Dear Chiara Borrelli,

We would like to thank you and the two reviewers for helpful comments on our manuscript. They have helped to improve our submission, and we are thankful for their critical assessment of our work. In our response, we have addressed each of their concerns and questions in the following format: Each question or comment is re-stated as in the original review of the manuscript in **bold font**. Our response to each comment/question is written in normal black font. Specific changes made in the manuscript are in blue font. Line numbers refer to the *revised track-changes version* of the manuscript.

Referee 1 – Adam Woodhouse (AW)

I very much enjoyed reading this paper and believe it is a fantastic contribution to Biogeosciences. I have left very minor comments which need addressing/acknowledging, but other than that, I look forward to the manuscript being published.

> We thank the reviewer for their positive assessment of the manuscript and appreciate their constructive feedback. We will address all minor comments as requested.

AW1. Redefine ECVs in the introduction.

> AC: Thank you for this suggestion. We redefined ECVs in the introduction in the revised track changes manuscript (**l. 42**).

AW2. "Would Polar Amplification be more appropriate as this study has global implications?"

> AC: We appreciate this point. While our sampling was conducted in the Arctic and initially framed in the context of Arctic Amplification, the broader implications indeed relate to Polar Amplification. We will clarify this in the revised manuscript. Specifically, we state (**l. 43-46**): Specifically, there are no observations of the physiological processes underlying the growth and response of these species to environmental stressors such as modern Polar Amplification (the faster warming of the high latitudes relative to the global mean) (Serreze and Barry, 2011; Rantanen et al., 2022).

AW3. Do you have temperature data for *T. quinqueloba* to compare with *N. pachyderma* and *N. incompta*?

> AC: Thank you for raising this. *T. quinqueloba* occurs across a wide temperature range (1–29.5 °C), depending on genotype. We will add this information and the relevant in the revised manuscript. Specifically we added (**ll. 73-76**): *T. quinqueloba* is typically associated with mid to higher latitude ecosystems at temperatures between 1°C and 21°C but predominates at temperatures colder than 12°C (Bé and Tolderlund, 1971), however, it has been found on the Great Barrier Reef and in the Arabian Sea at temperatures as high as 29.5°C (Darling et al., 2000; Seears et al., 2012).

AW4. Define MQ water.

> AC: We added a definition of MQ water in **ll. 219-220** of the revised manuscript

AW5. There are 7 types of *N. pachyderma* in Morard et al. (2024) - are all the specimens in this study the same type? This should be stated.

> AC: We did not genotype the specimens in this study. However, Darling et al. (2004 and 2007) found only one genotype of *N. pachyderma* (e.g., Type I) in the subpolar North Atlantic/Arctic Ocean. Bird et al. (2025) confirmed that there is only one genotype of *N. pachyderma* in the subpolar North Atlantic/Arctic Ocean. Specifically, we state (**ll. 464-467):** We did not genotype the specimens in this study. However, Darling et al. (2004 and 2007) found only one genotype of *N. pachyderma* (e.g., Type I) in the subpolar North Atlantic/Arctic Ocean, which was also confirmed by Bird et al. (2025).

AW6. *T. quinqueloba* also has many genotypes.

> AC: Of course, there are two Arctic-associated genotypes of *T. quinqueloba* (Type IIa and Type IIb (Darling et al., 2000). We will note this in our revised manuscript (**ll. 484-486**): While we did not genotype *T. quinqueloba*, it is likely that our specimens belong to one of the two Arctic-associated

genotypes, with type IIa occurring in both the subpolar Arctic and subpolar Antarctic, and type IIb restricted to the subpolar Arctic (Darling et al., 2000).

AW7. "This demonstrates that larger planktonic foraminifera exhibit higher metabolic rates, even when respiration is normalised to account for temperature effects using both species-specific and uniform Q_{10} values (Table 4)." The referee noted "Love this result! Great implications across the entire field"

AC: We thank the referee for their positive feedback on this result. We are pleased that the implications of the size–metabolism relationship were clear.

AW8. Figure 8 panel (b): This is quite difficult to see, could the crosses be made larger and have the images been colour-blind friendly-checked?

