurate taxa detection (Deagle et al. 2014). We found that the general invertebrate COI primer (ZBJ-Art, Zeale et al. 2011) did not amplify organisms within the order Hymenoptera. With our custom ant COI primer created to target this issue, the length of the region was not adequate to distinguish some species from each other. Thus, further primer development may need to be explored to get the suitable taxonomic resolution for uncovering ant diversity.

Secondly, approximately 14 of the OTUs detected from this data set could not be identified beyond order, indicating gaps within the current taxonomic databases available. However, this study represents the first use of this technique to characterise invertebrate taxa in south-western Australia, and in time, species identification will improve as more species are barcoded and added to databases. Also, we were able to detect ecosystem changes in a taxonomic independent space (using OTUs) and important indicator groups from the taxa that could be identified. Whilst identifying OTUs down to a species level is important for fine scale community composition monitoring and being able to identify changes in the environment early (Cristescu 2014), taxonomy independent analysis (using OTUS) still allows for basic ecological analyses to be conducted to answer relevant questions (Bik et al. 2012). Most importantly, the use of this methodology for identification reduces the man hours required for traditional studies (Ji et al. 2013), making it a more accessible and scalable for restoration monitoring practices (Fernandes et al. 2018).

Conclusions

To the best of our knowledge this is the first study to use DNA metabarcoding to investigate invertebrate community composition in a restoration context. Invertebrate community composition at two former sand mines appear to be recovering towards reference conditions over a thirteen-year time period. However, given the scarcity of data concerning the return of invertebrate communities in ecosystem restoration, we have no way of knowing how these changes will affect vegetation re-establishment or ecosystem functions including pollination and seed dispersal.

There is an increasing call to monitor restoration using the latest molecular approaches (Williams et al. 2014; Fernandes et al. 2018) but our findings suggest care must be taken in assay selection and optimisation to ensure that target groups are identified. Examining fine scale changes to

invertebrate communities may be an effective predictor for restoration success. However, experiments are needed where invertebrate assemblages in restoration of different standards are compared in order to establish whether consistent signals are detectable. Finally, this study highlights the need for exemplar restoration projects that include appropriate balanced design and multiple suitable reference sites. Such sites will enable rigorous tests of key ecological management questions and are critical in ecological restoration research, where imperfect study systems are often the location of research simply because they are all that exist (Prober et al. 2018). Good design of monitoring programs will also be important for stakeholders (e.g. mining companies) that need to demonstrate restoration success to regulators.

Literature Cited

Altschul SF, Gish W, Miller W, Meyers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215:403-410

Andersen AN, Hoffmann BD, Müller WJ, Griffiths AD (2002) Using ants as bioindicators in land management: Simplifying assessment of ant community responses. Journal of Applied Ecology 39:8–17.

Andersen AN, Fisher A, Hoffman BD, Read JL, Richards ROB (2004) Use of terrestrial invertebrates for biodiversity monitoring in Australian rangelands, with particular reference to ants. Austral Ecology 29:87–92

Andersen AN, Majer JD (2004) Ants show the way Down Under: invertebrates as bioindicators in land management. Frontiers in Ecology and the Environment 2:291–298

Bardgett RD, Whittaker JB, Frankland JC (1993) The effect of collembolan grazing on fungal activity in differently managed upland pastures: A microcosm study. Biology and Fertility of Soils 16:255–262

Beattle AJ, Oliver I (1994) Taxonomic minimalism. Trends in Ecology & Evolution 9:488–490

Beng KC, Tomlinson KW, Shen XH, Surget-Groba Y, Hughes AC, Corlett RT, Slik JWF (2016) The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. Scientific Reports 6:24965

Benigno SM, Dixon KW, Stevens JC (2013) Increasing Soil Water Retention with Native-Sourced Mulch Improves Seedling Establishment in Postmine Mediterranean Sandy Soils. Restoration Ecology 21:617–626

Bettenay E, McArthur WM, Hingston FJ (1960) The soil association of part of the Swan Coastal Plain, Western Australia. CSIRO Soils Division: Soils and Land Use Series.

Bik HM, Porazinska DL, Creer S, Caporaso JG, Knight R, Thomas WK (2012) Sequencing our way towards understanding global eukaryotic biodiversity. Trends in Ecology & Evolution 27:233–243

Bullock JM, Aronson J, Newton AC, Pywell RF, Rey-Benayas JM (2011) Restoration of ecosystem services and biodiversity: conflicts and opportunities. Trends in Ecology and Evolution 26:541–549

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2015) GenBank. Nucleic Acids Res. 44:D67–D72.

