nature portfolio

Peer Review File

Sequencing-guided re-estimation and promotion of cultivability for environmental bacteria

Corresponding Author: Dr Feng Guo

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Zheng et al. present a paper on sequencing-guided re-estimation and promotion of cultivability for environmental bacteria. First of all, the clarity of the paper is very low, which already starts with a very unclearly defined aim of the study. The description of the figures is also insufficient. Results of some experimental procedures are not shown or it is unclear why the procedures were conducted, e.g. the results do not speak about inoculum-derived MAGS. Furthermore, many of the conclusions of the study are not surprising, e.g. L. 361, 396, 426-31, 442, leaving behind doubts about whether the paper is sufficiently novel (I disagree with what is stated on lines 563-5) to be accepted in Nature Comms. Discussion is very insufficient and, in places, uneven, e.g., why r/K strategists are discussed so much when no results mention them. On the other hand, the topic is very interesting and, if presented more concisely and in a more logical manner, in my opinion, it could be significantly improved (incl. presenting the essential findings of the study while leaving the others at least for the Discussion) to meet the standards of Nature Comms.
Specific comments:
L. 71: Not really clear what is meant here, needs further clarification.
L. 86-9: I disagree that this is the sole definition of culuromics.
L. 111-123: This paragraph is not very clear, it should ideally be rewritten to more clearly highlight the objective of the study.
L. 112: missing word: of
L. 133, 135, 203 and others: fix grammar. Also, unified tenses should be used throughout the MS.
L. 149: What was the reason for such a selection of media? Why not both for both samples?
L. 150: In-situ media is a misleading term, use a different one.
L. 224: I do not seem to be able to find the results of the metagenomics of 6 inoculum samples.
L. 363: Is this one specific reason the only one the authors have considered?
L. 419: I am wondering why such thresholds were used. I would suggest using the common threshold of 98.65% similarity to delineate species as well.

Reviewer #2

(Remarks to the Author)
In this study, the authors used two types of samples, culture-enriched metagenomics, and many other experimental or evaluation methods, to reassess the cultivability of environmental bacteria and offers a promising path to optimize microbial cultivation. Overall, the study is of scientific interest but there are some major limitations of this study.
Major comments:
Introduction
Line 68-69: “As a result, a substantial proportion of prokaryotes was considered uncultured or uncultivable in the laboratory, including those referred to as "microbial dark matter" or candidate divisions, which are phylum-level lineages lacking any representative isolates.”. The definition of ‘microbial dark matter’ here seems too narrow, why are they only phylum-level lineages? The authors should search for the earliest proposed and authoritative references for this professional vocabulary.
Method
As a researcher who has been involved in both direct metagenomic sequencing, classical isolation and culture, and culture-enriched metagenomic sequencing, I really consider the approach to defining the proportion of cultivable taxa (PCT) in Methods of this manuscript is very crude and unscientific. The authors said “To calculate PCT, the number of ASVs detected

in the culturomic samples, which were also detected in the inoculum, was divided by the total number of ASVs detected in the inoculum." with no references. Indeed, a number of research on culture-dependent studies have focused on culturable taxa. Many previous studies have shown that the taxonomic composition of culture-enriched samples is often quite different from that under direct sequencing of the original sample (i.e., inoculum here). For culture-enriched metagenomic sequencing, many cells are detected as a result of enrichment expansion, and they are completely undetected in direct sequencing due to their extremely low abundance or other some reasons, which can be very high in number. In short, there is a fundamental difference between culture-dependent and culture-independent sequencing, and it is inappropriate to calculate the "PCT" by using the culture-independent results as the denominator.

Two types of environmental samples (soil and activated sludge) were employed for this study, why not feces, sediments, or other samples? What are the reasons for choosing these two types? The specific location, collection method, refinement types, and physicochemical parameters of soil should be described in detail, as well as the activated sludge samples.

What were the criteria for the selection of TSA and R2A conventional synthetic media? In general, the representation of taxa obtained with 2~3 media is not sufficient to assess the culturability of the microbes in the inoculum. This is one of the major limitations of this study. The authors should carefully consider the concept of Culturomics mentioned in this study, and should not consider multiple sequenced samples (multiple replicates) as utilizing Culturomics.

Line 192: The ASVs generated based on V4 region amplicon sequencing cannot serve as an information carrier reflecting species-level. The authors need to systematically evaluate the limitations of the application of short-read sequences of the V4 region in this study. In fact, for culture-enriched metagenomics, full-length 16S amplicon sequencing can better reflect changes at the species level than V4. In addition, the authors utilized the V4 region and performed a species novelty assessment based on different similarity thresholds (≥98%, 94-98%, 90-94%, and <90% sequence similarity, Line 419). It is quite surprising, since usually such analyses are based on 16S full-length sequences, and short-reads 16S cannot accurately define novelty, which may lead to very biased conclusions.

Line 195: Even though EzBioCloud database encompasses 16S rRNA gene sequences from type strains of described species and representative sequences of environmental clones, the official website has not maintained an update of this database since 2018.05 (https://www.ezbiocloud.net/resources/16s_download). From 2018 to 2023, there has been a significant increase in new species data and taxonomic changes, which may have a significant impact on the analysis results of this study. In contrast, using the latest bacterial and archaeal taxonomy training set No. 19 of RDP (update at 2023-08-23, https://sourceforge.net/projects/rdp-classifier/files/), SILVA and GreenGenes, etc. for this study would be better. I strongly recommend the classification system of the RDP database for culture-enriched 16S amplicon analysis, which has been proven to be more consistent with the classification results of NCBI.

Line 237: As shown in Figure 1A, it seems like two methods were utilized for the determination of PCC, i.e., fluorescent stained cell assay and plate-based CFU calculation, but from the method description I can only find the CFU method, is there a misrepresentation here? As mentioned by the author in the Introduction, the issue of "great plate count anomaly" needs to be noted. The disparity between the total cell count, typically determined through microscopy or flow-cell cytometry, and the colony-forming units (CFUs) on agar plates can differ by several orders of magnitude in typical environmental samples. So, why does the authors still use the agar plate-based CFU method to count cells in this study? Is there any other better way to replace it?

Line 223: How does genome-resolved metagenomics study contribute to the core theme of this article?

Line 224: From 515 (394 culture-enriched samples), only 6 inoculum samples and 17 derived from AS were used for metagenomic analysis. Why the authors selected only these few samples? What were the parameters used to select those?
Line 313-317: "For inoculum samples, MAGs with marker genes (DNA gyrase subunit B and RNA polymerase subunit B) … analysis (i.e., putatively cultivable though the genome was not obtained)", could the authors provide a more detailed explanation of this handling method for MAGs from inoculum samples?
Line 433-437: It is inappropriate to place this paragraph (characteristics of MAGs) here as it does not serve the theme of the results "A substantial number of cultivable taxa are novel but are present in low relative abundance". The results can be moved to the section in Line 508.

Discussion
I strongly recommend that the authors point out the key limitations, improvements, and future perspectives of this study in the Discussion, rather than smugly recounting only what they did.
Mirror comments:

Line 143: Change "filtrate" to "filtrates", and why is the volume of suspension for each sample in a large range of 50-100 mL?
Line 152: Change "MgSO4 ·7H2O" to "MgSO4·7H2O", delete the space before dot.
Line 157-158: It seems illogical for the authors to use a larger 0.22 μm filter membrane for sterilization after using a 0.1 μm filter membrane with a smaller pore size.
Line 162-163: "…autoclaved 6% agar (in ddH2O) was mixed in a 4:1 ratio with prewarmed (50°C)…", The method description doesn't seem very clear, so what is the final proportion of agar?
Line 177: Change "the kit" to "the same kit mentioned above"
Line 178: Change "To each plate" to "For each plate"
Line 223: The section "Metagenomes and MAGs annotation" should be moved before "Comparative genomics guiding the promotion of cultivation of a specific taxon"

Line 230-232: How to perform the dereplication of MAGs? using dRep? If so, please cite the related reference. In the METHOD section, the author uses a lot of methods, but the order of description and logic as a whole is puzzling…
Line 262: Add a space after “EBPR01”
Line 334-337: There are two ways to write VB12 here, “vitamin B12” and “Vitamin B12”, please confirm and unify it.
Line 341: “Rhodocyclaceae”, “Burkholderiales”, and other Latin names of prokaryotes in the whole manuscript are suggested to use italics.
My conclusion
The core theme, and value, of this paper is actually around the “an experimental approach to reassess the cultivability of environmental bacteria and a promising path to optimize microbial cultivation”. This I totally agree with.
However, I have a number of misgivings about the introduction of the PCT concept and calculation methodology, which may lead to a significant loss of credibility for many of the results (see also my comments about statistical validity). In addition, the manuscript suffers from another major problem - the authors attempt to solve too many problems at once using multiple methods, making this paper complex and confusing. Furthermore, the culture-enriched metagenomics of have been applied and interpreted in detail in both soil and activated sludge environment (e.g., Li et al., npj Biofilms and Microbiomes, 2023; Zhang et al., Environmental Science and Technology, 2022), which poses a challenge to the novelty of this manuscript.
In summary, I seriously evaluated this manuscript is not suitable for publication on Nature Communications, but this does not mean an underestimation of the important findings of this study.

Reviewer #3

(Remarks to the Author)
Zheng and coauthors present a well-written and well-explored assessment of the percent cultivable taxa in soil and wastewater. They find that the percentage of cultivable taxa is higher than one might guess from the low numbers of percent of cultivable cells. Furthermore, they explored features associated with uncultivability and found that smaller, less metabolically active cells, often with long doubling times, were over-represented in the uncultivable taxa. Then they identified genomic features associated with uncultivable Burkholdaria and showed through experiment that these features could be used to enhance cultivability. I think this is an important contribution to our understanding of what makes microbial taxa cultivable, and it is nicely done. I just have a few suggestions below:

Line 76: saying that Martiny based his conclusions on “the” 16S database is misleading. It would be better to say that he did is study on “a” 16S database. Using “the” suggests that Martiny considered the entire public database, instead of a handful of samples. The study that used the entire NR database was Lloyd et al., 2018 https://doi.org/10.1128/msystems.00055-18, that was the precursor to Steen et al., 2019 that is cited in this paper. Lloyd et al. found that the proportion of cultured taxa was also very low on average, although individual samples could have quite high proportions of taxa represented by cultures. Notably, human host environments. Given this paper’s findings, the assertion on line 71 that low numbers of uncultured taxa is a “misconception” is not fully supported by the literature. I think this section could just be tempered by pointing out the two options, rather than implying that Martiny’s view is widely-held to be true.

