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[4,8,11], Andrew Grigg[9], Mark Tibbett[4,10], 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]Alcoa of Australia Limited, Perth, Australia

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

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

Correspondence:

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

Email: craig.liddicoat@flinders.edu.au

Abstract

1. Soil microbiota are fundamentally linked to the restoration of degraded ecosystems, as they are central to important ecological functions including the support of plant communities. High throughput sequencing of environmental DNA used to characterise soil microbiota offers promise to monitor ecological progress towards reference states. 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.
2. We analysed patterns of surface soil bacterial community similarity to reference ('rehabilitation trajectory') data from three long-term (> 25 year) post-mining rehabilitation chronosequence case studies from south-west Western Australia. We examined the influence of different ecological distance measures, sequence grouping approaches, and eliminating rare taxa on rehabilitation trajectories and predicted recovery times. We also explored the issue of spatial autocorrelation in our rehabilitation trajectory assessments and trialled a first-pass approach for correcting its undue influence.
3. We found considerable variation in bacterial communities among reference sites within each case study minesite, providing valuable context for setting targets and evaluating recovery. Median Bray-Curtis similarities among references within each minesite ranged from 30–36%, based on amplicon sequence variant-level data. Median predicted times for rehabilitated sites to recover to these levels ranged from around 40 to over 100 years. We discuss strengths and limitations of the different approaches and offer recommendations to improve the robustness of this assessment method.
4. *Synthesis and applications.* We demonstrate a proof-of-concept, complexity-reducing application of soil eDNA sequence-based surveys of bacterial communities in restoration chronosequence studies to quantitatively assess progress towards reference communities

and corresponding rehabilitation targets. Our method provides a step towards developing microbiota-based SMART metrics for measuring rehabilitation success in post-mining, and potentially other, restoration contexts. Our approach enables prediction of recovery time, explicitly including uncertainty in assessments, and assists examination of potential barriers to ecological recovery, including biologically-associated variation in soil properties.

KEYWORDS:

Beta diversity, ecological distance, eDNA, mine closure assessment, restoration genomics, rehabilitation trajectory, soil microbiota, spatial autocorrelation

65 **1. INTRODUCTION**

66 Land degradation and transformation, with negative impacts to biodiversity and ecosystem
67 function, is estimated to impact 75% of the Earth's land surface, and this figure is projected to
68 rise to over 90% by 2050 (IPBES 2018). Ecological restoration—activity that supports
69 rehabilitation of locally representative, sustainable, biodiverse ecosystems (Gann *et al.*
70 2019)—is seen as integral to reversing these impacts, as highlighted by the UN declaration of
71 2021–2030 as the Decade on Ecosystem Restoration (https://www.decadeonrestoration.org/).
72 Restoration is technically challenging and requires considerable investment, without
73 guaranteed success (Tibbett 2015). With large investments in restoration (e.g. Menz, Dixon
74 & Hobbs 2013 estimate US$18 billion/yr is required to restore degraded lands globally;
75 BenDor *et al.* 2015 estimate US$9.5 billion/yr is spent in the USA alone), there is a need to
76 improve the evidence base to guide continuous improvement in restoration outcomes and to
77 underpin future investment.

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

definitions of completion criteria, Manero, Standish and Young (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; Turner *et al.* 2013; Delgado-Baquerizo *et al.* 2018). 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. 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., 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 techniques often characterise the complexity and site-specificity of soil ecosystems.

115 However, a focus on measuring 'similarity to reference' may help cut through the complexity
116 inherent to microbiota data. Along these lines, van der Heyde *et al.* (2020) visualised
117 temporal trends in ecological similarity to reference data in post-mining rehabilitation—
118 however, each rehabilitation sample was only compared to a single closest reference sample,
119 which potentially limited insight into variability and uncertainty in microbiota recovery.
120 Here we conduct a detailed exploration of a complexity-reducing application of
121 eDNA-based soil bacterial community data to assess post-mining rehabilitation in three long-
122 term (> 25 year) chronosequence case studies from south-west Western Australia.
123 Specifically, we aim to demonstrate the use of chronosequence-based rehabilitation
124 trajectories, using measures of percent similarity to ecological reference sites (hereafter
125 termed references), to assess progress of soil bacterial communities towards reference states
126 with increasing rehabilitation age. We note that further work that links microbiota to other
127 ecosystem components (e.g., vegetation, fauna) is important but beyond the scope of our
128 study.
129 Our *a priori* research questions were: (1) Can soil bacterial community data be used
130 to establish reference-based targets? (2) Can soil bacterial community rehabilitation
131 trajectory data be used to predict the time to recover to reference targets? (3) How are these
132 predictions of recovery influenced by different ecological distance/similarity measures and
133 sequence data resolution? For example, grouping bacterial taxa based on sequence similarity
134 might help reduce noise associated with DNA sequencing methods; taxonomic grouping
135 might assist interpretation if recognised groups can be discussed; and eliminating rare taxa (to
136 simulate reduced sequencing depths) might allow more cost-effective and rapid analyses. We
137 also recognise the potential for spatial autocorrelation—where measured outcomes are closer
138 in value due to closer spatial proximity—to confound the assessment of rehabilitation age in
139 chronosequence studies that lack appropriate spatial design and replication. Therefore, we

140 also undertake a preliminary, illustrative examination of spatial autocorrelation, and trial a
141 first-pass approach to highlight, and correct for, its excessive influence. We then discuss
142 limitations and synthesise our findings to inform future work.
143

144 ## 2. MATERIALS AND METHODS

145 ### 2.1 Data collection

146 We used surface soil bacterial 16S rRNA marker gene data from three case study minesites
147 (Figure 1; online Supporting Information (SI) Appendix, Tables S1–S3) from south-west
148 Western Australia. Soil sampling was undertaken in accordance with Australian Microbiome
149 (AM) protocols (Bissett *et al.* 2016; https://www.australianmicrobiome.com/protocols; SI
150 Appendix, Supplementary Methods). Each minesite experiences a Mediterranean-type
151 climate with hot, dry summers and cool, wet winters. Post-mining rehabilitation activities
152 typically involved deep-ripping, prior to the 'direct return' (where possible) of subsoil and
153 topsoil stripped from a separate pit about to be mined, followed by revegetation with locally
154 appropriate seed of diverse plant communities (Tibbett 2010). Precise soil handling and
155 storage techniques differed between the minesites and different pits within minesites.
156 Summary information for each minesite is provided below (see SI Appendix, Supplementary
157 Methods for more background information; other studies in-progress will provide expanded
158 analyses of surface and subsoil data from these minesites, including additional marker gene
159 datasets).