Clarke KR, Gorley 2015. PRIMER-E. Plymouth Routines in Multivariate Ecology Research. http://www.primer-e.com/ (Accessed 10 Aug 2017)

Cole RJ, Holl KD, Zahawi RA, Wickey P, Townsend AR (2016) Leaf litter arthropod responses to tropical forest restoration. Ecology and Evolution 6:5158–5168

Cristescu ME (2014) From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology & Evolution 29:566–571

Cross AT, Young R, Nevill P, McDonald T, Prach K, Aronson J, Wardell-Johnson GW, Dixon KW (2018) Appropriate aspirations for effective post-mining restoration and rehabilitation: a response to Kaźmierczak et al. Environmental Earth Sciences   77:256

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. Biology Letters 10:20140562

De Caceres M, Jansen F (2016) "Indicspecies" Package. http://r-project.org/pub/R/web/packages/indicspecies/indicspecies.pdf (Accessed 3 Sep 2017)

Dodd J, Heddle EM, Pate JS, Dixon KW (1984) Rooting patterns of sandplain plants and their

functional significance. Kwongan, plant life of the sandplain: biology of a south-west Australian shrubland ecosystem / editors J.S. Pate and J.S. Bear. University of Western Australia Press, Nedlands, Western Australia.

Dray S, Dufour A (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. Journal of Statistical Software 22:1-20.

Edgar RC (2013) UPARSE: Highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10:996–998

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194-2200

Eisenhauer N, Hörsch V, Moeser J, Scheu S (2010) Synergistic effects of microbial and animal decomposers on plant and herbivore performance. Basic Applied Ecology 11:23–34

Eisenhauer N, Sabais ACW, Scheu S (2011) Collembola species composition and diversity effects on ecosystem functioning vary with plant functional group identity. Soil Biology and Biochemistry 43:1697–1704

Fernandes K, van der Heyde M, Bunce M, Dixon K, Harris RJ, Wardell-Johnson G, and Nevill PG (2018) DNA metabarcoding—a new approach to fauna monitoring in mine site restoration. Restoration Ecology 26:1098-1107

Gellie NJCC, Mills JG, Breed MF, Lowe AJ (2017) Revegetation rewilds the soil bacterial microbiome of an old field. Molecular Ecology 26:2895–2904

Hawkins BA, Porter EE (2003) Does herbivore diversity depend on plant diversity? The case of California butterflies. American Naturalist 161:40–49

Herrick JE, Schuman GE, Rango A (2006) Monitoring ecological processes for restoration projects. Journal for Nature Conservation 14:161–171

Heterick B, Shattuck S (2011) Revision of the ant genus Iridomyrmex (Hymenoptera: Formicidae). Magnolia Press, Auckland

Heterick BE (2009) A guide to the ants of south-western Australia. West Australian Museum, Perth

Hobbs RJ, Suding KN (2009) New models for ecosystem dynamics and restoration. Island Press, Washington DC

Hoffmann BD, Andersen AN (2003) Responses of ants to disturbance in Australia, with particular reference to functional groups. Austral Ecology 28:444–464

Hopkin SP (1997) Biology of the Springtails: (Insecta: Collembola). OUP Oxford

Huson DH, Mitra S, Ruscheweyh H-J, Weber N, Schuster SC (2011) Integrative analysis of environmental sequences using MEGAN4. Genome Research 21:1552–1560

Ji Y et al. (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecology Letters 16:1245–1257

Leese F et al. (2018) Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived from the DNAqua-Net COST Action. Advances in Ecological Research 58:63–99

Lei K, Pan H, Lin C (2016) A landscape approach towards ecological restoration and sustainable development of mining areas. Ecological Engineering 90:320–325

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

Majer JD, Heterick BE, Gunawardene NR, Castalanelli M, Ledger J (2018) Sequencing the Ant fauna of a Small Island: Can Metagenomic Analysis Enable Faster Identification for Routine Ant Surveys? Sociobiology 65:422

Maher K (2009) Restoration of Banksia Woodland after the removal of pines at Gnangara: seed species requirements and prescriptions for restoration. Department of Environment and Conservation. Perth, WA

McDonald T, Gann GD, Jonson J, and Dixon KW (2016) International standards for the practice of ecological restoration – including principles and key concepts. Society for Ecological Restoration, Washington, D.C.