The fact that the authors find that PCT is greater than PCC is an expected outcome if the uncultivable taxa are in greater cell abundance than the cultivable ones. This much greater abundance of uncultivable taxa is shown by the very high percentages (44-84%) of the culturomics ASVs being too rare in the inocula to appear in their 16S libraries. This makes me wonder about the conclusion on lines 399-401 that more abundant ASVs are more likely to appear in the cultures. Is this still true if you include the high percentage of ASVs in cultures that are too rare to appear in the inoculum libraries at all?

The conclusions on 411-413 do not appear to be supported by Figure 2D. Here, it shows that cultivability decreases for increasingly phylogenetically distant OTUs. This, it seems to me, would suggest that lower phylogenetic similarity results in lower cultivability. However, this doesn’t seem like the best test to answer this question. It might be better to look at the distribution of similarity of the culturomics and the inoculum ASVs to their closest cultured relative.

Lines 447-449: I don’t know what it means to have a diameter below 94%.

Line 133 should read “the times ranged from 0.5 h to 1 d.

Line 220-221: What was the percentage range of contaminants across all samples? Were contaminate ASVs removed from both culturomics and inocula samples?

Lines 482: This is incorrect – the value for metabolic activity is the rRNA per rDNA, not rRNA per cell. This is an important distinction because I got very confused about how cell counting was achieved for uncultivable taxa and only figured it out after re-reading the methods.

563-565: I urge the authors to remove claims of priority and instead simply state the new things they have found, without pointing out that they are new.

Were any ASVs present from phyla or classes where no cultures have ever been obtained? If so, I think it would be helpful to put this study into context that while the authors new approach is extremely useful for obtaining broader taxonomic representation among cultures, there may be deeply-branching clades which are facing even steeper barriers to cultivability.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Review: NCOMMS-23-54252A - Sequencing-guided re-estimation and promotion of cultivability for environmental bacteria
I must admit that the MS is much improved and reads much better than before. However, I still have many comments that should be addressed before the MS can be published.
Specific comments:
- l. 114 - omit easy to be collected and transported - very poor reason
- l. 115 (and elsewhere) - oxygenic or aerobic - these two are very different
- l. 137-8 or Fig. 1 legend - a clear description of what plate, condition, and sample means (information in brackets is not clear)
- l. 142-151 - I think this information needs to be presented differently. As I was reading the text, I thought the following: this information is very crucial, however it makes me full of doubts about reaching the goal of the study - the number of both taxa and cells is always compared with the inoculum; I think the authors should consider presenting the results separately for the taxa/cells that can be compared with the inoculum and those that cannot be; in addition, the information on line 147-8 makes me wonder how similar these ASVa are - 75% is not a lot and 94.7% is still not sufficient to be the same species; finally, can authors discuss how likely it is that these ASVs that do not match represent a crosstalk between samples, etc.; did the authors use something like a mock community? I think it would be especially valuable in this case. THEREFORE, I suggest the information from the end of the paragraph be presented earlier. At the same time, % numbers of ASVs that were omitted this way need to be presented in the text, not just in supplementary material.
- l. 153 - the statistical difference is one thing, biological meaning is another - bearing this in mind, the information should be presented in a way that for soil the PTC is significantly higher than PCC with the difference in mean being x %; for the AS, however, these numbers are comparable, yet still significantly different. At the same time, at the end of this paragraph, the authors should explain what the differences between plate, condition, and sample mean. This information is completely missing.
- l. 160-1 - weird sentence - it is not true for PCT?
- l. 176 - be consistent with tenses, use past simple throughout the MS
- l. 179 - define at least briefly the extraction medium, maybe by saying based on… this is an issue of the results section being presented before the methods (things should be defined as they first appear in the text); Fig. 2 legend should also define the extraction medium; additionally, I would not call R2A artificial, maybe common laboratory defined medium is better
- oxygenic conditions is a term that needs to be replaced throughout the MS
- l. 192 - as well as phylogenetically related; what is meant here?
- l. 224 - the information in the brackets is given for species? It should be noted.
- l. 225 - is, not are
- l. 232 - what does this information say here?
- l. 251 - omit ‘therefore increasing our confidence in isolating this bacterium’
- l. 252 - fix wording - the isolate was not obtained by 16S rRNA gene sequencing
- l. 258 - I suggest including also a threshold of 98 % to perform this analysis
- l. 297 - to decipher; overall the grammar should be checked again throughout the MS
- l. 296 - I wonder if this whole chapter should not be placed before line 243; the text on line 243 and beyond already takes into account genome sequencing
- l. 328 - the all - grammatically wrong
- Fig. 6 - use relative abundance, not relativated
- l. 330 - abundance, not abundances; l. 332 - abuof? Once again, the grammar and typos need to be fixed throughout the MS! Other obvious mistakes: L. 345, 357, 362 (still)…
- l. 334 - to some extent is not a very scientific expression, be specific
- l. 358 - the phenomenon of dormancy should be noted here
- l. 367-70 - can authors discuss how likely it is that these ASVs that do not match represent a crosstalk between samples, etc.; did the authors use something like a mock community? I think it would be especially valuable in this case.
- l. 387 - use undescribed or hitherto uncultured rather than novel
- l. 389-97 - states of low metabolic activity should be discussed more along with how certain approaches can help with this phenomenon during cultivation, such as the use of resuscitation factors or signaling molecules; such a discussion can be well linked with the discussion on B12 vit in the medium.
- What I am missing in general is the number of CFUs that appeared on the plates before the colonies were used for sequencing. This information should appear in the text.
After these comments are addressed, I think the paper will make a valuable contribution to the journal.

Reviewer #2

(Remarks to the Author)
The manuscript entitled “Sequencing-guided re-estimation and promotion of cultivability for environmental bacteria”, was partially revised following reviewers’ feedback. The manuscript was somewhat enhanced by incorporating some of the modifications and addressing of concerns suggested by all the reviewers. However, in the process of revision, many

enhancements have not been substantially improved, and the manuscript still suffers from several inescapable problems. In short, I think this manuscript is not suitable for publication on Nature Communications.

Reviewer #3

(Remarks to the Author)
Thanks for the good discussion. I think I understand better that the authors have a different (equally valid) definition for cultivability that I do. The authors' definition is that a cel l or a taxon is cultivable in this experiment with these conditions, whereas I tend to think of it as an absolute designation - can this cell or taxon be cultivated by someone somewhere using any possible method. Seeing this distinction more clearly helps me understand the paper better. The revisions look great, and I congratulate the authors on their good work.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The revisions that have been made seem generally OK, but some issues are pending.
Consider changing microaerobic to microaerophilic.
L. 136: Present the aim/objective/novelty of the paper first, or at least some results, before getting to the point of contamination. That said, Fig. 1c should probably be presented before 1b because it is the main result!
I still think that either the English or the wording in general needs to be fixed in some places. See EXAMPLES of new text below.
Inoculums.
That is, as the number of plates and cultivation conditions increased, PCT gradually rises, with the extent to which it exceeded PCC progressively increasing. Above results demonstrated that PCC could significantly underestimate the cultivable diversity of environmental bacteria.
These low-abundance ASVs REMINDED US of the formation of microcolonies with high taxonomic novelty commonly reported on plates in previous literature.
Finally, the requirement of consistent tenses has yet to be met.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Zheng et al. present a paper on sequencing-guided re-estimation and promotion of cultivability for environmental bacteria. First of all, the clarity of the paper is very low, which already starts with a very unclearly defined aim of the study. The description of the figures is also insufficient. Results of some experimental procedures are not shown or it is unclear why the procedures were conducted, e.g. the results do not speak about inoculum-derived MAGS. Furthermore, many of the conclusions of the study are not surprising, e.g. L. 361, 396, 426-31, 442, leaving behind doubts about whether the paper is sufficiently novel (I disagree with what is stated on lines 563-5) to be accepted in Nature Comms. Discussion is very insufficient and, in places, uneven, e.g., why r/K strategists are discussed so much when no results mention them. On the other hand, the topic is very interesting and, if presented more concisely and in a more logical manner, in my opinion, it could be significantly improved (incl. presenting the essential findings of the study while leaving the others at least for the Discussion) to meet the standards of Nature Comms.

Re: Thank you for your constructive feedback. We have further clarified the aim of our study in the last paragraph of the Introduction (Line 112-128). Supplementary explanations have also been added to the figure legends. Significant modifications have been made to the experimental procedures and discussion sections, including the removal of those regarding r/k selection.

Regarding the concerns about the novelty of certain descriptions, we acknowledge that most of these results are not directly related to our main discoveries (i.e., re-estimation and promotion of cultivability). The sentence in question (lines 563 in the original version) has been deleted. As suggested by Reviewer #3, we have supplemented two relevant references (https://doi.org/10.1038/s41522-023-00439-8; https://doi.org/10.1128/aem.01151-23) in the sections of Introduction and discussion (Line 96 & Line382), which implemented cultivation guided by taxonomic information of cultivable bacteria on agar plates. However, neither of these studies successfully isolated novel strains at high taxonomic levels, nor did they re-estimate the cultivability of environmental bacteria.

Specific comments:

L. 71: Not really clear what is meant here, needs further clarification.

Re. Revised.

L. 86-9: I disagree that this is the sole definition of culuromics.

Re: Revised according to reference (https://doi.org/10.1038/s41579-018-0041-0).

L. 111-123: This paragraph is not very clear, it should ideally be rewritten to more clearly highlight the objective of the study.

Re: Thanks. This paragraph has been largely revised to make the aims of the study clearer.

L. 112: missing word: of

Re: Revised and thanks.

L. 133, 135, 203 and others: fix grammar. Also, unified tenses should be used throughout the MS.

Reply: Revised and thanks.

