Reviewer 2

The manuscript of Tan et al. presents a field phosphorus (P) fertilization experiment with measurements of P fractions, DOC and CO_2, as well as soil microbial community data. The study thus combines different relevant aspects in the context of C and P cycling in the studied tropical forest soils. The field experiment was thoroughly conducted, including many sampling time points for a field scale.

Comment

However, while I agree that P dynamics can influence CO_2 emissions, the discussed underlying mechanisms cannot be conclusively disentangled with the provided results. The experiment included only P addition treatments, no C addition or combined C and P addition, therefore, the discussed change in C cycling always occurred in parallel with the change in P cycling. Also, the effects of the P addition treatments on C dynamics are not as clear as suggested (e.g., DOC increase only in P3, not P1 and P2 (Figure 2a), no respiration increase in P1 (Figure3f)), while there was a clear increase in P availability (Figure 1a). The discussed relationships (e.g. DOC and MBP) are mostly minor/inexistent (very low R2 in Figure 4).

Response

We acknowledge your concerns raised. However, being a field-scale experiment on an acidic subtropical forest soil characterized by a significant amount of SOC, the experiment was designed to explore how this enormous SOC responds to P availability gradients. Adding a C source would be more suitable if it was a controlled experiment.

Parallel response of SOC following P is expected and has often been reported (in most cases). Specifically, our experiment was designed to explore how, and why these changes occur across increasing P supply gradients. Particularly, we sought to explore alternative mechanisms of these responses in acidic subtropical forest soils beyond what is currently known.

Currently, available studies using P addition on C dynamics (DOC or respiration) have reported contradicting reports: negative, positive, or insignificant changes.

While the majority of these studies explored one P addition level, varying P addition rates as currently done in our study would enable us to provide a more robust understanding of the linkage between the extent of P availability on SOC dynamics and the mechanisms driving these processes.

We acknowledge the insignificant response of DOC in P1 and P2, compared to P3. This is an important finding in our experiment, revealing that the responses of SOC to P addition largely depend on the concentration of P, as higher P concentration (P3) stimulated DOC compared to lower levels. These findings clarify that even though available P increased in P1 and P2, the concentration was not sufficient to drive the desorption of SOC and DOC release compared to P3. This supports our competitive sorption hypothesis, by proposing that P can only significantly displace and lead to significant desorption of SOC from its Fe-bound state at higher concentrations.

Although we observed a clear trend and a significantly positive relationship between DOC and MBP (Fig. S5), we acknowledge that several random factors at the field scale can result in the low predictive power of the fixed factor. The variation in sampling points at different sampling periods across the plots, variation in time of sampling number of samples collected, and other soil properties such as moisture content, difference in treatments, etc. may account largely for the low prediction. To verify the contribution of the main random factor (Sampling date) we subjected our dataset to the linear mixed-effect model and utilized the sampling date as the random factor. The results reveal the importance of sampling date in accounting for a significant variation among variables. For example, the relationship between DOC and MBP had a Marginal $R^2 = 0.17$ but a conditional R^2 of 0.37. This re-analysis clarified that the positive significant relationship between MBP and DOC was largely predicted by the sampling time, which cuts across different seasons (with varying soil and climatic conditions). Therefore, we have included as part of our recommendation that sampling for such large field-scale experiments should cut across numerous sampling points per plot, and sampling from points close to the previous sampling point, rather than randomly as currently done to limit large heterogeneity among samples. We have also further supported our results by an additional interpretation of Mantel's test where P cycling gene abundance had a significantly positive relationship ($p < 0.01$, $r >= 0.4$) with DOC compared to other variables. This indicated the significant need for DOC for P cycling compared to other soil processes following P addition.

Comment

As the provided results are not sufficient to support the current mechanism discussion, I suggest shifting the focus rather toward the discussion of the field setting and additional processes relevant at that scale. For example, I was missing a more profound discussion on the effect of tree presence (e.g. P uptake and C exudation) and leaching.

