Dear Editor,

Thank you for your feedback and the opportunity to revise our manuscript. We have revised the manuscript egusphere-2023-2973 entitled “Geochemical and microbial factors driving crustacean assemblages in adjacent groundwater bodies within the same aquifer”, addressing all the comments and suggestions of the referees. We have also prepared a point-by-point reply (in normal font) to their comments (in bold) and a marked-up version of the manuscript showing the changes made.

We look forward to your further feedback.

Best regards,

Tiziana Di Lorenzo

Reviewer C1

1) The MS tackles a scientifically interesting topic, i.e. environmental drivers of groundwater fauna biodiversity. In particular, groundwater ecosystems in volcanic settings have been approached rarely. It is appreciated that the study is an interdisciplinary apporach and takes multiple aspects ranging from hydrochemistry to microbiology and groundwater ecology into account. From my personal opinion, the data set is worth being published, but the paper first needs substantial revision. The study mainly suffers from the low number of sites and samples analysed in combination with an over-interpretation of the findings.

Thank you for your positive attitude towards improving our study. We have revised the manuscript following your suggestions, as detailed below.

In the following, weak points are discussed in detail.

2) The unit ‚groundwater body‘ is an established unit in groundwater regulation. All European countries have subclassified their aquifer into groundwater bodies. Sometimes, aquifers with similar hydrogeological and hydrochemical conditions are subdivided into different groundwater bodies. I wonder if the groundwater bodies mentioned in this study follow this official classification. If not, I suggest to use a different terminology.

We thank the referee for this suggestion. After verifying, we confirm that the three groundwater bodies in our study do not follow the official classification (i.e., are not included in the Water Framework Directive monitoring). We will revise the manuscript to use the term "aquifer unit" (as the referee suggested in points # 4 and 5) instead of "groundwater body.

3) A major drawback of he study is the limited number of samples. 10 wells have been sampled twice, in 2014 and 2015. The sampling sites distribute to 3 different geochemical conditions, i.e. 4, 3, and 3, respectively. Finally, there is 6-8 data points for each condition only. With 40% oft he samples not containing fauna, the data set is super small.

We understand the referee's concern. However, we have several points to support our analyses. The fact that 40% of the samples did not contain fauna was expected and is a common finding in ecological studies because stygofauna are rare. The rule-of-thumb “one species per one well” is widely observed globally, and so are azoic samples. We revised the discussion section to acknowledge this potential limitation (see lines 527-537).

Regarding the hydrochemical characterization, a preliminary survey was performed in March 2012 (see lines 77-79) to explore specifically the hydro-chemical settings. These data were not used in the statistics because the microbiology was analysed only in the 2014-2015 sampling campaigns. All data available are consistent and confirm the identification of facies as reported in the text.

4) At the end oft he introduction it is mentioned that permeability and porosity are considered as characteristics in comparative analysis of the 3 aquifer units studies. Later in the MS I did not find any data on porosity and its use in data analysis.

We thank the referee for this point. The referee is correct. We did not analyze the role of permeability and porosity on the crustacean assemblages due to the lack of detailed data at the site scale. We revised the text accordingly.

5. I am not sure if I missed the information, but are the aquifer units studied hydrologically connected with each other or not. Having a different hydrogeochemical signature does not exclude a hydrological connection that would allow invertebrates to migrate.

Yes, the aquifer units are likely hydrologically connected. Nevertheless, groundwater fauna show a strong preference for certain types of aquifer units and it seems that some species do not migrate into different units. This is why we often

observe a high rate of endemism in groundwater. We revised the text to improve clarity. Please, see our changes to lines 449-454.

6) Fig 1 indicates that some sites are located in a geological setting with a very restricted areal extention (<200m).
Thank you for your observation regarding Figure 1. We understand your concern about some sites being located in geological settings with a very restricted areal extent (<200m). However, this is not a problem for our study, as these small, localized geological features are essential for understanding the fine-scale variations in groundwater chemistry and the associated biological communities. These areas, despite their limited size, provide valuable insights into the heterogeneity and complexity of the aquifer system.

7) While it is appreciaed that numerous trace elements have been monitored, important nutrients such as ammonium and phosphate have not been determied accordigly. Here, IC analysis are generally not sensitive enough.
We apologize for any misunderstanding regarding ammonium and phosphate. In the original manuscript, we mentioned that these ions exhibited concentrations below the instrumental detection limit. This also applies to phosphate. The sensitivity of the methods (0.05 mg/L) was sufficient to evaluate the abundance of the two parameters in groundwater, but we could not investigate the correlation with stygofauna at lower concentrations. As a result, we could not include these ions in our analyses. However, we realized that we failed to provide this crucial information in the methodology section. We have revised the manuscript to include this important detail. Thank you for highlighting this point. Please, see our changes to lines 253-254 and 422-424.