160 Alcoa's *Huntly* bauxite-producing minesite is approximately 100 km south-east of
161 Perth, occurring in mixed open forest with dominant overstorey species of Jarrah (*Eucalyptus*
162 *marginata*) and Marri (*Corymbia calophylla*) on lateritic, nutrient poor soils. We consider
163 Huntly data sampled in 2016, with rehabilitation ages between 2–29 years old. Huntly's 36
164 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.

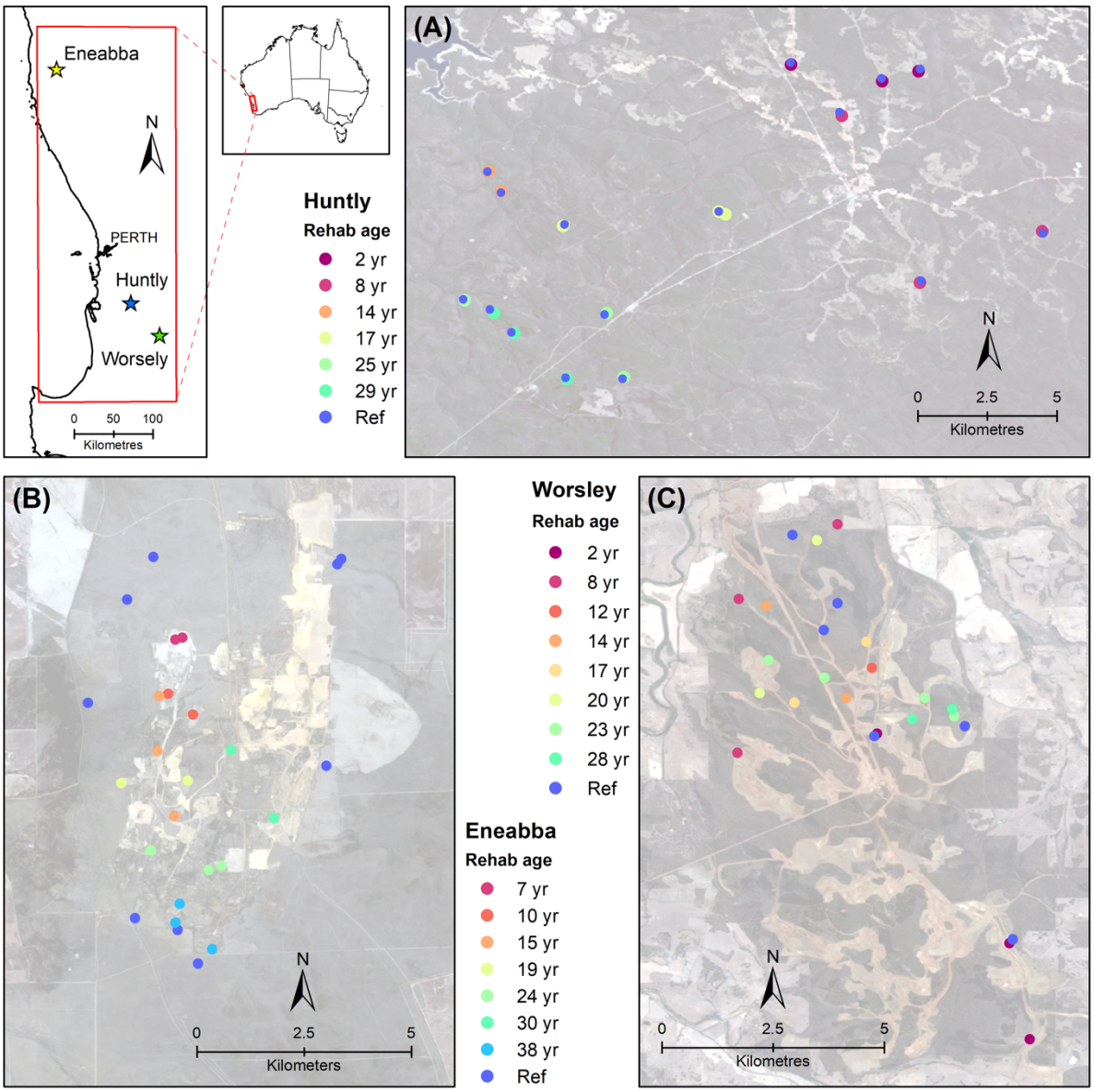


FIGURE 1. Locations of minesites and soil sampling sites: (A) Huntly, (B) Eneabba, (C) Worsley. (Imagery: Sentinel-2; https://eos.com/landviewer; EOS Data Analytics, Inc.)