Response

You are right. Field-based experiments in most cases, do not give robust predictions compared to controlled experiments in laboratories/greenhouses. We have emphasized in our manuscript that: despite the statistical significance in the proposed relationships underlying our mechanisms, the low R^2 values obtained were due to the complexity of the field scale experiment, particularly because our sampling dates cut across numerous sampling periods across different seasons, sampling (soil) from different points during sampling that led to larger heterogeneity, challenges in field-scale processes that could alter the integrity of gas sampling during and after measurement, etc. These may have led to several random factors accounting for a higher proportion of the variation than the fixed factors evaluated. However, the trend and significance of the relationship indicated the likelihood of our mechanisms proposed and the need for a more robust sampling scheme that will reduce field scale variability and increase the robustness of predictions.

We also acknowledge the limitation of our study by not exploring plant P uptake or C exudation. Our experiment was limited to how SOC in the bulk soil responds to soil available P (irrespective of the amount taken up by plants) and not the rhizosphere soil. Despite these limitations, we ensured to account for the contribution of the amount and DOC concentration of litterfall to the soil DOC pool. As shown in our results, there was no significant difference in these variables across the treatments, hence no significant contribution to the soil DOC that might influence the P-SOC relationship evaluated.

Comment

In addition, the statistics presented in the figures are unclear, and some methodological details are not explained or are inconsistent (e.g. P fraction methodology and naming/presentation in figures). Also, the manuscript would profit from streamlining the text (introduction and discussion are long and sometimes repetitive).

Response

We have taken time to clarify the presentation of the statistics used in the figures and provided details of these in the footnotes. We have also additional revisions and clarifications of the relevant P fractionation methods and presentation in the figures. Details of these changes are provided in the respective specific comments raised below.

We have also ensured that the manuscript is revised to streamline the relevant sections in the introduction and discussion by eliminating identified repetition to improve its readability.

Please find below some more specific comments:

Material & Methods

Comment

L. 134: "latosol red soil", in which soil classification system? Instead of giving the equivalent in the US system, it might be more universal to give the equivalent in the WRB system?

Response

We have provided the equivalent of the soil type (Acrisol) according to the World Reference Base for Soil Resources (WRB) classification system.

Comment

L. 143: Where can we find the basic nutrient concentration data?

Response

We regret omitting this information. The basic soil properties have now been provided in Table S1 in the supplementary material, as shown below:

Table S1. Basic soil properties of each treatment plot before treatment application

	SM (%)	pH	NO_3	NH_4	AP	TP	TC	TN	TK	BD
			← (mg kg^{-1}) →				← (g kg-1) →			(g cm^3)
C	59.51 ±5.23	4.22 ±0.12	5.01 ±2.47	2.42 ±1.41	2.59 ±0.65	275.13 ±51.72	18.75 ±7.22	1.33 ±0.45	12.01 ±1.82	1.05 ±0.20
P1	61.23 ±2.34	4.17 ±0.12	5.17 ±3.77	3.32 ±2.69	3.11 ±1.01	262.98 ±65.67	17.31 ±7.38	1.29 ±0.49	10.89 ±1.47	1.01 ±0.17
P2	61.89 ±5.95	4.22 ±0.14	6.62 ±3.28	2.74 ±0.67	2.63 ±0.61	301.54 ±64.14	20.57 ±7.59	1.47 ±0.50	11.2 ±2.19	1.04 ±0.25
P3	61.91 ±2.37	4.23 ±0.16	7.58 ±3.51	2.87 ±0.99	2.43 ±0.85	301.7 ±74.62	16.41 ±5.24	1.2 ±0.30	13.44 ±3.11	1.19 ±0.23

Values shown indicate the mean of 6 samples from the combined 0-10 and 10-20 cm (0-20 cm) depth. The numbers after the ± sign indicate the standard deviation of the means. SM: soil moisture, DB: bulk density, AP: available phosphorus, TC: total carbon, TN: total nitrogen, TP: total phosphorus, TK: total potassium

Comment

L. 184 Chapter title "Samples analyses" is very broad, I suggest dividing this part into several chapters, separating the quite different analyses.

Response

As suggested, we have separated the section on sample analyses into 3: Chemical analyses, enzymatic and microbial biomass analyses, as well as metagenomic analysis.