McGeoch MA (1998) The selection, testing and application of terrestrial insects as bioindicators. Biological Reviews 73:181–201

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

Miller BP et al. (2017) A framework for the practical science necessary to restore sustainable, resilient, and biodiverse ecosystems. Restoration Ecology 25:605–617

Moir ML, Brennan KEC, Koch JM, Majer JD, Fletcher MJ (2005) Restoration of a forest ecosystem: The effects of vegetation and dispersal capabilities on the reassembly of plant-dwelling arthropods. Forest Ecology and Management 217:294–306

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, Minchin PR, Hara RBO, Simpson GL, Solymos P, et al. (2016). Vegan: Community Ecology Package. https://Cran.r-Project.Org, Https://Github.Com/Vegandevs/Vegan 291 (accessed 12 Aug 2017)

Palmer MA, Filoso S (2009) Restoration of ecosystem services for environmental markets. Science 325:575–6

Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, et al. (2013) Essential biodiversity variables. Science 6177:277-278

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

Prober SM, Broadhurst L, Boggs G, Breed MF, Bush D, Lynch AJJ, Dickson F (2018) Discussion Paper: Achieving more with less – linking ecological restoration investments with ecological restoration research infrastructure. CSIRO, Australia.

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ (accessed 12 Aug 2017).

Rokich DP, Dixon KW (2007) Recent advances in restoration ecology, with a focus on the Banksia woodland and the smoke germination tool. Australian Journal of Botany. 55:375–389

Ritchie AL, Nevill PG, Sinclair EA, Krauss SL (2017) Does restored plant diversity play a role in the reproductive functionality of Banksia populations?. Restoration Ecology. 25:414-23.

Rosenberg DM, Danks H V., Lehmkuhl DM (1986) Importance of insects in environmental impact assessment. Environmental Management 10:773–783

RPS (2012) Environmental Assessment: Sand Quarries, Gnangara https://consultation.epa.wa.gov.au/seven-day-comment-on-referrals/sand-quarries-gnangara-pine-plantation/supporting_documents/CMS14120%20%20Environmental%20Assessment%20Report.pdf (accessed 7 December 2017)

Ruiz-Jaen MC, Mitchell Aide T (2005) Restoration Success: How Is It Being Measured? Restoration Ecology 13:569–577

Siddig AA, Ellison AM, Ochs A, Villar-Leeman C, Lau MK (2016) How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in Ecological Indicators. Ecological Indicators 60:223-30

Silva THSE, Karunarathna NB, Karunaratne WAIP (2017) Competence of Litter Ants for Rapid Biodiversity Assessments. International Journal of Biodiversity 2017:1–10

Smith MA, Fisher BL, Hebert PDN (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 360:1825–1834

Smith CMS, Bowie MH, Hahner JL, Boyer S, Kim YN, Zhong HT, Abbott M, Rhodes S, Sharp D, Dickinson N (2016) Punakaiki Coastal Restoration Project: A case study for a consultative and multidisciplinary approach in selecting indicators of restoration success for a sand mining closure site, West Coast, New Zealand. Catena 136, 91–103.

Sundermann A, Stoll S, Haase P (2011) River restoration success depends on the species pool of the immediate surroundings. Ecological Applications 21:1962–1971

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

Tizado EJ, Núñez-Pérez E (2016) Terrestrial arthropods in the initial restoration stages of anthracite coal mine spoil heaps in northwestern Spain: potential usefulness of higher taxa as restoration indicators. Land Degradation & Development 27:1131-1140.

Ward DF, New TR, Yen AL (2001) Effects of Pitfall Trap Spacing on the Abundance, Richness and Composition of Invertebrate Catches. Journal of Insect Conservation 5:47–53

Watts CH, Mason NWH (2015) If we build - they mostly come: partial functional recovery but persistent compositional differences in wetland beetle community restoration. Restoration Ecology 23:555–565

Williams A V, Nevill PG, Krauss SL (2014) Next generation restoration genetics: applications and opportunities. Trends in Plant Science 19:529–537

Wortley L, Hero J-M, Howes M (2013) Evaluating Ecological Restoration Success: A Review of the Literature. Restoration Ecology 21:537–543

