University of Reading

Next generation restoration metrics: using soil eDNA bacterial community data to measure trajectories towards rehabilitation targets

Article

Accepted Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Liddicoat, C. ORCID: https://orcid.org/0000-0002-4812-7524, Krauss, S. L. ORCID: https://orcid.org/0000-0002-7280-6324, Bissett, A. ORCID: https://orcid.org/0000-0001-7396-1484, Borrett, R. J. ORCID: https://orcid.org/0000-0001-8663-0844, Ducki, L. C., Peddle, S. D. ORCID: https://orcid.org/0000-0003-3464-3058, Bullock, P., Dobrowolski, M. P. ORCID: https://orcid.org/0000-0001-5586-4023, Grigg, A. ORCID: https://orcid.org/0000-0002-5818-2973, Tibbett, M. ORCID: https://orcid.org/0000-0003-0143-2190 and Breed, M. F. ORCID: https://orcid.org/0000-0001-7810-9696 (2022) Next generation restoration metrics: using soil eDNA bacterial community data to measure trajectories towards rehabilitation targets. Journal of Environmental Management, 310. 114748. ISSN 0301-4797 doi: https://doi.org/10.1016/j.jenvman.2022.114748 Available at https://centaur.reading.ac.uk/103441/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.jenvman.2022.114748

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Next generation restoration metrics: Using soil eDNA bacterial community data to measure trajectories towards rehabilitation targets

Author details:

Craig Liddicoat[1,2], Siegfried L. Krauss[3,4], Andrew Bissett[5], Ryan J. Borrett[6], Luisa C. Ducki[1,6], Shawn D. Peddle[1], Paul Bullock[7], Mark P. Dobrowolski[4,8,9], Andrew Grigg[10], Mark Tibbett[4,11], Martin F. Breed[1]

[1]College of Science and Engineering, Flinders University, Adelaide, Australia

[2]School of Public Health, The University of Adelaide, Adelaide, Australia

[3]Kings Park Science, Western Australia Department of Biodiversity Conservation and Attractions, Perth, Australia.

[4]School of Biological Sciences, University of Western Australia, Perth, Australia

[5]CSIRO Oceans and Atmosphere, Hobart, Australia

[6]College of Science, Health, Engineering and Education, Murdoch University, Perth, Australia

[7]South32 Worsley Alumina, Perth, Australia

[8]Iluka Resources Limited, Perth, Australia

[9]Harry Butler Institute, Murdoch University, Perth, Australia

[10]Alcoa of Australia Limited, Perth, Australia

[11]Department of Sustainable Land Management & Soil Research Centre, School of Agriculture, Policy and Development, University of Reading, Berkshire, United Kingdom

Corresponding author:

Craig Liddicoat, College of Science and Engineering, Flinders University, Sturt Road Bedford Park SA 5042, Adelaide, Australia

E-mail address: craig.liddicoat@flinders.edu.au

Next generation restoration metrics: Using soil eDNA bacterial community data to measure trajectories towards rehabilitation targets

Abstract

In post-mining rehabilitation, successful mine closure planning requires specific, measurable, achievable, relevant and time-bound (SMART) completion criteria, such as returning ecological communities to match a target level of similarity to reference sites. Soil microbiota are fundamentally linked to the restoration of degraded ecosystems, helping to underpin ecological functions and plant communities. High-throughput sequencing of soil eDNA to characterise these communities offers promise to help monitor and predict ecological progress towards reference states. Here we demonstrate a novel methodology for monitoring and evaluating ecological restoration using three long-term (> 25 year) case study post-mining rehabilitation soil eDNA-based bacterial community datasets. Specifically, we developed rehabilitation trajectory assessments based on similarity to reference data from restoration chronosequence datasets. Recognising that numerous alternative options for microbiota data processing have potential to influence these assessments, we comprehensively examined the influence of standard versus compositional data analyses, different ecological distance measures, sequence grouping approaches, eliminating rare taxa, and the potential for excessive spatial autocorrelation to impact on results. Our approach reduces the complexity of information that often overwhelms ecologically-relevant patterns in microbiota studies, and enables prediction of recovery time, with explicit inclusion of uncertainty in assessments. We offer a step change in the development of quantitative microbiota-based SMART metrics for measuring rehabilitation success. Our approach may also have wider applications where restorative processes facilitate the shift of microbiota towards reference states.

KEYWORDS:

eDNA, mine closure assessment, restoration genomics, rehabilitation trajectory, soil microbiota, spatial autocorrelation

1. INTRODUCTION

Land degradation and transformation, with negative impacts to biodiversity and ecosystem function, are estimated to impact 75% of the Earth's land surface, and this figure is projected to rise to over 90% by 2050 (IPBES, 2018). Ecological restoration—activity that supports rehabilitation of locally representative, sustainable, biodiverse ecosystems (Gann et al., 2019)—is seen as integral to reversing these impacts, as highlighted by the UN declaration of 2021–2030 as the Decade on Ecosystem Restoration (https://www.decadeonrestoration.org/). Restoration is technically challenging and requires considerable investment, without guaranteed success (Tibbett, 2015). With large investments in restoration (e.g. BenDor et al., 2015 estimate US$9.5 billion/yr is spent in the USA alone; Menz et al., 2013 estimate US$18 billion/yr is required to restore degraded lands globally), there is a need to improve the evidence base to guide continuous improvement in restoration outcomes and to underpin future investment.

Reference ecosystems provide an important basis for establishing targets and monitoring progress of restoration activities (Gann et al., 2019) (refer to online Supplementary Materials in Appendix A, Figure S1). In post-mining contexts, best practice guidelines require formal mine completion criteria to be prescribed in a matter that is specific, measurable, achievable, relevant and time-bound (SMART) (Australian_Government, 2016; Manero et al., 2021). To-date, completion criteria have largely focussed on vegetation community variables, with typical ecological measures including alpha and beta diversity reflecting the number of different taxa and community composition, respectively. For example, targets may be set at a minimum threshold similarity to a reference community. Despite available guidance, many completion criteria are ambiguous or ill-defined, and can result in unclear standards for regulators, unrealistic expectations for stakeholders, and represent a key barrier to the relinquishment of minesites

(Manero et al., 2021). To help move the industry towards improved definitions of completion criteria, Manero et al. (2021) suggest criteria for industry best practice, which include using multiple reference sites, monitoring and corrective actions (i.e., adaptive management), allowing innovation-guided completion criteria, and specific objectives and indicators.

Soil microbial communities (microbiota) have essential roles in organic matter decomposition, soil formation, and nutrient cycling, and therefore help regulate plant productivity and community dynamics (Harris, 2009). Patterns of land use, vegetation communities, and soil quality each help to shape soil microbiota (Bulgarelli et al., 2013; Delgado-Baquerizo et al., 2018; Turner et al., 2013). Microbiota depend on the resource and energy flows associated with aboveground biota, and therefore their monitoring may help indicate the impact of restoration interventions (Harris, 2009; Jiao et al., 2018; van der Heyde et al., 2020).

The development of low-cost, high-throughput sequencing of environmental DNA (eDNA) has enabled affordable, rapid and comprehensive assessment of soil microbiota; and these genomic techniques are now being used widely in a restoration context (Breed et al., 2019; Mohr et al., 2022). Applying recognised ecological assessment approaches to abundant eDNA-based microbiota data has potential to provide a novel tool for measuring trajectories and predicting time to recover towards restoration targets (Rydgren et al., 2019). Chronosequence study designs, while containing limitations (Walker et al., 2010), are commonly used to examine ecosystem recovery following restoration activities (Tibbett, 2010). However, there are few studies of soil microbiota from restoration chronosequences that explicitly visualise and evaluate patterns in ecological similarity to reference data with time since rehabilitation. It is customary for such studies (e.g., Fernandez Nuñez et al., 2021; Jiao et al., 2018; Schmid et al., 2020) to examine patterns in microbiota composition via analysis of taxonomic groups and ordination techniques which project multivariate

community data into lower dimensional space (e.g., 2-d plots). These popular techniques often characterise the complexity and site-specificity of soil ecosystems. However, a focus on measuring 'similarity to reference' may help cut through the complexity inherent to microbiota data. Along these lines, van der Heyde et al. (2020) visualised temporal trends in ecological similarity to reference data in post-mining rehabilitation—however, in their example each rehabilitation sample was only compared to a single closest reference sample, which potentially limited insight into variability and uncertainty in microbiota recovery.

Here we provide a proof-of-concept demonstration and detailed exploration of a new complexity-reducing application of eDNA-based soil bacterial community data to assess the progress of post-mining rehabilitation using three long-term (> 25 year) chronosequence case studies from south-west Western Australia. Specifically, we aim to demonstrate the use of chronosequence-based rehabilitation trajectories, using measures of percent similarity of bacterial community structure to ecological reference sites (hereafter termed references), to assess progress of soil bacterial communities towards reference states with increasing rehabilitation age. We note that further work that links microbiota to other ecosystem components (e.g., vegetation, fauna) is important but beyond the scope of our study.

Our intended audience includes microbiome researchers working in ecosystem restoration, as well as restoration managers who are considering new methods to add to their ecological monitoring toolkit. Our approach may also be adapted for monitoring and predicting microbiota recovery toward reference states in broader contexts, including microbiota-conscious urban design (Watkins et al., 2020) which represents an extension of ecosystem restoration in urban areas; and microbiota-mediated human health where the notion of diverse healthy reference states is well recognised (Lloyd-Price et al., 2016).

Due to the potential for alternative data processing options to cause varying impacts on our rehabilitation trajectory assessments, we compare outcomes from a range of potential

options relevant to microbiota data analyses. For example, compositional data analysis approaches are promoted to have greater statistical rigour compared to standard approaches (Gloor et al., 2017); grouping bacterial taxa based on sequence similarity (i.e., varying the resolution of operational taxonomic units, OTUs) might help manage noise associated with microbiome data; taxonomic grouping might assist interpretation if recognised groups can be discussed; and eliminating rare taxa (to simulate reduced sequencing depths) might allow more cost-effective and rapid analyses. We also recognise the potential for spatial autocorrelation—where measured outcomes are closer in value due to closer spatial proximity—to confound the assessment of rehabilitation age in chronosequence studies that lack appropriate spatial design and replication. Accordingly, our *a priori* research questions were: (1) can soil bacterial community data be used to establish reference-based targets? (2) can soil bacterial community rehabilitation trajectory data be used to predict the time to recover to reference targets? and (3) how are these predictions of recovery influenced by different ecological distance/similarity measures and sequence data resolution? (4) Additionally, we conduct a preliminary, illustrative examination of spatial autocorrelation, and trial an approach to highlight and 'correct' datasets where its influence appears excessive. We then discuss limitations and synthesise our findings to inform future work.