AC: We have thoroughly revised all figures in the revised manuscript with symbols and colours that are colour-blind friendly.

AW9. "Consequently, our findings alleviate concerns about physiological confounding due to respiration. However, this may not be the case for N. incompta and *T. quinqueloba*. The elevated Q_{10} values for these two species may call for the development of species-specific calibration equations that consider the influence of respiration for accurate proxy application." The referee noted that "Genetic types may have different results too."

AC: We agree that genotype may influence proxy reliability in different species. In the Arctic and North Atlantic, only one genotype (Type I) is known for both *N. pachyderma* and *N. incompta*, whereas two genotypes (Type IIa and Type IIb) are reported for *T. quinqueloba*. We will add a clarifying sentence in the revised manuscript in section 4.3. specifically we added (**ll. 569-572**): Whether genotype could have implications for accurate proxy applications is dependent on the species with both *N. pachyderma* and *N. incompta* only having one genotype (Type I) found in the Arctic/North Atlantic while *T. quinqueloba* has two genotypes (Type IIa and Type IIb) (Darling et al., 2000; Darling et al., 2006).

Referee 2 – Major Comments

The manuscript entitled "The impact of essential climate variables on respiration rates in subpolar and polar planktonic foraminifera" by Armitage et al. reported the relationship between environmental parameters and respiration rates of polar and subpolar planktonic foraminifera species. Temperature effects were discussed in detail, and the authors showed that *Neogloboquadrina pachyderma*, a polar species that is often utilized for paleoenvironment reconstruction, has relatively stable respiration rates over a wide range of temperatures with low Q_{10}. This finding alleviates our concerns on potential respiration effect on foraminiferal test geochemistry with regard to *N. pachyderma*. They also conducted micro-Xray scanning to calculate biovolume more precisely, which allowed discussion on allometric scaling of respiration for generalization.

This study is important to gain our understanding of the basic metabolic activity of foraminifera under different temperature conditions, as well as ground the validity of species to be used in paleoenvironmental reconstruction. The manuscript is overall well-written, with detailed methods used and carefully discussed. However, I have several major concerns regarding the statistical treatment of the data and, in particular, the interpretation of the results. In several places, the analyses rely on limited datasets or assumptions that are not fully justified, and some conclusions appear to extend beyond what can be robustly supported by the data. I believe that addressing the points raised below—especially by reconsidering the statistical approaches and tempering some of the broader interpretations—would substantially strengthen the manuscript.

AC: We thank the reviewer for their positive assessment of the manuscript and for recognising the value of our study for understanding foraminiferal metabolic responses and their implications for paleoenvironmental reconstruction. We appreciate the constructive nature of the comments and fully acknowledge the concerns raised regarding statistical treatment and interpretation. We have carefully considered each point in detail and provide responses and planned revisions in the sections below. These changes will strengthen the manuscript and ensure that our interpretations remain well supported by the available data.

RC1. Statistical analysis on respiration rates and other parameters

The relationships between respiration rate and essential climate variables (ECVs) are evaluated primarily through separate pairwise correlation analyses (reporting r^2 and p values for each parameter, Table 3). While this approach may be useful as an exploratory analysis, it has important limitations that should be acknowledged. Many of the environmental variables considered (e.g., temperature, nutrients, salinity, DIC) are likely to be intercorrelated due to shared environmental gradients, such as water mass structure or seasonality. As a result, the reported correlations do not allow the independent effects of individual parameters on respiration to be disentangled. In addition, testing multiple environmental variables separately raises concerns about multiple comparisons, which may inflate the likelihood of detecting spurious significant relationships. I recommend either applying a multivariate framework (e.g., multiple regression or related approaches).

As I explain in the next part, correlation analysis for N. incompta needs to be reconsidered, since the datasets (based on 3 stations) cover a narrow range of each variable.