8) I propose to either write, sulfate‘ or ‚SO42-‘ The later includes the charge.
Thank you for pointing this out. We have consistently used "sulfate" throughout the manuscript, except for one instance where a typo ("sulphate") occurred. We have corrected this typo.

9) In Tab 1. What ist he order oft he field parameters, major ions, and trace elements based on? Trace elements are sorte in an alphbethical order. Major ions couls be sorted into anions and cations? I tried to understand the logic of sorting.
Thank you for your valuable feedback regarding the organization of Table 1. The order of field parameters, major ions, and trace elements was not based on a specific sorting logic, which might have caused confusion. We acknowledge the importance of a clear and logical presentation of these parameters. To clarify, the field parameters were placed first because they are measured directly in the field. However, we understand the need for a more systematic approach to organizing the data. We have revised Table 1 as follows: field parameters are first listed to reflect their in-situ measurement; major ions are now categorized into anions and cations for clarity; trace elements are now consistently sorted in alphabetical order. This reorganization enhanced the readability and logical flow of the table, making it easier to understand the sequence of parameters. We appreciated your suggestion. Please, see our changes to Table 1.

10) There is not detailed information provided how the groundwater invertebrates were collected. Yes, by pumping. But more details are needed to understand possible limitations. Were the 1000L pumped by a succion pump or a submersile pump. From fully screened wells or from wells with a short screened stretch. Using a packer system or not. This information should be provided in a revised version of the MS.
We apologize for not providing sufficient clarity on this methodological aspect. Our sampling methodology was based on the manual published within the framework of the European project PASCALIS. This project addresses the biodiversity and ecosystem aspects of groundwater conservation in Europe. Two of our co-authors were members of the project, with one contributing to the development of the sampling manual. This manual can be considered the gold standard for groundwater biomonitoring and can be accessed at this link (https://www.researchgate.net/publication/267567541_Sampling_Manual_for_the_Assessment_of_Regional_Groundwater_Biodiversity).
The wells that we sampled are private wells used for domestic o irrigation purposes. They were equipped with in-place submersible pumps. We did not use any packer system. We do not have direct information about the screens but usually in the region wells are fully screened, with screen openings of about 1-2 mm. To clarify, the in-situ filtration involved pumping 1000 liters of groundwater and filer them through a 63 µm sieve. We acknowledge the potential concern that the animals might be minced by the pump; however, the sampling method outlined in the PASCALIS manual has been extensively tested and validated to minimize such biases. The 1000 liters were indeed taken during purging the well. The method is designed to capture organisms present in the well and its immediate surroundings, thereby providing a representative sample of the groundwater fauna. We have revised the text providing these details also including the size of the animals. Please, see our changes to lines 157-164.

11) To my personal opinion, it is over-interpretation if a subunit of an aquifer is called stable in conditions based on 6 samples from 3 sites and 2 years.
We thank the referee for this point. We revised the text throughout the manuscript to downsize our claims about the stability of the hydrogeochemical conditions. Please, see our changes to lines 191and 397-399.

12) I asked the authors to soften the interpretation of their findings. Based on the very small data set, I doubt that there is sufficient evidence of groundwater fauna diversity being driven by trace elements, or fauna abundane is driven by selected major ions.
We have revised the text to soften our claims. Low abundances are a common trait of groundwater assemblages. We respectfully suggest considering this natural characteristic in the context of our findings. Furthermore, we used a stringent cut-off criterion based on previous studies (Korbel and Hose, 2011, 2017; Di Lorenzo et al., 2020) to prevent unreliable and exaggerated claims about scientific phenomena. We revised the text to state that our results seem to suggest the correlations we have highlighted. It is important to consider our findings to encourage other researchers to perform similar analyses in their future studies, which can further support or refute our claims. Please, see our changes to lines 412-413, 415-416, 434-436, 450-451, 485-490.

13) The study, to my opinion, does not provide conclusive data on the sensitivity of groundwater crustaceans to individual ions and/or trace elements. The statement should be softend.
Thank you for your insightful feedback. We acknowledge that the study may not provide conclusive data on the sensitivity of groundwater crustaceans to individual ions and/or trace elements. We have softened our claims accordingly. Please, see our changes to lines 439-441, 546-547.

14) Similarly, there is not real conclusive evidence for a selective feeding of groundwater fauna on LNA prokaryotic cells. Please, do not oversell the results.
While there is not conclusive evidence for a selective feeding of groundwater fauna on LNA prokaryotic cells, it remains a possibility with a statistical significance. We have revised the manuscript to downplay our claims. Specifically, we have suggested that further studies could confirm the role of feeding behavior in structuring the groundwater microbial community and biomass, with direct consequences on subterranean carbon turnover and nutrient cycling. Please, see our changes to lines 485-494.