L. 149: What was the reason for such a selection of media? Why not both for both samples?
Re: The selection of a single medium for each type of sample was based on the widespread usage of these media and their relative simplicity in nutrient composition. Since we want to compare them with the extraction media, which were expected to be much more chemically complex than the synthetic media. Indeed, studies have used R2A medium for soil samples and TSA medium for activated sludge. There is evidence indicating poorer cultivability, such as only obtaining 10-20% of the CFUs on TSA medium compared to R2A medium for activated sludge (https://doi.org/10.1007/BF00185883). In addition, this also concerns the manageable workload. For instance, if we were to use four cultivation conditions (aerobic/anaerobic, synthetic/extraction medium) for one sample, with four gradients for each condition, and 10 plates per gradient, the total number of plates for one sample would be 160. We needed to consider the workload per experiment to avoid issues of intra-batch variation caused by excessive handling time (e.g., prolonged handling time might alter the microbial community in the inoculum).

Additionally, this study did not aim to examine diverse cultivation conditions but rather to establish a workflow to assess cultivability via sequencing methods. From the results, it is evident that the results obtained from a few cultivation conditions already indicate that previous estimates of bacterial cultivability in environmental samples were lower than the true cultivability. As discussed in the manuscript, expanding cultivation conditions may further increase the cultivability. Furthermore, we accepted the advice from the Editor and Reviewer #2 to minimize the use of the term "culturomics" in the article due to the limited range of cultivation conditions used in our study.

L. 150: In-situ media is a misleading term, use a different one.
Re: Thanks for your suggestion. It has been changed to “extraction medium” in the revision.

L. 224: I do not seem to be able to find the results of the metagenomics of 6 inoculum samples.
Re： We confirmed that the Table S1 provided the correct information for these inoculum samples with metagenomic data. The metagenomes of the six inoculum samples were used to obtain MAGs of uncultivable bacteria, which were referred to their genomic features comparing with cultivable MAGs. The corresponding results are presented in Table S3, Figure 6A, Figure S8, and Figure S9.

L. 363: Is this one specific reason the only one the authors have considered?
Re: Logically, bacteria growing on agar plates must originate from the inoculum because we included negative controls in each experiment to confirm that plating blanks did not form any colony. Following your comments, we conducted supplementary analyses: 1) The frequency of cultivable ASVs detected on agar plates but not in the inoculum samples (across all agar plates of a sample) was significantly lower than those detected in the original samples (all $P<0.001$, see the figure below). This supports their lower relative abundance in the inoculum samples. 2) Most of these cultivable ASVs could be found in the original samples with ASVs closely related in phylogeny (median values ranging from 96.1 to 98.8% for all samples, see the figure below). This implies their ecological relevance in the original samples, coexisting with ASVs potentially occupying similar ecological niches. However, it is noteworthy that most of these ASVs (76-90%) are not highly similar (>99% similarity) to their closest ASVs in the original samples, indicating artificial ASVs generated by PCR and sequencing errors may not be a major source of these ASVs.

These two supplementary analyses have been added as Figure S3&S4, and explanations have been provided in the revised manuscript in Line 145-151. It is worth mentioning that these ASVs do not involving PCT calculation, as already clarified in the original manuscript.

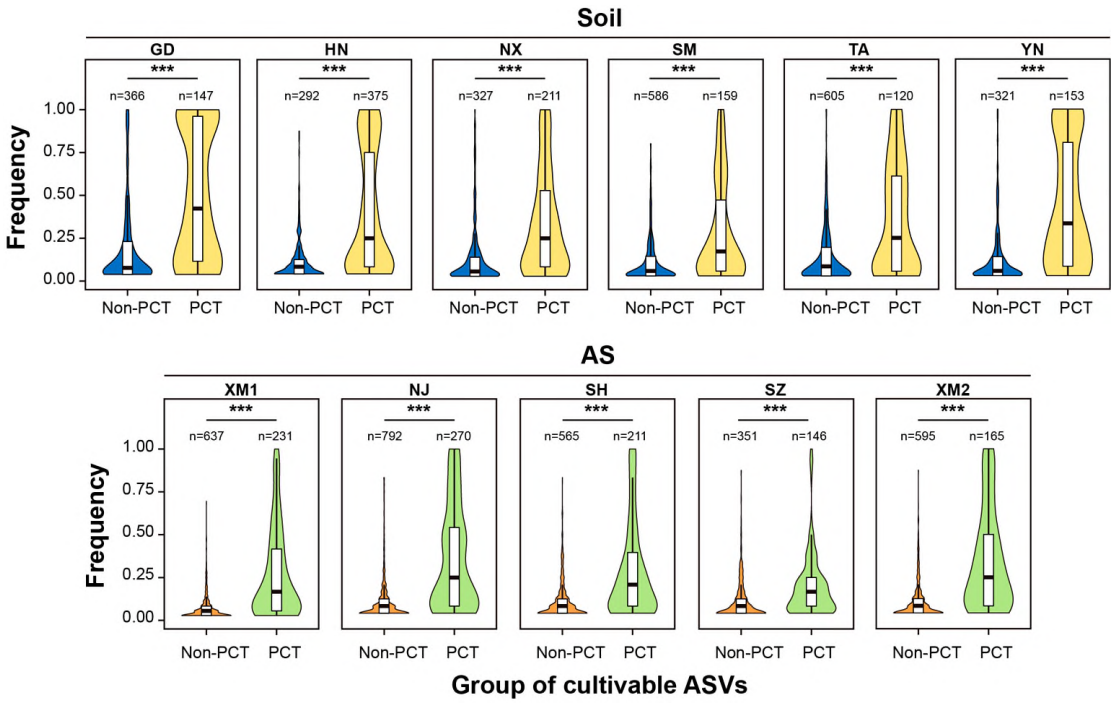


Figure S3 Comparison of the distribution frequencies on plates under specific sample-specific culture conditions between those involved in PCT calculation (detected in the inoculum) and those not involved (non-PCT, not detected in the inoculum) among cultivable ASVs detected on plates. The comparison was conducted using the Wilcoxon test (three asterisks denote P <0.001).

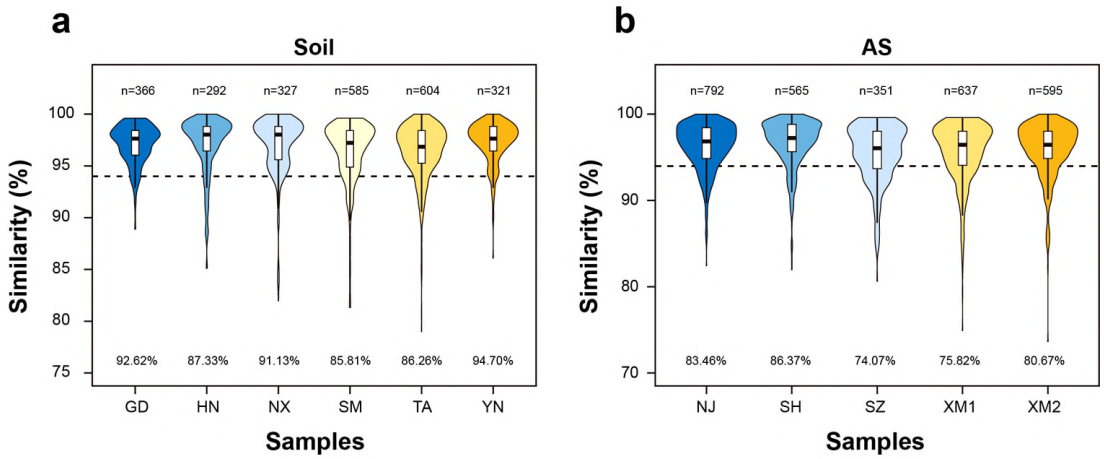


Figure S4. Distribution of phylogenetic similarity between the cultivable ASVs that are not detected in the inocula and the most closely related ASVs in the corresponding inoculum sample. The dashed line in the figure represents the division at 94% similarity, with the percentage below indicating the proportion of ASVs with >94% similarity. These values indicate that these undetected cultivable ASVs very likely originate from the inocula and their absence in the inoculum may be due to their low abundance.

L. 419: I am wondering why such thresholds were used. I would suggest using the common threshold of 98.65% similarity to delineate species as well.

Re: The choice of these thresholds was primarily based on the length of the V4 amplicon. For the V4 amplicon after primer removal, its length is approximately 253 nt. A single nucleotide difference corresponds to a 0.4% difference. It would not align well with the commonly used species threshold of 98.65% similarity derived from full-length 16S analysis (98.4% or 98.8%). We acknowledge that employing full-length 16S rRNA gene may enhance the accuracy of PCT and the thresholds we applied has some subjectivity. However, the main purpose here is to establish a rough novelty range for comparing different treatment groups. We believe it might be challenging or impossible to establish a universally applicable taxonomic standard for V4 similarity.

To answer this question and additional one from Reviewer #2, we conducted a simple analysis of the RDP type strain database. Among the 16S sequences from 18,939 different species type strains, 10,292 (54.3%) species shared a similarity of greater than 98.65% with at least one other species, while 9,351 (49.4%) species had identical V4 region sequences with at least one other species. The numbers are approximately equal, indicating a lack of unified standards for species delineation based on the 16S rRNA gene.

Therefore, we can only state that our use of V4 ASVs serves as a marker representing taxa, but a certain threshold does not guarantee a direct correspondence to a taxonomic rank. This is also why we opted for PCT (proportion of cultivable taxa) instead of "PCS" (proportion of cultivable species).

Reviewer #2 (Remarks to the Author):

In this study, the authors used two types of samples, culture-enriched metagenomics, and many other experimental or evaluation methods, to reassess the cultivability of environmental bacteria and offers a promising path to optimize microbial cultivation. Overall, the study is of scientific interest but there are some major limitations of this study.

Re：Thanks. We have tried addressing your comments below.

Major comments:

Introduction

Line 68-69: "As a result, a substantial proportion of prokaryotes was considered uncultured or uncultivable in the laboratory, including those referred to as "microbial dark matter" or candidate divisions, which are phylum-level lineages lacking any representative isolates.". The definition of 'microbial dark matter' here seems too narrow, why are they only phylum-level lineages? The authors should search for the earliest proposed and authoritative references for this professional vocabulary.

