

Round #1

Author's Reply:

by Tatiana Giraud, 16 Mar 2022 05:36

Manuscript: <https://doi.org/10.1101/679472>

Revision

The manuscript has been evaluated by two referees, who agree that this interdisciplinary study on differences in terms of species composition between sourdoughs from artisanal and farmer bread is sound and interesting for the evolutionary biology community. Both referees nevertheless provide a list of excellent additional suggestions, such as improving the readability of figures, adding discussion on a broader context and be more cautious about the term domestication.

Thank you for giving us the opportunity to submit a revised draft of our manuscript. We appreciate the time and effort that you and the reviewers have dedicated to providing your valuable feedback on our manuscript. We have done our best to answer all the comments. We have submitted a new version of our manuscript on BioRxiv :

<https://www.biorxiv.org/content/10.1101/679472v3.full>

We have also submitted a tracked changes document on the PCI, with our modification in red.

In addition, I recommend to add a genuine discussion and not just a brief conclusion lacking context. An extended discussion should replace the findings in a larger context, e.g. comparing with other systems where industry decreases micro-organism diversity (eg *Penicillium camemberti* DOI : 10.1016/j.cub.2020.08.082 and *Penicillium roqueforti* <https://doi.org/10.1111/mec.15359>) and highlighting better the importance of the results and the insights gained. Some points of uncertainty should also be discussed, eg whether the silage strains could be feral rather than wild representatives?

We have replaced our result & discussion session, by two separate sessions, one on the results and the other on the discussion. We have also replaced our findings in a larger context and better highlighted their interests.

The figures can be much improved and this is very important for rendering a paper understandable and appealing. The figure 2 is much too small, colors would help clarifying; "artisanal" and "farmer" are hard to distinguished from point labels; placing the three panels as lines rather than columns would help make them larger and readable. The figure 3 should show the bread types. The figure 4 is also hard to read and understand, in particular the

black and white rectangle do not make easy to distinguish bread types. I cannot read labels in the figure 5. Redefine all abbreviations in all figure legends.

We have improved figures.

We have improved Figure 2, 3, 4, 5. Figure 2: artisanal and farmer groups are now distinguished by different colors. The same colors are used Figure 3, Figure 4 and Figure 5 to distinguish the two types of bread practices. All abbreviations are now defined in legends.

I would encourage resubmission if you are able to revise the manuscript along all these lines.

In addition, please also address the comments below:

-In the title and in several places in the manuscript, there may be a confusion between species diversity and composition. If I understand correctly, the differences were tested in terms of species composition rather than diversity? Please clarify the terms and add some more direct analyses of diversity, such as Shannon or Simpson index.

Thank you for pointing this out. We went through the word “diversity” along the MS and change it when it was not accurate. We analyzed both alpha and beta diversity (lines 196-200, 311-320). To estimate alpha diversity, we calculated species richness, Chao1, Shannon and Simpson (lines 311-315). To analyze beta diversity, we calculated weighted Unifrac distances between sourdough fungal communities and clustered sourdoughs on the basis of the PCoA of their weighted Unifrac distances (lines 376-387). We also compared species composition between sourdoughs (lines 395-396).

-In the title and elsewhere: no plural when a word is used as an adjective before another word (e.g. no s for bread in the title, fermentation L159, sourdough L315)

-L41: delete “their”, the reference is two sentences away.

corrected

-L61: roqueforti not roquefortii

corrected

-L68 “that” should not be omitted after suggest in written text

corrected

-L70, L74 : bread-making

We replaced all “bread making practices” by “bread-making practices”

-L75: delete the space before the point

corrected

-L89: delete may

corrected

-L96: farmers

corrected

-L160: no number at the beginning of a sentence, or only written in full letters

corrected

-L164, 165: silage not ensilage (French)

corrected

-L220 habitats

corrected

-L315: if you compare sourdoughs, it should be difference and not diversity, or a word is missing

corrected

-L331, L352, 355, 357, 470 and elsewhere: no abbreviation at the beginning of a sentence

corrected

-L346: I missed information about known abilities of *Kazachstania* species for fermentation?