Yeeles P, Lach L, Hobbs RJ, van Wees M, Didham RK (2017) Woody plant richness does not influence

invertebrate community reassembly trajectories in a tree diversity experiment. Ecology 98:500–511

Yan D, Mills JG, Gellie NJ, Bissett A, Lowe AJ, Breed M (2018) High-throughput eDNA monitoring of fungi to track functional recovery in ecological restoration. Biological Conservation 217:113–120

Young TP (2000) Restoration ecology and conservation biology. Biological Conservation 92:73–83

Yu DW, Ji Y, Emerson BC, 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 3:613–623

Zeale MRK, Butlin RK, Barker GLA, Lees DC, Jones G (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. Molecular Ecology Resources 11:236–244

Tables and Figures

Table 1: Mean OTU richness values and sample dispersion from the centroid for age plots at the two sites

Plot	OTU Richness	Centroid Dispersion
Gingin	**11.2**	-
2 years	10.67	25.4
4 years	12	33.81
6 years	11.67	27.68
8 years	12.34	27.47
Reference	9.3	44.5
Gnangara	**11.4**	-
2 years	9.33	26.54
6 years	12.67	18.4
10 years	12.0	37.27
13 years	10.33	43.13
Reference	12.67	32.84

Table 2: Identified OTU of significance from the indicator species calculation through the 'indicspecies' function (De Caceres & Jansen, 2016) and the corresponding age of significance.

OTU Taxon ID	Stat	p	Age of Restoration (Years)	Notes
Entomobrya	0.87	<0.001	2	Genus of springtails. Play an important role in soil nutrient cycling (Hopkin, 1997).
Entomobryidae OTU	0.82	<0.01	2, 4, & 6	Family of springtails. Play an important role in soil nutrient cycling (Hopkin, 1997).
Chloropidae OTU	0.82	<0.01	4, 10, 13 & Reference	Family of small flies, known as grass flies.
Pterygota OTU	0.82	<0.01	4, 10, 13, & Reference	Large subclass of insects including orders of winged insects, and those that are secondarily wingless.
Salticidae OTU	0.73	<0.05	4, 6, 8 & 13	Jumping spiders. Represent predators in the environment.
Hymenoptera OTU	0.89	<0.05	2, 4, 6, 8 & 13	Order of all bees, wasps, and ants. Ecologically important order.
Rhytidoponera metallica	0.96	<0.001	All Restoration Plots	Green-head ant. Endemic Australian species. Disturbance indicator species (Hoffman & Andersen, 2003)
Dolichoderinae OTU	0.72	<0.05	4 & 8	Large subfamily of ants. Generally, predators or scavengers.
Melophorous	0.91	<0.05	8	Genus of ant. Endemic to Australia. Known to be omnivorous and some species will eat seeds. Active in high temperatures (Heterick, 2009)
Julida OTU	0.71	<0.05	8 & Reference	Order of introduced millipedes. Detritivore.

Ectobiidae OTU	0.71	<0.05	Gingin only: Reference	Family of cockroaches. Sometimes known as wood cockroaches.
Iridomyrmex splendens	0.87	<0.05	Gnangara only: Reference	Species of native Australian ant. Known for inhabiting forested areas (Heterick & Shattuck, 2011).

Figure 1: Sampling locations at both study sites a) Gingin, and b) Gnangara. Maps show vegetation restoration plots, sampling locations, years of sampling, and pictures from the sampling plot at the different years. All reference plots abutted the restoration project and were from the same plant community on which the restoration was based.

Figure 2: A cladogram featuring all OTUs identified to family level classification for all two assays. In total, there were 77 OTUs detected.

Figure 3: a) Relationship between average plot similarity (Bray Curtis) to reference communities and plot age (regression analysis, $F_{1,6}$=10.61, p=0.016, R^2= 0.645) b) NMDS plot based on Bray-Curtis dissimilarity from presence-absence data showing trends for both sites in two dimensional space (p<0.001, stress=0.149, non-metric fit R^2= 0.978)