Re: Thank you for your comment. The article that first introduced the term "Microbial Dark Matter" to describe uncultivated microorganisms was Marcy et al. (2007) (https://doi.org/10.1073/pnas.0704662104), while the most influential paper popularizing this concept is the one we cited (Rinke et al. 2013, https://doi.org/10.1038/nature12352). Both papers

focused on uncultivated microbial entities at the phylum level (e.g., TM7 phylum for Marcy et al., 2007). Our description primarily aligns with the latter because it explicitly defined the term as follows: "Candidate phyla that represent a major unexplored portion of microbial diversity have been referred to as microbial dark matter (MDM) (cite Marcy et al., 2007)." We also recognize that some scholars extend the term MDM beyond candidate phyla to encompass all uncultivated microorganisms. However, starting from the earliest definition, we decided to make slight modifications as follows: “lineages lacking any representative isolates at high taxonomic ranks”.

Method

As a researcher who has been involved in both direct metagenomic sequencing, classical isolation and culture, and culture-enriched metagenomic sequencing, I really consider the approach to defining the proportion of cultivable taxa (PCT) in Methods of this manuscript is very crude and unscientific. The authors said “To calculate PCT, the number of ASVs detected in the culturomic samples, which were also detected in the inoculum, was divided by the total number of ASVs detected in the inoculum.” with no references. Indeed, a number of research on culture-dependent studies have focused on culturable taxa. Many previous studies have shown that the taxonomic composition of culture-enriched samples is often quite different from that under direct sequencing of the original sample (i.e., inoculum here). For culture-enriched metagenomic sequencing, many cells are detected as a result of enrichment expansion, and they are completely undetected in direct sequencing due to their extremely low abundance or other some reasons, which can be very high in number. In short, there is a fundamental difference between culture-dependent and culture-independent sequencing, and it is inappropriate to calculate the “PCT” by using the culture-independent results as the denominator.

Re: Regarding the definition of PCT, to our knowledge, previous studies have not proposed a standardized method based on experimental and sequencing data to evaluate PCT, although there are indeed many studies comparing the microbial communities of original samples with those after enrichment or cultivation (generally referring to beta diversity index), with most focusing on enrichment of liquid mixed microbial communities in various conditions. Only a very few studies have studied on the general cultivable taxa on agar plates (https://doi.org/10.1038/s41522-023-00439-8; https://doi.org/10.1128/aem.01151-23; https://doi.org/10.1111/1462-2920.15798). In the article, we emphasize that this direct analysis of cultivable microbial communities is not novel, but the method of calculating cultivability has not been studied in previous reports.

We cannot fully agree with your opinion that "there is a fundamental difference between culture-dependent and culture-independent sequencing." We believe that for most liquid enrichment systems, such situation may be more common, as liquid culture systems tend to be homogenous and well mixed, and it is known that this condition is more likely to lead to competitive exclusion, even for those dominant taxa in the inoculum. In other word, microorganisms that were previously abundant in the system may not be able to adapt to the enrichment conditions and then disappear. However, the agar plate isolation results in relatively less competition for each cell inoculated (at least in the initial phase), reducing the likelihood of competitive exclusion (https://doi.org/10.1038/nature00823). The agar plate culture method serves the purpose of this

study, as with most methods used to determine PCC, and it aligns with our goal, that is, to compare PCC and PCT.

Additionally, in Figure 2c, we compared the PCT of ASVs with different relative abundances in the original samples. This preliminary result suggests that there is no significant difference in the cultivability rate between low-abundance ASVs (<0.01%) and those relatively higher (0.01-0.1%).

Two types of environmental samples (soil and activated sludge) were employed for this study, why not feces, sediments, or other samples? What are the reasons for choosing these two types? The specific location, collection method, refinement types, and physicochemical parameters of soil should be described in detail, as well as the activated sludge samples.

Re: The reasons for choosing these two types of samples are manifold. 1) Both soil and activated sludge samples harbor high microbial diversity and density, with a substantial presence of uncultivated microbial taxa and typically 0.1-10% cultivation in literature-based PCC calculations. 2) Considering multiple geographically dispersed samples, convenience of sample collection and transportability were important factors. Soil and activated sludge are relatively easy to collect and transport compared to feces or sediments, as they may undergo rapid microbial inactivation upon removal from their original environment (anaerobic condition). 3) Our laboratory has long been engaged in studying the microbial ecology of activated sludge. We are familiar with the microbial communities therein. Additionally, we also want to recover potential important microbial taxa in activated sludge. A brief description is also included in the main text (Line 112-114). We have provided additional information about the samples in Table S1 as your request.

Certainly, covering more types of samples would be more meaningful for our conclusions. However, considering the number of samples required for each type and the workload per sample (as discussed in our response to Reviewer #1, at least 160 agar plates are needed per sample), we chose 6 and 12 samples of soil and AS, respectively.

What were the criteria for the selection of TSA and R2A conventional synthetic media? In general, the representation of taxa obtained with 2~3 media is not sufficient to assess the culturability of the microbes in the inoculum. This is one of the major limitations of this study. The authors should carefully consider the concept of Culturomics mentioned in this study, and should not consider multiple sequenced samples (multiple replicates) as utilizing Culturomics.

Re: Thanks. We agree with your opinion. In the revised manuscript, we have minimized the usage of the term of culturomics. For determining PCT, we primarily utilized two types of culture media (i.e., conventional synthetic media R2A or TSA, and extraction media) and three oxygenic conditions. From the perspective of culturomics, indeed, the limited variety of culture media used is a drawback. This limitation is also attributed to the workload (also replied to reviewer #1). Under the current design of culture media, typically each sample requires 160-300 agar plates, and adding more culture conditions would result in a substantial increase in this number. Additionally, we mentioned in the discussion that increasing culture conditions would further enhance PCT (Line 366-370).

Line 192: The ASVs generated based on V4 region amplicon sequencing cannot serve as an information carrier reflecting species-level. The authors need to systematically evaluate the limitations of the application of short-read sequences of the V4 region in this study. In fact, for culture-enriched metagenomics, full-length 16S amplicon sequencing can better reflect changes at the species level than V4. In addition, the authors utilized the V4 region and performed a species novelty assessment based on different similarity thresholds (≥98%, 94-98%, 90-94%, and <90% sequence similarity, Line 419). It is quite surprising, since usually such analyses are based on 16S full-length sequences, and short-reads 16S cannot accurately define novelty, which may lead to very biased conclusions.

Re: Firstly, we agree that the resolution of the V4 region of the 16S gene cannot reliably serve as a measure of species-level novelty. While the full-length 16S sequencing based on third-generation sequencing technologies has been reported in recent years, it has not been widely adopted yet, whereas the V4 region is still extensively used due to its mature analysis workflow. Considering that our study began in 2018, we decided to sequence the V4 region from the outset. Of course, combining our approach with results from full-length sequencing can have higher resolutions of taxonomy and phylotype, therefore provide a more precise determination of PCT.

According to Edgar (2016, https://doi.org/10.1093/bioinformatics/bty113), 100% similarity in the V4 region serves as a criterion for species delineation to some extent. However, it is important to note that we never designated the taxonomic units represented by ASVs as species representatives in the results. In PCT, the 'T' represents taxa rather than species. To answer this potential bias of species-level delineation by V4 and full-length sequences, we conducted a simple analysis of the RDP type strain database. Among the 16S sequences from 18,939 different species type strains, 10,292 (54.3%) species shared a similarity of greater than 98.65% with at least one other species, while 9,351 (49.4%) species had identical V4 region sequences with at least one other species. The numbers are approximately equal, indicating a lack of unified standards for species delineation based on the 16S rRNA gene.

For setting different thresholds to describe the novelty of ASVs, as mentioned in our response to Reviewer #1, we did not intend to correlate different thresholds with various levels of taxonomic ranks, as this is impossible to achieve. The main purpose for setting these thresholds is to establish a rough novelty range for comparing different treatment groups.

Line 195: Even though EzBioCloud database encompasses 16S rRNA gene sequences from type strains of described species and representative sequences of environmental clones, the official website has not maintained an update of this database since 2018.05 (https://www.ezbiocloud.net/resources/16s_download). From 2018 to 2023, there has been a significant increase in new species data and taxonomic changes, which may have a significant impact on the analysis results of this study. In contrast, using the latest bacterial and archaeal taxonomy training set No. 19 of RDP (update at 2023-08-23, https://sourceforge.net/projects/rdp-classifier/files/), SILVA and GreenGenes, etc. for this study would be better. I strongly recommend

the classification system of the RDP database for culture-enriched 16S amplicon analysis, which has been proven to be more consistent with the classification results of NCBI.

Re: Thank you for your suggestion. The original intention of choosing the EzBioCloud database was its inclusion of type strains and clone representatives, which provided reliable references for the novelty of our described ASVs. As far as we know, the SILVA and GreenGenes databases do not provide reliable information on whether sequences come from type strains, although this does not affect their authority as general classification annotations (especially at the genus and higher taxonomic levels). We appreciate the information you provided about the RDP database, which we had not noticed before that it included annotations of type strains. The downloadable database from EzBioCloud has not been updated since 2018 (containing 15,925 sequences of type strains and 48,675 clone representative sequences), In contrast, taxonomy training set No. 19 of RDP includes annotations from nearly 19,000 sequences of type strains. We also notice that the RDP database has a limited number of sequences from non-type strains (especially those from cloned uncultured organisms), with fewer than 6,000 sequences, which may result in insufficient general classification annotations for the entire microbial community (especially when classifying "dark matter" taxa to lower taxonomic ranks). However, since our results did not display overall microbial community annotation results, using the RDP database is appropriate. The modified figures include Figure 3-a, c, e, f, g, 4-a and 6-b,c. It appears that there are only minor changes between the new results and the original ones.

Line 237: As shown in Figure 1A, it seems like two methods were utilized for the determination of PCC, i.e., fluorescent stained cell assay and plate-based CFU calculation, but from the method description I can only find the CFU method, is there a misrepresentation here? As mentioned by the author in the Introduction, the issue of “great plate count anomaly” needs to be noted. The disparity between the total cell count, typically determined through microscopy or flow-cell cytometry, and the colony-forming units (CFUs) on agar plates can differ by several orders of magnitude in typical environmental samples. So, why does the authors still use the agar plate-based CFU method to count cells in this study? Is there any other better way to replace it?