2. MATERIALS AND METHODS

2.1 Data collection

We used surface soil bacterial 16S rRNA marker gene data from three case study minesites (Figure 1; Appendix A, Tables S1–S3) from south-west Western Australia. Soil sampling was undertaken in accordance with Australian Microbiome (AM) protocols (Bissett et al., 2016; https://www.australianmicrobiome.com/protocols; Appendix A, Supplementary Methods). Each minesite experiences a Mediterranean-type climate with hot, dry summers

and cool, wet winters. Post-mining rehabilitation activities typically involved deep-ripping, prior to the 'direct return' (where possible) of subsoil and topsoil stripped from a separate pit about to be mined, followed by revegetation with locally appropriate seed of diverse plant communities (Tibbett, 2010). Precise soil handling and storage techniques differed between the minesites and different pits within minesites. Summary information for each minesite is provided below (see Appendix A, Supplementary Methods for more background information; other studies in-progress will provide expanded analyses of surface and subsoil data from these minesites, including additional marker gene datasets).

Alcoa's *Huntly* bauxite-producing minesite is approximately 100 km south-east of Perth, occurring in mixed open forest with dominant overstorey species of Jarrah (*Eucalyptus marginata*) and Marri (*Corymbia calophylla*) on lateritic, nutrient poor soils. We consider Huntly data sampled in 2016, with rehabilitation ages between 2–29 years old. Huntly's 36 samples correspond to rehabilitation years: 1987 (n = 3), 1991 (n = 3), 1999 (n = 3), 2002 (n = 3), 2008 (n = 3), 2014 (n = 3), reference (n = 18), where each reference site was paired with an adjacent rehabilitation site.

Iluka Resource's *Eneabba* mineral-sand minesite is approximately 280 km north of Perth, occurring in sandplain heath vegetation comprising low shrubland on undulating infertile siliceous sandplains, predominantly featuring perennial woody species from the Proteaceae, Myrtaceae, and Fabaceae families. We consider Eneabba data sampled in 2019, with rehabilitation ages between 7–38 years. Eneabba's 26 samples correspond to rehabilitation years: 1981 (n = 3), 1989 (n = 2), 1995 (n = 3), 2000 (n = 2), 2004 (n = 3), 2009 (n = 2), 2012 (n = 2), reference (n = 9).

South32's *Worsley* bauxite-producing minesite is located approximately 150 km south of Perth, occurring in Jarrah (*Eucalyptus marginata*) forest on lateritic, nutrient poor soils. We consider Worsley data sampled in 2019, with rehabilitation ages between 2–28 years old.

Worsley's 25 samples correspond to rehabilitation years: 1991 (n = 2), 1996 (n = 4), 1999 (n = 2), 2002 (n = 2), 2005 (n = 2), 2007 (n = 1), 2011 (n = 3), 2017 (n = 3), reference (n = 6).

Each soil sample had physico-chemical analyses performed at CSBP Laboratories (Perth, Western Australia) to quantify key soil abiotic variables as prescribed by AM protocols, including soil texture, organic carbon, ammonium, potassium, sulphur, calcium, pH, nitrate, phosphorous, and electrical conductivity.

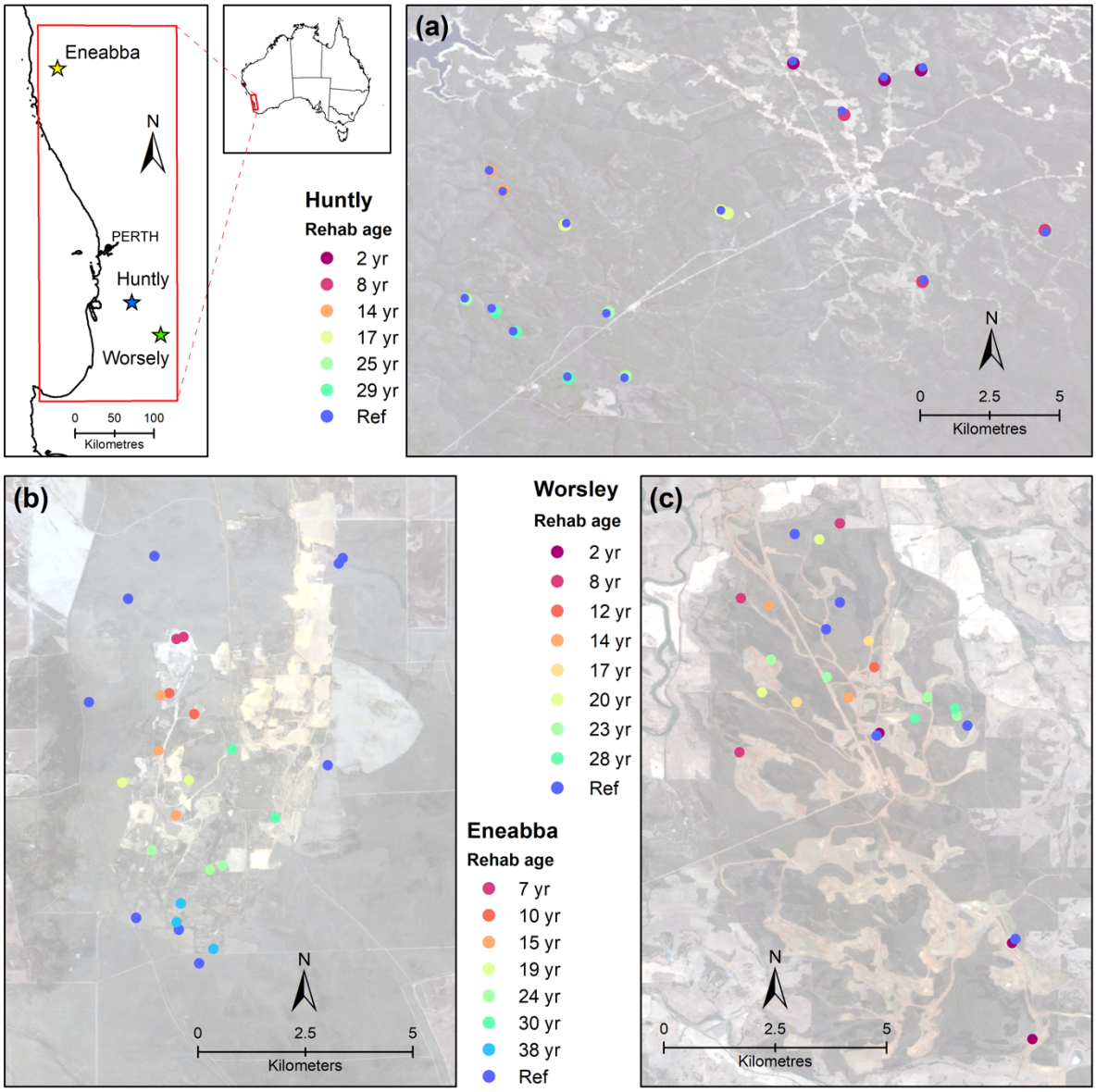


FIGURE 1. Locations of minesites and soil sampling sites: (a) Huntly, (b) Eneabba, (c) Worsley. (Imagery: Sentinel-2; https://eos.com/landviewer; EOS Data Analytics, Inc.)

194 **2.2 eDNA sequencing, bioinformatics, and data preparation**

195 DNA extraction, PCR and preliminary bioinformatic analyses were undertaken in accordance
196 with AM workflows (Bissett et al., 2016; see Appendix A, Supplementary Methods). From
197 this workflow, denoised 16S rRNA gene amplicon sequence variant (ASV) level abundance
198 data were produced for all minesites. Note, in this study ASVs are equivalent to zero radius
199 OTUs (zOTUs). Further data preparation and analyses were largely undertaken in R version
200 4.0.3 (R-Core-Team, 2020) utilising the framework of the R phyloseq package (McMurdie
201 and Holmes, 2013) to manage the datasets (see Appendix A, Supplementary Methods for
202 number of sequences and ASVs studied in each minesite, initial data cleaning steps, and
203 preparation of phylogenetic trees).
204

205 **2.3 Data visualisation and statistical analyses**

206 We visualised the sequence depth of samples using rarefaction curves (Appendix A,
207 Figure S2). We performed exploratory data analyses to visualise ASV alpha diversity,
208 evenness, and relative abundance via heatmaps of phyla, classes, and orders in each minesite
209 (Appendix A, Supplementary Methods, Figures S3–S13). Alpha diversity and evenness were
210 based on rarefied ASV abundances (as below), while relative abundances were computed
211 using non-rarefied data. We used alternative normalisation approaches consistent with
212 common practice in the literature (Gloor et al., 2017; Weiss et al., 2017). Further exploratory
213 data analyses included preliminary visualisations of soil and landscape variables that
214 associated with the soil bacterial community samples within each minesite (see Appendix A,
215 Supplementary Methods, Supplementary Data, Figures S14–20).
216 To prepare for the computation of 'standard' ecological distance measures (as
217 described by Gloor et al., 2017; e.g. Bray-Curtis, Jaccard, UniFrac), we normalised the
218 sequence data for sampling effort by rarefying abundances of ASVs, and other taxonomic

levels investigated (see below), to the minimum sample sequence depth within respective minesites (Huntly, n = 17,485 sequences; Eneabba, n = 10,142 sequences; Worsley, n = 54,122 sequences) using the *rarefy_even_depth()* function from R phyloseq.