Figure 4: Mantel correlograms for the Gingin and Gnangara mine sites showing the significant (p<0.05), positive spatial correlation of the community similarity.
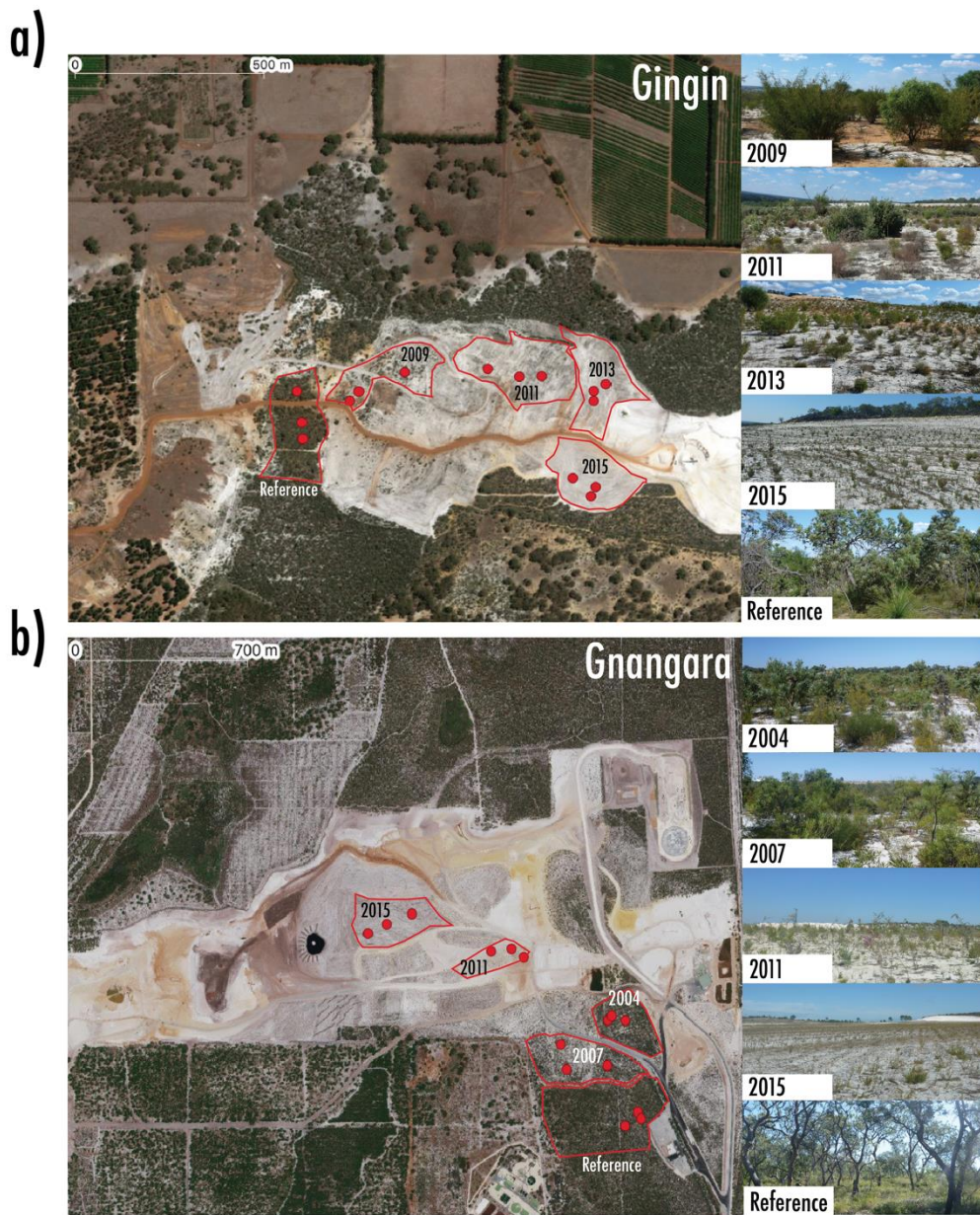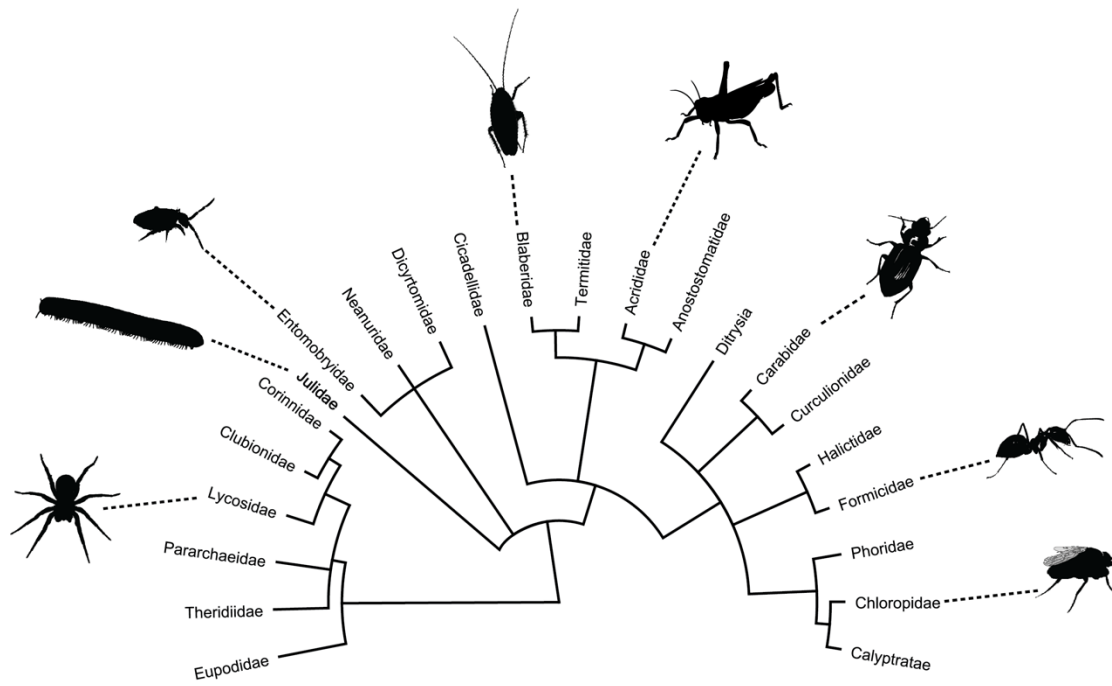
Fig. 1