Re: Thank you for your feedback. Regarding fluorescent staining microscopic counting, we have already described the methodology and cited references (L140 in the original version, Line 431-433 in the revision) of the original manuscript. To be more readable, we have supplemented this information in this section (Line 521-523). It is important to note that while we statistically calculated the PCT, we also employed the conventional method of calculating the agar plate count (PCC) for comparison. The results indicate that our calculated PCC is comparable to that of many previous studies, while the PCT is several times higher than PCC. Our findings suggest that the conventional cultivability calculated based on PCC may not represent the values of PCT, which are of greater concern from the perspectives of cultivable resource and diversity.

Line 223: How does genome-resolved metagenomics study contribute to the core theme of this article?

Re: Thanks. We have supplemented the reasons for conducting the genome-resolved metagenomics (Line 297-299).

Line 224: From 515 (394 culture-enriched samples), only 6 inoculum samples and 17 derived from AS were used for metagenomic analysis. Why the authors selected only these few samples? What were the parameters used to select those?

Re: Thank you for this question. Apparently, the cost of genome-resolved metagenomic sequencing is much higher compared to amplicon sequencing. More importantly, we believe that conducting metagenomic sequencing on all or a large number plates is unnecessary. This is because, for specific samples, the taxa with the high abundance on the plates are similar across all plates of the same culture conditions, while the differences mainly lie in low-abundance taxa.

To answer this question, we conducted an analysis for ASVs that is over 0.5% in at least one plates (assuming that we would expect to obtain their corresponding MAGs when sequencing all plates) for a specific growth condition of a sample (the plate number is six in most cases). The probability of detection of these ASVs on any single plate of the corresponding growth condition was 61.6% (median, ranging from 45.6% to 80.3% for 10 analyzed datasets). Therefore, sequencing more replication samples may not yield significant gains in new genomes. During the experimental design, we selected a small subset of representative samples primarily based on their alpha diversity (particularly avoid plates where few taxa dominate), beta diversity (avoid repeatedly generating the same MAGs) and distribution of novel taxa as determined by the amplicon sequencing results. Explanations are provided in the revised version at Line 609-611.

Line 313-317: “For inoculum samples, MAGs with marker genes (DNA gyrase subunit B and RNA polymerase subunit B) … analysis (i.e., putatively cultivable though the genome was not obtained)”, could the authors provide a more detailed explanation of this handling method for MAGs from inoculum samples?

Re: Thank you for your question. We realize that the description here was not very clear. We revised related content as (Line 631-642): “To categorize MAGs into cultivable or uncultivable phenotypes (Table S3), the following steps were carried out. MAGs obtained from agar plate samples were designated as cultivable phenotypes. However, not all MAGs derived from inoculum samples were directly defined as uncultivable ones for two reasons. First, there were highly similar MAGs (ANI>99%) can be obtained from both cultivation sample and inoculum samples. Second, MAGs can only represent nearly 40% reads (ranging from 14-70%) for each agar plate metagenome. The remaining reads may match MAGs that were obtained in the inoculum metagenomes. Therefore, the uncultivable MAG was defined as following. If selected marker genes (DNA gyrase subunit B and RNA polymerase subunit B) of a MAG extracted from inoculum samples were absent from or only be detected in the agar plate metagenomes with < 1× coverage per 10 GB (based on read mapping and average of the two genes), this MAG was defined as uncultivable MAGs.”

Line 433-437: It is inappropriate to place this paragraph (characteristics of MAGs) here as it does

not serve the theme of the results “A substantial number of cultivable taxa are novel but are present in low relative abundance”. The results can be moved to the section in Line 508.
Re: We cannot fully agree to move this paragraph to later. This part of the results cannot align with the content in Line 508. However, it fits the current section title of " Substantial cultivable taxa are novel but rare on plates," as it defines novelty based on the taxonomy of MAGs.

Discussion
I strongly recommend that the authors point out the key limitations, improvements, and future perspectives of this study in the Discussion, rather than smugly recounting only what they did.
Re: Thank you for your suggestion. We have made significant changes to the Discussion section, including removing the content on r/k selection and adding supplements on the limitations of this study.

Mirror comments:

Line 143: Change “filtrate” to “filtrates”, and why is the volume of suspension for each sample in a large range of 50-100 mL?
Re: Thank you for correcting the grammar. It has been correcting. The filtration volume is determined by whether it is blocked and unable to continue filtering. There are differences in the degree of blockage among different samples

Line 152: Change “MgSO4 ·7H2O” to “MgSO4·7H2O”, delete the space before dot.
Re: Revised. Thanks.

Line 157-158: It seems illogical for the authors to use a larger 0.22 μm filter membrane for sterilization after using a 0.1 μm filter membrane with a smaller pore size.
Re: This is a matter of experimental trick. The purpose of this filtration step is to prepare the extraction medium, ensuring that the final filtrate collected is sterile. Please note that we did not subject the extraction liquid to high-temperature and high-pressure sterilization afterward, but merely heated it to 50°C and mixed it with sterilized agar. The main consideration for doing so is that sterilization may cause irreversible changes to the components of the extraction liquid, thereby affecting cultivable taxa.

We were not confident in sterile operation using large-volume filtration devices. However, if we were to use conventional manual sterile syringe filters (0.2 μm pore-size), it would require excessive manpower (considering the high concentration of particles in the extraction liquid), and some ultramicroorganisms may pass through the 0.2 μm membrane. Therefore, we first used a 2.7-0.45-0.1 μm membrane for non-sterile operation with a large-volume filtration pump, and finally, we used a 0.2 μm membrane for manual sterile operation, reducing workload, ensuring the filtrate is sterile, and effectively eliminating the influence of bacteria smaller than 0.2 μm.

Line 162-163: “…autoclaved 6% agar (in ddH2O) was mixed in a 4:1 ratio with prewarmed (50°C)…”, The method description doesn’t seem very clear, so what is the final proportion of agar?
Re: The final concentration of agar is 1.2%. Supplemented.

Line 177: Change “the kit” to “the same kit mentioned above”
Re: Revised.

Line 178: Change “To each plate” to “For each plate”
Re: Revised and thanks.

Line 223: The section “Metagenomes and MAGs annotation” should be moved before “Comparative genomics guiding the promotion of cultivation of a specific taxon”
Re: Thanks. Revised.

Line 230-232: How to perform the dereplication of MAGs? using dRep? If so, please cite the related reference. In the METHOD section, the author uses a lot of methods, but the order of description and logic as a whole is puzzling…
Re: It should be noted that we have already performed intrasample dereplication of MAGs using dRep as part of the metaWRAP processing of a metagenomic dataset. The secondary dereplication mentioned here is intended to remove highly similar MAGs from different metagenomic datasets (e.g., MAGs with ANI>99% from agar plates under different culture conditions of the same sample). As described, we directly conducted all-versus-all average nucleotide identity (ANI) calculations for all MAGs (using the FastANI algorithm) and retained the higher-quality MAG in any MAG pair with an ANI greater than 99%. The description has been revised accordingly. The reference of dRep is cited in the revision (Line 619).

Line 262: Add a space after “EBPR01”
Re: Revised. Thanks.

Line 334-337: There are two ways to write VB12 here, “vitamin B12” and “Vitamin B12”, please confirm and unify it.
Re: Thanks. We unify it as vitamin B_{12}。

Line 341: “Rhodocyclaceae”, “Burkholderiales”, and other Latin names of prokaryotes in the whole manuscript are suggested to use italics.
Re: Thanks. We also noticed the recently published article (Https://doi.org/10.1186/s43008-020-00048-6). Revised.

My conclusion
The core theme, and value, of this paper is actually around the “an experimental approach to reassess the cultivability of environmental bacteria and a promising path to optimize microbial cultivation”. This I totally agree with.
However, I have a number of misgivings about the introduction of the PCT concept and calculation methodology, which may lead to a significant loss of credibility for many of the results (see also my comments about statistical validity). In addition, the manuscript suffers from another major problem - the authors attempt to solve too many problems at once using multiple methods, making this paper complex and confusing. Furthermore, the culture-enriched metagenomics of have been applied and

interpreted in detail in both soil and activated sludge environment (e.g., Li et al., npj Biofilms and Microbiomes, 2023; Zhang et al., Environmental Science and Technology, 2022), which poses a challenge to the novelty of this manuscript.

In summary, I seriously evaluated this manuscript is not suitable for publication on Nature Communications, but this does not mean an underestimation of the important findings of this study.

Re: Thank you for your assessment of the importance of our research. Here, I would like to summarize two conceptual advances of this study for your reference, both of which are reflected in the title, namely the re-estimation and promotion of cultivability for environmental bacteria. First, we proposed an experimental method for calculating PCT and compared it with the widely determined PCC, demonstrating the underestimation of taxonomic cultivability by PCC. Second, we consider the ability to grow and detect under specific culture conditions as a phenotype of environmental bacteria, and this phenotype-genotype association information can guide isolation cultivation. Regarding the first point, although Martiny et al. (2019) introduced the concepts of PCT and PCC, no one has yet provided an experimental pipeline for calculating PCT. While we acknowledge that there is room for improvement in this method, such as using full-length sequencing as you mentioned, we believe this is a step in the right direction towards assessing and improving the cultivability of environmental bacteria. The second point enables omics information to be better utilized to guide isolation cultivation, moving away from analyzing and metabolically reconstructing specific MAGs alone (usually lacking specificity). It is also worth noting that, unlike many complex methods, the methods used in this paper are generally applicable.

We appreciate your for providing these references. We have reviewed them and compared them with our research. We also included another reference guiding isolation through amplicon sequencing, which focused on AS and was just published after our submission. Here is the information for these references:

1. Li et al., 2023. "Capturing the microbial dark matter in desert soils using culturomics-based metagenomics and high-resolution analysis." https://doi.org/10.1038/s41522-023-00439-8

2. Zhang et al., 2022. "Using Culture-Enriched Phenotypic Metagenomics for Targeted High-Throughput Monitoring of the Clinically Important Fraction of the β-Lactam Resistome." https://doi.org/10.1021/acs.est.2c03627

3. Verhoeven et al., 2023. "Amplicon-guided isolation and cultivation of previously uncultured microbial species from activated sludge.". https://doi.org/10.1128/aem.01151-23

Among these references, the relationship between reference 2 and our study is relatively week, as it focuses on antibiotic-resistant bacteria and resistance genes. References 1 and 3 are cited in the revision. Both studies used either amplicon sequencing or metagenomic sequencing to characterize cultivable microbiota on agar plates and even described some novel taxa. Li et al., (2023) attempted guided isolation cultivation using taxonomic information, while Verhoeven et al. (2023) compared the effect of adding in-situ nutrients to the culture medium for cultivation improvement and attempted isolation. However, from the results, they did not ultimately achieve the isolation of species at high taxonomic ranks, whereas our study obtained a novel class-level representative.