To prepare 'compositional' data analysis distance measures we followed the recommendations of Gloor et al. (2017) and Quinn et al. (2019), using non-rarefied data. However, we took the pragmatic initial step of excluding taxa that contained zero counts in more than 90% of samples within each minesite, to help limit the potential for artefactual influences (as discussed later) to be introduced by the subsequent steps of zero replacement and centred log ratio transformation. Then, following Quinn et al. (2019) we used the default geometric Bayesian multiplicative model in the *cmultRepl()* function of the R zCompositions package (Palarea-Albaladejo and Martín-Fernández, 2015) to replace zeros with small numbers; before computing centred log ratio transformations using the *propr()* function from the R propr package (Quinn et al., 2017). Further steps are outlined in section 2.3.1 below.

We examined a range of alternative qualitative and quantitative beta diversity (i.e., distance or community dissimilarity) measures which were converted to similarity, to model rehabilitation trajectories and time to reach reference targets (as described further below). For the minesite with the largest number of samples (Huntly), we also investigated data pre-processing options of grouping by sequence similarity, taxonomic grouping, and excluding rare taxa. Details of the number of samples, taxa and sequences considered for all minesites, distance measures and data processing options (see below) are provided in Appendix A, Table S4. Supporting data and R code used in our study are available online (see data availability statement).

2.3.1 Comparison of alternative ecological and compositional similarity measures

For each minesite, we used the cleaned and rarefied ASV-level bacterial community data to derive standard ecological distance matrices using distance measures commonly employed in microbiota studies—i.e., Jaccard, Bray-Curtis, Unweighted UniFrac and Weighted UniFrac (Lozupone et al., 2007)—via the *vegdist()* function from the R vegan package (Oksanen et al., 2020). We also compared results from the Bray-Curtis measures with the compositional data analysis approach from computing Aitchison distances via *vegdist()* (i.e., these were derived from Euclidean distances between samples after centred log ratio transformation; Gloor et al., 2017). For each minesite, Bray-Curtis distances were visualised using principal coordinates analysis (PCoA) ordination, while Aitchison distances were visualised using principal components analysis (PCA) (Gloor et al., 2017) (Figure 2). For the comparison between Bray-Curtis and Aitchison measures at Worsley we used the spatially filtered dataset which excluded the southernmost samples as described in section 2.3.6.

The rehabilitation trajectory analyses presented here were then derived from a subset of data contained in the above distance matrices. Specifically, only pairwise distances between samples and all reference samples within a minesite were considered (including distances among reference samples). That is, any pairwise distances not involving a reference were not included in these analyses.

For standard measures (i.e. Bray-Curtis, Jaccard, Weighted UniFrac and Unweighted UniFrac), data were then expressed as percent similarity to reference values using (adapted from Legendre and Legendre, 2012):

$$\%\ Similarity\ to\ reference_{Standard\ (b/w\ sample\ i\ and\ Ref\ j)} = 100 * (1 - distance_{ij})$$

For the compositional Aitchison measures, similarity to reference was calculated using (adapted from Legendre and Legendre, 2012):

$$\%\ Similarity\ to\ reference_{Aitchison\ (b/w\ sample\ i\ and\ ref\ j)} = 100 * (1 - \frac{distance_{ij}}{distance_{max}})$$

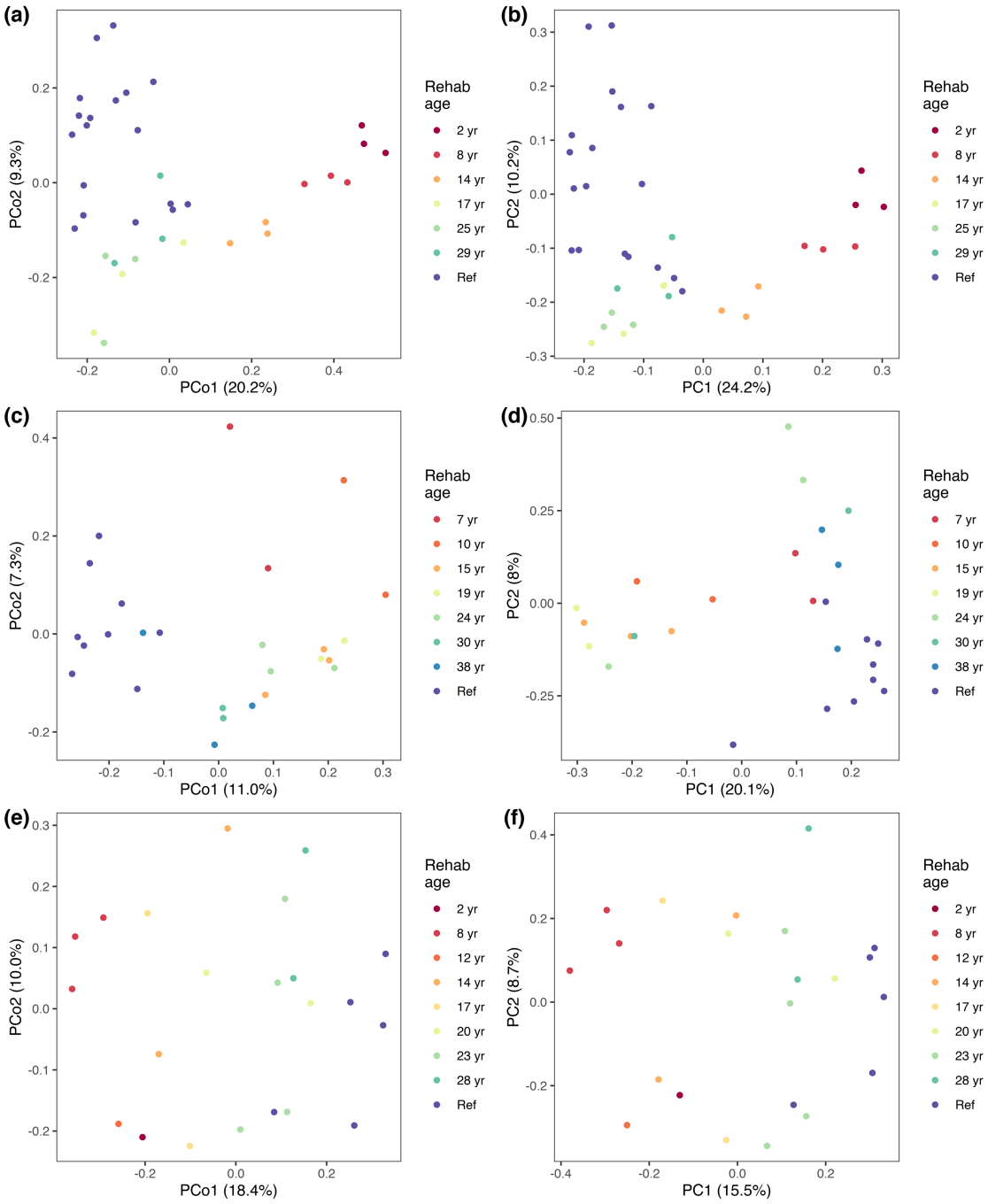


FIGURE 2. PCoA and PCA visualisations of differences in soil bacterial communities for: Huntly (n = 36 samples) using (a) Bray-Curtis distances (30,751 ASVs; 629,460 sequences) and (b) Aitchison distances (25,720 ASVs, 1,723,759 sequences); Eneabba (n = 26 samples) using (c) Bray-Curtis distances (27,115 ASVs; 263,692 sequences) and (d) Aitchison distances (24117 ASVs; 2,042,214 sequences); Worsley (excluding southernmost samples, n

277 = 22 samples) using (e) Bray-Curtis distances (53404 ASVs; 1,190,684 sequences) and (f)
278 Aitchison distances (43,598 ASVs; 1,782,724 sequences).

279

280 **2.3.2 Grouping by sequence similarity**

281 For Huntly data only, separate R phyloseq objects were generated to represent soil bacterial
282 community data with sequences clustered into 99%, 97%, 95%, and 90% identity OTUs (see
283 Appendix A, Supplementary Methods). 99% and 97% identity OTU clustering have been
284 used widely in recent years (prior to the emergence of zOTUs or ASVs), and accordingly, we
285 included OTUs with a range of clustering thresholds to examine whether consistent patterns
286 emerged. For these analyses, OTUs were formed, abundance data were rarefied, and then
287 Jaccard and Bray-Curtis distances and similarity to references were calculated.

288

289 **2.3.3 Taxonomic grouping**

290 For Huntly data only, we examined the influence of taxonomic grouping (i.e., ASV, genus,
291 family, order, class, and phylum) on the assessments of recovery. We also tested the
292 influence of discarding versus retaining (at the next available classified grouping) taxa that
293 were unclassified at each taxonomic rank, which we termed 'pruned' and 'non-pruned' data
294 respectively. Grouping was undertaken using *tax_glom()*; and in 'pruned' datasets,
295 unclassified taxa were removed using *prune_taxa()* from R phyloseq. For these analyses, taxa
296 were grouped, abundance data were rarefied, then Jaccard and Bray-Curtis distances and
297 similarity to references were calculated. Richness and evenness of sequences at the order,
298 class and phylum level were also visualised based on rarefied data and plotted together with
299 composite estimates within rehabilitation age groups from merged-sample bootstrap
300 resampling (Liddicoat et al., 2019) (B=100).

301

302 **2.3.4 Excluding rare taxa**

303 For Huntly data only, we examined the influence of excluding rare taxa, by considering all
304 ASVs, then ASVs with >0.001 %, > 0.01%, and > 0.1% relative abundance within each
305 minesite. For these analyses, ASVs with below the respective relative abundance threshold
306 were removed, abundance data were rarefied, then Jaccard and Bray-Curtis distances and
307 similarity to references were calculated.

308

309 **2.3.5 Rehabilitation trajectory modelling**

310 The progress of rehabilitation was then visualised using boxplots and logarithmic models
311 based on the similarity to reference data. Boxplots were generated from the series of
312 similarity to reference data on the y-axis and increasing rehabilitation age on the x-axis,
313 concluding with reference samples (e.g., Figure 3). Testing for differences in similarities to
314 reference at each rehabilitation age (as visualised with boxplots) was performed using the
315 Kruskal-Wallis rank sum test, followed by post-hoc Dunn tests for multiple comparisons,
316 with Bonferroni adjusted threshold *P*-values. The multiple comparison testing used default
317 two-sided *P*-values and alpha = 0.05 nominal level of significance.