Thank you for your comment. We have added the following sentence (line 347): « The *Kazachstania* genus is one of the closest genetically related genus to *Saccharomyces* and contained Crabtree positive yeasts, able to ferment glucose even when oxygen is present if the amount of sugar is sufficient ».

-L370, 381, 528: diversity and/or composition?

We corrected our sentence according to what we meant

-L319: associated with instead of related to

corrected

-References: check format, eg species name italics, no capital at words, ...

Reviews

Reviewed by anonymous reviewer, 15 Mar 2022 23:04

I very much enjoyed reading this ms that reports on an interdisciplinary and participatory research concerning the investigation of the yeast communities associated with sourdough in France. This is a well-designed and carefully conducted investigation. Having no major concerns about this ms, my comments should be understood as a contribution to improve this fine work.

Major

Fig 3 depicts very important results for the study but I think it is very difficult to locate a particular species given that small differences in colour tones. For example, I cannot distinguish *S. cerevisiae* from *P. kudriavzevii* on a printed image. My suggestion is that you color-code a restricted number of the more relevant species. The others could be grouped in genera (or other groups) and given the same colour. For example, the same colour for the genus *Cryptococcus* (a supplemental figure could provide the entire set of species or even a series of detailed images for those species that were grouped in the main fig. For example, a supplemental fig where *Cryptococcus* species are coloured while the remaining are in shades of grey.

We have now highlighted the three main yeast species by black surrounding and have changed some color to make the differences more visible

Fig. 4. Again some species - *S. cerevisiae* vs *Pichia* and species of *Kazachstania* are undistinguishable.

We have changed some colors to make the differences more visible

Fig 5 is very interesting but the lettering is so small that it cannot be read. For the first clustering indicate some species for each of the 3 groups so that they have a more relevant meaning.

We made some changes in figure 5 to make it clearer. We did not add species for each of the three groups for the first clustering as several species explained the clustering. The species are shown figure 5C.

Line 315 – 320. Could you please be more specific when analysing unifracs distances. – which communities (samples) are similar, and which are not? Perhaps adding also a few more sentences on the approach and data interpretation would also help.

The details on Unifrac distances and which community are similar are presented line 380-384 and figures 5A and 5C. We have now added a sentence line 315 “the clustering of sourdoughs according to their Unifrac distances is shown Figure 5

Finding that several species of *Kazachstania* can dominate sourdough fermentations is a very interesting result. Given that this genus has many species, it would be relevant to know what the sourdough species have in common? Are they phylogenetically related?

Thank you for this comment. According to one of the reviewer comments, we have separated the results and discussion. We now discuss more this result (lines 506-513) and give more information on the *kazachstania* genus in the discussion (line 496-519).

The authors have shown that the sourdough community can change. Similarly, the human microbiome project has shown that the gut microbiome can change among healthy individuals. Nevertheless, it appears that, at a general level, the main (gene) functions that the different communities perform are similar. Perhaps this point could be addressed with respect to the sourdough communities – different in composition but perhaps similar in function.

We thank the reviewer for this interesting comment. We discussed this partly lines 431-449. Unfortunately, in our paper we missed statistical power to compare sourdough community function. Indeed, it will be great to study the convergence of sourdough community function.

The section “phenotypic signatures of domestication” appears to require some caution in the interpretation of the results. First the number of sourdough isolates is disproportionately high when compared with non-sourdough strains. Moreover, you are comparing fresh isolates with collection strains. Secondly, given that these phenotypes were not investigated at the genetic level, assuming domestication appears to be premature. Domestication must have a genetic basis and studying phenotypes without studying genotypes appears is a problem if one wants to make statements about domestication (which, in fact are not that much relevant for the study). Moreover, you devote two figures to this unbalanced comparison and conclude that the strains have been domesticated without providing, as mentioned above, any genetic evidence.