More importantly, these references did not cover the two innovative points we mentioned above.

Reviewer #3 (Remarks to the Author):

Zheng and coauthors present a well-written and well-explored assessment of the percent cultivable taxa in soil and wastewater. They find that the percentage of cultivable taxa is higher than one might guess from the low numbers of percent of cultivable cells. Furthermore, they explored features associated with uncultivability and found that smaller, less metabolically active cells, often with long doubling times, were over-represented in the uncultivable taxa. Then they identified genomic features associated with uncultivable Burkholdaria and showed through experiment that these features could be used to enhance cultivability. I think this is an important contribution to our understanding of what makes microbial taxa cultivable, and it is nicely done. I just have a few suggestions below:

Re: We appreciate that you summarized the innovative points of our study.

Line 76: saying that Martiny based his conclusions on "the" 16S database is misleading. It would be better to say that he did is study on "a" 16S database. Using "the" suggests that Martiny considered the entire public database, instead of a handful of samples. The study that used the entire NR database was Lloyd et al., 2018 https://doi.org/10.1128/msystems.00055-18, that was the precursor to Steen et al., 2019 that is cited in this paper. Lloyd et al. found that the proportion of cultured taxa was also very low on average, although individual samples could have quite high proportions of taxa represented by cultures. Notably, human host environments. Given this paper's findings, the assertion on line 71 that low numbers of uncultured taxa is a "misconception" is not fully supported by the literature. I think this section could just be tempered by pointing out the two options, rather than implying that Martiny's view is widely-held to be true.

Re: We appreciate your suggestion regarding Martiny's article. However, upon careful review of Lloyd et al., 2018, we found that this study focused on the proportion of cultured or uncultured taxa relative to sequenced data, which is somewhat closer to what we described as PCC rather than PCT. It should be noted that both articles analyzed cultivability using sequencing data, but Lloyd et al., 2018 focused on PCC, emphasizing the predominance of uncultivated cells in natural ecosystems, while Martiny calculated both PCT and PCC. Since our introduction here pertains to PCT, we cited Martiny's study here, but we do not imply it is widely regarded as correct. The methodological concerns raised by Steen et al., 2019, which we cited, have cast doubt on it.

The results of Lloyd et al., 2018 seems align reasonably well with previous experimental conclusions on PCC calculation. In the revised discussion, we address possible reasons underlying the differences between PCC and PCT, citing Lloyd et al. accordingly (Line 357).

The fact that the authors find that PCT is greater than PCC is an expected outcome if the uncultivable taxa are in greater cell abundance than the cultivable ones. This much greater abundance of

uncultivable taxa is shown by the very high percentages (44-84%) of the culturomics ASVs being too rare in the inocula to appear in their 16S libraries. This makes me wonder about the conclusion on lines 399-401 that more abundant ASVs are more likely to appear in the cultures. Is this still true if you include the high percentage of ASVs in cultures that are too rare to appear in the inoculum libraries at all?

Re: It's an interesting question about how the cultivability of rare taxa that were not detected in the inoculum. It must be acknowledged that our data cannot determine this value because our calculations must be based on taxa detected in the original samples. If cultivable taxa not detected in the inoculum are included, the cultivability rate of low abundance taxa must significantly increase. However, such output would be biased as the number of ASVs that are undetected and uncultured is unclear.

We can only infer from the trend in Figure 3c that these taxa may have a similar PCT to low abundance taxa (e.g., detected but below 0.01% taxa). Two aspects, which we did not further discussed in the manuscript, can influence this result. First, the number of cells plated on agar plates is a bottleneck; most of our inoculation plated ～10^4 cells on each agar plates, so the probability of being plated might not be high for taxa with even lower rarity, while relatively higher abundance taxa (e.g., >0.1% taxa) were almost 100% plated on every plate, and their number of plated cells might be high. This leads to the second point. Microbial cells in the environment vary individually, some may be dormant with low metabolism, or be in a viable but non-cultivable state, or even dead. Cells with higher abundance in the original sample are more likely to have active cells that can grow on agar plates.

The conclusions on 411-413 do not appear to be supported by Figure 2D. Here, it shows that cultivability decreases for increasingly phylogenetically distant OTUs. This, it seems to me, would suggest that lower phylogenetic similarity results in lower cultivability. However, this doesn't seem like the best test to answer this question. It might be better to look at the distribution of similarity of the culturomics and the inoculum ASVs to their closest cultured relative.

Re: The data of Figure 2d means: when at least one ASV within an OTU (defined by different thresholds) is cultivable, what is the probability that the other ASVs of the OTU are cultivable. Regrettably, upon carefully checking, we found two faults in this figure: 1) Although mentioned in the main-text, in the figure legend we overlooked explaining the filtering condition for OTUs, i.e., they must contain 5 or more ASVs to reduce the statistical unreliability of OTUs containing only few ASVs; 2) The original vertical axis should be quartiles (0-25-50-75-100%), not the previous 0-40-60-80-100%. We have made corrections for these issues.

The purpose of this figure is to illustrate the likelihood that taxa phylogenetically related to a cultivable ASV can be cultured under similar or identical conditions if we determine that a certain ASV can be cultivated under specific conditions. For example, if a 0.97-OTU contains 10 ASVs and among them, 5 are cultivable, then the percentage of cultivable ASVs within the OTU is calculated as (5-1)/(10-1) = 4/9 = 44.4%. The basis for our conclusion here is that even using the results of

0.97-OTUs, less than 25% (median value, both soil and activated sludge) of closely related ASVs can be cultured and detected. This figure demonstrates the low reliability of predicting cultivability based on ASVs, which is not concluded from the trend of predictability decreasing as similarity decreases (only valid for soil samples) shown in the figure.

Unlike our emphasis on cultivability predictions under specific cultivation conditions, your suggested method, which compares the cultivable and uncultivable ASVs with a database of cultivable strains, is the approach used by Martiny (2019) and Lloyd et al., (2018). This method does not consider specific cultivation conditions but rather offers a more generalized prediction. We tried your suggestion. The results are shown in the Figure S6 in the revision and below. Clearly, this result indicates that cultivable taxa indeed have higher similarity to already isolated strains compared to uncultivable taxa. This suggests that, in general, easy-to-culture taxa are preferentially cultured and described, as well as phylogenetically related. However, this does not imply that using the same cultivation conditions reliably cultivates strains phylogenetically close to each other within samples. Above description is supplemented in the main text of the revision (Line 189-194).

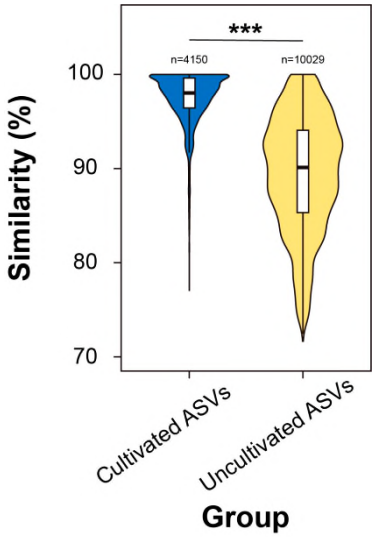


Figure S6. The phylogenetic relatedness (similarity of 16S rRNA gene) of cultivated and uncultivated ASVs with the type strain databases (RDP taxonomic training set version No. 19). Comparison analysis was performed using Wilcoxon-test (three asterisks for P<0.001).

Lines 447-449: I don't know what it means to have a diameter below 94%.
Re: Thanks. It is a grammatic mistake and revised.

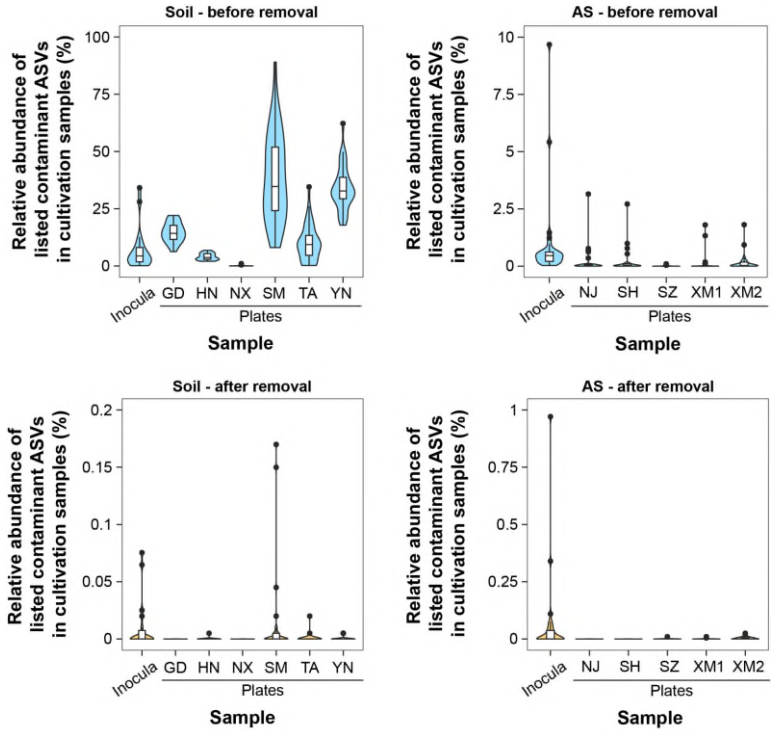
Line 133 should read "the times ranged from 0.5 h to 1 d.
Re: Thanks, and revised.

Line 220-221: What was the percentage range of contaminants across all samples? Were contaminate ASVs removed from both culturomics and inocula samples?