318 After observing the variation in similarity to reference values among references
319 within each minesite (e.g., Figure 3), we defined rehabilitation targets for the purpose of this
320 study as the median (= the central value) of among-reference similarities. This target median
321 value varied by minesite, distance/similarity measure, and pre-processing option.

322 We predicted the time to reach a restoration target (= recovery time) by modelling the
323 trend in similarity to reference with increasing rehabilitation age using bootstrapped (B =
324 100) logarithmic models. The median, 2.5th and 97.5th percentiles of predicted recovery time
325 were evaluated. Our use of logarithmic models was consistent with the approach of Rydgren

et al. (2019), except we used similarity not distance measures. Each iteration of the bootstrap involved random sampling with replacement from the available chronosequence similarity to reference data, excluding outliers identified via the *boxplot()* function in base R, and developing a predictive logarithmic model for similarity to reference out to a maximum rehabilitation age of 500 years, or until the target was reached. Models that failed to reach the target were reported with a prediction time of '>500 years'. Rectangular hyperbola and negative exponential models were also trialled but were abandoned after many cases failed to produce model fits. During our early analyses, we also uncovered example data that highlighted a distorting influence on our trajectory (and recovery time) modelling that appeared to be consistent with the application of 'direct return' soils in young rehabilitation sites. Specifically, this soil material was more similar to references than older rehabilitation sites. Including these samples with elevated similarity to references in the logarithmic modelling appeared to bias models towards flatter, longer trajectories of recovery. Therefore, to reflect the likely onset of recovery towards reference states we decided to only commence logarithmic models (via our automated modelling algorithm) from the youngest rehabilitation age group that had a next older group with increased median similarity to references. As discussed later, this check on model commencement was designed to avoid likely distortions in the modelling of recovery, in particular, due to potential biological inertia in direct return soils (Janzen, 2016).

2.3.6 Exploring spatial autocorrelation

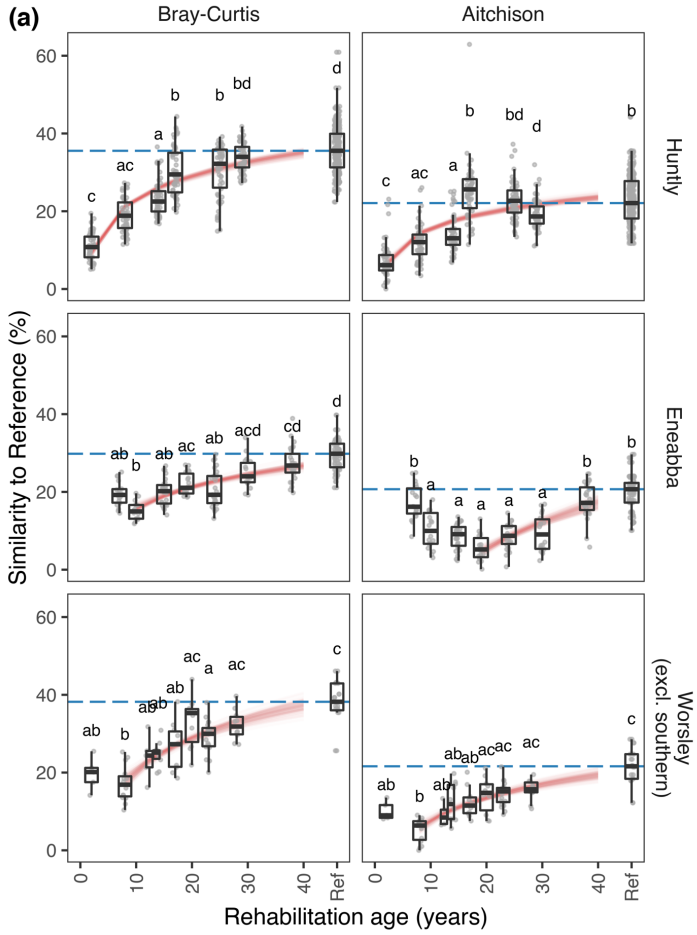
To explore the influence of spatial autocorrelation on our trajectory analyses, we produced variogram-like plots (adapted from Webster and Oliver, 2007) using Bray-Curtis ecological distances (between samples and references) on the y-axis, and geographic distances (between samples and references) on the x-axis. Each rehabilitation age group was modelled as a

351 second-order polynomial, allowing the possible expression of curvilinear trendlines that
352 mimicked variogram-like relationships (i.e., increasing then flattening). Assuming reference
353 curves offered a natural baseline trend for spatial autocorrelation within each minesite
354 environment, we applied a 'correction' to the curvilinear trendline for each rehabilitation age
355 group by calculating the difference in mean-centred model curves (= rehabilitation age group
356 minus reference), such that 'corrected' data for rehabilitation age groups expressed the same
357 ecological distance-geographic distance curvilinear trend as seen for references (see
358 Appendix A, Supplementary Methods for further details of the rationale and approach for this
359 preliminary analysis). Rehabilitation trajectories and predicted recovery times were compared
360 between 'original' and 'corrected' data, for the Bray-Curtis similarities. For the Worsley
361 minesite, a filtered dataset, and corresponding correction, were also prepared which excluded
362 the three southernmost samples (i.e., two 2-year old samples and an adjacent reference),
363 which were geographically separate from the other Worsley samples (see Figure 1, and
364 Appendix A Table S3).
365

366 **3. RESULTS**

367 **3.1 General findings**

368 We found remarkable variability among reference samples within each minesite (Figure 3;
369 Appendix A, Table S5, Figures S21, S23–S25). Median among-reference similarities ranged
370 from <20% to >95% across all measures, and between approximately 30–40% for Bray-
371 Curtis measures, with variation depending on the specific distance measure, pre-processing
372 option, and minesite. All rehabilitation trajectory plots indicated recovery, displaying the
373 general pattern of increasing similarity to references with increasing rehabilitation age
374 (Figure 3; Appendix A, Figures S21, S23–S25), although the logarithmic models and
375 predicted recovery times varied with distance measures, pre-processing and minesite.



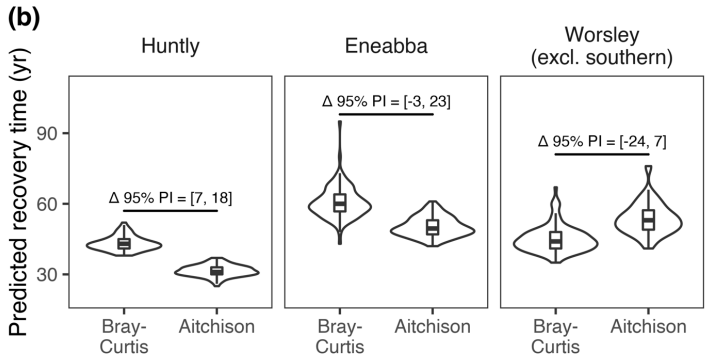


FIGURE 3. Modelled rehabilitation trajectories (a) and predicted recovery times (b) for Huntly, Eneabba, and Worsley (excluding southernmost samples) based on surface soil bacterial community similarity to reference data using Bray-Curtis and Aitchison measures. Plots are derived from the same data that underpin Figure 2. In (a), blue dotted lines denote the target median similarity among reference soils, and red lines represent logarithmic models

384 for changing similarity to reference with rehabilitation age based on bootstrap resampling and
385 modelling (B=100). Boxplots display the distribution of similarity to reference values across
386 rehabilitation ages (groups not sharing a letter are significantly different). In (b), violin plots
387 with boxplot inlays depict the distribution of recovery times from the 100 bootstrap model
388 runs. Δ 95% prediction intervals (PI) indicate whether differences in recovery times predicted
389 using alternative measures (Bray-Curtis versus Aitchison) are significantly different.
390

391 **3.2 Alternative ecological and compositional measures**

392 Despite some differences in the expression of rehabilitation trajectories using Bray-Curtis
393 versus Aitchison measures at Huntly, Eneabba, and Worsley (excluding southernmost
394 samples), these measures produced comparable predictions for recovery time within each
395 minesite (Figure 3; Appendix A, Table S6). At Huntly, predicted recovery times differed by
396 around 12 years, with a median recovery of 43 years for Bray-Curtis measures and 31 years
397 for Aitchison measures. At Eneabba, the median Bray-Curtis recovery time was 60 years,
398 while the median Aitchison recovery time was 50 years, however due to the spread of model
399 outcomes, predictions from these measures were not significantly different (i.e., Δ 95%
400 interval contains zero; Figure 3). Similarly, at Worsley (excluding southernmost samples),
401 the median Bray-Curtis recovery time was 44 years, while the median Aitchison recovery
402 time was 53 years, however predictions from these measures were not significantly different
403 (i.e., Δ 95% interval contains zero; Figure 3).

404 Among standard measures we found a general increase in similarity to reference
405 values across the ecological measures, from Jaccard (generally lowest similarities), Bray-
406 Curtis, Unweighted UniFrac, to Weighted UniFrac (generally highest similarities) (Appendix
407 A, Figure S21, Table S5). The greatest y-axis span, and therefore greatest sensitivity to detect
408 change, in similarity to reference values between the youngest rehabilitation ages and

409 references occurred with Bray-Curtis measures (Appendix A, Figure S21). The smallest span
410 (or flattest curves) in similarity to reference values between the youngest rehabilitation ages
411 and references occurred with Weighted UniFrac measures.

412 Except for the Unweighted Unifrac result at Huntly, Jaccard measures generally
413 returned the longest predicted recovery times, followed by reduced or similar recovery times
414 predicted using Bray-Curtis, Unweighted Unifrac and Weighted UniFrac measures
415 (Appendix A, Figure S22, Table S6). Low sample sizes (and corresponding low numbers of
416 distance measures) represent a limitation in our analysis, and the ecologically-distant samples
417 in the 17-year and 25-year rehabilitation age group at Huntly (Figure 2a) are likely
418 contributing to the reduced similarity and longer rehabilitation trajectory in Unweighted
419 UniFrac data. These 17-year and 25-year rehabilitation age group data at Huntly express
420 reduced alpha diversity and evenness compared to other samples, however reasons for this
421 are unclear (Appendix A, Figures S3–S4).