AC: Following the reviewer's suggestion, we conducted a suite of additional multivariate analyses to evaluate whether respiration responds to combined environmental gradients. As noted by the reviewer, many of the ECVs are likely to exhibit covariance due to shared hydrographic parameters. Specifically, we added **(ll. 353-376)**: However, strong covariance between ECVs in the modern ocean may obscure biological responses of *N. pachyderma* and *N. incompta* to complex environmental gradients. Consequently, we conducted a suite of multivariate analyses to evaluate whether size-normalised respiration rates respond to combined environmental gradients or individual variables. Principal component analysis (PCA) of the complete ECVs dataset relevant for each species revealed structured environmental axes, with PC1–PC4 capturing 95% of environmental variance for the *N. pachyderma* dataset and PC1–PC2 capturing 100% of variance for the *N. incompta* dataset (Supplementary Table S2). Evaluation of the PCA loadings indicates that these principal components correspond to coherent environmental gradients rather than isolated variables. For example, in the *N. pachyderma* dataset, PC1 primarily reflects a water-mass and nutrient gradient, characterized by strong positive loadings of nutrients (PO_4^{3-}, TON, SiO_2) and DIC and negative contributions from temperature and oxygen, consistent with colder nutrient-rich waters. PC2 is dominated by carbonate chemistry variables, including alkalinity, pH, and calcite saturation state (ΩCa), reflecting variability in the carbonate system. PC3 and PC4 capture secondary productivity-related signals, largely associated with fluorescence, oxygen, and temperature. Despite four clear multivariate gradients, size-normalised respiration rates for *N. pachyderma* do not show a significant relationship with any of the four principal components (all $p > 0.42$), and scatterplots of size-normalised respiration rates across PC1–PC4 space are not significantly different from zero (Supplementary Figure S1). Principal Component Regression explained 0% of respiration variance (adjusted r^2=0.005) with all PC slopes non-significant except PC4 ($p = 0.01$) but with negligible explanatory power (Table S4). Partial least squares regression similarly showed minimal predictive skill (Y-variance ≤ 22.7%; RMSEP comparable to the raw standard deviation; all 95% coefficient intervals overlapping zero; Table S5). Redundancy analysis detected a statistically significant constrained fraction ($p = 0.001$), but this accounted for only 37.4% of variance and did not improve performance relative to null models (Table S7). These results indicate that respiration remained invariant across the full multivariate environmental space sampled, supporting physiological robustness rather than the masking of univariate relationships by collinearity.

In the revised manuscript we also fully acknowledge that we have a more limited dataset for *N. incompta* Specifically we added **(ll. 377-385):** For *N. incompta* we have a more limited dataset measured over a narrower range in environmental gradients. Here we find significant linear correlations notable for temperature, SiO_2, salinity, fluorescence, dissolved O_2, pH and Ω_{Ca} (Table 3, Table S1). Furthermore, size-normalised respiration rates for *N. incompta* were significantly correlated to the environmental gradients represented by both PC1 and PC2 using Principal Component Analysis ($p<0.01$; see also supplementary Table S4 and Figure S2). PLSR results were consistent, with 75.25% of Y-variance explained and coefficient intervals not overlapping zero (Table S5). Redundancy analysis also identified a strong environmental signal (73.4% constrained variance; $p = 0.001$; Table S6). This suggests that unlike *N. pachyderma* size-normalised respiration rates in *N. incompta* are sensitive to integrated environmental structure. These contrasting responses suggest fundamental differences in metabolic plasticity between the two species.

RC2. For this species, respiration rates were measured at only three temperatures (10, 13, and 14 °C, in situ), covering a very narrow temperature range. In addition, the respiration rates recorded the highest at 13°C, and declined at 14°C. The calculation of Q10 based on the present data set appears problematic. Q10 assumes a monotonic, approximately exponential increase in metabolic rate over a sufficiently wide temperature range, under conditions where temperature is the primary limiting factor. Moreover, respiration peaked at 13 °C and declined at 14 °C, indicating a non-monotonic response and suggesting that the measurements may already span an optimal temperature or the onset of thermal stress. Under these conditions, the fundamental assumptions underlying Q10 are not met, and the resulting values are difficult to interpret physiologically. I therefore suggest either refraining from calculating Q10 or clearly stating that any estimated Q10 values are highly tentative and limited to a restricted temperature interval.