15) It is appreciated that several tests have been used to estimate the ‚fauna diversity tob e expected'. However, with the liited data set, I wonder how reliable these estimations are.
Thank you for your comment and for appreciating the various tests we used to estimate the expected fauna diversity and functional traits. We understand your concern regarding the reliability of these estimations given the limited data set. We employed rigorous and widely accepted statistical methods to provide the most accurate estimations possible. These methods are permutational, i.e., designed to work with small sample sizes and have been validated in numerous ecological studies on freshwater meiofauna. We believe that our approach provides sound (though preliminary) insights that can guide future research in this area. Please, see our changes to lines 527-529.

16) The authors discuss the possiblity that the sulfate-depleted aquifer unit lack sufficient energy to host a divers groundwater fauna. What is this statement based on. The DOC levels are similar tot he other aquifer units. There is more than enough nitrate in all 3 units. Phosphate, which couls limit microbial production, have not been measured. Is there a clear hint that the sulfate-depleted aquifer unit is energy poor?
We thank the referee for this point. Our intention was to convey that the sulfate-depleted aquifer unit showed the lowest values of total cell count and no groundwater-dependent species. It is not a matter of diverse assemblages, but rather that we did not find any groundwater-dependent species in the sulfate-depleted aquifer unit, except for *Moriaria poppei*, which, however, is a surface water species that accidentally entered that aquifer unit. We do not wish to overstate our claims, but it is important to stimulate a discussion about the energy requirements for groundwater-dependent species to survive and thrive in their environment. This is an issue that has not yet been thoroughly investigated, and we aim to bring attention to it. We have revised the text to enhance clarity by ensuring that we are not overselling the results.Please, see our changes to lines 518-523.

17) Please soften the statements not only in the discussion but also in the conclusion section.
Done. Thank you very much for the time you've spent in the revision. Your suggestions helped a lot and improved pour manuscript. Please, see our revised conclusions.

Reviewer C2
The paper explores environmental and biotic drivers of invertebrate taxonomic and functional diversity in volcanic aquifer. The paper is multidisciplinary and combines ecology with geochemistry and microbiology. The topic is interesting and underexplored so I think the paper is worth considering. However, it has several drawbacks that I believe need to be addressed before final acceptance. First, I think that methodology, especially in relation to invertebrates is poorly described and potentially not appropriate. Second, I believe that sample size is (too) small to draw such strong conclusions as the authors did, and that one would need to address this issue more carefully.

Finally, I am missing raw data that would enable appropriate assessment of approach and reproducibility of the study. Below, I am addressing few main issues, mostly connected to analyses of invertebrate assemblages.
We thank the referee for her/his positive attitude about our manuscript. We have fully revised it to address most of the reviewer's concerns.

Sampling design

1) 10 sites were sampled twice, in Nov-Dec 2014 and Oct-Dec 2015. I am not sure if seasonality and stability of GW can be inferred from two sampling events, both performed at the same time of the year (late autumn-early winter). I suggest that authors explain why this specific design is sufficient to »confirm the occurrence and stable nature" (165-166).
Thank you for raising this important point. We acknowledge that seasonality in groundwater recharge can cause variability in the mixing of different waters from the unsaturated and saturated zones, which can be relevant for geochemical processes. While previous studies have shown that seasonality has relatively little influence on the distribution of biota in groundwater ecosystems, we recognize that this could be a potential flaw in our study. We have stated the possibility of this flaw in the last paragraph of the discussion. We address the limitations of our study design and emphasize the need for further research to confirm our findings. Please, see our changes to lines 536-541, 695-696.

2) Out of 20 samples, 11 were without invertebrates. Three of empty samples were from E-A GWB, while S_D GWB was consistently without invertebrates. Next, in three samples only one individual of one species was found. Is sample size of 9 samples with invertebrates sufficient to infer assemblage composition? If yes, why do you think so? Given the low abundance of many groundwater species this needs better elaboration. As study tests for three different communities in three different GWBs, rarefaction curves should be calculated per each group and not for pooled sample of all groups (fig S1, lines 179-184).
We thank the referee for this point. We are afraid there has been a misunderstanding. We collected 20 biological samples, where 8 (not 11 as mentioned by the referee) did not contain any invertebrates. So, 12 samples showed invertebrates, belonging to 9 species. At the aquifer scale, we have evaluated the completeness of crustacean sampling effort by using five non-parametric (Chao1, Chao2, Jackknife1, Jackknife2, and Bootstrap) and one parametric (Michaelis-Menten) estimators (Magurran, 2021). Three non-parametric estimators (out of five) and the non-parametric one indicated that 100% of the expected biodiversity was collected during the sampling survey. So, we could be sure enough that our monitoring had been sufficient to infer the crustacean assemblage composition. On the other hand, the remaining two estimators suggested that we likely missed out 1 - 3 species. Following the referee suggestion, we also performed the analysis per each aquifer unit. The results indicated that we collected 100% of the expected biodiversity in the sulfate-depleted aquifer unit, while 1-3 species were likely missed out from the earth-alkaline and K-rich aquifer units. We discussed this uncertainty, stating that this finding is commonly encountered because groundwater sampling is essentially a blind process (Mammola et al., 2021). Wells serve as windows through which we gain insight into the subterranean biodiversity in the portions of the aquifer surrounding them, but groundwater bodies extend more extensively (Ficetola et al., 2019). The rule-of- thumb "one species per one well" is a common finding and we revised our text to make it clearer. We have revised the final paragraph of our discussion to highlight that incomplete sampling may be a limitation of our study. Consequently, our findings should be considered preliminary, yet informative, as complete sampling is rarely achieved in studies concerning groundwater ecosystems. Please, see our changes to lines 215, 311-315, 531-535.