a) Gingin

2009
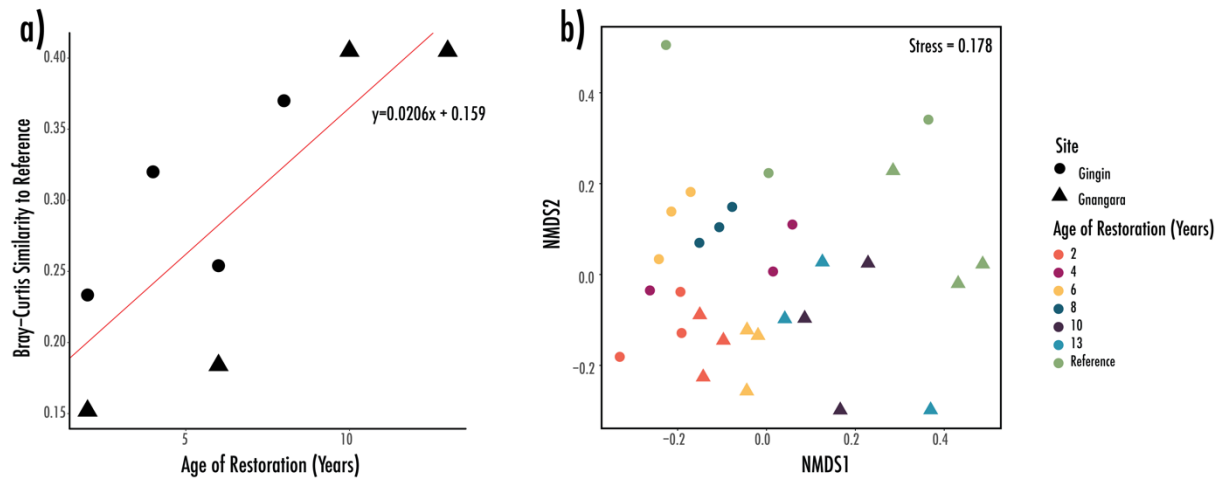2011
2013
2015
Reference

b) Gnangara

2004
2007
2011
2015
Reference

Fig. 2

Dicyrtomidae
Neanuridae
Cicadellidae
Blaberidae
Termitidae
Acrididae
Anostostomatidae
Ditrysia
Carabidae
Curculionidae
Halictidae
Formicidae
Phoridae
Chloropidae
Calyptratae
Entomobryidae
Julidae
Corinnidae
Clubionidae
Lycosidae
Pararchaeidae
Theridiidae
Eupodidae

Fig. 3

Fig 4.