2.2 eDNA sequencing, bioinformatics, and data preparation

DNA extraction, PCR and preliminary bioinformatic analyses were undertaken in accordance with AM workflows (Bissett *et al.* 2016; see SI Appendix, Supplementary Methods). From this workflow, denoised 16S rRNA gene amplicon sequence variant (ASV) level abundance data were produced for all minesites. Note, in this study ASVs are equivalent to zero radius

194 operational taxonomic units (zOTUs). Further data preparation and analyses were largely
195 undertaken in R version 4.0.3 (R-Core-Team 2020) utilising the framework of the R phyloseq
196 package (McMurdie & Holmes 2013) to manage the datasets (see SI Appendix
197 Supplementary Methods for number of sequences and ASVs studied in each minesite, initial
198 data cleaning steps, and preparation of phylogenetic trees).
199

200 **2.3 Data visualisation and statistical analyses**

201 We visualised the sequence depth of samples using rarefaction curves (SI Appendix, Figure
202 S2). For the majority of subsequent analyses, we normalised the sequence data for sampling
203 effort by rarefying abundances of ASVs, and other taxonomic levels investigated below, to
204 the minimum sample sequence depth within respective minesites (Huntly, n = 17,485
205 sequences; Eneabba, n = 10,142 sequences; Worsley, n = 54,122 sequences) using the
206 *rarefy_even_depth()* function from R phyloseq. For ASV-level data, the total rarefied
207 sequences comprised at Huntly, n = 629,460 sequences (30,751 ASVs); Eneabba, 263,692
208 sequences (27,115 ASVs); and Worsley, 1,353,050 sequences (54,327 ASVs). Exploratory
209 data analyses to visualise ASV alpha diversity, evenness, and relative abundance via
210 heatmaps of phyla, classes, and orders in each minesite are presented in the SI Appendix,
211 Supplementary Methods, and Figures S3–S13. Exploratory data analyses also included
212 preliminary visualisations of soil and landscape variables that associated with the soil
213 bacterial community samples within each minesite (refer to SI Appendix Supplementary
214 Methods, Supplementary Data, Figures S14–20).
215 We examined a range of alternative qualitative and quantitative beta diversity
216 measures (i.e., ecological distance or community dissimilarity, converted to similarity) to
217 model rehabilitation trajectories and time to reach reference targets (described below). For
218 the minesite with the largest number of samples (Huntly), we also investigated data pre-

219 processing options of grouping by sequence similarity, taxonomic grouping, and excluding
220 rare taxa. Details of the number of taxa considered and percentage of sequences remaining
221 after grouping, rarefying, and exclusions (see below) are indicated in the SI Appendix, Table
222 S4.
223

224 **2.3.1 Alternative ecological distance measures**

225 For each minesite, we used the cleaned and rarefied ASV-level bacterial community data to
226 derive ecological distance matrices using distance measures commonly employed in
227 microbiota studies—i.e., Jaccard, Bray-Curtis, Unweighted UniFrac and Weighted UniFrac
228 (Lozupone *et al.* 2007)—via the *vegdist()* function from the R vegan package (Oksanen *et al.*
229 2020).
230

231 **2.3.2 Grouping by sequence similarity**

232 For Huntly data only, separate R phyloseq objects were generated to represent soil bacterial
233 community data with sequences clustered into 99%, 97%, 95%, and 90% identity OTUs (see
234 SI Appendix, Supplementary Methods). For these analyses, OTUs were formed, abundance
235 data were rarefied, and then Jaccard and Bray-Curtis distances were calculated.
236

237 **2.3.3 Taxonomic grouping**

238 For Huntly data only, we examined the influence of taxonomic grouping (i.e., ASV, genus,
239 family, order, class, and phylum) on the assessments of recovery. We also tested the
240 influence of discarding versus retaining (at the next available classified grouping) taxa that
241 were unclassified at each taxonomic rank, which we termed 'pruned' and 'non-pruned' data
242 respectively. Grouping was undertaken using *tax_glom()*; and in 'pruned' datasets,
243 unclassified taxa were removed using *prune_taxa()* from R phyloseq. For these analyses, taxa

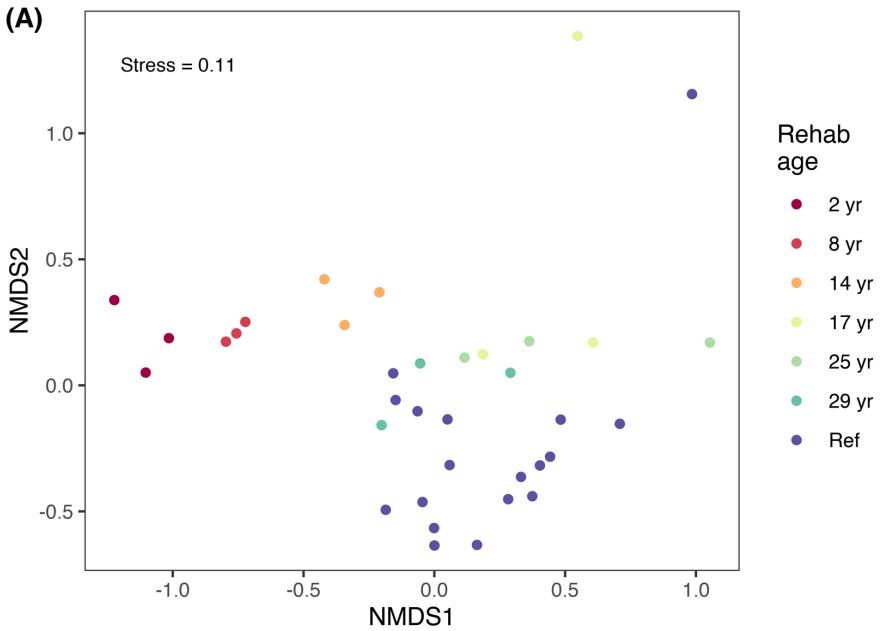
were grouped, abundance data were rarefied, then Jaccard and Bray-Curtis distances were calculated. Richness and evenness of sequences at the order, class and phylum level were also visualised based on rarefied data and plotted together with composite estimates within rehabilitation age groups from merged-sample bootstrap resampling (Liddicoat *et al.* 2019) (B=100).