Invertebrate taxonomy

3) Invertebrates were not determined to species level as the text suggests (line 157) but to level of morphological taxonomic units. Several taxa are determined to the level of genus. Given the high presence of cryptic species in GWs, COI barcoding would be welcome. If not, then it should be clear from the text that morphospecies were the units of interest.
Thank you very much for your insightful comments regarding the taxonomic resolution of our invertebrate analyses. We completely agree with your observation that our analyses were based on morphological characteristics. We recognize that several taxa were identified only to the genus level and that the use of morphological taxonomic units might overlook cryptic species diversity. To address this, we revised the manuscript to clarify that our study primarily utilized morphological-based analyses. Additionally, we acknowledge the potential of DNA barcoding, specifically COI barcoding, in uncovering cryptic diversity. However, as highlighted in recent studies by Korbel et al. (2024), there are limitations to the current eDNA approaches. These studies suggest that eDNA methods, which commonly use biomarkers such as 16S and COI, may miss a number of crustacean species and should not yet be fully decoupled from morphological analyses. In light of your suggestion, we revised the text to reflect the importance of integrating eDNA methods as a future perspective for this study. We believe that combining morphological and molecular approaches will enhance the

accuracy and comprehensiveness of biodiversity assessments in groundwater ecosystems. Please, see our changes to lines 170-173, 533-541.

Functional traits

4) Methodology is poorly described. Authors classified organisms based on three functional traits: locomotion (burrowers, interstitial, swimmers) feeding habits (fine sediments-microorganisms-living microphytes, deposit-feeders, collectors, grazers) and thermal tolerance (eurythermal, moderately stenothermal, stenothermal). Since not all taxa were studied in cited literature, and also the variables measured and analysed in this study are not clear, I suggest the authors expand the methods section with explanation, and include raw data / literature supporting classification.
You are right. We agree with your observation that the methodology description should be more detailed regarding the classification of organisms based on their functional traits. In response to your suggestion, we revised the methods section to include a more detailed explanation of the classification criteria for locomotion, feeding habits, and thermal tolerance. We also included references to the literature that supports our classifications in Table 2. Please, see our changes to lines 176-188 and Table 2.

5) Now, one needs to combine two different supplementary tables to figure out that one specimen of *Niphargus sp*. was classified as a: 1 burrower, 0.5 as deposit-feeder and 0.5 as collector, and 0.5 as eurythermal and 0.5 as stenothermal. Another issue, visible from this case is that the decision process for species where their status is not known is not described at all. I would guess that 0.5 represents lack of knowledge and not actual composition of assemblage of 1 specimen. This has major impact on subsequent analyses, especially given the small sample size.
Thank you for your insightful comments regarding the classification and decision process for species traits, particularly for *Niphargus* sp.. We acknowledge that the current method of combining supplementary tables to determine the classifications is cumbersome. Additionally, we recognize the need to better describe the decision process for species where specific trait information is lacking. We revised Table 2 to provide full information about traits' attribution. We revised the methodology (please, see our response to your previous point). In addition, you are right. The value of 0.5 represents an equal likelihood of *Niphargus* sp. exhibiting traits across multiple categories due to the variability observed within the genus and the fact that our specimens were not identified to the species level. This was done to reflect the known diversity in their feeding habits and thermal tolerance (Fišer, 2019). Field data support this view to some extent. *Niphargus virei* is intolerant to cold temperatures in the laboratory, which agrees with its distribution that is limited with presumed borderline of Alpine Pleistocene glaciers (Folquier et al., 2008). By contrast, *Niphargus rhenorhodanesis* shows remarkable tolerance to low temperatures in the laboratory, exceeding the tolerance of surface *Gammarus*. Since we did not identify the specimens down to the species level, we preferred to share the percentages equally per each category to reflect this variability. Maxillae I in *Niphargus* are the most variable appendages in the mouthpart apparatus. The number of spines on its outer lobes and density of denticles along these spines perhaps indicate different modes of feeding. As for the potential major impact on subsequent analyses, we respectfully dissent from your opinion. *Niphargus* sp. was represented by only one specimen out of a total of 210 crustacean individuals collected. Therefore, its influence on the overall analysis is minimal. Please, see our changes to lines 176-188 and Table 2.

