

Most of the manuscript reads well and it has interesting findings that are worth being told. However, several aspects require clarification and revision. In particular, it is unclear how changes in the community within six weeks of soil disturbance can be characterized without pre-disturbance samples for comparison. If such data exist, this should be clearly explained. The findings of biotic influence being more important than edaphic influences is quite interesting.

Thank you for your time, and your constructive comments. We have addressed the comments made, and believe this has given the manuscript a clearer focus. We have included our responses to the comments below, as well as in a supplement .pdf with red font for clarity. The updated line numbers in our responses refer to the line numbers in the updated manuscript.

Scientific remarks:

Line 59: 2018 and 2014 are not very “recent”.

This has been removed.

Line 64: It would be great to define microbial activity before stating the examples in this paragraph, there is a plethora of different approaches (ATP quantification, enzymatic activity as in FDA assays, growth, viability,...). You are showing examples for different approaches (respiration etc.), but in the following it's not always clear which measurements are used.

A definition of microbial activity has been added. At lines 65-66: "Soil microbial activity refers to the growth and metabolic processes undertaken by soil microbial communities (Nazir et al., 2014 )."

Line 66: What is meant with transcriptionally active? Does this mean that they have a higher total amount of RNA? Or just more of the gene of interest is transcribed?

This refers to a section of the Coolen & Orsi manuscript where authors describe different taxonomic abundances in transcriptional reads. There was an increased representation in Firmicutes, Bacteroidetes, Euryarchaeota, and ascomycetous Fungi metatranscriptomes relative to thawed samples. This has been corrected to read the following on lines 68 - 71: "While specific taxa upregulate broad transcriptional activity (with increased total RNA production) during permafrost thaw, only a fraction of the total community participates in biogeochemical cycling (Coolen & Orsi, 2015). "

There is even evidence that RNA levels go down with warming as in Söllinger et al. 2024 <https://pmc.ncbi.nlm.nih.gov/articles/PMC11126301/> , although this doesn't seem to be from frozen soils. Therefore, these soils will respond differently to long-term warming.

Yes, unfortunately these are not permafrost affected soils, and while their findings are of interest are not directly relevant to our study.

Line 71: Does this refer to (seasonally) frozen temperate soils or just "regular" thawed temperate soil?

This refers to thawed temperate soils under typical non-frozen conditions. We have clarified this in the manuscript to avoid confusion. The sentence at line 74 now reads: “much lower than the 91.3% viability ratio found in thawed temperate soils (Janssen et al., 2002).”

Lines 82 ff: The phrasing suggests that the Tibetan Plateau is in the Arctic. Of course, both are examples of permafrost harboring environments, but it could have been made a little bit more clearly here.

We have clarified the geographic distinction by specifically stating that the Qinghai-Tibetan Plateau refers to permafrost-affected environments outside of the Arctic. The sentence from lines 84-86 now reads: “In other permafrost affected environments outside of the Arctic, artificial disturbances – namely infrastructure development – are the most common causes of thermokarst lake formation in the Qinghai-Tibetan Plateau (Lin et al., 2016).”

Line 90 ff: Present the Canadian study site a little bit more in a manner that it is exemplary for permafrost studies. It is only implied through context that it has permafrost.

We clarified and revised the manuscript to explicitly state that the site is underlain by permafrost at line 94: “...soils and vegetation were stripped in preparation for mining at a permafrost-affected site near Dominion Creek, Yukon, Canada”

Line 96: While I am sure that the study will help with understanding future findings, the statement of this study being crucial overstates the study’s implications.

We have reduced the strength of the original statement and revised the sentence to avoid overstating the findings. Lines 100-101 now reads: “This information helps clarify how permafrost communities restructure upon thaw.”

Lines 100ff: The paragraph doesn’t explain how/that the samples were taken after disturbance, which is the premise of the manuscript.

We have clarified that sampling occurred after the disturbance event. An introductory sentence on lines 104-106 now reads: “To assess the impacts of disturbance and thaw on permafrost-affected soils, we sampled after a disturbance along a gradient from undisturbed to a thawed thermokarst pond (Figure 1).”

Line 108: The names of the sites are confusing. First, Core site 2 and Core site 3 are mentioned, but no Core site 1. Later on, surface samples from sites A to F are mentioned. It would be good to have an overview about how many sites were sampled and what is meant by Core site 1-3 and A-F, because they are all called sites. Figure 1 also does not help much in the distinguishing between sites 1-3 and A-F.

