Response to reviewers

Dear recommender and reviewers,

We would like to thank you for handling our MS and for reviewing our paper. Here are the answers to the pertinent questions you may have had while reading it. We hope you find our answers and clarifications satisfactory.

Reviewer 1

- The relevance of figure 1 is questionable. It does not bring a substantial amount of information. Moreover, data in the text below figure 1 do not seem to confirm the data represented in figure 1: for example, for meat, the text states that 4 species per genus were found for Yarrowia and Cladosporium and 2 species for Candida only. However, in figure 1, only a single dot can be found at 4 species/genus and 3 dots at 2 species/genera. Either there's a consistency problem between the figure and the text (especially for meat) or the data are confusingly expressed.

Thank you for this comment. We would like to keep this figure as it illustrates well the difference in phylogenetic diversity across mock and fermented products. We correct the figure that indeed contained an error for meat data. We are gratefull for pointing it out.

- The ITS region being subjected to significant size polymorphisms (insertion/deletions) as shown in fig2 and fig 6, It is often difficult to interpret/make sense of phylogenetic trees built on sequence alignments. It might be interesting for the authors to elaborate and discuss on the relevance of the trees obtained and shown in fig3

Thank you for this relevant comment. Figure 3 was included to illustrate the phylogenetic diversity of the fungal species in our mock samples. We now explain it at the beginning of the Results section: "*In this study, we compare the efficiency of 4 barcodes (ITS1, ITS2, D1D2, RPB2) and seven bioinformatic workflows to detect the species in microbial community of 4 fermented products (bread, wine, cheese, fermented meet) using mock and real samples. The phylogenetic diversity of fungal species analysed is illustrated Figure 3.*"

We fully agree that the difference in ITS length can hamper the quality of sequence alignments and the estimation of phylogenetic distances. As proposed by Reviewer 2, we replaced the ITS phylogenetic tree by the D1/D2 phylogenetic tree on Figure 3. In addition, we have added the following sentence in the Discussion section: "*Although ITS1-2 seem to be the best barcodes for distinguishing between species, their difference in size may hinder sequence alignment and therefore beta diversity estimates that take phylogenetic distances into account.*"

- Panels should be numbered or labeled (A, B, C, D)
- Precision should be provided in the legend to clarify the difference between the small dots and the bigger dots

We have followed your advice and modified the figures accordingly.

- Figure 6 legend stats "ITS1 and ITS2 amplicon size…". However, it seems that only ITS1 data are presented.

Thank you very much for noticing the mistake. Only ITS1 is represented. The legend has been corrected.

- Figure 8: chosen colors make it difficult to distinguish partially reconstructed and perfectly reconstructed sequences. A better choice of colors would greatly benefit the readability of the figure

The colors of Fig. 8 have been changed to better distinguish the three categories.

Reviewer 2

- Ad Fig. 3: To my knowledge, phylogenetic relationships derived from ITS1 and ITS2 sequences become unreliable above class level. If the authors want to include this figure, I would suggest to use LSU sequences instead.

Thank you for this suggestion. We replaced the ITS phylogenetic tree by the D1/D2 phylogenetic tree on Figure 3.

- Ad "Analysis of real samples" (lines 449 ff.): Results are only discussed in the context of marker choice. This might be missleading. In the case of missing Yarrowia and Candida species when using the ITS1 marker, this can more likely be attributed to the primers used (ITS1F and ITS2) which have known missmatches to those taxa (see Tedersoo and Lindahl, 2016).

Thank you for this relevant remark. We now also discuss primer mismatches issues.

- The preprint is very detailed on the technical side of things but rather brief concerning the biological backgrouns and the presentation of the results for real food samples. Thus, the motivation for this study is not very clear. The comparison of the results of fermented food samples is purely descriptive and and limited to a few selected findings which are only attributed to targeted marker sites. As mentioned briefly by the authors, primer bias is another important factor and should be included. A graphical representation of these results is missing.

Thank you for this comment. Our motivation in including real samples was to validate our conclusions from MOCK communities and to test the effect of the fermented product matrice. We have added sentences in the Results and Discussion sections to precise this aim. The limited number of real samples studied did not allow us to make a comparison of microbial diversity between fermented products. That's the raison why we did not add any figure.

- Lines 125 ff.: A key component of this study are fungal mock communities representative for fermented meat, wine, cheese and sourdough. The authors write that the selection of those species was "based on an inventory of the most frequently described species in the literature". Yet, neither references for the literature sources used for that purpose nor further methodological insight how these species were selected are given.

We completely agree with this comment. Explanations have been added in the Methods section on how species have been chosen and the corresponding references too.

- Lines 294 ff.: Which algorithms were used here? UPARSE or UNOISE (i.e. OTU- or ZOTU delineation)? If UPARSE the comparison of perfectly reconstructed sequences (Fig. 5) would not be sound since OTUs would be compared to ASVs. Similarly, was QIIME2 used with the DADA2 plugin or using a OTU clustering algorithm?

Thank you for this interesting question. Tools benchmarking is always tricky concerning the choice of the parameters/algorithms for each tool. The settings of the usearch and qiime2 tools were chosen taking into account the guidelines (tutorials) provided by these tools on their respective websites and following what was done in the study by Bernard et al., 2021. Here, we use USEARCH with the *unoise* option (ZOTUs) and the *vsearch* option (de novo clustering) for qimme2, as described in https://doi.org/10.57745/109NNP.
The section dedicated to perfectly constructed sequences aims to check whether the tools are capable of outputting the expected sequence at the end of each pipeline. We therefore compare the sequences obtained by each pipeline with those expected, regardless of how they were defined (clustering or denoising).

- Ad Fig. 5: In my opinion, the outcomes of the benchmark of the pipelines used would be more clear and transparent if the results from all four marker sites would be presented separately and not mixed together.

The results from all four marker sites are mixed together to give a synthetic view of the results for all markers combined. However, detailed results for each ecosystem are available in one of the reports mentioned in section Data, scripts, code and supplementary information availability (https://doi.org/10.57745/109NNP)

- Consider re-running some of the analyses using the current version of UNITE (v9.0). The version used here is >2 years old and since then ~9x more fungal sequences and >25% more fungal reference sequences have been added to the database.

The analyses were re-run with the latest version of Unite (v9.0). The manuscript and the reports have been updated. The results for the real samples remain unchanged, as we have manually assessed the ASVs obtained. However, as you can see in Figure 4, many of the ITS2s of our 118 species are much better represented.

- Colours with better contrast could be picked in the Fig. 8 to discriminate partially and perfectly reconstructed sequences.

Your opinion is shared by reviewer 1. The colors of Fig. 8 have been changed to better distinguish the three categories.