2.3.4 Excluding rare taxa

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

2.3.5 Rehabilitation trajectory modelling

The rehabilitation trajectory analyses presented here were derived from a subset of data contained in the abovementioned ecological distance matrices. Specifically, only pairwise distances between samples and reference samples were considered (including distances among reference samples within minesites). Data were then expressed as percent similarity values (i.e., 100 × (1 - distance) [%]). Rehabilitation trajectory boxplots were then generated from the series of similarity to reference data on the y-axis and increasing rehabilitation age on the x-axis, concluding with reference samples (e.g., see Figure 2B). Testing for differences in similarities to reference at each rehabilitation age (as visualised with boxplots) was performed using the Kruskal-Wallis rank sum test, followed by post-hoc Dunn tests for multiple comparisons, with Bonferroni adjusted threshold *P*-values. The multiple comparison testing used default two-sided *P*-values and alpha = 0.05 nominal level of significance.



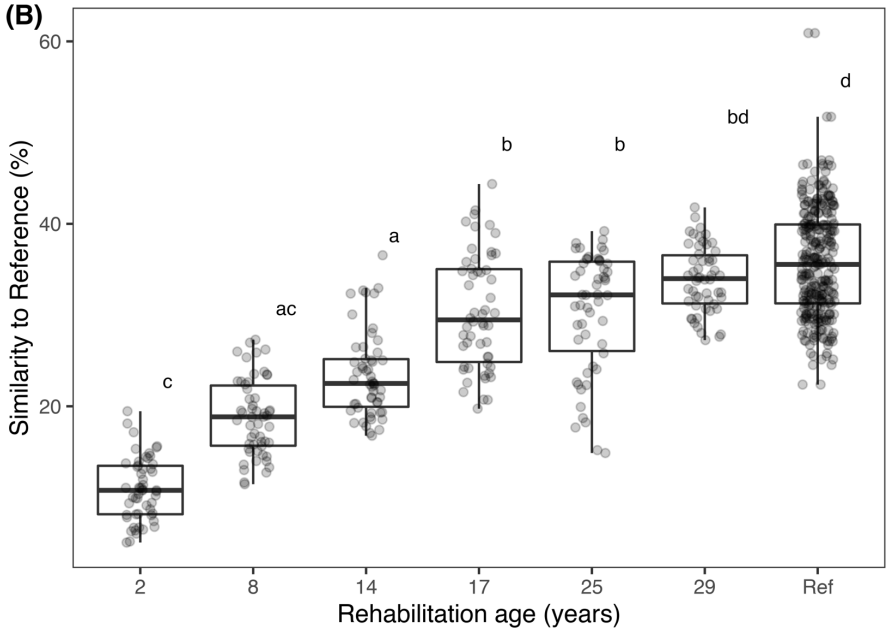


FIGURE 2. Example data for the Huntly minesite chronosequence soil bacterial communities. (A) NMDS ordination to visualise beta diversity or compositional differences based on ASV-level data and Bray-Curtis distances. (B) Rehabilitation trajectory boxplots express a trend in ecological similarity to references with increasing rehabilitation age. The similarity data in (B) are based on a subset of the distance matrix data that underpins (A), i.e. (B) uses only those distances that involve reference samples, and the value is expressed as

similarity to reference (%). Groups not sharing letters are significantly different. Note the x-axis is presented here as a categorical (not numerical) scale. Sample sizes for the rehabilitation age groups (used to produce distance data) are described in section 2.1.

After observing the variation in similarity to reference values among references within each minesite (e.g., Figure 2B), we defined rehabilitation targets for the purpose of this study as the median (= the central value) of among-reference similarities. This target median value varied by minesite, distance/similarity measure, and pre-processing option. We predicted the time to reach a restoration target (= recovery time) by modelling the trend in similarity to reference with increasing rehabilitation age using bootstrapped (B = 100) logarithmic models. The median, 5th and 95th percentile of predicted recovery time were evaluated. Our use of logarithmic models was consistent with the approach of Rydgren *et al.* (2019), except we used ecological 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.

2.3.6 Exploring spatial autocorrelation

To explore the influence of spatial autocorrelation on our trajectory analyses, we produced variogram-like plots with ecological distance (i.e., Jaccard, Bray-Curtis; between samples and references) on the y-axis, and geographic distances (between samples and references) on

303 the x-axis. Each rehabilitation age group was modelled as a second-order polynomial,
304 allowing the possible expression of curvilinear trendlines that mimicked variogram-like
305 relationships (i.e., increasing then flattening). Assuming reference curves offered a natural
306 baseline trend for spatial autocorrelation within each minesite environment, we applied a
307 'correction' to the curvilinear trendline for each rehabilitation age group by calculating the
308 difference in mean-centred model curves (= rehabilitation age group minus reference), such
309 that 'corrected' data for rehabilitation age groups expressed the same ecological distance-
310 geographic distance curvilinear trend as seen for references (see SI Appendix Supplementary
311 Methods for further details of the rationale and approach for this preliminary analysis).
312 Rehabilitation trajectories and predicted recovery times were compared between 'original'
313 and 'corrected' data, considering Jaccard and Bray-Curtis similarities. For the Worsley
314 minesite, a 'filtered' dataset, and corresponding correction, were also prepared which
315 excluded the three southernmost samples, which were geographically separate from the other
316 Worsley samples (see Figure 1, and SI Appendix Table S3).
317