The methods section 3.1 has been reorganized to better delineate where samples originated. This includes moving a paragraph from the results to the methods section which describes our definition of permafrost and active layer soils.

Line 115: Was the spade also disinfected? How else was assured that no contamination took place?

Yes, all tools were disinfected between samples. Surfaces were cleaned with 70% ethanol followed by bleach to minimize contamination. We have added this to the methods section. Lines 134-135 now read: “A separate spade was cleaned with 10% bleach and 70% ethanol before sampling...”

Line 124: How deep was the active layer before it was removed?

The active layer at Core 1 was 14cm, at Core 2 it was 34cm, at Core 3 it was 85cm. A sentence at lines 144-145 was added to clarify this: “The thaw depth was 14 cm, 34 cm, and 85 cm for Cores 1, 2, and 3, respectively (Figure S1).”

Line 126: A depth of 1m measured from where?

The depth was measured from the soil surface. This has been added in the manuscript to the end of this sentence at line 148.

Line 131: Haven't they already been thawed because of the season? Or were they seasonally frozen? The maximum active layer depth is typically in September. How did you distinguish between seasonally frozen soil and permafrost?

The soils above maximum thaw depth were seasonally thawed, and are defined as the active layer. Soils below this depth remain perennially frozen and are defined as permafrost. We determined these layers by measuring gravimetric water content at 1 cm intervals across each of the soil cores taken. As expected, a sharp peak in water content marked the ice-rich transition zone at the base of the active layer. In our study, this occurred at approximately 45 cm below the soil surface. This explanation has been added to the methods and materials section of this manuscript at lines 157-163. This section reads: “Active layer soil was distinguished from permafrost by measuring gravimetric water content at 1 cm intervals along each core. A pronounced increase in water content was observed at the transition zone at 45 cm as shown in Figure S1: Core 1 - 45 cm, Core 1 75 cm, Core 1 95 cm, Core 2 75 cm, Core 2 95 cm, Core 3 75 cm, and Core 3 95 cm. Soil water accumulates at the bottom of the active layer, allowing for determination of the maximum active layer depth through the determination of gravimetric water content (Chang et al., 2024).”

Line 133: How long have the samples been stored frozen before being processed? Assembling has been done in 2016 and the paper is now under review it makes you wonder how long everything has been frozen before being analyzed.

Samples were processed shortly after collection: DNA extraction and chemical analyses were completed in 2016-2017. All samples were stored at -20°C prior to processing. Delays in manuscript preparation occurred later and did not affect sample integrity.

Line 142: I would replace "section 1" with "the first section", as they are not specifically named. Likewise with Section 2 and 3 later on. I think it would be clearer to just describe and not name everything.

We have renamed Section 1, 2, and 3 with descriptive phrases ie. first section, second section.

Lines 181ff: I think this paragraph can be shortened with the reference to the source that you've already put in.

We have removed the PCR reagent concentrations, while leaving the PCR primers and description of our control results.

Line 202: How many samples were without the negative extractions and blanks? I think this is the more important number.

There were 18 samples sequenced, while 5 controls were sequenced. Lines 228-231 has been altered to: "Before sequence processing, a total of 552,252 reads were recovered from 18 samples including surface samples, core samples, as well as 5 blank extractions, negatives, and a mock community."

Line 214: The usage of OTUs in contrast to ASVs should be explained.

Although ASVs and DADA2 are current best practice, this dataset was created 8 years ago. Early MiSeq chemistries predate modern ASV methods, and OTU methods remain sufficient for legacy datasets such as this. Importantly, OTU based analyses are still sufficient to identify ecologically relevant changes in microbial communities while also supporting reproducibility across previous studies of similar context. Kerrigan & D'Hondt (2022) (Patterns of Relative Bacterial Richness and Community Composition in Seawater and Marine Sediment Are Robust for Both Operational Taxonomic Units and Amplicon Sequence Variants) demonstrated that ASV and OTU methods showed broadly similar results, particularly for trends in beta-diversity, alpha-diversity and class-level relative abundance, all of which were the basis of our manuscript.

Line 235: It is not necessary to explain which specific plots were done, because they can just be presented in the results. This sentence also doesn't seem to have a good syntax.