AC: We agree that the narrow temperature interval limit the interpretability of the Q_{10} estimate for *N. incompta*. We note however that the additional PCA analysis revealed a significant monotonic response of respiration to PC2 (e.g., representing temperature, salinity and alkalinity) which would support our interpretation and support calculation of preliminary Q10 for N incompta. Acknowledging the reviewers point though we will explicitly state that the Q_{10} value is preliminary and needs to be confirmed over a larger environmental gradient. Specifically, we state in the revised (track changes) manuscript (**ll. 393-396**): We caveat that Q_{10} values for *N. incompta* are only tentative due to the small temperature gradient (e.g., ΔT=4°C) covered in this dataset. Specifically, the assumptions required for calculating Q_{10} are not fully met for this species, and any resulting Q_{10} values should be interpreted with caution as a preliminary result that needs to be confirmed over a larger temperature gradient (e.g. ΔT=10°C).

RC3. Spinose vs non-spinose interpretation. The authors discuss the difference in Q_{10} values between *N. pachyderma* and *T. quinqueloba* relating the morphology and trophic mode of the species. It is true that *T. quinqueloba* is a spinose species, but this species is a "short-spined" species that has completely different ecology and physiology from typical spinose-species like *Globigerinoides*, *Globigerina*, *Globigerinella*, *Orbulina*, etc. Specifically, *T. quinqueloba* is not a carnivorous species, nor a symbiont-bearing species, nor an oligotrophic-adapted species. Presence of spines is an adaptation for planktonic lifestyle, but since the non-spinose Neogloboquadrina species also share the shallow habitat as is presented in the sample metadata (Table 1), morphological difference (presence or absence of spines) is not meaningful to explain the Q_{10} difference, I would say. As the authors noted at L443–445, it is true that non-spinose *N. incompta* showed relatively high Q10 (although it needs reconsideration as I pointed out above), which already collapses the validity of spinose/non-spinose comparison. I would say it's just species-specific.

AC: We thank the reviewer for this valuable clarification. We agree that our previous interpretation placed undue emphasis on the spinose/non-spinose distinction and may not accurately reflect the ecology of *T. quinqueloba*. Specifically, we state in the revised (track changes) manuscript (**ll. 486-496**): *T. quinqueloba* displayed a high Q_{10} value of 4.54 in our experiments, indicating a strong temperature dependence that may influence its metabolic performance, growth, and survival (Mundim et al. 2020). This contrasts with the lower Q_{10} of 1.41 observed in *N. pachyderma*, highlighting clear interspecific differences in thermal sensitivity. T. quinqueloba is a spinose species, however its ecology and physiology differ markedly from symbiont-bearing tropical taxa for which elevated Q_{10} values have been reported previously (Lombard et al. 2009). Although preliminary, the relatively high Q_{10} observed for the non-spinose *N. incompta* suggests that spine presence or absence does not reliably predict thermal sensitivity. Instead, we posit that temperature responses are species-specific, reflecting distinct ecological strategies rather than broad morphological categories. Such species-level differences likely contribute to the latitudinal partitioning of planktonic foraminiferal assemblages, with taxa occupying thermal niches that reflect their individual physiological tolerances (Bé & Tolderlund, 1971; Ying et al., 2023).

A brief note on the two Arctic genotypes (Type IIa and Type IIb) of *T. quinqueloba* will be included for ecological context in section 4.1. (**ll. 484-486**). While we did not genotype *T. quinqueloba*, it is likely that our specimens belong to one of the two Arctic-associated genotypes, with type IIa occurring in both the subpolar Arctic and subpolar Antarctic, and type IIb restricted to the subpolar Arctic (Darling et al., 2000).

RC4. Metabolic allometry and "crossover point". Representing the biovolume–respiration scaling relationship of planktonic foraminifera with data from other publications is interesting and potentially valuable. However, I don't fully understand the discussion on "crossover point" in L463–471. What exactly does the crossover point in Fig. 8 mean?