318 **3. RESULTS**

319 **3.1 General findings**

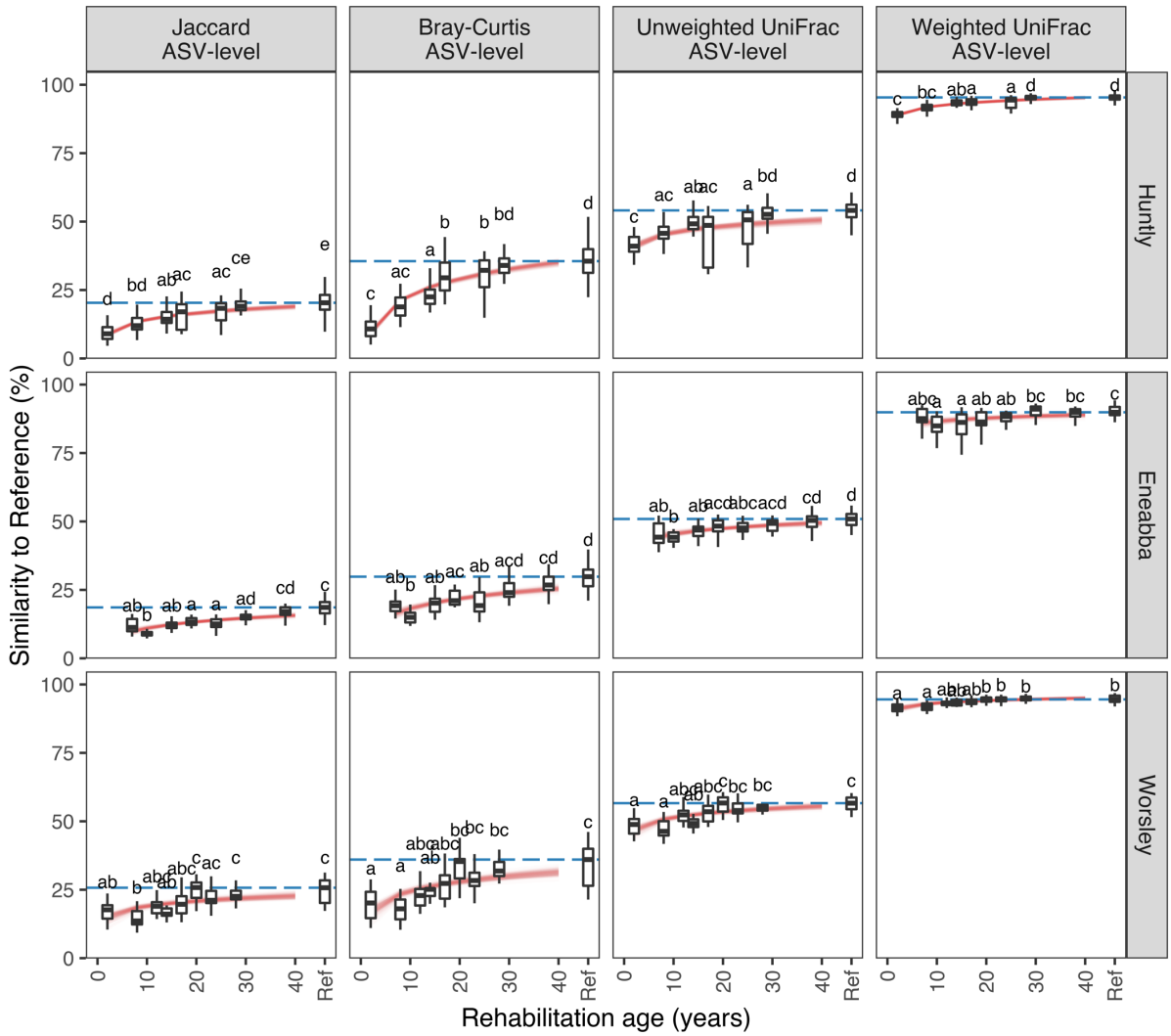
320 We found remarkable variability among reference samples within each minesite (Figure 3; SI
321 Appendix, Table S5, Figures S21–S23). Among-reference similarities ranged from <20% to
322 >95% depending on ecological measures, pre-processing, and minesite. All rehabilitation
323 trajectory plots displayed the general pattern of increasing similarity to references with
324 increasing rehabilitation age (Figure 3; SI Appendix, Figures S21–S23), although the
325 logarithmic models and predicted recovery times varied by ecological measures, pre-
326 processing and minesite.
327

328 **3.2 Alternative ecological measures**

329 We found a general increase in similarity to reference values across the ecological measures,
330 from Jaccard (generally lowest similarities), Bray-Curtis, Unweighted UniFrac, to Weighted
331 UniFrac (generally highest similarities) (Figure 3; SI Appendix, Table S5). The greatest y-
332 axis span, and therefore greatest sensitivity to detect change, in similarity to reference values
333 between the youngest rehabilitation ages and references occurred with Bray-Curtis measures
334 (Figure 3). The smallest span (or flattest curves) in similarity to reference values between the
335 youngest rehabilitation ages and references occurred with Weighted UniFrac measures.

336 Except for the Unweighted Unifrac result at Huntly, Jaccard measures generally
337 returned the longest predicted recovery times, followed by reduced or similar recovery times
338 predicted using Bray-Curtis, Unweighted Unifrac and Weighted UniFrac measures (Figure 4;
339 SI Appendix, Table S6). Low sample sizes (and corresponding low numbers of distance
340 measures) represent a limitation in our analysis, and the ecologically-distant samples in the
341 17-year and 25-year rehabilitation age group at Huntly (Figure 2A) are likely contributing to
342 the reduced similarity and longer rehabilitation trajectory in Unweighted UniFrac data. These
343 17-year and 25-year rehabilitation age group data at Huntly express reduced alpha diversity
344 and evenness compared to other samples, however reasons for this are unclear (SI Appendix,
345 Figures S3–S4).

346



347

348 FIGURE 3. Rehabilitation trajectory plots based on surface soil bacterial community

349 similarity to reference samples, for the Huntly, Eneabba, and Worsley minesites. Boxplots

350 display the distribution of similarity values across rehabilitation ages (groups not sharing a

351 letter are significantly different). Blue dotted lines denote the median similarity among

352 references. Red lines represent logarithmic models for the change in similarity to reference

353 with rehabilitation age, based on bootstrap resampling and modelling (B=100).