This sentence has been removed.

Line 260: How did this not result in the lysis of many cells? But I guess at least they are all lysed in the same manner and thus comparable. Is there a reference that shows that this method is proven?

This method has been published in Saidi-Mehrabad et al., (2020) (Permafrost Microbial Community Structure Changes Across the Pleistocene-Holocene Boundary).

A very similar method was also used in Burkert et al., (2019) (Changes in the Active, Dead, and Dormant Microbial Community Structure across a Pleistocene Permafrost Chronosequence), where soils were sonicated for 1 minute at 20 V in a tetrasodium pyrophosphate solution. While we concur there may be loss of viable cells due to sheering forces, these should be minimized due to the short duration of sonication and should be consistent across all samples. This limitation is acknowledged, and should be consistent across samples, minimizing bias.

Line 275: How were they grouped? Per core or per depth, or something else?

We have clarified this section at lines 305-309 to: Permafrost samples included Core 1 (45 – 50 cm, 75 – 80 cm, 95 – 100 cm), Core 2 (75 – 80 cm and 95 – 100 cm) and Core 3 (75 – 80 cm and 95 – 100 cm) (n = 7). Active layer samples included Core 1 (15 – 20 cm, 30 – 35 cm), and Core 2 (30 – 35 cm, 45 – 50 cm) (n = 4).

Lines 280-290: This is what I meant in line 90. This also seems to belong to a separate paragraph in the Material and Methods section, which you could name "study site". It is not a result.

We have moved this section to the methods.

Line 289: Is there a reference that a water table correlates with the permafrost depth?

Yes, Chang et al., 2024 (Unraveling the non-linear relationship between seasonal deformation and permafrost active layer thickness), where soil moisture increases at the bottom of the active layer when soils are not saturated. This reference has been added to lines 163-165.

Line 310: Which specific edaphic parameters were used for this statement? Also state which NMDS method was used for this.

We have clarified that the analysis was performed using redundancy analysis (RDA) and that it included the following edaphic parameters: P (%), K (%), S (%), Mg (%), Ca (%), Fe (%), Cu (mg/kg), Mn (mg/kg), Zn (mg/kg), Na (mg/kg), NO<sub>3</sub> (mg/kg), NH<sub>4</sub> (mg/kg), TN (%), TC (%), SOM (%), pH, and water content. This information has been added to the manuscript at lines 272-275.

Lines 316f: This would be part of the discussion.

This has been moved to the discussion at lines 479-481.

Line 330: Does that mean that the active layer and sample C and D are the same grouping, because of the letter a that is on top of them? This is not clear enough for me.

Surface samples A-F were treated as a single statistical grouping to test how surface samples compared to frozen active layer and permafrost cores. This set up allows us to visualize how diversity changed with disturbance, while comparing to frozen samples. This has been clarified in the manuscript at line 365: "Surface samples were assessed as a single statistical grouping (n = 6)".

Line 372: How do you know that it changes after being disturbed? If I understood this correctly, it has only been sampled once.

Although we do not have samples from before disturbance, we have paired our disturbed site alongside undisturbed sites which allows us to determine disturbance-related differences. This space-for-time substitution design is standard when pre-disturbance baselines are unavailable, and although this is not a direct longitudinal study design, it supports our inferences regarding the impact of disturbance on microbial communities. Sentences have been included to address this limitation have been added to lines 156-158 "Surface samples along a disturbance gradient were collected to assess the impact of thaw as a result of the disturbance, allowing for direct comparison to undisturbed frozen soil cores beneath them." as well as 167-170: "Adjacent undisturbed areas were used as controls to assess the impact of disturbance on microbial

community composition. These soils were collected as nearby as possible to minimize the impact of soil spatial heterogeneity, however, some underlying spatial variation cannot be fully removed.”

Line 387: I don’t think it makes sense to name the Clusters (1 – 3) if you only use the cluster names in this figure description. Rather describe which samples cluster together.

The cluster names have been removed from the results body and from the figure description.

Line 396: Do you mean Zn concentration?

Yes, we referred to Zn concentration. This has been corrected.

Line 405: “This indicates” should be part of the discussion.

Given that we have a section expanding upon the impacts of biotic interactions in the discussion, we have opted to remove this sentence entirely.

Line 419: Which samples have been the basis for assigning the indicator OTUs?