Moreover, I am concerned that the interpretation of the resulting scaling exponent may be overstated. The authors' statement that foraminiferal metabolism is somewhat "intermediate" between protists and metazoans more complex metazoans relies on cross-study comparisons that involve heterogeneous data sets, differing methodologies, and taxonomically broad groups. Given these uncertainties, the observed position of the foraminiferal scaling exponent relative to other organisms may reflect dataset composition or methodological differences rather than fundamental differences. I therefore suggest toning down this interpretation and framing it more explicitly as a hypothesis or conceptual possibility

> **AC:** We thank the reviewer for these helpful observations. We will clarify that the "crossover point" in Fig. 8 represents the intersection of fitted size-normalised respiration lines after temperature normalisation and does not imply a physiological threshold. Specifically, we state in (**ll. 529-533**): Comparison of respiration–biovolume relationships for *N. pachyderma* and *N. incompta* in Fig. 8a (4 °C) and Fig. 8b (15 °C) indicates that *N. incompta* surpasses *N. pachyderma* at the higher temperature of 15 °C, showing a crossover between 4 °C and 15 °C that reflects the point at which the relative size-normalised respiration of the two species changes under temperature normalisation
>
> We will also revise the interpretation of the metabolic scaling exponent and frame the comparison with protists and metazoans as a conceptual possibility rather than a firm conclusion, acknowledging dataset heterogeneity and methodological differences. Nevertheless, we note here that there are several studies regarding Kleiber´s Law that compare metabolic scaling between different taxonomically broad groups, including small prokaryotes, protists and metazoans (including larger animals, see also DeLong et al., (2010)). By necessity the methodologies to determine metabolic rates for different species are diverse but the trends are consistent. Furthermore, while each method might have limitations they are usually tested, internationally peer-reviewed, and well established. If we would not allow ourselves to compare the resulting rates, we would have to assume that a certain method is producing "wrong" or "inferior" results. Nevertheless, to address the reviewers' concerns, we state (**ll. 550-555**). The position of foraminiferal scaling within this framework should be interpreted cautiously, as cross-study comparisons involve heterogeneous datasets and methodologies. However, the pattern raises the possibility that planktonic foraminifera may exhibit metabolic behaviour intermediate between protists (which they are) and metazoans, potentially reflecting their large cell size that overlaps with the size of small metazoans, structural complexity, calcification, or ecological specialisation. This remains a hypothesis that warrants further targeted investigation.

RC5. Symbiotic ecology of *T. quinqueloba*.In Hemleben et al. (1989), it is indeed written that *T. quinqueloba* possesses symbionts, but no data are presented. Stangeew (2001) interpreted this species as symbiotic, based on the statement in Hemleben et al. (1989), which also does not show any evidence for the presence of symbionts on this species. Takagi et al. (2019) classified *T. quinqueloba* as a non-symbiotic species based on the absence of active chlorophyll fluorescence (photosynthetic activity). In this sense, "…their presence remains elusive (e.g., Takagi et al. 2019, ….)" is not appropriate. As far as I know, no positive data/evidence of the presence of symbionts for *T. quinqueloba* is available. Since the authors' observation is also in alignment with the absence of symbionts for *T. quinqueloba*, I think it's safe to say the specimens they used were non-symbiotic.

> **AC:** We thank the reviewer for this clarification. We agree that earlier references to symbionts in *T. quinqueloba* were unsupported and that more recent work (e.g., Takagi et al., 2019) demonstrates the absence of a systematic active chlorophyll fluorescence. Our own observations are consistent with a non-symbiotic ecology in the Arctic. We will revise the manuscript to clearly state that the specimens used in this study were non-symbiotic. Specifically, we state (**ll. 335-341**): While early studies by Hemleben et al. (1989) and Stangeew et al., (2001) noted the possible presence of symbionts in *T. quinqueloba*, these studies did not present direct evidence, and subsequent studies have found no

indication of active photosymbiosis (e.g., Takagi et al., 2019; Hoogakker et al. (2022); Kanbur, 2025). Consistent with this none of the specimens analysed in this study bore symbionts. Thus, calculated respiration rates were not for adjusted for photosynthesis as in Lombard et al. (2009) for *O. universa, G. ruber* and *G. siphonifera.*