Thank you for this comment. We are now more cautious. We do not speak about domestication signature anymore. We just explain that we are looking for any evidence of a domestication syndrome. We have added in the discussion that it would be interesting to analyze genetic signatures of domestication in *Kazachstania* (line 522-524).

“Therefore, our results suggest that populations of *K. humilis* have been selected by humans and can therefore be considered domesticated”. Here and in the conclusion section, I think that these type of statements should be softened

We have deleted this sentence and replaced it line 527 by “This species is therefore frequently found in bakeries, where short fermentations are often favored. This may explain why sourdough strains of *K. humilis* seem to have been selected for increased fermentation speed”

Minor

Line 55. Perhaps rephrase since the Neolithic did not “emerge”. It is just a time period designated by humans

Thank you, we rephrased by “Humans started to ferment food before the Neolithic using naturally fermenting microbial communities.”

Line 58. I think “selected strains” were not in use in the 19th century.

Thank you, we replaced “selected strains” by “starters”

Line 58. “reduction in species”. Please explain.

We are meaning reduction in species diversity

“reduction in genetic diversity” Indeed this is typically observed in plant crops but can you cite studies in microbes where this was observed?

Yes, we now cite studies that have shown a reduction in genetic diversity in the discussion line 576.

Line 82. I know the cited work but cannot avoid two comments. First, you are claiming that a domestication process is contributing for a species diversity. However, a few lines earlier you claimed that domestication reduces diversity. Secondly, “indicating that sourdough bread making contributes to the conservation of *S. cerevisiae* diversity” entails a complex rationale. In the first place this reasoning assumes that diversity was already present (before sourdough) and this practice helps preserving it. The problem is that sourdough strains arose after domestication and therefore do not occur in nature. If one wants to measure the diversity of *S. cerevisiae*, considering the diversity of its truly wild populations is the relevant thing to do. Anything else would mean that the diversity of this organism is a human construction, due to domestication. I think this reasoning is true for any microbe. In brief I agree that sourdough practices are important to maintain the diversity of bread-related lineages (in a given species) but I have doubts that an artificial fermentation can maintain or expand the diversity of a species.

Thank you for this interesting argument. I think we simply miss-use the term diversity here. As you suggest, we replaced it by “indicating that sourdough use contributes to the conservation of bread related *S. cerevisiae* lineages”

Line 280. You mention that you isolated 20 to 40 yeast strains per sourdough sample and cite previous works. I presume these were aleatory isolations and that this approach allowed you to calculate relative frequencies. I understand you cite two other studies but a brief description should perhaps be included as this part must be made clear as it is essential for understanding your work.

Thank you for your suggestion. We have now added “by picking colonies randomly”

Line 289. "T. aestivum". Please do not abbreviate the genus name the first time you use it.

corrected

Line 289. "... assigned to plant pathogen fungi, such as Alternaria, Aspergillus or Fusarium, Gibberella". Is the labelling "plant pathogenic" the most appropriate, given that these genera also include many saprobes? Perhaps "... assigned to genera that include plant pathogen species together with saprobic ones"

Thank you for your suggestion, we corrected it

Line 345. I have mixed feelings about the title of this section. Clearly finding several species of Kazachstania as dominant is a very interesting result. The problem is that the genus has lots of species, more than 30, and therefore the title is not very accurate.

We changed the title line 345 to "Sourdough yeast species mostly belong to the *Kazachstania* genus"

Still on this topic in line 349: "In addition, eight species of the *Kazachstania* genus were found in sourdough, a much higher figure than for the *Saccharomyces* genus, represented by only two species (*S. uvarum* and *S. cerevisiae*)" If the genus *Kazachstania* is 4 times larger than the genus *Saccharomyces* (approx 40 species vs approx 10 species) then these numbers are not so special.