422

423 **3.3 Grouping by sequence similarity (Huntly only)**

424 Grouping by sequence similarity resulted in progressive overall shifts towards increasing
425 similarity to reference values from ASV-level (generally lowest similarities), 99%, 97%,
426 95%, to 90%-identity clustered OTUs (generally highest similarities) (Appendix A, Figure
427 S23). Predicted recovery times with more broadly clustered OTUs followed continuous and
428 seemingly predictable patterns of: (i) increasing recovery times with Jaccard measures (i.e.,
429 medians of 61 to 134 years), and (ii) decreasing to steadying recovery times with Bray-Curtis
430 measures (i.e., medians of 43 to 37 years) (Appendix A, Figure S26a, Table S6).

431

432 **3.4 Taxonomic grouping (Huntly only)**

433 Moving from ASV to genus-level data resulted in a pronounced shift towards increasing
434 similarity to reference, with similar although somewhat flatter rehabilitation trajectory curves
435 at higher taxonomic groupings (Appendix A, Figure S24). Visually, there appeared to be little
436 effect on the rehabilitation trajectory plots from pruning unclassified taxa (Appendix A,
437 Figure S24). Using Jaccard measures, moving from ASV-level to grouping at genus-level or
438 higher groupings dramatically increased predicted recovery times, compared to other
439 measures (Appendix A, Figure S26b, Table S6). Also, pruning of unclassified groups reduced
440 the smoothness or continuity in Jaccard-predicted recovery times (Appendix A, Figure S26b).
441 Using Bray-Curtis measures, we found a non-linear pattern of recovery times across the
442 taxonomic groupings, with shorter times to reach the target in genus, family, and order-level
443 groups, and longer recovery times in other groupings (Appendix A, Figure S26b; see
444 Appendix A, Figures S5–S13 for relative abundances of order, class, and phylum-level taxa
445 for each minesite). Richness and evenness of bacterial communities varied across
446 rehabilitation age groups and taxonomic groupings (e.g., data for phylum, class, and order-
447 level are shown in Appendix A, Figure S27), which may help explain the somewhat erratic
448 results from taxonomic grouping.
449

450 **3.5 Excluding rare taxa (Huntly only)**

451 Removing rare taxa to the point of retaining ASVs with >0.01% relative abundance produced
452 results from the Jaccard analysis that appeared to mimic results from the Bray-Curtis analysis
453 (Appendix A, Figure S25). When only more common ASVs with >0.1% relative abundance
454 were retained, both the Jaccard and Bray-Curtis results appeared to reflect over-simplified
455 communities, resulting in shorter predicted recovery times. However, including only ASVs
456 with >0.001% relative abundance resulted in a dataset with approximately 60% of the
457 original taxa and 95.8% of total sequences after rarefying (i.e., 17,941 compared to 30,751

458 ASVs and 603,072 compared to 629,460 sequences; Appendix A, Table S4) and produced
459 only a small increase in predicted recovery times for both Jaccard and Bray-Curtis measures
460 (Appendix A, Figure S26c, Table S6).
461

462 **3.6 Correcting for spatial autocorrelation**

463 We modelled the slope-trends of the relationships between ecological distance to references
464 and geographic distance to references, within rehabilitation age classes, for each of the
465 minesites using Bray-Curtis measures (see Appendix A, Huntly and Eneabba: Figures S28–
466 S29; Worsley: Figure 4). We also applied a 'correction' for the spatial autocorrelation, such
467 that rehabilitation age groups were adjusted to display the same ecological-geographic slope
468 trend as found in references (refer to the 'c' panels in Appendix A, Figures S28–S29; Figure
469 4). Figure 4d–f also includes the Worsley 'filtered' dataset and corresponding correction,
470 where the three southernmost geographically separate samples were excluded. Rehabilitation
471 trajectory plots, and predicted recovery times, using corrected data were compared to the
472 original uncorrected data (see Figure 5 and Appendix A, Table S6). In the case of Huntly and
473 Eneabba, only minor differences were found between original and corrected predicted
474 recovery times (Huntly: medians of 43 vs. 41 years, Δ 95% prediction interval = [1, 4];
475 Eneabba: medians of 60 vs. 56 years, Δ 95% prediction interval = [-1, 11]).
476 However, results from the Worsley data are featured because of the illustrative signal
477 we found there. Worsley displayed a strong ecological distance-geographic distance trend in
478 among-reference data indicating excessive spatial autocorrelation (note the upward sloping
479 'Ref' line in Figure 4a), and the greatest divergence of all the minesites in predicted recovery
480 times between original and corrected data (i.e., medians of 39 vs. 50 years, Δ 95% prediction
481 interval = [-49, 5]; Figure 5; Appendix A, Table S6). Notably, the spatial autocorrelation
482 correction at Worsley caused such an adjustment in similarity to reference values that the

youngest rehabilitation age group was included in the logarithmic trajectory models in the corrected data, but not in the original data. However, with exclusion of the southernmost Worsley samples (i.e., the filtered dataset), the signal of spatial autocorrelation disappeared (i.e., absence of upward sloping lines in Figure 4d, f) and predicted recovery times for filtered and filtered-corrected data displayed almost identical distributions (i.e., median recovery times were equivalent at 44 years in each scenario, Δ 95% prediction interval = [-9, 9]; Figure 5; Appendix A, Table S6).

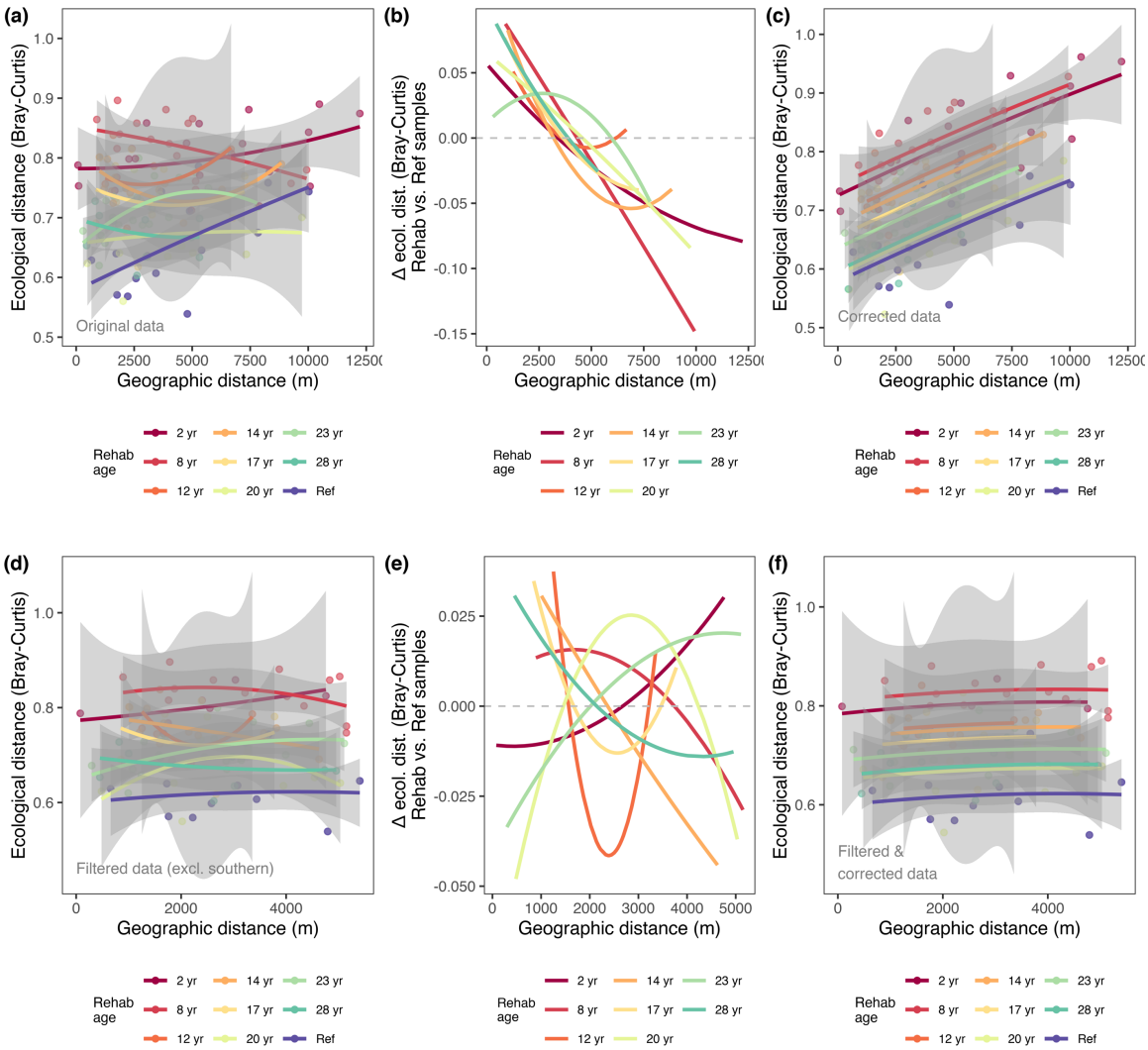
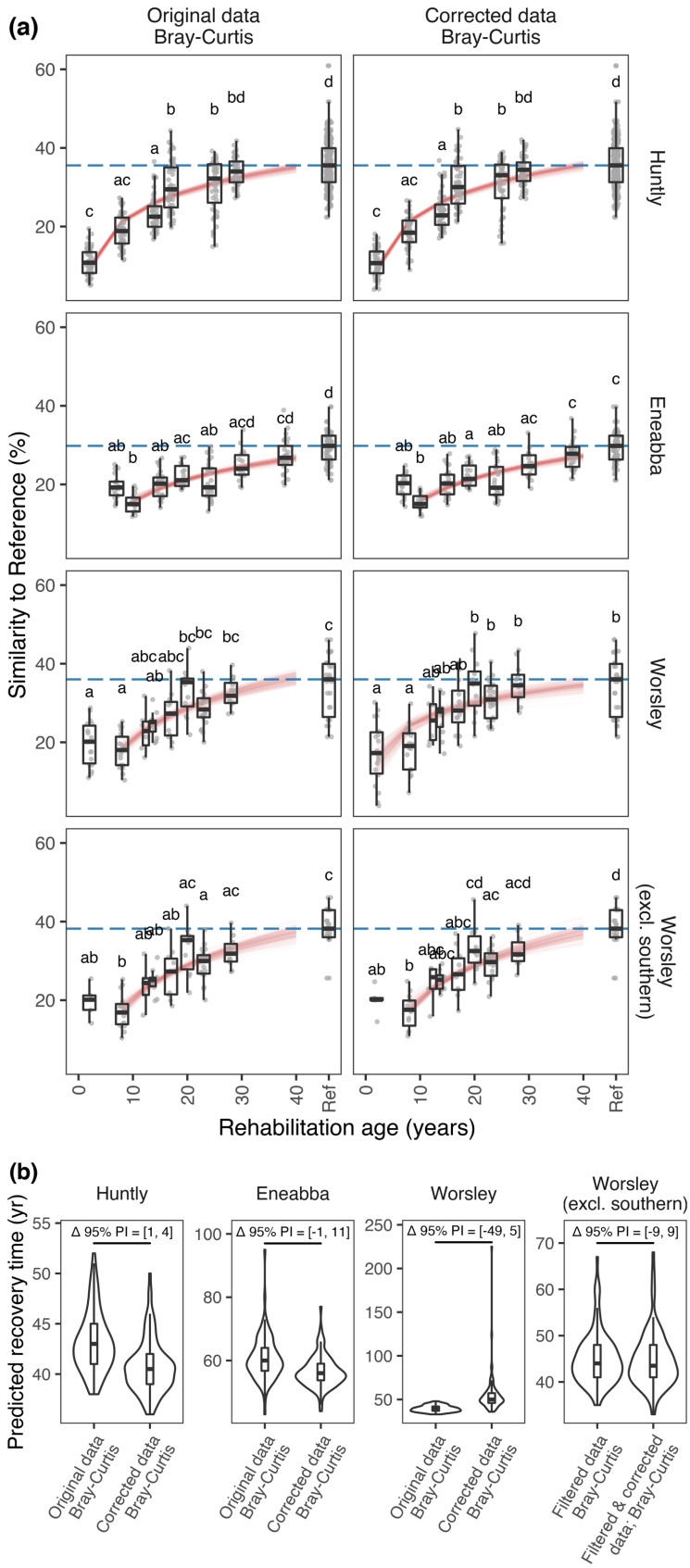