Comment

L. 193 and 207: Please be more precise regarding the methods. Malachite green method is used to determine phosphate/inorganic P in different extracts. This can be done for different pools and is not necessarily available P. E.g. instead "Inorganic P in xx and xx extracts was determined using Malachite green". Also, for microbial P there needs to be an extraction step before measuring P by the Malachite green method. How was this done and was this step part of the sequential extraction or not?

Response

We apologize for this poor presentation of the P extraction process. We have carefully revised the presentation of the methods accordingly, as follows:

"Soil available P concentration was extracted using the Bray-1 method. The extracted available P and the sequentially fractionated inorganic P were measured using the malachite-green method (Ohno and Zibilske, 1991) and read on a spectrophotometer (UV-3802H, UNICO, Shanghai, China) at a wavelength of 640 nm".

We have further expatiated on the methods used for the extraction and measurement of MBP as follows:

"Because of the weathered nature of the soils and its high P sorption potential, we utilized NaH_2PO_4 as a spike during the determination of MBP and measured its recovery to adjust for P sorption. The P in the sample, control, and recovery was extracted using the Bray-1 method and measured using the malachite-green method."

Comment

How did you determine the organic P concentrations (NaHCO3 Po and NaOH Po)?

Response

Due to the need to be concise, we provided the reference for the extraction method (the original method by Hedley and the modification by Hou et al. used in our study). Our study focused on the available P and the NaOH-extracted inorganic P, representing the Fe-bound P. Therefore, the mentioning of the organic P fractions was to indicate that they were products of the sequential extraction process, part of which provided the NaOH-extracted inorganic P.

Results

Comment

L. 252 and Figure 1: In the method section you described a modified Hedley extraction including resin Pi but not the Bray extraction that seems to be behind the "available P" discussed here. Why are you not showing the resin Pi when you then use the other pools from the sequential extraction (NaOH Pi)? If the Bray extraction was used this needs to be added/adapted in the method section.

Response

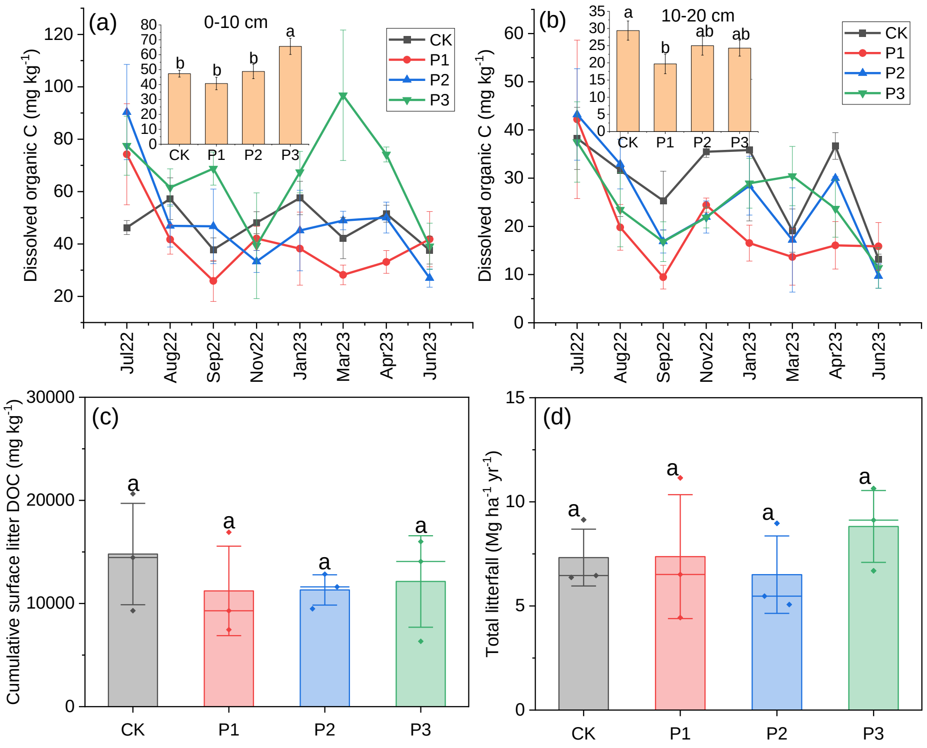
We are sorry for the mix-up. We have clarified in the methods section to show that the available P was extracted using the Bray-1 method. Our target from the sequential P extraction was to obtain the NaOH-extracted inorganic P (which represents the Fe-bound P) relevant to our current investigation. Other P forms including the resin P were part of the extraction outcome and not necessarily our target fraction.

