Data Paper

Shu-Kuan Wong, Ryo Kaneko, Shota Masumoto, Ryo Kitagawa, Akira S. Mori and Masaki Uchida. Soil Microbial Composition and Diversity in the Low Arctic Tundra of Salluit. Polar Data Journal. 2021, 5, p.54–59. https://doi.org/10.20575/00000026.

(Received 12/21/2020; Accepted 3/6/2021)

1st submission

Editor Start Date: 12/22/2020

Editor Stop Date: 1/22/2021

Reviewer #1 (1/3/2021–1/22/2021)

Reviewer #2 (1/6/2021–1/11/2021)

Reviewer #1: Anonymous

The authors report data on soil microbiome collected in Salluit at the low Arctic in Canada. Since the polar microbiome can be sensitive to the effects of climate change, the data provided by this manuscript will be important baseline data. This submission to the Polar Data Journal seems appropriate. The manuscript is well written, but there are two minor comments.

1. First, the main bacterial phyla detected are described, but if there are minor but rare/unique lineages (e.g. uncultivated candidate divisions), it would be better to include them as well.
2. Second, in my understanding, the accession number (PRJNA 541059) in the manuscript is the BioProject ID. I think there are also BioSample and DRA accession numbers, but are they not available? I think these numbers should be included in your manuscript.
3. Furthermore, this number (PRJNA 541059) has already been mentioned in the online-published paper (Masumoto et al., Polar Science, 2020; https://doi.org/10.1016/j.polar.2020.100562), but it would also be better to carefully describe its relevance to the data of this manuscript.

Reviewer #2: Anonymous

This paper reports on the metabarcoding data of soil bacteria from a low arctic tundra in Canada. The dataset provides a baseline for the understanding of the soil microbial diversity in arctic tundra, but the descriptions at present are generally too poor to convey the significance of the dataset. The authors need to describe and summarize the data more in detail by including a new figure and/or table.

1. The first paragraph of Background & Summary is too short to convey the significance of the data. The descriptions in the second paragraph of Background & Summary are not the summary but parts of the results of metabarcoding.

These should be described in the subsection 3.2 with a summarized figure or table. Consulting Mise et al. (2019) will help the authors to see what are needed in metabarcoding data papers. Mise K, Moro H, Kunito T, Senoo K, Otsuka S. 2019. Prokaryotic community structure of long-term fertilization field Andisols in central Japan. Microbiol Resour Announc 8: e01551-18. https://doi.org/10.1128/MRA.01551-18.

2. Please explain why the study site was chosen in the Section 2 Location.

3. Regarding 3.1 Soil Sampling, collecting samples at 6 m intervals along 150 m transect will make a total of 26 samples.

4. The description in the subsection 3.2 is unsatisfactory and should be more detailed. Overall picture of dataset should be indicated in a figure or table using a representative bioinformatic tool.

5. What do you mean by 'etc.' in 4 Data Record? Specify.

Authors Response:

First and foremost, we would like to thank the two reviewers for their constructive comments and suggestions to improve the manuscript. We appreciate your time and effort in reviewing this manuscript. Our responses to the comments are as below:

Response to reviewer #1;

The authors report data on soil microbiome collected in Salluit at the low Arctic in Canada. Since the polar microbiome can be sensitive to the effects of climate change, the data provided by this manuscript will be important baseline data. This submission to the Polar Data Journal seems appropriate. The manuscript is well written, but there are two minor comments.

1. First, the main bacterial phyla detected are described, but if there are minor but rare/unique lineages (e.g. uncultivated candidate divisions), it would be better to include them as well.

There were many rare or unique lineages detected from our study, in fact, too many to mention each and every single one of them separately. Furthermore, we refrain from dwelling too much on rare or minor sequences as they may be present due to sequencing artefacts or too low in abundance to contribute to the biogeochemical functions in the area. In our case, these minor groups, with relative abundance of less than 1 %, are grouped together (Figure 3).

2. Second, in my understanding, the accession number (PRJNA 541059) in the manuscript is the BioProject ID. I think there are also BioSample and DRA accession numbers, but are they not available? I think these numbers should be included in your manuscript.