Re: Thank you for this question. According to your question, we found that the listed contaminant ASVs indeed exist at higher abundances (10-30% for median values) in soil cultivation samples (see the figure below). They were indeed removed from the downstream analysis. We hadn't noticed

this phenomenon before, since the two inner standard libraries were developed only two for activated sludge samples, which means that our previous criterion of "For each ASV, if its relative abundance in the 3×10^3 internal standard sample (the average value of triplicates) exceeded the maximum relative abundance observed in all experimental samples from the same library, it was considered as contamination within our experimental system" might have removed some ASVs that actually exist, even dominate, in soil samples, although this seems not significantly affect the results of PCT values. Nevertheless, we've implemented new criteria for removing contaminant ASVs as follows: "For each candidate contaminant ASV (listed in Table S2), if its relative abundance in the 3×10^3 internal standard sample (the average value of triplicates for each library, referring to the higher value in the two libraries if one ASV was detected in both) exceeded the maximum relative abundance observed in each sample, it was considered as contamination only for the corresponding sample (or library). Under this new criterion, the percentage range of contaminants across all samples is shown below (<1% in all samples, <0.2% in >99% samples), which is much lower than before. In the revision, Figures 1b, c, d, Figures 2a, b, c, d, Figures 3a, and c have been revised accordingly, while the results are almost unchanged. The Figure S1 was revised accordingly.

It's important to note that defining and removing pollution sequences remains a major challenge in microbial community analysis, especially when inter-sample cross-talking may exist and is difficult to be defined. We cannot be entirely confident that our method completely removes all pollution ASVs or correctly retains all true ASVs. However, since the bacterial biomass of agar plate samples and inoculum samples are usually high (several to tens of μg DNA can be obtained for each sample), we believe the contamination is not severe in our cases and the it has little



The relative abundance of determined contaminant ASVs in samples by using the previous criterion (upper panels) and current version (bottom panels).

Lines 482: This is incorrect – the value for metabolic activity is the rRNA per rDNA, not rRNA per cell. This is an important distinction because I got very confused about how cell counting was achieved for uncultivable taxa and only figured it out after re-reading the methods.

Re: We performed a conversion from rDNA to cell counts by normalization using the genomic copy number information provided by the rrnDB (https://doi.org/10.1093/nar/gku1201). Specific details can be found in the original version at Line 302-306 or the revised version at Line 599-604. Assuming a bacterium has three rDNA copies in its genome, its cell count is determined by dividing the number of its absolute rDNA copies (in one unit volume of the sample) by 3.

563-565: I urge the authors to remove claims of priority and instead simply state the new things they have found, without pointing out that they are new.
Re: Thanks. This sentence was removed in the revision. The discussion section was largely revised according to the comments from you and other reviewers.

Were any ASVs present from phyla or classes where no cultures have ever been obtained? If so, I think it would be helpful to put this study into context that while the authors new approach is extremely useful for obtaining broader taxonomic representation among cultures, there may be deeply-branching clades which are facing even steeper barriers to cultivability.

Re: Sorry we are not able to locate this feedback in the manuscript. Following up on the last question, we believe this might be in the discussion section. As evidenced by the results of this study, which yielded a pure culture representing a novel class, this method demonstrates generalizability. Discussion was added in Line 382-392 in the revision. Of course, it also requires improvements in cultivation methods, including culturomics, cultivation devices, and the phenotype-genotype association-based mining used in this study.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Review: NCOMMS-23-54252A - Sequencing-guided re-estimation and promotion of cultivability for environmental bacteria.

I must admit that the MS is much improved and reads much better than before. However, I still have many comments that should be addressed before the MS can be published.
Re: Thank you for your comments. We have attempted to address each of your points below.

Specific comments:
- l. 114 - omit easy to be collected and transported - very poor reason
Re： Deleted.

- l. 115 (and elsewhere) - oxygenic or aerobic - these two are very different
Re: Revised here and elsewhere. We use the term “the availability of oxygen” here. It includes aerobic, anaerobic and microaerobic conditions.

- l. 137-8 or Fig. 1 legend - a clear description of what plate, condition, and sample means (information in brackets is not clear)
Re: Thanks. Revised (see L147-150 and L1013-1015 in the revision).

- l. 142-151 - I think this information needs to be presented differently. As I was reading the text, I thought the following: this information is very crucial, however it makes me full of doubts about reaching the goal of the study - the number of both taxa and cells is always compared with the inoculum; I think the authors should consider presenting the results separately for the taxa/cells that can be compared with the inoculum and those that cannot be; in addition, the information on line 147-8 makes me wonder how similar these ASVs are - 75% is not a lot and 94.7% is still not sufficient to be the same species ; finally, can authors discuss how likely it is that these ASVs that do not match represent a crosstalk between samples, etc.; did the authors use something like a mock community? I think it would be especially valuable in this case. THEREFORE, I suggest the information from the end of the paragraph be presented earlier. At the same time, % numbers of ASVs that were omitted this way need to be presented in the text, not just in supplementary material.

Re: Thank you for these comments. We have re-adjusted the description of this part of the results. In short, before presenting the PCT and PCC comparison results, we first evaluated the impact of contamination sequences and cross-talk, then explained the PCT and PCC calculation methods, pointed out the judgment situation of ASVs not detected in original samples.

You mentioned the following description: "In addition, these ASVs exhibit significantly lower frequencies in the agar plates compared to other cultivable ASVs (Figure S3), with most of them having phylogenetically related ASVs (74.1-94.7% of them have a >94% similarity relative

detected in the inoculum) in the inoculums (Figure S4)." The intended meaning here is that most of these undetected ASVs have closely related phylotypes already detected in the same inoculum, suggesting their presence is reasonable (or can be explained by niche selection), rather than coming from unknown sources. To make this clearer, appropriate modifications have been made. At the same time, we have moved previous Figure S2, which describes the proportion of these ASVs, to the main text as Figure 1b to emphasize the importance of these results.

Regarding the overall assessment of cross-talk (not just specifically for those cultivable ASVs undetected in the inoculum), we previously did not explain in detail due to length considerations. Based on your suggestion, we have supplemented this in the main text at L140-145 and L501. In our experimental design, unique dual barcodes were applied, meaning all different samples mixed in one sequencing library do not share any barcodes. According to literature, this greatly reduces the possibility of cross-talk of amplicon sequencing (https://link.springer.com/article/10.1186/s12864-017-4428-5). Our laboratory has applied this method in recently published studies (https://doi.org/10.1007/s00253-019-10183-9; https://doi.org/10.1186/s40168-021-01214-7).

More importantly, amplicon sequencing data also support that the impact of cross-talk is very low. Firstly, although we did not use a mock community, the sequencing results of the single internal standard bacteria showed that in the two sequencing libraries involved, we only found 2 reads of the internal standard were detected in all samples (=0.0007%, see Figure R1a below). Considering that there were ~50% of all cleaned reads of the library 2 is NS6-2, such low cross-talk indicates its impact was negligible. Conversely, when copy number of the internal standard is relatively high (see Figure S1, $>10^5$ copy of 16S rDNA/reaction, which is obviously lower than the copy number of all experimental samples, considering that 10-20 ng genomic DNA were added into each reaction and 1 ng genomic DNA may contain 10^5-10^6 copies as estimated by referring to *E. coli*), the sum of contamination sequences and cross-talk sequences (we cannot discriminate the two types of errors) is less than 0.05-0.1% in average (Figure R1a). It also supported that cross-talk was extremely low in the experimental samples. Secondly, we can find several uncultured phyla (e.g., *Parcubacteria* and *Candidatus* Saccharibacteria) with high relative abundance in certain inoculum (>2% for maximum values), but these phyla are completely undetected in all culture plate samples (see Figure R1b below). If cross-talk occurs randomly, this phenomenon also supports that it has almost no effect of cross-talk.

We have placed the descriptions about cross-talk in supplementary Figure S2, and the description in the main text has been mentioned earlier (L140-145). Thank you again for these comments.

a

	Library 1			Library 2		
	Experimental samples	Internal standard samples	Internal standard samples (*NS6-2* copy number >10⁵)	Experimental samples	Internal standard samples	Internal standard samples (*NS6-2* copy number >10⁵)
Total reads	528753	237378	71873	289142	268584	91443
Reads of *NS6-2* (proportion)	0	228726 (96.36%)	71837 (99.95%)	2 (0.0007%)	261156 (97.23%)	91344 (99.90%)

b

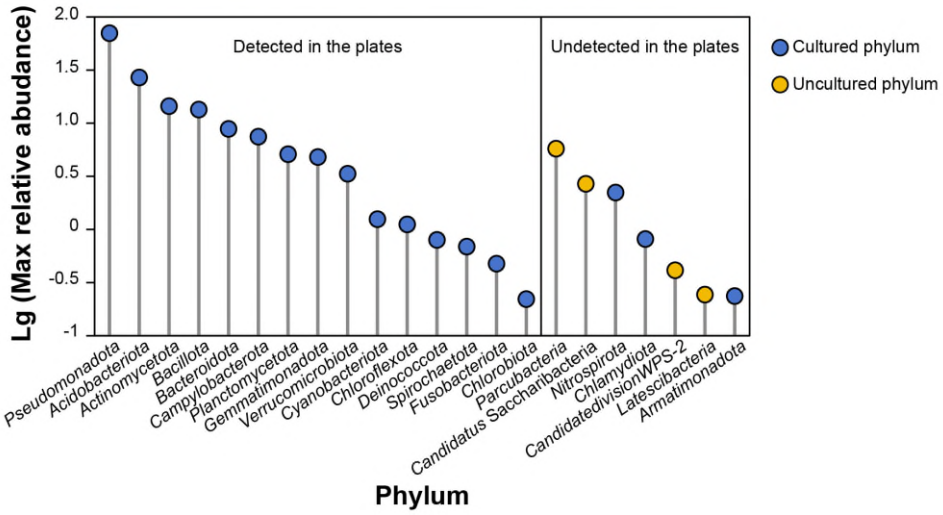


Figure R1. Evidence supporting that the inter-sample cross-talk had a negligible impact. a Almost no internal standard reads were detected in the experimental samples. **b** The detection for phyla with maximum relative abundance >0.2% across all the 11 inoculum samples in the agar plate samples. Although some uncultured phyla, such as *Parcubacteria* and *Candidatus* Saccharibacteria are dominant in certain inoculum, they are completely undetected in the agar plate samples. If the cross-talk was severe and occurred randomly, these taxa should have been detected in the agar plate samples. Their absence supported that the cross-talk should not be serious.