We have replaced our sentence to delete any superlative (lines 346-347): "In addition, eight species of the *Kazachstania* genus were found in sourdough, while the *Saccharomyces* genus was represented by two species (*S. uvarum* and *S. cerevisiae*)"

Line 456 "Previous studies on the bacteria content of the same sourdoughs showed that..." Perhaps changing the sentence to "same type of sourdough".

We did not change the sentence as the same sourdough as the ones studied in the paper were previously analyzed.

The first reference in the reference list is incomplete. See also ref #25.

Reviewed by anonymous reviewer, 08 Mar 2022 21:05

Michel et al. integrate reported data from farmer-bakers and artisanal bakers in France coupled with sequencing and lab experiments to identify what practices influence yeast species diversity and community in French sourdough breads. They identify a number of differences that separate farmer and artisanal bakers. Using a culture based and culture-free sequencing approach, they identify species diversity in each of the sourdoughs. For the most part, the culture based method does a good job at identifying the dominant species in the sourdough, but the culture free method picks up more rare species and some unculturable

species that are more prevalent in a couple of starters. Their most notable finding is that farmer-bakers sourdoughs do differ from artisanal bakers in species composition, which is mostly driven by whether the baker uses commercial yeast. I also found the results differentiating *Kazachstania humilis* and *bulderi* really interesting – the authors show that the two species are differentiated in a number of fermentation metrics, which follows their usage in either farmer-bakers or artisanal bakers. The *K. humilis* sourdough strains perform better than *K. humilis* strains isolated from other environments (although there are only a couple samples here). As the authors mention, follow up work with dough is needed, and the language about domestication could be tempered slightly, but I felt the work done was adequate for the scope of this manuscript.

I found the interdisciplinary approach of this study to be elegant. While other factors may be contributing to their findings differentiating species composition via farmer-bakers and artisanal-bakers (as they acknowledge), this study supports that preserving cultural practices can preserve/cultivate microbial diversity in our foods. It also catalogues a huge number of variables in breadmaking that have been hypothesized to contribute to a particular species community in sourdough. This is a very nice addition to the literature. A few minor comments below:

Line 279-280 isolated 20-40 yeast strains? Does this mean colonies? Throughout the manuscript there was reference to strains, but to me, a strain means a distinct genetic entity, which I don't think is known in this context.

In yeast, we call strain a population derived from a single cell, i.e. from a single colony. We use "strain" even in the case where no genetic data is available.

I had a couple of comments on the figures. First, I would suggest an additional figure or figure panel, which combines the Unifrac tree with the metabarcoding sequencing data to help visualize the species composition for each branch.

We made figure 5 clearer and this is what we do Figure 5C.

Figure 1 – I do not see the center of the groups denoted with a triangle (and not sure what this would represent). The labels next to the triangles in 1B are not always clear from comparison to Figure 1C.

We have improved Figure 2, 3, 4, 5. Figure 2: artisanal and farmer groups are now distinguished by different colors. The same colors are used Figure 3, Figure 4 and Figure 5 to distinguish the two types of bread practices. All abbreviations are now defined in legends.

Figure 5- please label each panel with letters for reference in the text. Also just visually, the scale of the font differs very dramatically between the legends and titles and is a bit hard to read the small font sizes.

corrected

Figure 6 – could you include the trait data for the non-sourdough strains on this figure as well? Also, I would be curious to know how *K. humilis* and *K. bulderi* differ in these fermentation metrics from *S. cerevisiae* sourdough and commercial strains.

Sourdough and non-sourdough strains are shown Figure 6. The traits differentiating them are shown by the correlation circle on the top right. Unfortunately, we are not able to compare *S. cerevisiae* fermentation performance with *Kazachstania* ones, as we have block effect and we did not initially design our experiment for comparing *S. cerevisiae* and *Kazachstania*. We are planning to do it.