6) Second, in this particular case I am wondering what morphological traits were used to determine *Niphargus* as borrower and to define its feeding habits, given the fact that burrowing is not recognised as usual locomotion of *Niphargus* species, and that few studies that addressed feeding habits of the genus showed high diversity in diet of different *Niphargus* species? Again, given the small sample size, erroneous classification can have high impact on final results – I suggest to review and expand this part of Methods section and include relevant literature, data and clarifications.
You are correct; we did not provide sufficient information in our original draft on the criteria used to classify the traits of each species. We acknowledge that this omission may have led to confusion, particularly in the case of the genus *Niphargus*. To address this, we have revised the Methods section to include a detailed explanation of the morphological traits and literature used to determine the classifications for locomotion and feeding habits. For example, while it is true that burrowing is not typically recognized as a common locomotion trait for all *Niphargus* species, there is evidence (including our personal observations and studies such as Hose and Stampp, 2019) that some *Niphargus* species do exhibit burrowing behavior. Regarding the diversity in the diet of different *Niphargus* species, we agree that this variability needs to be clearly addressed. We have revised and refined our classification criteria, ensuring they are well-supported by relevant literature and data. Please, see our changes to Table 2.

Hose, G.C., Stumpp, C. Architects of the underworld: bioturbation by groundwater invertebrates influences aquifer hydraulic properties. Aquat Sci 81, 20 (2019). https://doi.org/10.1007/s00027-018-0613-0.

Results

7) Table 2: how were the mean values calculated and how were zero abundances treated? I find the reporting of results a bit confusing. For example, how can *Niphargus sp*. have a mean abundance value 2 for K-rich, while *Morraria popei* has mean abundance value 1 for S-D while both morphospecies were represented with only 1 individual in total, according to Table A1e? Also, I would suggest different metric, as mean value rounded to integer higher than 0 (if this was the case) is not informative with such small abundances. The same goes for functional composition. If it was indeed calculated from equal probabilities that morphospecies belong to specific class, this is not appropriate approach. In my opinion, morphospecies for which you do not have info about their functional trait should be excluded. Also, were abundances of each species incorporated into calculation of trait profile per sample and per GW? This is not clear from line 162.
You are absolutely right. We have revised Table 2 to provide the number of individuals per each species and functional category for each aquifer unit. About how zero abundances were treated, in the original version of our manuscript we wrote that "We added a dummy variable equal to 1 to all data to allow the analysis of [abundance] values equal to zero (Clarke and Gorley, 2005)." We revised the text to ensure that the method for calculating trait profiles per sample and per groundwater unit is clearly described. Please, see our changes to Table 2 and lines 190-192.

Data availability (lines 520-521)

8) Appendix 1 does not contain raw data of functional traits for each morphospecies, but only derived functional composition per site (Table A1f). Reported data is not sufficient to fully reproduce the results. Issues and questions related to invertebrate composition and functionality should be resolved and better explained before one can evaluate accuracy of subsequent methodological approaches, results and interpretation.
In line with our previous responses, we have revised the manuscript to include the raw data of functional traits for each morphospecies. This revision is reflected in Table 2 that provides detailed trait data for each species, ensuring transparency and reproducibility of our results. We respectfully disagree that the reported data is insufficient to fully reproduce the results. The tables in the appendix are exactly those we used in our analyses. We encourage our readers to reproduce the analysis using these tables to verify their accuracy and completeness.

9) Additionally: Figure 1 caption is not complete (Geology types 44 and 46 are missing).
Thank you for pointing out the incomplete caption for Figure 1. We apologize for the oversight. We have revised the caption to include the missing geology types.
We appreciate your attention to detail and your valuable feedback.

Reviewer C3
The manuscript follows a multidisciplinary approach which nicely links groundwater fauna to biogeochemical data.
The research questions are stated in the introduction.
The methods and statistics section are sound and fairly clear to the reader.
However, there are several issues that remain unclear and should be revised.
Thank you very much for your positive feedback and for recognizing the multidisciplinary approach of our manuscript. We appreciate your acknowledgment of the clear research questions and the sound methodology and statistics sections. We have revised the manuscript to address any issues that were unclear.