Comment

Figure 1 (also applicable to other Figures): For the inserted small plot (barplots in general), it is not clear from the plot which treatments differed from each other. The statistics were done only for the average over time (inserted small plot), or? Also, the p-value in the plot is different from the p-value in the Figure description.

Response

The inserted bar plots within each line graph represent the mean annual average of the respective data presented in the main line graph. We have ensured to provide the statistics of the plots using mean separations and provided additional information on the statistics in the footnotes of the figures to clearly explain the contents of the figures. The initial p-values presented were for the inserted bar plots. However, to avoid confusion, a proper mean separation was done for the bar plots across all the relevant figures, and the p-values initially shown were removed to ensure clarity. A revised figure (Fig. 2) is shown below for reference.



Comment

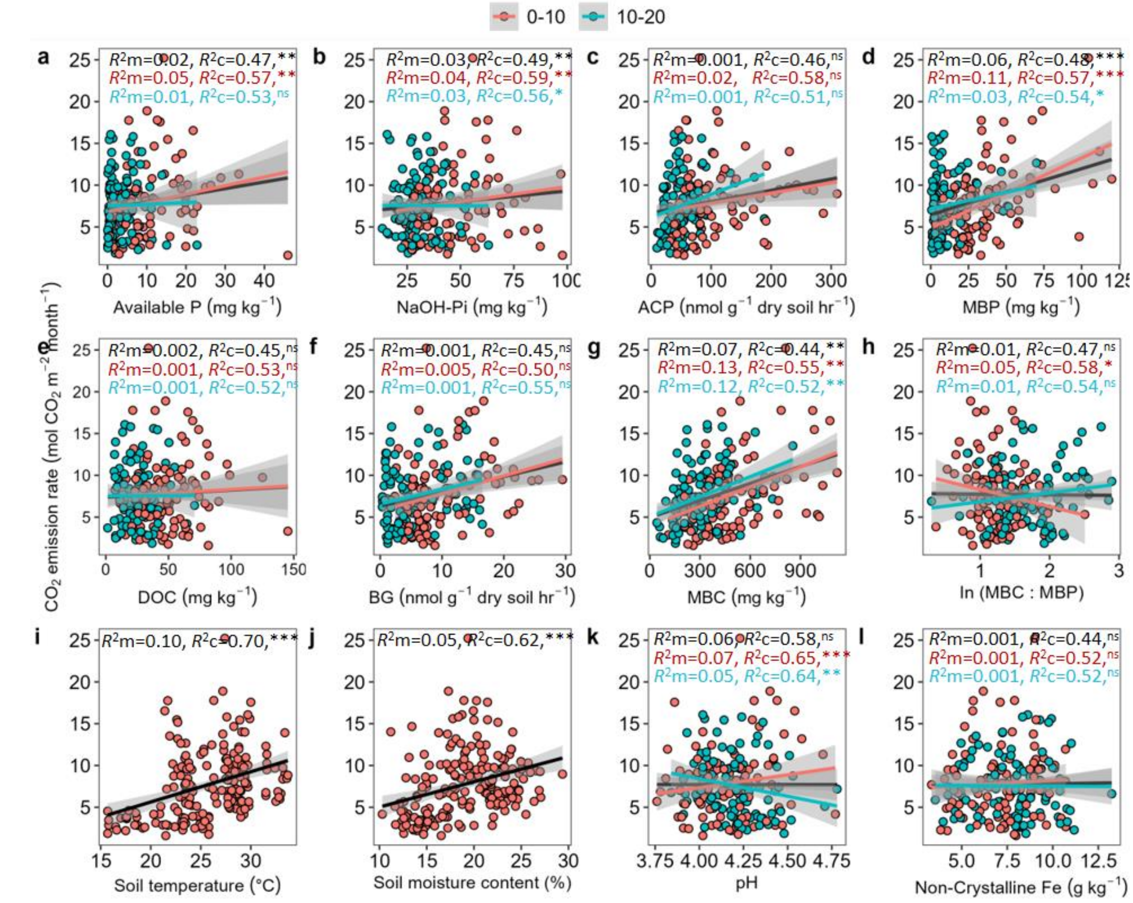
L. 316- 341 and Figure 4: Even if the relationships have significant p-values, some of the R2 are very low (< 0.1), and I would be more careful interpreting these relationships. To me, these figures mainly show that there is a lot of variability, as expected in a field setting. Also, if you discuss the soil horizons separately, please give the R2 for the relationships per horizon, not only the p-value.