FIGURE 4. Exploring spatial autocorrelation in the Worsley (a–c) and filtered Worsley (excluding southernmost samples) (d–f) datasets, based on Bray-Curtis distance measures.

(a, d) Ecological distance to reference versus geographic distance to reference for rehabilitation age groups. (b, e) Mean-centred difference in ecological distance to reference between rehabilitation age groups and among references. (c, f) Corrected ecological distance to reference versus geographic distance to reference for rehabilitation age groups, to match the slope-trend of ecological to geographic distances as found among references.


(a)
Original data
Bray-Curtis
Corrected data
Bray-Curtis
Huntly
Eneabba
Worsley
Worsley
(excl. southern)
Similarity to Reference (%)
Rehabilitation age (years)
(b)
Huntly
Eneabba
Worsley
Worsley
(excl. southern)
Predicted recovery time (yr)
Δ 95% PI = [1, 4]
Δ 95% PI = [-1, 11]
Δ 95% PI = [-49, 5]
Δ 95% PI = [-9, 9]
Original data
Bray-Curtis
Corrected data
Bray-Curtis
Filtered data
Bray-Curtis
Filtered & corrected
data: Bray-Curtis

502 FIGURE 5. Modelled rehabilitation trajectories (a) and predicted recovery times (b) for
503 Huntly, Eneabba, Worsley, and Worsley (excluding southernmost samples) based on surface
504 soil bacterial community similarity to reference data using Bray-Curtis measures, with and
505 without correction for spatial autocorrelation. Other features are as described in Figure 3.
506

507 **4. DISCUSSION**

508 **4.1 Standard vs. compositional data analysis**

509 Our rehabilitation trajectory models produced comparable predictions for recovery times
510 using Bray-Curtis (standard) and Aitchison (compositional) measures. At Huntly, Eneabba,
511 and Worsley (excluding southernmost samples) median recovery times differed by around a
512 decade (i.e., 43 vs. 31 years, 60 vs. 50 years, 44 vs. 53 years respectively), however for two
513 out of three minesites the distribution of bootstrap model predicted recovery times was not
514 significantly different. We suspect that both Bray-Curtis (standard) and Aitchison
515 (compositional) measures will provide slightly different perspectives to the trajectory
516 modelling (discussed below), while neither method is perfect.

517 Compositional data analysis has been recently promoted as a more robust approach
518 for analysing microbiome datasets (Gloor et al., 2017; Quinn et al., 2019), however it is not
519 without limitations particularly for sparse datasets (containing many zeros), and where low
520 sequence counts are commonly encountered (Lovell et al., 2020). In particular, replacement
521 of zeros with small positive numbers has potential to cause distortions in data that will affect
522 the relative abundance of small counts to a greater degree than large counts (Lovell et al.,
523 2020). Distortions in data due to zero replacement are also increased where there are large
524 numbers of zeros present (Martín-Fernández et al., 2015). Therefore, our approach to exclude
525 taxa that contained zero counts in more than 90% of samples within each minesite represents
526 a compromise between losing representation of less common taxa and potentially introducing

527 spurious log ratio abundance patterns within the compositional data analysis. Following log
528 ratio analysis, only the relative information is of interest; for example, counts of 1,2,3
529 become equivalent to counts of 100, 200, 300. However, advocates of these approaches have
530 suggested that it is up to the analyst to decide whether the relative, rather than the absolute,
531 structure of the parts is of primary interest (Martín-Fernández et al., 2015). Also, the
532 replacement of absolute zeros (representing true absences; as opposed to zeros due to
533 rounding or resulting from insufficiently large samples) with small numbers is potentially
534 inappropriate (Martín-Fernández et al., 2015), and creates a theoretical challenge to
535 performing log ratio analyses on soil microbiota data from diverse environments where many
536 absolute zeros (true absences) are likely.
537

538 **4.2 Alternative standard ecological measures**

539 Bray-Curtis measures produced the greatest range in similarity values between young
540 rehabilitation and reference samples, and therefore are likely to offer the greatest sensitivity
541 to quantify the progress of recovery of soil bacterial communities towards reference states. In
542 contrast, Weighted UniFrac offered limited sensitivity to detect changes with rehabilitation
543 age (i.e., shallow trajectory curves) and may result in under-prediction of recovery times.
544 Low variation in Weighted Unifrac similarities likely reflects a level of consistency of high
545 proportions of somewhat closely related organisms across the samples. Jaccard distances
546 represent the proportion of unshared taxa out of the total number of taxa recorded in two
547 groups (Anderson et al., 2006). Unweighted UniFrac uses phylogenetic information and
548 calculates the fraction of the branch length in a phylogenetic tree that leads to descendants in
549 either, but not both, of the two communities (Lozupone et al., 2007). These qualitative
550 measures reflect the survival and presence of taxa (Jaccard) and related lineages (Unweighted
551 UniFrac), where loss of sequences may reflect extreme or limiting environmental conditions

552 (e.g., soil abiotic factors) or limited geographic distribution. Meanwhile, Bray-Curtis and
553 Weighted UniFrac measures emphasise abundant organisms. Similarity to reference generally
554 increased with increasing abundances of shared taxa for Bray-Curtis, and shared lineages of
555 related sequences for Weighted UniFrac. The quantitative measures often reflect the growth
556 or decline of certain organisms due to factors such as nutrient availability and variation in
557 environmental conditions (Lozupone et al., 2007).
558

559 **4.3 Grouping by sequence similarity**

560 Grouping near identical sequences will reduce the denominator used in calculating Jaccard
561 distances. For a given number of unshared taxa between samples, using broader OTU clusters
562 will make the proportion of unshared taxa (compared to all taxa) larger when there are a
563 smaller number of total taxa present. Our data suggest this shifting Jaccard calculation can
564 impact some samples strongly (e.g., note the 17-year age group in Appendix A, Figure S23)
565 resulting in a gradual increase in predicted recovery times with broader (reduced identity
566 threshold) OTU clusters. On the other hand, broader OTU clusters will aggregate some
567 sequences into already large groups and will tend to further emphasise abundant groups.
568 Consequently, our Bray-Curtis data suggest broader OTU clustering will make the target
569 similarity easier to reach and predicted recovery times reduced accordingly.
570

571 **4.4 Taxonomic grouping**

572 We do not recommend grouping 16S rRNA data by taxonomy to quantify recovery in soil
573 bacterial communities due to the erratic behaviour of predicted recovery times.
574

575 **4.5 Excluding rare taxa**

576 We show that filtering out of rare taxa to a limited extent (>0.001% relative sequence
577 abundance) produces a relatively small increase in predicted recovery times for both Jaccard
578 and Bray-Curtis measures. In our case, this filtering removed many ASVs but only a low
579 percentage of total sequences. Interestingly, this low level of exclusion of rare taxa does not
580 appear to moderate the assessment by producing reduced recovery times. At the low level of
581 exclusion, our analysis using rarefied data and similarity to reference measures may help
582 mitigate some of the impacts and concerns of removal of rare sequences experienced
583 elsewhere (e.g., Schloss, 2020). This raises the prospect to reduce sequencing depth, and
584 potential for shifting investment towards more robust assessments that incorporate a larger
585 number of samples with reduced sequencing depth and cost per sample.
586