1) In general, the research questions are only partly answered and the dataset is limited to do so. Some methods, especially the collection of the crustaceae, are not clearly described or seem to introduce a large bias.
Thank you for your constructive feedback on our manuscript. We apologize if our explanation of the research questions and their answers was not sufficiently clear. We have revised the manuscript accordingly to ensure our research purpose is more explicitly communicated. The goal of our study was to assess whether the environmental differences that we highlighted in the three aquifer units were also mirrored by the biological assemblages in adjacent groundwater bodies within the same aquifer. Additionally, we aimed to identify the main drivers of these disparities, focusing on hydrological, chemical, and microbial factors. We have revised the manuscript to ensure that these research questions are clearly stated and to explain how our findings address each of them. Additionally, we have provided a more detailed description of our methods, particularly regarding the collection of crustaceans, to address any potential biases and improve the clarity of our study design. Our sampling methodology was based on the manual published within the framework of the European project PASCALIS. This project addresses the biodiversity and ecosystem aspects of groundwater conservation in Europe. Two of our co-authors were members of the project, with one contributing to the development of the sampling manual. This manual can be considered the gold standard for groundwater biomonitoring and can be accessed at this link (https://www.researchgate.net/publication/267567541_Sampling_Manual_for_the_Assessment_of_Regional_Groundwater_Biodiversity).
To clarify, the in-situ filtration involved pumping 1000 liters of groundwater and filter them through a 63 µm sieve. We acknowledge the potential concern that the animals might be minced by the pump; however, the sampling method outlined in the PASCALIS manual has been extensively tested and validated to minimize such biases. The 1000 liters were indeed

taken immediately after purging the well to ensure collecting the animals residing in the aquifer and not in the water column of the well. Regarding the filter screen of the borehole, we ensured that the wells were fully screened at their bottom and that the diameter of the filter holes was appropriate relative to the size of the animals (< 1 mm) we collected. Please, see our changes in the Methods.

In the following there is a more detailed feedback:
2) L41: The differentiation of aquifers into groundwater bodies is not apparent from Aquilina et al. 2023. To my knowledge, it rather describes a high spatial variability of hydrological and geological factors within aquifers.
Thank you for your comment regarding the differentiation of aquifers into groundwater bodies. We acknowledge that Aquilina et al. (2023) primarily describes the high spatial variability of hydrological and geological factors within aquifers. Our intent was to highlight the complexity and variability within aquifers, which can be characterized into distinct aquifer units based on specific criteria. Following the suggestion of a previous referee, we have revise the text using "aquifer units" in the place of "groundwater bodies".

3) L54: I would at least note the difficulty of inferring filtering because competition can give rise to patterns identical to those caused by environmental filtering (Cadotte and Tucker, 2017)
Thank you for this comment. We have revised the manuscript to mention the difficulties in accurately distinguishing environmental filtering from competitive interactions, as both can produce similar patterns in species distribution. We addressed this aspect in the introduction. However, the pattern we observed, particularly the complete absence of species in the sulfate-depleted aquifer unit (with the exception of one specimen of an epigean copepod species), is unlikely to be due to competition. Please, see our changes to lines 64-72.

4) L58: How do you define trophic filters? Availability of food? But then you would that than still be considered environmental filtering or competition?
We define trophic filters as factors related to the availability and type of food resources in an environment that influences the distribution and composition of organisms. Trophic filters can encompass both environmental filtering and competition. We have include this definition in the revised manuscript to clarify the concept of trophic filters. Please, see our changes to lines 64-72.

5) L59: I doubt that groundwater species compete for space.
We respectfully disagree with your assertion that groundwater species do not compete for space. Studies by Culver and colleagues in Appalachian cave streams in the United States provide compelling evidence of interspecific interactions influencing dispersal (Culver et al., 1991; see review in Culver and Pipan, 2019). These studies illustrate that amphipods and isopods are highly aggregated in riffles, where they reside on the underside of rocks to avoid being washed out by the current. In these environments, individuals compete for space, leading to encounters that result in one individual moving to another stone or being displaced by the current.

6) L113: is "volcanic apparatus" the correct terminology?
Thank you for noting the point. Volcanic apparatus, system, and edifices are used to refer to a volcanic complex with multiple eruption centers. In this study case, the term "system" is likely more appropriate. We have revised the text accordingly. Please, see our changes to line 120.