Response

We agree with your observations.

Despite the significant relationships shown by our obtained p-values and the apparent trends, the R^2 values were low. We acknowledge that, unlike controlled experiments, several random factors at the field scale can result in the low predictive power of the fixed factor. The variation in sampling points at different sampling periods across the plots, treatment variation, variation in time of sampling, and number of samples collected, as well as other soil properties such as moisture content, temperature, etc. may account largely for these variations. To verify the effect of the sampling date on the relationships, we re-analyzed our data by subjecting it to the linear mixed-effect model using the sampling date as a random factor. The results reveal the importance of sampling date in accounting for a significant proportion of the variation among variables. For example, the relationship between DOC and available P had a marginal R^2 = 0.25 but a conditional R^2 of 0.44. Also, the relationship between CO_2 and MBP had a Marginal R^2 = 0.11, but a conditional R^2 of 0.57 for the 0-10 cm depth. This re-analysis clarified that the positive significant relationship between MBP and DOC or CO_2 was largely dependent on the sampling time, which cuts across different seasons (with varying soil and climatic conditions). Due to these random factors associated with the field settings, the linear regression model initially used did not account for these random factors (low R^2). As stated we suggested earlier, sampling for such large field-scale experiments should cut across numerous sampling points per plot, and sampling from points close to the previous sampling point, rather than randomly as currently done to limit large heterogeneity among samples.

As suggested, we have provided the R^2 values of each horizon, in addition to the indication of their p-values, as shown below. In addition, we provided the marginal R^2 (for the fixed factor) and conditional R^2 (influence of the random factor) values on each panel. To ensure clarity, we also provided adequate explanations of the figure elements in the figure captions.



Comment

Figure 5: What are the additionally added R2 representing here, they seem not to match the r given for the connections. Fe-P is mentioned in the description but not shown.

Response

The values between the variables in the SEM figure represent the correlation coefficient (r), which measures the strength and direction of the linear relationship between the two variables. The additional R^2 values shown beside the main response variables represent the coefficient of determination that indicates how well the model explains the variance of the dependent variable.

The mention of Fe-P in the footnote of the SEM figure was in error. We have deleted it accordingly.

Comment

L. 389-392: Yes, but in the detailed analysis with linear models in Figure 4 these patterns are not conclusive.

Response

Quite understood. As explained earlier, the complexity of the field settings may have contributed to the low R^2 values reported in the linear models between CO_2 and MBP, available P or Fe-bound P, despite the apparent trend and significant p-values. We have clarified that the sampling date (which encompasses the method of sampling, sampling time, and plot heterogeneity) was a key random factor affecting these relationships. The r values shown in the Mantel test indicate the autocorrelation coefficient of the relationships among the variables.

Comment

Figure 6: Why did you in the RDA (Figure 6h) include CO_2 together with the genes as target variable while the other chemical variables are used as explanatory variables?

Response

This is because the response of microbial genes and CO_2 (microbial respiration) is primarily regulated by the other soil chemical variables explored.

Comment

L.387: The mantel test seems mostly useful to target the relationship between different datasets (genes and soil properties), so to me it is unclear why this analysis is used here to focus again on the correlations between CO_2 and MBP, Ap, Fe-P (more detailed in Figure 4) and not to explain the links between genes and soil properties.