Indicator OTUs were identified separately for permafrost samples and active-layer samples. Permafrost and active layer samples were defined in section 3.1. This description has been added to section 3.5 of the methods at lines 282-285: Active layer OTUs were defined from Section 3.1, where active layer samples included Core 1 - 30 cm, Core 2 - 30 cm, and Core 2 - 45 cm while permafrost samples included Core 1 - 45 cm, Core 1 75 cm, Core 1 95 cm, Core 2 75 cm, Core 2 95 cm, Core 3 75 cm, and Core 3 95 cm.

Line 444: All bacterial communities combined?

Yes, this has been corrected.

Line 460: axis labels

Axis labels have been added.

Lines 497ff: Link the expected changes (pH, mineral N, C,...) better to the literature that you cite. How do the other findings relate to your assumptions?

We have revised our manuscript to directly relate our expected changes to the literature, and have compared these findings to our original hypotheses. This can be found in lines 566-570: “We therefore expected that disturbance and thaw would result in rapid changes to pH, water content, micronutrient concentration, and exchangeable nitrogen content which would correlate to microbial community composition. These anticipated changes reflect documented biogeochemical responses to thaw (Scheel et al., 2022; Chen et al., 2017; Tripathi et al., 2018).”

Lines 511f: How can this be said without knowing how the soil has been before disturbance? There might have been general spatial variation. This possibility should at least be acknowledged or refuted – maybe in the M&M section.

We acknowledge that without pre-disturbance samples, some of our results may reflect spatial heterogeneity. This study uses adjacent undisturbed samples as controls to reduce this effect, however, we have still added this caveat in the methods in lines 167-170.

Line 520-523: Reference? Or do you refer to the presented study?

This refers to the presented study. As such, we have altered these sentences at lines 524-527 to: "During this in situ thaw experiments, microbial community successional patterns within the communities are were observed; however, interpretations of function from taxonomic assignment must be made with caution, requiring further study into the genetic potential and activity of the community."

Lines 524 ff: It would be nice to have more than one reference for the hole paragraph.

Two additional references have been added, discussing the impact of soil fauna and soil mixing on permafrost microbial community shifts: "Physical mixing may also contribute to these community shifts, which has previously been shown in laboratory experiments (Doherty et al., 2025)." And "Soil fauna, such as collembola, may be responsible for the transport of soil microorganisms into thawing permafrost layers (Monteux et al., 2022)." at lines 540-542 and 549-550.

Are there more in situ thaw experiments from other study sites that you refer to in the text? Because they are in plural.

This has been corrected to be singular.

Line 556ff: "permafrost and active layer community clusters could not differentiate between permafrost and active layer community clusters." This doesn't make sense

This sentence has been removed.

Lines 565-576: Really interesting findings!

Thank you for your comment!

Lines 609-613: Please link this more to your findings. Also, be careful with the statement of "activity", because I don't recall you measuring something in the direction of activity of the cells. In general, what counts as "active" is also not that clear.

We have changed this to "Future studies may investigate whether activity changes result from a small but viable community in thawing permafrost which becomes more abundant with thaw." At lines 629-630.

References: Could possibly be updated with more recent literature.

References have been added throughout this revision, including Doherty et al., 2025, Monteux et al., 2022, Scheel et al., 2022, and Burn et al., 2025.

Formal remarks:

General: Sentences shouldn't be started with abbreviations.

This has been corrected throughout the manuscript.

Line 110: Numbers until twelve should be written as words and not digits

This has been corrected.

Line 165: No brackets around the reference, but only the year. "performed as described by Saidi-Mehrabad et al. (2020)."

This is meant to be an in-text citation where the author's name is in the sentence, and this format follows the Biogeosciences format found at <https://www.biogeosciences.net/submission.html#references>.

Lines 392ff: Avoid putting crucial explanations into brackets.

This sentence has been edited to read: "Both Bray-Curtis and Jaccard presence/absence distances within clusters correlated significantly with differences in pH ( $r^2 = 0.4316$  and  $0.5000$ , respectively,  $p < 0.05$ )..." at lines 396-400.

Line 584: aerobic in lower case

This has been corrected.

Line 596: functionality?

Yes, corrected.

Line 648: Space missing after comma.

Corrected.

Lines 482ff: awkward syntax

We have removed this sentence.