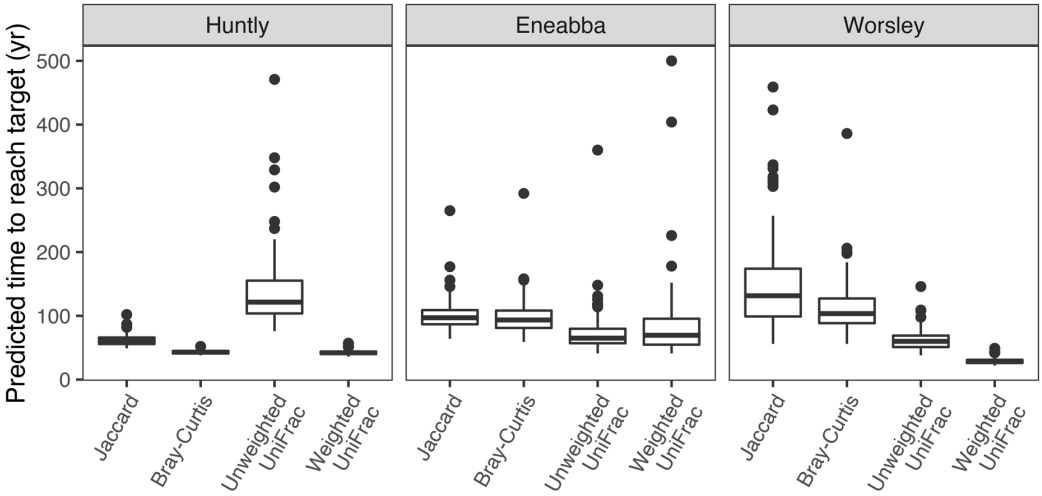


FIGURE 4. Predicted recovery times for soil bacterial ASVs to reach the target similarity to reference (= median of among-reference similarity values), for Huntly, Enneabba, and Worsley, considering Jaccard, Bray-Curtis, Unweighted- and Weighted-UniFrac measures, based on bootstrap (B=100) logarithmic models (see SI Appendix, Table S6 for values).

3.3 Grouping by sequence similarity (Huntly only)

Grouping by sequence similarity resulted in progressive overall shifts towards increasing similarity to reference values from ASV-level (generally lowest similarities), 99%, 97%, 95%, to 90%-identity OTUs (generally highest similarities) (SI Appendix, Figure S21). Predicted recovery times with more broadly clustered OTUs followed continuous and seemingly predictable patterns of: (i) increasing recovery times with Jaccard measures, and (ii) decreasing to steadying recovery times with Bray-Curtis measures (SI Appendix, Figure S24A, Table S6).

3.4 Taxonomic grouping (Huntly only)

371 Moving from ASV to genus-level data resulted in a pronounced shift towards increasing
372 similarity to reference, with similar although somewhat flatter rehabilitation trajectory curves
373 at higher taxonomic groupings (SI Appendix, Figure S22). Visually, there appeared to be
374 little effect on the rehabilitation trajectory plots from pruning unclassified taxa (SI Appendix,
375 Figure S22). Using Jaccard measures, moving from ASV-level to grouping at genus-level or
376 higher groupings dramatically increased predicted recovery times, compared to other
377 measures (SI Appendix, Figure S24B, Table S6). Also, pruning of unclassified groups
378 reduced the smoothness or continuity in Jaccard predicted recovery times (SI Appendix,
379 Figure S24B). Using Bray-Curtis measures, we found a non-linear pattern of recovery times
380 across the taxonomic groupings, with shorter times to reach the target in genus, family, and
381 order-level groups, and longer recovery times in other groupings (SI Appendix, Figure S24B;
382 see SI Appendix, Figures S5–S13 for relative abundance of order, class, and phylum-level
383 taxa for each minesite). Richness and evenness of bacterial communities varied across
384 rehabilitation age groups and taxonomic groupings (e.g., data for phylum, class, and order-
385 level are shown in SI Appendix, Figure S25), which may help explain the somewhat erratic
386 results from taxonomic grouping.
387

388 **3.5 Excluding rare taxa (Huntly only)**

389 Removing rare taxa to the point of retaining ASVs with >0.01% relative abundance produced
390 results from the Jaccard analysis that appeared to mimic results from the Bray-Curtis analysis
391 (SI Appendix, Figure S23). When only more common ASVs with >0.1% relative abundance
392 were retained, both the Jaccard and Bray-Curtis results appeared to reflect over-simplified
393 communities, resulting in shorter predicted recovery times. However, including only ASVs
394 with >0.001% relative abundance produced a small increase in predicted recovery times for
395 both Jaccard and Bray-Curtis (SI Appendix, Figure S24C, Table S6).

396

397 **3.6 Young rehabilitation sites with 'direct return' soils**

398 During our analyses, we uncovered example data that highlighted a distorting influence on
399 our trajectory modelling from young rehabilitation sites with 'direct return' soils.
400 Specifically, these fresher soil materials were more similar to references than older
401 rehabilitation sites. At Eneabba and Worsley, we compared trajectories with and without the
402 youngest rehabilitation age groups (i.e., excluding 7-year old sites at Eneabba; and 2-year old
403 sites and associated reference X138404 at Worsley) (SI Appendix, Figure S26–S29; Table
404 S6). Excluding these samples reduced predicted recovery times, e.g., for Bray-Curtis
405 similarities from median (and 5th percentile, 95th percentile) values of 94 (69, 131) years to 60
406 (53, 71) years at Eneabba; and from 104 (69, 174) to 41 (36, 47) years at Worsley.

