Anonymous Referee #1

In the manuscript by Wang et al., the authors claim that they have characterized the archaeal tetraether membrane lipids (GDGTs) produced by the marine euryarchaeota group II (MGII) and assess the effect of their synthesized lipids on the GDGT-based paleotemperature proxy TEX_{86}. For that purposes, they analyzed core lipids (CL) and intact polar lipids (IPL) GDGTs, they performed 454 pyrosequencing analyses to address the archaeal diversity and quantified the abundance of the MGII by quantitative PCR (qPCR) in suspended particulate matter (SPM) and surface sediments collected along the salinity gradient of the lower Pearl river and its estuary to the coastal South China Sea.

General comments:

I have great concerns about this manuscript. First, the authors do not characterize the GDGTs of the MGII but rather analyze the GDGT content of samples in which MGII is present (but far from being dominant). I do not say that MGII could not produce GDGTs, which is possible (maybe the authors are right that they synthesize GDGT-1, 2 and 3) but this study does not prove that at all (this could only be proven by cultures). I do have to recognize though that the authors based their assumptions in the change in the relative abundance of phospho IPLs (more labile) observed in the Pearl River Estuary (mixed water) site respect to the seawater location, which is the only way to connect lipid biomarkers to their sources avoiding potential preservation issues of CL GDGTs and IPL-GDGTs with glycolipid headgroups. Unfortunately, the samples chosen to make these assumptions are far from being the most ideal as in the best of the cases the MGII was only 30% of the total archaeal population. Even if the samples analyzed would contain 90% of MGII according to genetic analyses it wouldn't be possible to make such an assumption as we also have to take into account other factors such as differences in the extractions efficiencies in the pool of lipid and DNA extracted. This manuscript is also not novel if we compared it with the manuscript recently published by the same authors (Wang et al., Chemical Geology 2015), in which they already suggested that a change in the archaeal community composition (contribution by MGII) could be responsible for the differences in GDGT distribution and thus of the unusually low TEX_{86} values in the same system described by the current manuscript. Therefore, the only novelties in the present manuscript are the correlations between the different phosphor IPLs GDGTs and the ratio of MGII vs total archaea. However, these are mere correlations. For all we know the Thaumarchaeota Marine Group I (which makes more than 30% of the total archaeal population in the river estuary (mixing water station) could also be responsible for the change in the GDGT distribution detected in this sample (MGII are barely 17% of the total archaea in this sample as seen in Figure 2). It is even possible that the Thaumarchaeota population in the river estuary are phylogenetically different from the seawater station (which actually would make sense as the ecosystems are quite different) and synthesize GDGTs in different proportions accounting for the differences seen here. In any case, we can't conclude neither of these hypotheses with the experimental data provided in this manuscript. The tone of the title and abstract makes you assume that the contribution of MGII to the GDGT pool has been demonstrated in this study, which is far

from being the case and is very misleading. Besides, I have other concerns regarding the design of the experiment (especially regarding the molecular data) that I will list below, which discourage the publication of the manuscript in its current form.

We acknowledge that this reviewer's comments are important and thoughtful. We have done makeup experiments to quantify the abundances of MG-I (AOA) and total DNA from the same filter samples (Table 1 and Fig. 5). The qPCR date exhibited that MG-II was statistically higher (duplicate experiments) than MG-I in the mixing water and seawater, which suggests that MG-II predominantly occurred in the water column of sampling stations at the Pearl River Estuary and coastal South China Sea. These results yet seem to be different from the 454 sequencing data (Fig. 2), which showed that MG-I were dominated in both mixing water (estuary) station and seawater (coastal SCS) station. This might be due to the different amplification efficiencies for the two groups (97% for MG-II and 87% for MG-I). Since result from qPCR is more straightforward than 454 sequencing to reflect the abundance of archaeal 16S genes, it is reasonable to indicate that the samples chosen to make these comparisons are appropriate. On the other hand, linear regression analysis showed that there is no correlation between the ratio of MG-I/total Archaea (%MG-I) and the fractional abundance of GDGTs (%GDGTs) (data no shown); however, a significant correlation existed between the ratio of MG-II/total Archaea (%MG-II) and %ringed-GDGTs (Fig. 6), which suggests that MG-II may be a significant source of GDGT-1, -2 and -3 in the PR estuary and coastal SCS.

We do recognize the differences in the extraction efficiencies for lipids and DNA. Although the absolute quantification might be affected by the extraction method, the ratio of individual item to the total, such as the ratio of MG-II/Archaea and fractional abundance of GDGTs, could avoid systematic error and reflect the relative distribution of MG-II and GDGTs.

This study is indeed a follow-up work after Wang et al. (2015, Chemical Geology), but represents an important increment toward a better understanding of sources of GDGTs in marine environment. The two papers also have different focuses. In Wang et al. (2015), the main propose was to evaluate the factor(s) causing the unusually low TEX_{86} in the coastal area; MG-II was tentatively hypothesized to be a factor since this group of archaea did exist in the research region. But we didn't provide any evidence to support the hypothesis. Therefore, emphasis of this study was trying to explore MG-II-produced GDGTs and to further evaluate how these GDGTs influence the TEX_{86}. So it is an important and valuable step forward from the study of Wang et al. (2015). Since no MG-II culture exists to show the profile of its membrane lipid, a comparison between lipid and DNA would be the best way to evaluate the relationship between MG-II and GDGTs. In order to have a convincing comparison, we chose lipid data with only phosphate-head groups as they can represent living biomass. Despite the co-existence of MG-I and MG-II in the water column of the studying area, the linear relationship between %MG-II and %phospho IPL-GDGTs is able to at least suggest that MG-II (rather than MG-I) have the potential to produce ringed-GDGTs in situ in the water column of the study area.