RC6. Biovolume and empty final chambers. It is usually the case that the final chamber of collected foraminifera specimens is empty. In that case, biovolume estimation from the whole test would cause overestimation, since the final chamber generally has the largest volume. In this study, was this point considered? Since the experiments were conducted at different time points from collection (within 24 hrs without food supply for CE23011 and fed specimens within 11 days for 2024 samples), specimens conditions might have been different. Ideally, filled or not needs to be checked, and the biovolume needs to be corrected by excluding the empty chambers. If this is not possible, at least, please make remarks on the cytoplasm volume, that it is not always equal to the cavity volume. In Burke et al. (2025), 75% of cavity volume was applied as biovolume. This is an alternative way to take into account the void part.

AC: We thank the reviewer for raising this important point. In the revised manuscript, we will define the term "maximum biovolume" for clarity. Specifically, we state **(ll. 307-319):** In this study, we define the term "test biovolume" as the internal cavity volume bound by the calcite test, representing the space theoretically occupied by living cytoplasm in planktonic foraminifera. While cytoplasmic density and space occupancy may vary within chambers, this metric captures the maximum possible total living volume of the cell inside its shell rather than carbon-equivalent biomass. In our definition, test biovolume is thus confined to the interior volume of the test and does not capture the "catchment volume" or the rhizopodial network which can increase the effective cell volume by several orders of magnitude compared to the test alone (Gaskell et al., 2019). We choose not to scale reported biovolumes at 75 % of the test biovolume (as in Burke et al 2025) because the majority of specimens analysed in this study exhibited full chambers prior to measurements. Furthermore, we note that the suggested 75% method arose from the limitations of estimating biovolumes from total test volumes that included the shell (e.g. Hannah et al. 1994). Now, using micro-CT internal volume reconstructions we have more advanced tools to directly measure the cavity volume. Finally, there is no empirical evidence that supports a 75% cell occupancy of the cavity volume for asymbiotic planktonic foraminifera, nor does a validated technique exist that would allow us to measure it.

Therefore, when we compare our results, we use 100% of the cavity volume as biovolume and adjust all previously published datasets accordingly. This ensures a like-for-like comparison across studies. We will clarify this distinction in the revised manuscript. Specifically, we state in (**ll. 424-430**): Test bio volumes for *N. pachyderma, N. incompta* and *T. quinqueloba* were computed using the relationship between cavity volume and maximum diameter (Fig. 3) derived in this study. We compare our results to previous studies by adjusting biovolumes reported in Burke et al (2025) (e.g., 75% of cavity volume) to 100%. This ensures a like-for-like comparison across studies.

Minor Comments

RC: Hemleben et al. (1989) and Stangeew (2001) citation corrections.

AC: These corrections were made.

RC: Fig. 2 The illustration of the Unisense logos are confusing. Since it resembles planktonic forams (maybe the logo derives from forams), I thought, at first glance, the specimens are located in those boxes. Please delete the logo. In addition, the cable of the "calibration chamber" is not connected anywhere. Is this correct?

AC: We removed the Unisense logos and corrected the calibration chamber illustration in the revised figure.

RC: Fig. 5 Why is the y-axis for panel (a) (Temperature) alone on a log scale whereas the others are in linear scale?

AC: We appreciate the reviewer's comment. Temperature is well established to scale logarithmically with respiration rates (e.g., Burke et al., 2025; Lombard et al., 2009), whereas comparable empirical relationships do not exist for the other variables. We will clarify this in the revised manuscript (**ll. 342-**

347): The sensitivity of size-normalised respiration rates and temperature for tropical and subtropical planktonic foraminifera is best described by an exponential or Arrhenius relationship, rather than a simple linear one (Lombard et al., 2009). We therefore plot the log, size-normalised respiration rates against temperature for our analysis, as temperatures are well established to scale logarithmically with respiration rates (e.g. Burke et al., 2025; Lombard et al., 2009). Conversely, empirically supported functional relationships are not currently available for the other environmental variables measured in this study.