Yes, you are correct. In addition to the BioProject ID, there are also BioSample accession numbers. It is usually sufficient and common to only provide Bioproject accession in publications. Additionally, we also received an email by NCBI to cite the BioProject accession number (SRA accession) starting with PRJNA XXXX in our publications

upon completion of our sequence submission to the SRA. As per your suggestion, we have added these extra details to the 'Data Records' section.

3. Furthermore, this number (PRJNA 541059) has already been mentioned in the online-published paper (Masumoto et al., Polar Science, 2020; https://doi.org/10.1016/j.polar.2020.100562), but it would also be better to carefully describe its relevance to the data of this manuscript.

Masumoto et al. (2020) research was carried out under the same project as ours, which was to map the biodiversity in the Arctic areas. However, their main focus was on fungi and ours was the bacterial communities. Nevertheless, the sequences for fungi (LSU) and bacterial (16S) communities were deposited under the same BioProject because they originate from the same project.

Response to reviewer #2;

This paper reports on the metabarcoding data of soil bacteria from a low arctic tundra in Canada. The dataset provides a baseline for the understanding of the soil microbial diversity in arctic tundra, but the descriptions at present are generally too poor to convey the significance of the dataset. The authors need to describe and summarize the data more in detail by including a new figure and/or table.

1. The first paragraph of Background & Summary is too short to convey the significance of the data. The descriptions in the second paragraph of Background & Summary are not the summary but parts of the results of metabarcoding. These should be described in the subsection 3.2 with a summarized figure or table. Consulting Mise et al. (2019) will help the authors to see what are needed in metabarcoding data papers. Mise K, Moro H, Kunito T, Senoo K, Otsuka S. 2019. Prokaryotic community structure of long-term fertilization field Andisols in central Japan. Microbiol Resour Announc 8: e01551-18. https://doi.org/10.1128/MRA.01551-18.

We have added more details to the background (Line 8 – 12) to make it more succinct and detailed. As per the summary, we have referred to previously published papers in PDJ and found that the findings of the research are often summarized in that section and not in the methodology section (subsection 3.2) as suggested. Therefore, we would like to retain the summary for our metabarcoding results as it is in the Background and Summary section.

2. Please explain why the study site was chosen in the Section 2 Location.

At the end of the paragraph, we have previously mentioned that the sites were chosen based on their vegetation coverage and environmental conditions for this study.

3. Regarding 3.1 Soil Sampling, collecting samples at 6 m intervals along 150 m transect will make a total of 26 samples.

Technically, your comment is correct and a 144 m transect would have 25 samples for every 6 m interval for our case. We mentioned 150 m transect because we set up a line with a distance of 150 m for every sampling but collected 25

samples from the 150 m-long transect. Since this is rather confusing, we have decided to remove "150 m" from the manuscript.

4. The description in the subsection 3.2 is unsatisfactory and should be more detailed. Overall picture of dataset should be indicated in a figure or table using a representative bioinformatic tool.

We have included a summary of the metabarcoding results in Figure 2 and Figure 3, containing all the different phyla found in our samples.

5. What do you mean by 'etc.' in 4 Data Record? Specify.

There are many downstream bioinformatics tools to process these sequences and it would be too long to list everything so we gave examples of the two most common tools (MOTHUR and QIIME) but there are a wider variety of choices available depending on the user and hence, the "etc.".

2nd submission

Editor Start Date: 2/15/2021

Editor Stop Date: 3/5/2021

Reviewer #1 (2/16/2021–2/17/2021)

Reviewer #2 (2/16/2021–3/2/2021)

Reviewer #1: Anonymous

The authors have responded appropriately to my concerns in the first version of this manuscript.

Reviewer #2: Anonymous

The authors take the comments into consideration and revise the manuscript. I have a few additional comments that will help the authors preparing the final draft.

Please explain why the study was conducted in Salluit. Is this a part of a large-scale research project?

The authors need to define 'relative abundance' appeared in Figure 2.

Corrected according to above comments.