407

408 **3.7 Correcting for spatial autocorrelation**

409 We modelled the slope-trends of the relationships between ecological distance to references
410 and geographic distance to references, within rehabilitation age classes, for each of the
411 minesites with Bray-Curtis and Jaccard measures (see SI Appendix, Figures S30–S35). We
412 also applied a 'correction' for the spatial autocorrelation, such that rehabilitation age groups
413 were adjusted to display the same ecological-geographic slope trend as found in references
414 (refer to the 'C' panels in SI Appendix, Figures S30–S35). Rehabilitation trajectory plots, and
415 predicted recovery times, using these corrected data were compared to the original
416 uncorrected data (see SI Appendix, Figures S36–S37 and Table S6). Worsley displayed a
417 strong ecological distance-geographic distance trend in among-reference data (= spatial
418 autocorrelation), and the greatest divergence of all the minesites in predicted recovery times
419 between original and corrected data (SI Appendix, Figure S36–S37). However, with
420 exclusion of the southernmost Worsley samples (i.e., the filtered dataset), the spatial

421 autocorrelation signal disappeared and predicted recovery times for filtered-original and
422 filtered-corrected data displayed comparable distributions (SI Appendix, Figure S38–S40,
423 Table S6).
424

425 **4. DISCUSSION**

426 **4.1 Alternative ecological measures**

427 Bray-Curtis measures produced the greatest range in similarity values between young
428 rehabilitation and reference samples, and therefore are likely to offer the greatest sensitivity
429 to quantify the progress of recovery of soil bacterial communities towards reference states. In
430 contrast, Weighted UniFrac offered limited sensitivity to detect changes with rehabilitation
431 age (i.e., shallow trajectory curves) and may result in under-prediction of recovery times.
432 Low variation in Weighted Unifrac similarities likely reflects high proportions of somewhat
433 closely related organisms across the samples. Jaccard distances represent the proportion of
434 unshared taxa out of the total number of taxa recorded in two groups (Anderson, Ellingsen &
435 McArdle 2006). Unweighted UniFrac uses phylogenetic information and calculates the
436 fraction of the branch length in a phylogenetic tree that leads to descendants in either, but not
437 both, of the two communities (Lozupone *et al.* 2007). These qualitative measures reflect the
438 survival and presence of taxa (Jaccard) and related lineages (Unweighted UniFrac), where
439 loss of sequences can reflect extreme or limiting environmental conditions (e.g., soil abiotic
440 factors) or limited geographic distribution. Meanwhile, Bray-Curtis and Weighted UniFrac
441 measures emphasise abundant organisms (or abundant sequences). Similarity generally
442 increased with increasing abundances of shared taxa for Bray-Curtis, and shared lineages of
443 related sequences for Weighted UniFrac. The quantitative measures often reflect the growth
444 or decline of certain organisms due to factors such as nutrient availability and sublethal
445 variation in environmental conditions (Lozupone *et al.* 2007).

446

447 **4.2 Grouping by sequence similarity**

448 Grouping near identical sequences will reduce the denominator used in calculating Jaccard
449 distances. For a given number of unshared taxa between samples, using broader OTU clusters
450 will make the proportion of unshared taxa (compared to all taxa) larger when there are a
451 smaller number of total taxa present. Our data suggest this shifting Jaccard calculation can
452 impact some samples strongly (e.g., note the 17-year age group in SI Appendix, Figure S21)
453 resulting in a gradual increase in predicted recovery times with broader (reduced identity
454 threshold) OTU clusters. On the other hand, broader OTU clusters will aggregate some
455 sequences into already large groups and will tend to further emphasise abundant groups.
456 Consequently, our Bray-Curtis data suggest broader OTU clustering will make the target
457 similarity easier to reach and predicted recovery times reduced accordingly.

458

459 **4.3 Taxonomic grouping**

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

462

463 **4.4 Excluding rare taxa**

464 We show that filtering out of rare taxa to a limited extent (>0.001% relative sequence
465 abundance) produces a relatively small increase in predicted recovery times for both Jaccard
466 and Bray-Curtis measures. Interestingly, this low level of exclusion of rare taxa does not
467 appear to moderate the assessment by producing reduced recovery times. At the low level of
468 exclusion, our analysis using rarefied data and similarity to reference measures may help
469 mitigate some of the impacts and concerns of removal of rare sequences experienced

470 elsewhere (e.g., Schloss 2020). This raises the prospect to reduce sequencing depth, and
471 potential for shifting investment towards more robust assessments that incorporate a larger
472 number of samples with reduced sequencing depth and cost per sample.
473

474 **4.5 Influence of 'direct return' soils in young rehabilitation sites**

475 For reasons discussed here and below, we suggest it may be prudent for future similarity to
476 reference trajectory assessments to exclude young rehabilitation sites with 'direct return' soils
477 that display elevated similarity to reference. As observed at Eneabba and Worsley, the
478 inclusion of young rehabilitation samples that were overly similar to references resulted in
479 seemingly biased, longer predictions of recovery time. The industry best practice of 'direct
480 return' of topsoil to new rehabilitation sites is based on objectives to minimise soil
481 degradation and expedite ecosystem recovery. However, our use of monotonic logarithmic
482 models applied to a data series that contains young rehabilitation sites with elevated
483 similarity to reference values, followed by older sites with reduced similarity to reference
484 values, results in the seemingly perverse outcome of a flatter, longer modelled trajectory of
485 recovery. The enhanced ecological similarity to reference in young rehabilitation sites with
486 'direct return' soils reflects a biological inertia, or temporary carryover effect, from unmined
487 areas where the soils originate, and confounds the relationship between soil microbiota
488 development and rehabilitation age. For 'direct return' soils, we speculate the time taken for
489 local influences to become dominant in shaping the resident microbiota may be in the order
490 of 1-10 years, varying on a case-by-case basis, e.g., due to soil factors including organic
491 matter and clay content, as well as the magnitude of environmental influences. Soil
492 microbiota will be shaped by influences including local rainfall, temperature, aspect, soil
493 water availability and transport (e.g. run-on, lateral flow), and vegetation communities via
494 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 & Cardon 2012); potential microbiota uptake and transfer via xylem into the phyllosphere (Fausto *et al.* 2018; Deyett & Rolshausen 2019) and subsequent leaf litter; and capillary rise in heavier textured soils under conditions of soil water evaporation.