7) L154: To me the in-situ filtration is not clear. You pumped 1000L and collected it with a 63 um sieve? Are the animals not minced by the pump? I would assume that this method introduces considerable a bias. I assume the 1000L are taken immediately before purging the well. This is not clear from the method part. Are the animals already in the well before pumping or do you "suck" them into the well. Previous net sampling could have answered this question. Do you have any information on the filter screen of the borehole? Are they fully screened. What is the diameter of the filter holes - compared to the size of the animals you find?
We apologize for not providing sufficient clarity on this methodological aspect. Our sampling methodology was based on the manual published within the framework of the European project PASCALIS. This project addresses the biodiversity and ecosystem aspects of groundwater conservation in Europe.
(https://www.researchgate.net/publication/267567541_Sampling_Manual_for_the_Assessment_of_Regional_Groundwater_Biodiversity).
To clarify, the in-situ filtration involved pumping 1000 liters of groundwater and filter them through a 63 µm sieve. We acknowledge the potential concern that the animals might be minced by the pump; however, the sampling method outlined in the PASCALIS manual has been extensively tested and validated to minimize such biases. The 1000 liters were indeed taken during purging the well. The method is designed to capture organisms present in the well and its immediate surroundings, thereby providing a representative sample of the groundwater fauna. We have revised the text providing these details also including the size of the animals. Please, see our changes to lines 161-170.

8) L160: I like the idea of differentiating functional traits but how did you get this information?

We have revised the text to provide detailed information on how we classified the functional traits of each species, including the literature and data sources used for these classifications. Each trait was assigned based on existing literature, supplemented by our own observations from previous preliminary investigations. Each species was attributed to only one category per trait, with the exception of *Niphargus* sp. In our study, each species was attributed to only one category per trait, except for *Niphargus* sp., which showed a range of traits across different categories. For instance, temperatures in subterranean environment are presumably stabile and subterranean species are presumably stenothermal. Field data support this view to some extent. *Niphargus virei* is intolerant to cold temperatures in the laboratory, which agrees with its distribution that is limited with presumed borderline of Alpine Pleistocene glaciers (Folquier et al., 2008). By contrast, *Niphargus rhenorhodanesis* shows remarkable tolerance to low temperatures in the laboratory, exceeding the tolerance of surface *Gammarus*. Since we did not identify the specimens down to the species level, we preferred to share the percentages equally per each category to reflect this variability. Maxillae I in *Niphargus* are the most variable appendages in the mouthpart apparatus. The number of spines on its outer lobes and density of denticles along these spines perhaps indicate different modes of feeding. Please, see our changes to Table 2 and lines 176-192.

9) L200: Missing citation for the BEST procedure.
We apologize for this oversight. The BEST procedure reference for the DistLM model is:

Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. Ecological Monographs, 69(1), 1-24. We will revise the manuscript to include this citation.

Table1
10) The statistical differences are not clearly indicated.
e.g. Potassium: K the lowest and the highest value both are labeled with superscript letter b, but the medium value (from sulfate-depleted GWB) is labeled a. This doesn't make sense to me (also when looking at the plot). There are several others where it doesn't make sense (e.g. U)
We apologize for the inconvenience and confusion caused by these typos. You are correct that the labeling of superscript letters, for potassium (K). However, for uranium (U), it is so. Please, see the results of the analysis in Table S1. We have carefully double-checked the data and revised the table.

11) Regarding the substrates of the Ecolog plate it is not clear what values are stated. Are these percentages of degraded substrates?
The percentage values were calculated as a fraction of the total absorbance of a single sample in the plate. The absorbance is proportionally related to the substrate degradation. We added these methodological details to clarify the issue. Please, see our changes to lines 160-161.

Figure2
11) Wouldn't it make more sense to report the ratio of HNA to LNA. As shown now it seems to be redundant information.
Thank you for your suggestion. While reporting the ratio of HNA to LNA cells could provide useful information, we believe it is important to present the data for HNA and LNA cells separately. This approach allows us to capture and analyze the distinct contributions and behaviors of these two cell types within the microbial community. By examining them separately, we can better understand their individual roles and interactions within the groundwater ecosystem. Additionally, we tested their effects on the composition of the crustacean assemblages using linear models, and they are not redundant.

12) The n. ind. for Crustaceans is calculated is given per 1000L?
Yes, thank you for this point. In each sample, abundances are reported as number of individuals per 1000 L of groundwater. We have revised the caption of Table A2e to make it clearer.

13) L283 - 288: Difficult to understand
There was a typo in the sentence (repetition of part of the sentence). We have revised the manuscript. Please, see our changes to lines 291-293.

14) Table2: I assume the traits are given as relative abundances but don't add up to 1. What are the rest? not determined? Is there any reference for the trait types?
You are correct; the traits are given as mean values, which is why they do not add up to 1. We acknowledge that this has caused confusion, and we apologize for the lack of clarity. The same issue has been raised by a previous referee, and we have revised the table. We have provided a full explanation for the traits in the revised manuscript, ensuring that it is clear which traits were calculated. We have also included references for the trait types to support our classifications.

15) L337: Make clear to which multivariate models you refer.