587 **4.6 Influence of 'direct return' soils in young rehabilitation sites**

588 For reasons discussed here and below, we suggest it is prudent for these similarity to
589 reference trajectory assessments to exclude young rehabilitation sites with 'direct return' soils
590 that display elevated similarity to reference—as we implemented in our automated trajectory
591 modelling algorithm. In earlier preliminary work at Eneabba and Worsley, we observed that
592 the inclusion of young rehabilitation samples that were overly similar to references resulted
593 in seemingly biased, longer predictions of recovery time. The industry best practice of 'direct
594 return' of topsoil to new rehabilitation sites is based on objectives to minimise soil
595 degradation and expedite ecosystem recovery. However, our use of monotonic logarithmic
596 models applied to a data series that contains young rehabilitation sites with elevated
597 similarity to reference values, followed by older sites with reduced similarity to reference
598 values, results in the seemingly perverse outcome of a flatter, longer modelled trajectory of
599 recovery. The enhanced ecological similarity to reference in young rehabilitation sites with
600 'direct return' soils reflects a biological inertia, or temporary carryover effect, from unmined

areas where the soils originate, and confounds the relationship between soil microbiota development and rehabilitation age. For 'direct return' soils, we speculate the time taken for local influences to become dominant in shaping the resident microbiota may be in the order of 1-10 years, varying on a case-by-case basis, e.g., due to soil factors including organic matter and clay content, as well as the magnitude of environmental influences. Soil microbiota will be shaped by influences including local rainfall, temperature, aspect, soil water availability and transport (e.g., run-on, lateral flow), and vegetation communities via plant-soil feedbacks. Existing deeper soil and substrate may also influence rehabilitation surface soils via upward movement of water, nutrients, and some microbiota through mechanisms including: hydraulic redistribution by plant root systems (Neumann and Cardon, 2012); potential microbiota uptake and transfer via xylem into the phyllosphere (Deyett and Rolshausen, 2019; Fausto et al., 2018) and subsequent leaf litter; and capillary rise in heavier textured soils under conditions of soil water evaporation. Other factors affecting the similarity to reference of direct return soils include their source location (are they taken from sites that are generally closer to other reference sites or adjacent to rehabilitation sites?), the depth of fresh topsoil applied, the condition of subsurface layers (e.g., fresh vs. stockpiled), and the depth and method of tillage or mixing of the soil surface and subsurface layers following soil return. Our approach to automate the commencement of logarithmic models once there is at least an initial increase in similarity to reference values provides an objective approach to help overcome the potential model-biasing effect of biological inertia that is found in some direct return soils.

4.7 Spatial autocorrelation

As observed in the Worsley data, we found signals of excessive spatial autocorrelation where strong slopes were detected in plots of ecological distance to reference versus geographic

626 distance to reference, and where substantial differences were detected in the logarithmic
627 models and/or predicted recovery times between original and corrected datasets. Excluding
628 geographic outliers in the filtered Worsley analysis also removed a clear spatial
629 autocorrelation signal in the data, which indicates the importance of sampling designs. If
630 rehabilitation sites reflect environmental settings or imported soils that are overly similar or
631 dissimilar to references (i.e., different to natural background rates of spatial autocorrelation),
632 this may unduly bias predicted recovery times. Where possible, we recommend a sampling
633 approach that resembles the approach used at Huntly, where each reference site was spatially
634 paired with an adjacent rehabilitation site. This approach helps capture variation among
635 references (within a given minesite) relevant to the broader range of rehabilitation sites; and
636 provided there is adequate spatial replication and geographic outliers are avoided, then undue
637 influence from spatial autocorrelation should be avoided.

638 Our analysis of spatial autocorrelation should be viewed as introductory and
639 illustrative. For 'direct return' soils at young rehabilitation sites, our approach is deficient
640 because we do not account for their previous location. Although, we anticipate localised
641 influences would dominate the shaping of resident soil microbiota in rehabilitation sites after
642 a few years, as discussed above.

643 Plant-soil-microbiota feedbacks represent a complicating factor for disentangling
644 effects of soil abiotic condition, rehabilitation age, and residual/unexplainable spatial
645 autocorrelation in restoration chronosequence studies. This is because chronosequence
646 studies (which presume a 'space-for-time' proxy relationship between treatments and
647 outcomes) typically do not collect sufficient data to determine whether soil conditions have
648 influenced rehabilitation outcomes, plants have conditioned soils, or both situations have
649 occurred. Studies that have considered plant-soil feedbacks in restored Jarrah forest (Huntly)
650 sites have shown differential correlative effects of rehabilitated soil biotic and abiotic

651 properties (Orozco-Aceves et al., 2017). Also, plant-soil feedbacks behave differently in
652 unmined versus rehabilitated soils (Orozco-Aceves et al., 2015). Further work is required to
653 build understanding of this topic (e.g., via longitudinal studies).

654 Study designs should account for geographically variable factors that may influence
655 soil bacterial communities to build robust evidence. Microbial communities are influenced by
656 a range of often complex edaphic, physiographic, climatic and biotic factors (e.g., vegetation
657 and land use history, soil texture, available nutrients and moisture, drainage, pH, salinity,
658 sunlight; Brown et al., 2018; Delgado-Baquerizo et al., 2018; Zhu et al., 2021). The need to
659 reflect such environmental variation through choosing appropriate references is well
660 recognised by industry (Australian_Government, 2016; Manero et al., 2021).

661 Accordingly, where restoration study locations are characterised by distinct modes of
662 geographically distributed soil and landscape conditions or pre-disturbance ecosystem types
663 (e.g., uplands vs. lowlands; riparian vs. non-riparian; dune vs. swale) it would be appropriate
664 to reflect this striking contrast in reference environmental conditions and desired
665 rehabilitation outcomes by undertaking separate rehabilitation trajectory assessments for each
666 major representative target ecosystem type. The optimal distribution of each set of
667 rehabilitation and reference sites should reflect the spatial composition of major landform- or
668 post-restoration target ecosystem-types specific to each study area, and might be informed via
669 pre-existing mapping or imagery, digital elevation models, and other land resource
670 assessment tools and techniques (e.g., digital clustering of landscape types; de Bruin and
671 Stein, 1998).

672

673 **4.8 Other limitations**

674 There are important limitations in our study, in addition to those already discussed. The
675 robustness of our study would be improved with more samples per minesite to help better

676 capture minesite-wide variation. We did not consider soil microbiota patterns at depth, which
677 are also important. Also, major changes to rehabilitation practices over time will disrupt the
678 'space-for-time' substitutive modelling approach that is relied upon in chronosequence
679 studies such as ours. For any restoration chronosequence study careful sample selection is
680 required to avoid confounding factors as much as possible (Walker et al., 2010). There are
681 potential limitations in our study associated with the phylogenetic trees we used to generate
682 UniFrac distances (see Appendix A, Supplementary Methods for details). Tree-building often
683 represents a compromise between accuracy in representing phylogenetic relationships and
684 computing time, and it was beyond the scope of our study to test the sensitivity of our
685 UniFrac-based analyses to the quality of trees used. We used logarithmic models which
686 assume a monotonic recovery function, however other models that account for variable trends
687 over time, and varying success for different rehabilitation techniques or sites, may offer
688 improved estimates of recovery time. We suggest these limitations should be investigated in
689 future studies.
690

691 **5. CONCLUSIONS**

692 We provide a proof-of-concept demonstration of an innovative, chronosequence-based,
693 similarity to reference trajectory assessment method, to quantitatively track progress in soil
694 microbiota with post-mining rehabilitation. Through incorporating microbiota survey data
695 from multiple reference sites of varying character, we revealed substantial variation among
696 reference ecosystems within each minesite that can inform realistic rehabilitation targets. Our
697 method reduces the complexity associated with microbiota data and enables prediction of
698 recovery time to reach reference-based targets with explicit inclusion of uncertainty in
699 assessments. Also, the use of soil microbiota data provides another line of evidence, which in
700 conjunction with wider minesite information, could assist in the examination of potential

701 impediments to the progress of rehabilitation, thereby helping to inform adaptive
702 management. From our investigations, we recommend using ASV-level Bray-Curtis
703 similarities which appear to offer a relatively sensitive and stable basis for modelling
704 rehabilitation trajectories. We recommend wherever possible to maximise sample sizes,
705 employ spatial pairing of reference and rehabilitation sites, and to exclude geographically-
706 distant, non-representative sampling areas. We used an automated modelling routine to
707 exclude young rehabilitation sites with 'direct return' soils that displayed elevated similarity
708 to reference values, which would have biased the trajectory modelling. Further fine-tuning to
709 identify possible minor reductions in sequencing depths (eliminating some rare taxa) offers
710 promise to reduce per sample costs, enabling investment in more samples, to help deliver
711 more robust assessments. This work represents an important step towards a reduced-
712 complexity microbiota-based monitoring and evaluation framework consistent with many
713 best practice principles for setting, monitoring and managing towards mine completion
714 criteria recommended by (Manero et al., 2021). We anticipate that our approach could be
715 expanded to other eDNA sequence-based survey data (e.g., fungal ITS and eukaryote 18S
716 rRNA marker genes, functional potential from shotgun metagenomic data), and may have
717 application in wider contexts where there is interest in monitoring restorative processes that
718 facilitate a shift in microbiota towards reference states.

ACKNOWLEDGEMENTS

We acknowledge the contribution of the Australian Microbiome consortium in the generation of data used in this publication. The Australian Microbiome is supported by funding from Bioplatforms Australia and the Integrated Marine Observing System (IMOS) through the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS), Parks Australia through the Bush Blitz program funded by the Australian Government and BHP, and the CSIRO. This research was also supported by the Australian Research Council (LP190100051).

AUTHOR CONTRIBUTIONS

CL, SLK, MT and MFB conceived the ideas and designed the study; SLK, RJB, LCD, PB, MPD, AG collected the data; CL, SLK, AB, MFB analysed and interpreted the data with contributions from all authors; CL led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

Data and code are available at: https://data.bioplatforms.com/organization/about/australian-microbiome and on figshare at https://doi.org/10.25451/flinders.16920985

ORCID

Craig Liddicoat – https://orcid.org/0000-0002-4812-7524

Siegfried L. Krauss – https://orcid.org/0000-0002-7280-6324

Andrew Bissett – https://orcid.org/0000-0001-7396-1484

Ryan J. Borrett – https://orcid.org/0000-0001-8663-0844

Shawn D. Peddle – https://orcid.org/0000-0003-3464-3058

Mark P. Dobrowolski – http://orcid.org/0000-0001-5586-4023

Andrew Grigg – https://orcid.org/0000-0002-5818-2973

Mark Tibbett – https://orcid.org/0000-0003-0143-2190

Martin F. Breed – https://orcid.org/0000-0001-7810-9696

REFERENCES

Anderson, M.J., Ellingsen, K.E., McArdle, B.H., 2006. Multivariate dispersion as a measure of beta diversity. Ecology Letters 9, 683-693.