Any decision to exclude young rehabilitation sites with direct return soils from the modelling should be made on a case-by-case basis. In particular, this decision should reflect whether these data display elevated similarity to reference values; and consider factors such as the source location of direct return soils (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 following soil return.

4.6 Spatial autocorrelation

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

519 Our analysis of spatial autocorrelation should be viewed as introductory and
520 illustrative. For 'direct return' soils at young rehabilitation sites, our approach is deficient
521 because we do not account for their previous location. Although, we anticipate localised
522 influences would dominate the shaping of resident soil microbiota in rehabilitation sites after
523 a few years, as discussed above.
524 Plant-soil-microbiota feedbacks represent a complicating factor for disentangling
525 effects of soil abiotic condition, rehabilitation age, and residual/unexplainable spatial
526 autocorrelation in restoration chronosequence studies. This is because chronosequence
527 studies (which presume a 'space-for-time' proxy relationship between treatments and
528 outcomes) typically do not collect sufficient data to determine whether soil conditions have
529 influenced rehabilitation outcomes, plants have conditioned soils, or both situations have
530 occurred. Studies that have considered plant-soil feedbacks in restored Jarrah forest (Huntly)
531 sites have shown differential correlative effects of rehabilitated soil biotic and abiotic
532 properties (Orozco-Aceves, Tibbett & Standish 2017). Also, plant-soil feedbacks behave
533 differently in unmined versus rehabilitated soils (Orozco-Aceves, Standish & Tibbett 2015).
534 Further work is required to build understanding of this topic (e.g., via longitudinal studies).
535

536 **4.7 Other limitations**

537 There are important limitations in our study, in addition to those already discussed. The
538 robustness of our study would be improved with more samples per minesite to help better
539 capture minesite-wide variation. We did not consider soil microbiota patterns at depth, which
540 are also important. Also, major changes to rehabilitation practices over time will disrupt the
541 'space-for-time' substitutive modelling approach that is relied upon in chronosequence
542 studies such as ours. As for restoration chronosequence studies elsewhere, careful sample
543 selection is required to avoid confounding factors as much as possible (Walker *et al.* 2010).

544 There are potential limitations in our study associated with the phylogenetic trees we used to
545 generate UniFrac distances (see SI Appendix, Supplementary Methods for details). Tree-
546 building often represents a compromise between accuracy in representing phylogenetic
547 relationships and computing time, and it was beyond the scope of our study to test the
548 sensitivity of our UniFrac-based analyses to the quality of trees used. We used logarithmic
549 models which assume a monotonic recovery function, however other models that account for
550 variable trends over time, and varying success for different rehabilitation techniques or sites,
551 may offer improved estimates of recovery time. We suggest these limitations should be
552 investigated in future studies.
553

554 **5. CONCLUSIONS**

555 We provide a proof-of-concept demonstration of an innovative, chronosequence-based,
556 similarity to reference trajectory assessment method, to quantitatively track progress in soil
557 microbiota with post-mining rehabilitation. Through incorporating microbiota survey data
558 from multiple reference sites of varying character, we revealed substantial variation among
559 reference ecosystems within each minesite that can inform realistic rehabilitation targets. Our
560 approach reduces the complexity associated with microbiota data and enables prediction of
561 recovery time to reach reference-based targets. The use of soil microbiota data (including
562 alpha diversity, evenness, compositional data, and microbiota-associated soil variables; SI
563 Appendix, Figures S3–S20) provides another line of evidence, which in conjunction with
564 other information, could assist in the examination of potential impediments to the progress of
565 rehabilitation, thereby helping to inform adaptive management. From our investigations, we
566 recommend using ASV-level Bray-Curtis similarities which appear to offer a relatively
567 sensitive and stable basis for modelling rehabilitation trajectories. We recommend wherever
568 possible to maximise sample sizes, employ spatial pairing of reference and rehabilitation

sites; and to exclude geographically-distant, non-representative sampling areas. We also recommend considering, on a case-by-case basis, the exclusion of young rehabilitation sites with 'direct return' soils that display elevated similarity to reference values, which may unduly bias the trajectory modelling. Further fine-tuning to identify possible minor reductions in sequencing depths (eliminating some rare taxa) offers promise to reduce per sample costs, enabling investment in more samples, to help deliver more robust assessments. This work represents an important step towards a reduced-complexity microbiota-based monitoring and evaluation framework consistent with many best practice principles for setting, monitoring and managing towards mine completion criteria recommended by (Manero, Standish & Young 2021). We anticipate that our approach could be expanded to other eDNA sequence-based survey data (e.g., fungal ITS and eukaryote 18S rRNA marker genes, functional potential from shotgun metagenomics), and may have broader applicability for evaluating rehabilitation progress beyond post-mining contexts.

583 **ACKNOWLEDGEMENTS**

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

591

592 **AUTHOR CONTRIBUTIONS**

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

597

598 **DATA AVAILABILITY STATEMENT**

599 Data and code are available at: https://data.bioplatforms.com/organization/about/australian-
600 microbiome and https://github.com/liddic/resto_traj

601

602 **ORCID**

603 Craig Liddicoat – https://orcid.org/0000-0002-4812-7524
604 Siegfried L. Krauss – https://orcid.org/0000-0002-7280-6324
605 Andrew Bissett – https://orcid.org/0000-0001-7396-1484
606 Ryan J. Borrett – https://orcid.org/0000-0001-8663-0844
607 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.

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.

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.

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.

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. (ed. L. Montanarella, Scholes, R., and Brainich, A. (eds.)), pp. 744. Bonn, Germany.

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.

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.

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.

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.

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.

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.