Response

You are right. Just as you stated, the mantel test as shown in Fig. 6g was used to establish the relationships between the P and C-cycling genes and soil properties. We have streamlined our interpretation to focus on the links between these genes and soil properties and avoided repeating the results on the relationships between soil chemical properties previously detailed in Figure 4.

Discussion

Comment

L. 453-455: Please check sentence logic; if BG activity decreased, it was not responsible for the breakdown of organic C and increase in DOC?

Response

Thank you for pointing this out. The confusion in the statement was due to missing words that altered the meaning of the sentence. The statement should read: "With the increase in DOC following P addition, there was a significant reduction in beta-glucosidase activity involved in the breakdown of organic C to release glucose for microbial use". We have corrected it accordingly.

Comment

L.461-464: There might have been no increased uptake/use of C by microbes but rather leaching of DOC?

Response

We acknowledge the possibility of DOC leaching aside its microbial uptake for metabolic processes. This assumption has been included as a potential reason for the insignificant change in MBC, aside from microbial uptake across the treatments. Our results on microbial use of DOC were proposed because the increase in MBP incorporation requires enormous C which was mostly derived from the DOC, since no concurrent significant microbial decomposition of SOC was observed.

Comment

L. 464-466: "MBP increased with increasing DOC concentration": I was not able to find a visualization of this relationship, except for the rank correlation in Figure 6g, where the correlation seems inexistent/minor and the SEM model in Figure 5 (r=0.14). This relationship would need to be shown more clearly for this conclusion. I agree that there is a cost of C for microbes, but MBP also increased with P addition (Figure 1), and this pattern seems stronger, so MBP might just follow P availability. Indeed, in your SEM model, the link between available P and MBP is stronger (r= 0.38) than between DOC and MBP (r=0.14). To be able to conclude that DOC increase is the main pathway increasing MBP, a treatment with DOC increase but no P addition would be needed.

Response

The visualization of the positive relationship between MBP and DOC is shown in the supplementary material (Fig. S5) (and shown below for reference).

We acknowledge the weak relationship between DOC and MBP in the SEM and correlation. This could primarily be because these relationships were linear and do not take into account, random factors that drive them. As shown in Fig. S5, where the linear mixed-effect model utilized the sampling date as a random factor, we show the large variation in marginal and the conditional R^2 values between this relationship. This indicates the significant role of the sampling date in predicting the relationship between DOC and MBP. On the other hand, the positive relationship between available P and MBP would largely be significant irrespective of the sampling times as P availability is expected to increase the amount of P taken up into the microbial biomass.

Being a field-based experiment, including a treatment with increasing DOC might be challenging, except in a controlled investigation (which was not part of this study). However, our revised model showed that this relationship was largely governed by the sampling dates, highlighting the complexity of this relationship which could be influenced by soil and climatic changes during the different sampling times, unlike previous reports that take no consideration of seasonal influence.