- l. 153 - the statistical difference is one thing, biological meaning is another - bearing this in mind, the information should be presented in a way that for soil the PTC is significantly higher than PCC with the difference in mean being x %; for the AS, however, these numbers are comparable, yet still significantly different. At the same time, at the end of this paragraph, the authors should explain what the differences between plate, condition, and sample mean. This information is completely missing.

Re: Thank you for this valuable comment. We have made modifications according to your suggestions. Regarding the differences between plate, condition, and sample-level, our explanation is as follows (L175-178): “That is, as the number of plates and cultivation conditions increased, PCT gradually rises, with the extent to which it exceeded PCC progressively increasing. Above results clearly demonstrated that PCC could significantly underestimate the cultivable diversity of environmental bacteria.”

- l. 160-1 - weird sentence - it is not true for PCT?

Re: Please see the revision (L175-178). Here, we emphasize the comparison between PCC and PCT in our study. As stated in the article, a prevailing misconception is the confusion between the proportion of cultivable cells (PCC) and the proportion of cultivable taxa (PCT). Nearly all previous studies measured PCC, rather than considering the cultivable taxa (e.g., species) in the context of total taxa present in the inoculum. Our results demonstrate that PCC is significantly lower than PCT, indicating that the previously used PCC can substantially underestimate the cultivable diversity of environmental bacteria.

- l. 176 - be consistent with tenses, use past simple throughout the MS
Re: Revised throughout the MS.

- l. 179 - define at least briefly the extraction medium, maybe by saying based on… this is an issue of the results section being presented before the methods (things should be defined as they first appear in the text); Fig. 2 legend should also define the extraction medium; additionally, I would not call R2A artificial, maybe common laboratory defined medium is better
- oxygenic conditions is a term that needs to be replaced throughout the MS
Re: We have briefly introduce the extraction medium in the Introduction section (L118) and here (L195). The information of extraction medium is also briefly supplemented in the legend of Figure 2. R2A and TSA media are categorized as common laboratory media, because the strict definition of a defined medium is a culture medium without complex components, which they do not meet. The oxygen conditions have been revised to the availability of oxygen.

- l. 192 - as well as phylogenetically related; what is meant here?
Re: We deleted this phrase.

- l. 224 - the information in the brackets is given for species? It should be noted.
Re: Yes, modified (L242). Thanks.

- l. 225 - is, not are
Re: Modified. Thanks.

- l. 232 - what does this information say here?
Re: See L250 in the revision. We revised the sentence as: “These low-abundance ASVs remind us of the phenomenon of microcolonies commonly reported on plates in previous literature (https://doi.org/10.1126/science.1070633, https://doi.org/10.1128/aem.71.12.8714-8720.2005). Indeed, we identified unneglectable numbers of microcolonies (diameter <300 µm, approximately 20% to the number of microcolonies, see Figure 3d and Figure S7).”

- l. 251 - omit ‘therefore increasing our confidence in isolating this bacterium’
Re: Deleted.

- l. 252 - fix wording - the isolate was not obtained by 16S rRNA gene sequencing

Re: Deleted the phrase. Thanks.

- l. 258 - I suggest including also a threshold of 98 % to perform this analysis
Re: Thank you for your comment. We have redone this part of the analysis. However, if we use a 98% threshold for filtering, the number of total matching sequences drops dramatically from over 70,000 to 228, and in no sample does the relative abundance of this bacterium exceed 0.07%. More importantly, upon reanalysis, we realized an issue we had not previously considered: although the Earth Microbiome Project sequences used for analysis are from the V4 region, they are only 90-150 nt incomplete short fragments (https://doi.org/10.1038/nature24621), which may not allow for reliable taxonomic annotation through BLASTN alignment with a single EBPR01 reference sequence.

Evidence supporting this concern is as follows: if we take the first or the end of 100 nt of the V4 region, some sequences from other classes (e.g., *Ignavibacteria* and *Kapabacteria*) within *Bacteroidota* (Figure 4b) will have over 95% similarity with sequences in *c_JABWAT01* (3 sequences, EPBR01 and the other two extracted from MAGs JABWAT01 and JAFLBE01). This result indicate that we cannot conduct robust analysis for the ecological distribution of this lineage based on the EMP 16S amplicon dataset, which represents the most extensive ecosystem sampling. We also attempted to search the NCBI databases using all full-length 16S sequences of *c_JABWAT01*, but the results showed that almost no deposited full-length sequences had over 92% similarity with these sequences (<5 hits with >90% similarity). This is also consistent with the result that no EMP samples contained EBPR01-related phylotypes at a relatively high abundance. Therefore, we had to abandon the 16S-based ecological distribution analysis for this elusive lineage. As a result, we have removed Figure 4d and all related content, retaining only the analysis based on the source of the MAGs affiliated with this class.

- l. 297 - to decipher; overall the grammar should be checked again throughout the MS
Re: Revised. Thanks. We made corrections to grammar problems during revising the MS.

- l. 296 - I wonder if this whole chapter should not be placed before line 243; the text on line 243 and beyond already takes into account genome sequencing
Re: Indeed, some of the information about genomes has already appeared earlier. However, it is necessary to provide additional explanation about the relevant genomes here. Therefore, we have modified the first few sentences of this paragraph as follows: “For the MAGs from 6 AS inoculum samples and 12 corresponding agar plate samples (detailed in Table S1), we defined MAGs with cultivable and uncultivable phenotypes (see Methods). This discrimination allows us to perform comparative genomics to screen for genotypes associated with the phenotypes.” See L312 in the revision.

- l. 328 - the all - grammatically wrong
Re: Revised.

- Fig. 6 - use relative abundance, not relativated
Re: Modified. Thanks.

- l. 330 - abundance, not abundances; l. 332 - abuof? Once again, the grammar and typos need to be fixed throughout the MS! Other obvious mistakes: L. 345, 357, 362 (still)…
Re: Modified. Thanks.

- l. 334 - to some extent is not a very scientific expression, be specific
Re: We deleted the phrase.

- l. 358 - the phenomenon of dormancy should be noted here
Re: Added. Thanks.

- l. 367-70 - can authors discuss how likely it is that these ASVs that do not match represent a crosstalk between samples, etc.; did the authors use something like a mock community? I think it would be especially valuable in this case.
Re: Thank you for your comment. The issue of cross-talk has been addressed in detail earlier. Considering that other researchers will adopt the methods used in this study, we fully understand the rationality of your suggestion to strengthen the methodological descriptions. As per your earlier suggestion, we have already explained the potential impacts of contamination and cross-talk, as well as the assessment results, at the beginning of the *Results* section (L140-145).

- l. 387 - use undescribed or hitherto uncultured rather than novel
Re: Modified. Thanks.

- l. 389-97 - states of low metabolic activity should be discussed more along with how certain approaches can help with this phenomenon during cultivation, such as the use of resuscitation factors or signaling molecules; such a discussion can be well linked with the discussion on B12 vit in the medium.
Re: Thank you. We have added a discussion on improving cultivability through approaches that activate bacterial metabolism, such as using signaling molecules, resuscitation factors, and iron carriers. (L406-408)

- What I am missing in general is the number of CFUs that appeared on the plates before the colonies were used for sequencing. This information should appear in the text.
Re: In L458 and 472 (the last version) of the Methods section, we have mentioned the number of cells (typically 5,000-500,000) plated for CFU counting and the approximate range of CFU counts per plate (100-800 in most cases) used for PCC and PCT calculation. In this revision, the related information can be found in L473 and 488.

After these comments are addressed, I think the paper will make a valuable contribution to the journal.
Re: We appreciate your comments, which have helped to further improve this paper.

Reviewer #2 (Remarks to the Author):

The manuscript entitled “Sequencing-guided re-estimation and promotion of cultivability for environmental bacteria”, was partially revised following reviewers’ feedback. The manuscript was somewhat enhanced by incorporating some of the modifications and addressing of concerns suggested by all the reviewers. However, in the process of revision, many enhancements have not been substantially improved, and the manuscript still suffers from several inescapable problems. In short, I think this manuscript is not suitable for publication on Nature Communications.

Re: We appreciate your previous comments, which have helped to further improve this paper.

Reviewer #3 (Remarks to the Author):

Thanks for the good discussion. I think I understand better that the authors have a different (equally valid) definition for cultivability that I do. The authors' definition is that a cell or a taxon is cultivable in this experiment with these conditions, whereas I tend to think of it as an absolute designation - can this cell or taxon be cultivated by someone somewhere using any possible method. Seeing this distinction more clearly helps me understand the paper better. The revisions look great, and I congratulate the authors on their good work.

Re: We appreciate your previous comments, which have helped to further improve this paper.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The revisions that have been made seem generally OK, but some issues are pending.
Consider changing microaerobic to microaerophilic.
Re: Thanks. The word has been changed throughout the manuscript.

L. 136: Present the aim/objective/novelty of the paper first, or at least some results, before getting to the point of contamination. That said, Fig. 1c should probably be presented before 1b because it is the main result!
Re: Thank you for your feedback. In our original submission, we placed the current Fig. 1b in the supplementary materials rather than in the main text, adhering to the convention of presenting important and novel results first. However, in the last round of review (R2), you suggested moving Fig. 1b to the main text, as the contamination issue indeed requires careful explanation and attention. We believed it was right and we made the change accordingly. Given that in Nature Communications, the methods section is positioned later, we believe the current order of the figures and text is more logical. Specifically, it is necessary to first evaluate the reliability of the data (Fig. 1b & Fig. S1, S2, S3 & S4) before presenting the statistical results, placing the key findings later then. Therefore, we prefer to retain the current order.

I still think that either the English or the wording in general needs to be fixed in some places. See EXAMPLES of new text below.

Inoculums.
That is, as the number of plates and cultivation conditions increased, PCT gradually rises, with the extent to which it exceeded PCC progressively increasing. Above results demonstrated that PCC could significantly underestimate the cultivable diversity of environmental bacteria.
These low-abundance ASVs REMINDED US of the formation of microcolonies with high taxonomic novelty commonly reported on plates in previous literature.
Finally, the requirement of consistent tenses has yet to be met.
Re: Thank you for your valuable suggestions. We have revised the manuscript for language clarity and ensured consistency in the use of tenses.