We apologize for any confusion caused. To clarify, the multivariate models we referred to are distance-based linear models (DistLM) and distance-based redundancy analyses (dbRDA). These models were used to assess the relationships between environmental factors (hydrogeological, physical-chemical, and microbial variables) and the taxonomic and functional composition of crustacean assemblages. We have revised the text to clearly specify these models and provide a detailed description of the factors they analyze. Please, see our changed to lines 350-375.

16) In the results and discussion there is little information on the first research question: "if the hydrogeological factors affect the crustacean assemblages", but the explanation is mostly based on the physio-chemical and microbial data. Since this is directly stated as an objective in the introduction you should elaborate on this question.
Thank you for your valuable feedback. We appreciate your attention to this important aspect of our study and made the necessary revisions to improve the clarity of our manuscript. Please, see our changes to results and discussions.

Reviewer C4

General comments

1) Good research with a multidisciplinary angle that fits well into the EGU philosophy. Please, flow my specific comments that can improve your manuscript.
Thank you for your positive attitude towards our study. We sincerely appreciate your comments and the time you spent revising our paper.

Specific comments
2) Lines 19-36. Make clear in the abstract the size of your crustaceans. The pore size of some volcanic lithofacies is large.
Done. We have revised adding that the species ranged in size from 0.036 to 1 mm.

3) Lines 19-36. Make clear in the abstract if you're looking at the life of organisms as indicators of groundwater quality or issues transport and flow velocities in the subsurface.
We apologize for not making our aim clear enough in the original manuscript. Our goal was to assess the taxonomic composition and functional diversity of crustacean assemblages in the three groundwater bodies. Specifically, we wanted to determine if the hydrogeochemical differences we identified were also reflected in the biological assemblages. We have revised the abstract and final paragraph of the introduction to make this objective clearer.

4) Line 60. "Research on the effect of environmental filtering on groundwater crustaceans has primarily focused on karst aquifers". Check if relevant literature exists on the Chalk in southern England which is karstic. I have seen recent and relevant presentations at conferences at the Geological Society, London.
Thank you very much for this observation. Indeed, Maurice et al. (2016) have conducted several studies on the invertebrates of chalk aquifers. We intend to expand the literature to include these studies.
Maurice, Louise, et al. "The invertebrate ecology of the Chalk aquifer in England (UK)." Hydrogeology Journal 24.2 (2016): 459-474.

5) Line 70. "Permeability and porosity of the rock matrix". Knowledge of matrix properties of aquifers are poor in Central Italy, but laboratories are currently setting up new equipments. Clarify current status of knowledge.
We regret that we cannot accommodate the referee's suggestion regarding the use of new equipment to assess the permeability and porosity of the rock matrix, as it falls outside the scope of our study. We do not have direct information on these parameters apart qualitative estimations from the literature. So, we have revised the text deleting the reference to these two parameters and leaving only the depth to water table. We appreciate the referee's comment, which helped us refine our objectives. We aimed to assess the effects of water table depth (as a proxy for aquifer units' isolation from the surface) as the main hydrogeological factor. We have cited the paper by Brancelj et al. (2016) in this context.
Brancelj, A., Žibrat, U., & Jamnik, B. (2016). Differences between groundwater fauna in shallow and in deep intergranular aquifers as an indication of different characteristics of habitats and hydraulic connections. Journal of limnology, 75(2).

6) Lines 77-113. Which is the driving force of flow in your aquifer units? Matrix, fractures or both?
In the volcanic formations of Sabatini Mounts, porosity is due to both fractures and matrix. Primary porosity, related to matrix, is more important in the pyroclastic fall products while pyroclastic flows and lava flow have coexistence of primary and secondary porosity, the latter due to either cooling fractures or fragile deformation. See Manca et al 2017 for a more detailed description.
F. Manca, S. Viaroli & R. Mazza (2017) Hydrogeology of the Sabatini Volcanic District (Central Italy), Journal of Maps, 13:2, 252-259. http://dx.doi.org/10.1080/17445647.2017.12977407)

7) Lines 81-83. Specific volcanoes? Sabatini?
Yes, we are referring to the Sabatini Volcanic complex, in particular to its eastern border where the volcanic products overlap on the sedimentary layers of Pleistocene. We have revised to better specify this in the text.

8) Lines 219-384. The results are accurate.
Thank you for your positive attitude toward our study.

Figures and tables
9) Figure 1. "N" for the north above the arrow?
We have revised the figure. Thank you.

10) Figure 2. Good figure. There is room to make it larger.
Thank you for your positive feedback on Figure 2. We appreciate your suggestion to enlarge the figure. However, we believe that the current size of the figure effectively conveys the information without overwhelming the layout. We prefer to keep it as it is to maintain the balance and clarity of the overall presentation.

11) Figure 3. I have here information on the size of this crustacean species in the caption. Make the information clear in the text.
Done. Thank you.