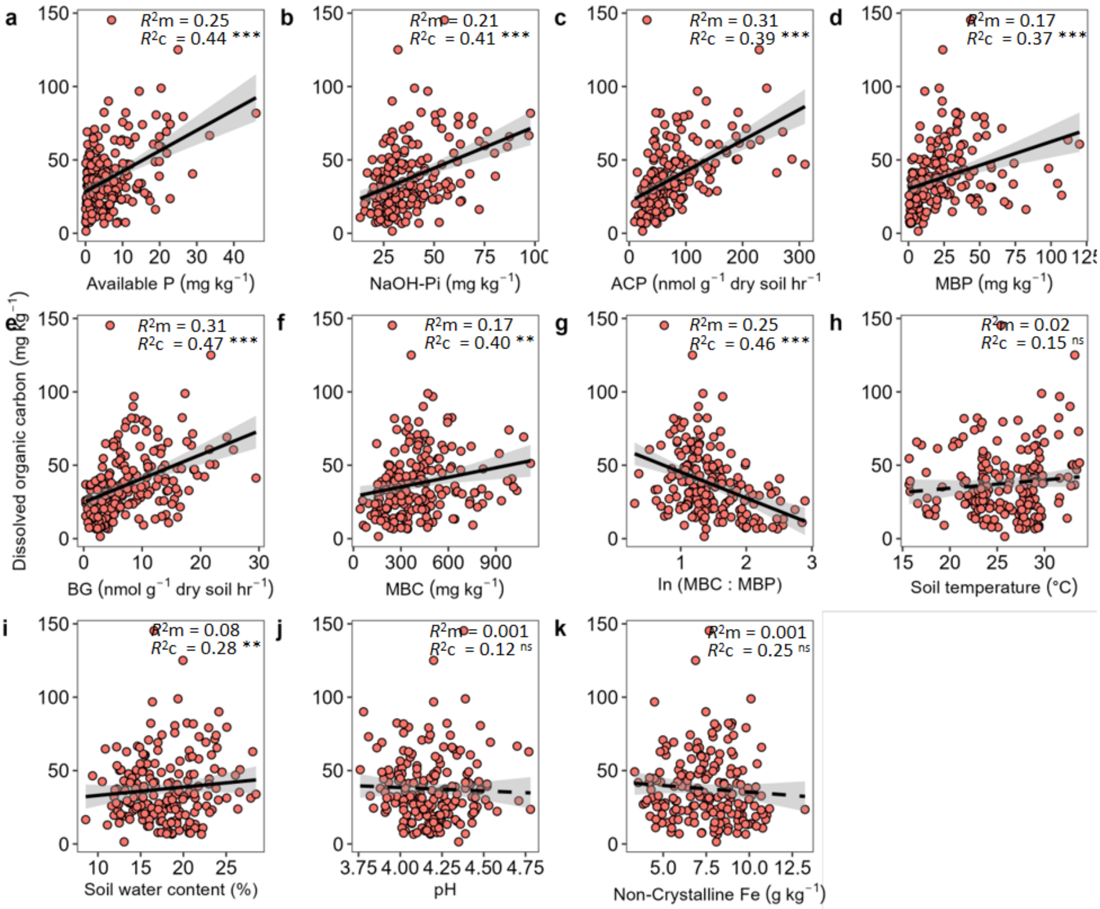


Figure S5

Comment

L. 487-489: From the presented results, it is not clear that the microbes were C-limited; you would need to explicitly check the effect of C addition to conclude that.

Response

While it is not clear if the microbes were C-limited, the increase in MBP infers the enormous C needed for this process. However, since no notable microbial SOC breakdown was observed, we propose that the C needed for MBP incorporation was derived largely from the desorbed C. The enormous C need for MPB incorporation could result in C-limitation, which was satisfied by the increased concentration of desorbed SOC. This is evident from the potential tradeoff of C incorporation into MBC for C utilization needed for MBP incorporation. The relationship between DOC and MBC:MBP shown in Figure S5, panel g (shown above) shows the significance of DOC for MBP incorporation, compared to MBC. Perhaps, we may consider including Figure S5 in the main text, rather than the supplementary material. Additionally, our Mantel's test showed a P significantly positive relationship ($p < 0.01$, $r >= 0.4$) between P cycling gene abundance and DOC compared to other variables. This indicated the significant need for DOC (which could have arisen from C-limitation) for P cycling.

Comment

L. 505: If you refer to the SEM model in Figure 5, Fe-P is not included there.

Response

The statement refers to the linear mixed-effect model shown in Fig. S5 (a revised version of the figure is shown above).

Comment

L. 519-521: Or again, this has to do with the higher P availability itself.

Response

The statement suggested that the negative relationship that ensued between DOC and the MBC:MBP ratio indicates that the labile organic C was preferentially used for building MBP instead of MBC. This is shown in Fig. S5.[panel g] (as DOC increases, the ratio of MBC:MBP reduces, indicating that the increase in DOC is significantly associated with increased MBP relative to MBC).

Comment

L. 594-595: Where do you show the leaching of DOC? The results focus on 0-20 cm DOC contents in soil; leaching itself was not shown in the manuscript/figures.

Response

We acknowledge the erroneous representation, especially as DOC leaching was not investigated, rather, its concentration was limited to the top 20 cm of soil explored. We have revised and removed the reference to DOC leaching to ensure that our conclusions are in line with our results obtained and devoid of speculations.