Reviewer 1: Approved with Reservations
In this manuscript, Sobel et al. present fungal microbiome data from 39 different beers as the culmination of a crowdfunded citizen science campaign. These data will be of interest to citizen scientists and financial backers of the project, as well as those in the fermentation (especially beer) industry. Overall, the data seem sound, but I have some concerns:

Major comments
Were any controls for contamination used, i.e., are all of the fungi identified actually from the beer samples? The sequencing of a non-beer sample such as water that had been handled in the same way as the beer samples would help to determine if fungal DNA contamination occurred during sample processing.

In line with the comment above, the manuscript states that "…microorganisms, or their DNA, could be carried over from the ingredients to the final product." Can the authors comment on whether they know if they are detecting the fungi themselves or DNA remnants from fungi that came from the raw ingredients of the beer? Again, one could add purified control DNA to a mash, brew and bottle a beer, and then try to detect that DNA by PCR in the end product (or even at various stages along the brewing and fermentation process). Attempting this with various concentrations of DNA would also yield information on how many cells of a particular species would be necessary on malted barley, for instance, to be detected in the final beer.

Minor comments:
In the introduction, the authors state that "…sour beer…[is] produced without the controlled addition of known yeast cultivates." Although this may be true for some types of sour beer like lambic and gueuze, many sour beers made in the U.S. are inoculated with known strains of yeast. In those cases, the souring bacteria are usually the unknowns.

Why was a fecal DNA prep kit used for DNA extraction?

The authors collected 120 beers from 20 countries but only sequenced the fungi from 39 (mostly from Switzerland). Is there an explanation for this attrition?
Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Partly

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: I also participated in a crowdfunding campaign to use next-gen enunciating to analyze beer samples (https://experiment.com/projects/mapping-the-sour-beer-microbiome).

Reviewer 2: Approved with Reservations
This article describes how innovative, participant-driven research projects can create an interesting data set outside the traditional Academy. This is an extremely laudable goal and the resulting data will be of interest to a broad audience. In its current state the article is a hybrid between data note, methods article and research article that delivers preliminary results. Regardless of the form of the article, the methods section would benefit a lot from a more detailed description of the methodology (see details below). Similarly, the analyses and results are a bit lacking at this stage, especially with respect to the basic metrics of the data sets (again, see below).

Major comments

Methods
the methods section should be significantly improved/extended for a better understanding. I'm aware that most/all the things are on GitHub, but having to crosslink these makes it hard to follow. Specifically the following things should be improved upon:
What modifications were done to the protocol of the ZR Fecal DNA MiniPrep kit? (suggestion: putting the modified protocol to https://www.protocols.io/ if deemed useful)
Which parameters were used to perform the bwa alignment?
What was the size of the reference database that was used for the mapping?
How was the hierarchical clustering done that is described in the methods section? Which clustering method was used? Which distance measure was used?
Results
"We obtained an average library size of 600K reads (min 350K, max 2400K)" This is a rather large difference between the different libraries. Does the number of species found correlate with the sequencing depth? I.e. would you have found more species if you had sequenced more data for the smaller libraries? A rarefaction analysis would be useful to understand the impact of sequencing depth on species recovery. A minor, related suggestion: Having a table of sequencing statistics so that the reader can compare the samples.
A major thing that is not mentioned in the results/discussion is the number of reads which could not be mapped against the reference database. How many reads of each library did not belong to any of the reference ITS sequences? And are the non-mapping reads similar to each other or can be clustered into OTUs? This would be needed to understand how many species/OTUs are in a given sample but could not be classified due to a lack of reference database. Without this the ITS diversity in a sample cannot be correctly estimated.

Minor comments:
"we built the proof of concept for a targeted metagenome analysis pipeline for beer samples that

can be used in high schools, citizen science laboratories, craft breweries or industrial plants" It would be good to at least briefly discuss how this is currently limited by the need to have access to a high-throughput sequencer.
It would be great if "terroir" could be defined in the introduction for those not too familiar with oenology
"a total of 88 fungal species were identified, including 52 unique occurrences" are unique occurrences those species which are only found in a single beer? I'd suggest rephrasing it for a better understanding.
"Interestingly, most brews were found to contain low to medium presence of multiple other yeast species, including Saccharomyces bayanus (used in winemaking and cider fermentation), Saccharomyces kudriavzevii and Saccharomyces pastorianus (used in lager manufacturing), Saccharomyces eubayanus (a probable parent of Saccharomyces pastorianus) and Brettanomyces bruxellensis (typically used for the production of the Belgian beer styles)" please include citations for these explanations of the different taxa.
It's a matter of taste, but I recommend rethinking the use of "microbial dark matter", c.f. http://merenlab.org/2017/06/22/microbial-dark-matter/ for an explanation of why.
Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Partly

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

Referee Expertise: Fungal metagenomics & bioinformatics