Australian_Government, 2016. Mine Rehabilitation: Leading Practice Sustainable Development Program for the Mining Industry.

BenDor, T., Lester, T.W., Livengood, A., Davis, A., Yonavjak, L., 2015. Estimating the Size and Impact of the Ecological Restoration Economy. PLOS ONE 10, e0128339.

Bissett, A., Fitzgerald, A., Meintjes, T., Mele, P.M., Reith, F., Dennis, P.G., Breed, M.F., Brown, B., Brown, M.V., Brugger, J., Byrne, M., Caddy-Retalic, S., Carmody, B., Coates, D.J., Correa, C., Ferrari, B.C., Gupta, V.V.S.R., Hamonts, K., Haslem, A., Hugenholtz, P., et_al, 2016. Introducing BASE: the Biomes of Australian Soil Environments soil microbial diversity database. GigaScience 5, 21.

Breed, M.F., Harrison, P.A., Blyth, C., Byrne, M., Gaget, V., Gellie, N.J.C., Groom, S.V.C., Hodgson, R., Mills, J.G., Prowse, T.A.A., Steane, D.A., Mohr, J.J., 2019. The potential of genomics for restoring ecosystems and biodiversity. Nature Reviews Genetics 20, 615-628.

Brown, A.G., Lespez, L., Sear, D.A., Macaire, J.-J., Houben, P., Klimek, K., Brazier, R.E., Van Oost, K., Pears, B., 2018. Natural vs anthropogenic streams in Europe: History, ecology and implications for restoration, river-rewilding and riverine ecosystem services. Earth-Science Reviews 180, 185-205.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., Themaat, E.V.L.v., Schulze-Lefert, P., 2013. Structure and Functions of the Bacterial Microbiota of Plants. Annual Review of Plant Biology 64, 807-838.

de Bruin, S., Stein, A., 1998. Soil-landscape modelling using fuzzy c-means clustering of attribute data derived from a Digital Elevation Model (DEM). Geoderma 83, 17-33.

Delgado-Baquerizo, M., Reith, F., Dennis, P.G., Hamonts, K., Powell, J.R., Young, A., Singh, B.K., Bissett, A., 2018. Ecological drivers of soil microbial diversity and soil biological networks in the Southern Hemisphere. Ecology 99, 583-596.

Deyett, E., Rolshausen, P.E., 2019. Temporal Dynamics of the Sap Microbiome of Grapevine Under High Pierce's Disease Pressure. Frontiers in Plant Science 10.

Fausto, C., Mininni, A.N., Sofo, A., Crecchio, C., Scagliola, M., Dichio, B., Xiloyannis, C., 2018. Olive orchard microbiome: characterisation of bacterial communities in soil-plant compartments and their comparison between sustainable and conventional soil management systems. Plant ecology & diversity 11, 597-610.

Fernandez Nuñez, N., Maggia, L., Stenger, P.-L., Lelievre, M., Letellier, K., Gigante, S., Manez, A., Mournet, P., Ripoll, J., Carriconde, F., 2021. Potential of high-throughput eDNA sequencing of soil fungi and bacteria for monitoring ecological restoration in ultramafic substrates: The case study of the New Caledonian biodiversity hotspot. Ecological Engineering 173, 106416.

Gann, G.D., McDonald, T., Walder, B., Aronson, J., Nelson, C.R., Jonson, J., Hallett, J.G., Eisenberg, C., Guariguata, M.R., Liu, J., Hua, F., Echeverría, C., Gonzales, E., Shaw, N., Decleer, K., Dixon, K.W., 2019. International principles and standards for the practice of ecological restoration. Second edition. Restoration Ecology 27, S1-S46.

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome Datasets Are Compositional: And This Is Not Optional. Frontiers in microbiology 8, 2224-2224.

Harris, J., 2009. Soil Microbial Communities and Restoration Ecology: Facilitators or Followers? Science 325, 573-574.

IPBES, 2018. The IPBES assessment report on land degradation and restoration, in: Montanarella, L., Scholes, R., and Brainich, A. (eds.) (Ed.), Bonn, Germany, p. 744.

Janzen, H.H., 2016. The Soil Remembers. Soil Science Society of America Journal 80, 1429-1432.

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q., Wei, G., 2018. Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. Microbiome 6, 146.

Legendre, P., Legendre, L., 2012. Chapter 7 - Ecological resemblance, in: Legendre, P., Legendre, L. (Eds.), Developments in Environmental Modelling. Elsevier, pp. 265-335.

Liddicoat, C., Weinstein, P., Bissett, A., Gellie, N., Mills, J., Waycott, M., Breed, M., 2019. Can bacterial indicators of a grassy woodland restoration inform ecosystem assessment and microbiota-mediated human health? Environment International 129, 105-117.

Lloyd-Price, J., Abu-Ali, G., Huttenhower, C., 2016. The healthy human microbiome. Genome Medicine 8, 51.

Lovell, D.R., Chua, X.-Y., McGrath, A., 2020. Counts: an outstanding challenge for log-ratio analysis of compositional data in the molecular biosciences. NAR Genom Bioinform 2.

Lozupone, C.A., Hamady, M., Kelley, S.T., Knight, R., 2007. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Applied and environmental microbiology 73, 1576-1585.

Manero, A., Standish, R., Young, R., 2021. Mine completion criteria defined by best-practice: A global meta-analysis and Western Australian case studies. Journal of Environmental Management 282, 111912.

Martín-Fernández, J.-A., Hron, K., Templ, M., Filzmoser, P., Palarea-Albaladejo, J., 2015. Bayesian-multiplicative treatment of count zeros in compositional data sets. Statistical Modelling 15, 134-158.

McMurdie, P.J., Holmes, S., 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8, e61217.

Menz, M.H.M., Dixon, K.W., Hobbs, R.J., 2013. Hurdles and Opportunities for Landscape-Scale Restoration. Science 339, 526.

Mohr, J., Harrison, P., Stanhope, J., Breed, M.F., 2022. Is the genomics 'cart' before the restoration ecology 'horse'? Insights from qualitative interviews and trends from the literature. Philosophical Transactions of the Royal Society B: Biological Sciences.

Neumann, R.B., Cardon, Z.G., 2012. The magnitude of hydraulic redistribution by plant roots: a review and synthesis of empirical and modeling studies. New Phytologist 194, 337-352.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2020. vegan: Community Ecology Package. R package version 2.5-7.

Orozco-Aceves, M., Standish, R.J., Tibbett, M., 2015. Soil conditioning and plant-soil feedbacks in a modified forest ecosystem are soil-context dependent. Plant and soil 390, 183-194.

Orozco-Aceves, M., Tibbett, M., Standish, R.J., 2017. Correlation between soil development and native plant growth in forest restoration after surface mining. Ecological engineering 106, 209-218.

Palarea-Albaladejo, J., Martín-Fernández, J.A., 2015. zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. Chemometrics and Intelligent Laboratory Systems 143, 85-96.

Quinn, T.P., Erb, I., Gloor, G., Notredame, C., Richardson, M.F., Crowley, T.M., 2019. A field guide for the compositional analysis of any-omics data. GigaScience 8.

Quinn, T.P., Richardson, M.F., Lovell, D., Crowley, T.M., 2017. propr: An R-package for Identifying Proportionally Abundant Features Using Compositional Data Analysis. Scientific Reports 7, 16252.

R-Core-Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rydgren, K., Halvorsen, R., Töpper, J.P., Auestad, I., Hamre, L.N., Jongejans, E., Sulavik, J., 2019. Advancing restoration ecology: A new approach to predict time to recovery. Journal of Applied Ecology 56, 225-234.

Schloss, P.D., 2020. Removal of rare amplicon sequence variants from 16S rRNA gene sequence surveys biases the interpretation of community structure data. bioRxiv, 2020.2012.2011.422279.

Schmid, C.A.O., Reichel, R., Schröder, P., Brüggemann, N., Schloter, M., 2020. 52 years of ecological restoration following a major disturbance by opencast lignite mining does not reassemble microbiome structures of the original arable soils. Science of The Total Environment 745, 140955.

Tibbett, M., 2010. Large-scale mine site restoration of Australian eucalypt forests after bauxite mining: soil management and ecosystem development, Ecology of Industrial Pollution, pp. 309-326.

Tibbett, M., 2015. Mining in Ecologically Sensitive Landscapes. Victoria: CSIRO Publishing, Victoria.

Turner, T.R., Ramakrishnan, K., Walshaw, J., Heavens, D., Alston, M., Swarbreck, D., Osbourn, A., Grant, A., Poole, P.S., 2013. Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. The ISME Journal 7, 2248-2258.

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 749, 142262.

Walker, L.R., Wardle, D.A., Bardgett, R.D., Clarkson, B.D., 2010. The use of chronosequences in studies of ecological succession and soil development. Journal of Ecology 98, 725-736.

Watkins, H., Robinson, J.M., Breed, M.F., Parker, B., Weinstein, P., 2020. Microbiome-Inspired Green Infrastructure: A Toolkit for Multidisciplinary Landscape Design. Trends in Biotechnology 38, 1305-1308.

Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists, 2nd Edn. ed. John Wiley & Sons, Ltd, West Sussex, UK.

Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J.R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E.R., Knight, R., 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5, 27.

Zhu, M., Li, Y., Zhang, W., Wang, L., Wang, H., Niu, L., Hui, C., Lei, M., Wang, L., Zhang, H., Yang, G., 2021. Determination of the direct and indirect effects of bend on the urban river ecological heterogeneity. Environmental Research, 112166.