The title and abstract have been improved to better reflect the revised content of this paper based on the reviewer's comment. In particular, the title is changed to "Evaluation of Marine Group-II *Euryarchaeota*'s contribution to cyclopentyl tetraethers in the Pearl River estuary and coastal South China Sea: Potential impact on the TEX_{86} paleothermometer".

Specific comments:

Abstract

Line 8: as mentioned above, in this manuscript you don't characterize the GDGTs produced by MGII.

It was changed to " we assessed the relationship between MG-II *Euryarchaeota* and GDGTs…".

Line 10: would be better to talk about 16S rRNA gene pyrosequencing rather than 454 (which is merely the machine).

Changed. Thank you.

Line 15: "MGII euryarchaeota as the second dominant group": fine, MGII are the second dominant group (*17% of the total) but Thaumarchaeota make up more than 30% of the reads (as seen in Figure 2 (PR estuary mixed water). This sentence (and the whole study) is biased towards what the authors want to demonstrate but the rest of the archaeal community (which we know they make GDGTs) are excluded from the conclusions!

Thank you for the comment. Based on the qPCR results, the abundance of MG-II was significantly higher than MG-I (avg. 23% for MG-II & < 1% for MG-I in the PR estuary mixing water; see details in Table 1). So this sentence is changed to "with MG-II Euryarchaeota being one of the dominant groups of archaea in the mixing water and seawater stations".

Line 16: "qPCR data indicated that the abundance of MGII euryarchaeota in the mixing water was three to four orders of magnitude higher than in the river water and in the seawater": Yes, this is correct, but still taking the data of the qPCR analysis in Table 1, MGII range from 0.2-30% of the total archaeal population in the mixing water. Why is not the impact of other "more dominant" archaeal populations in this sample being discussed here or in the rest of the manuscript?

The identified groups of Archaea with notable proportions are MG-I, MG-II, Methanogens, MBGB, MCG and YLA114. According to the above analysis, MG-I was not the significant GDGT-producers in the studying area (no relationship between GDGT and DNA). Methanogens produce GDGTs without rings. MBGB, MCG and YLA 114 kept a significantly smaller abundance than MG-I & II in this area. We cannot exclude the possible contribution of GDGTs from other groups except for MG-II; however, the linear

relationship between phospho IPL-GDGTs and MG-II DNA data is able to strongly indicate a significant contribution of ringed GDGTs from MG-II.

Line 18-line 22: For the reasons mentioned above I strongly disagree to this statement: the existence of correlation does not does not suggest that MGII produce GDGT in the water column. . .

The original sentence has been changed to "Significant linear correlations were observed between the gene abundance ratio of MG-II *Euryarchaeota* vs. total archaea and the fractional abundance (%) of GDGTs-1, -2, -3, or -4 as well as the ring index based on these compounds, whereas no relationship was observed between the ratio of MG-I/total archaea and %GDGTs, which collectively suggest that MG-II *Euryarchaeota* may actively produce GDGTs in the water column."

Introduction

Lines 20-21: "However the lack of direct link between archaeal lipids and DNA prevented the drawing of a more concrete conclusion". . ..which the current manuscript has not been able to address neither Material and Methods.

The materials (filters with dominant group of MG-II) and methods (total phospho IPL-GDGTs) have the potential to address this issue. Please see the general response above for more details.

Line 25 (page 5): The range of liters filtered is quite broad. It is essential that the authors report the total GDGT and DNA content that was extracted from these filters. Otherwise it's impossible to asses if enough material was extracted and analyzed.

The total GDGTs and DNA content are listed in Table 1. The results suggest that the materials (lipid and DNA) extracted from the samples are enough.

Line 26 (page 5): The filters used were GF/F 0.7 um. This is always an issue for this kind of studies as we don't know if the archaeal population is biased by the diameter of the filter pore. MGII have been seen to be prevalent in particles (Galand et al., 2010) and genome analyses suggest that they have a particle-attached lifestyle (Iversson et al., 2012). Considering this, the 0.7um could potentially select for MGII rather than Thaumarchaeota and completely invalidating the results. The authors cannot assess this point with the data presented here but at least they should account for this possibility.

Thank you for the comments. During the cruise, we also collected 0.2 μm filters. In preparation of this study, this issue has already been considered. However, the whole 0.2-μm filters have already been extracted for lipids and no extra samples left for DNA quantification (unpublished data and a manuscript prepared including these data). So we could not estimate the amount of MG-I 16S gene that went through 0.7 μm filter but collected by the 0.2 μm filter. However, based on a comparison experiment showing different yields of GDGTs from 0.2 μm and 0.7 μm filters, we noticed that the 0.7 μm filter collected significant abundance of GDGTs with only phosphate heard groups

(phospho IPL- GDGTs; 90 ~ 95%), which is well accepted as a biomarker to represent living organisms. This results indicate that the 0.7 μm filter may have collected significant amount of source organisms of the GDGTs. Therefore, based on the lipid results, it is convincible and reasonable to use 0.7 μm filter evaluating the relationship between lipids and DNA.

Lines 8-9 (page 8): I am puzzled with the idea that the authors extracting the DNA contained in the filters by washing frozen filters 3x with PBS filter and centrifuge the supernatants to continue with the DNA extraction. This is insufficient. We regularly extract DNA from glass fiber filters and the DNA is way more attached to the filter than in the case of polycarbonate filters therefore a bed beating step in lysis buffer is essential to get the DNA from the cells (needless to say that this is extremely important for a proper extraction of DNA from archaeal cells). I just can't imagine that you can get representative DNA extracts by washing frozen filters. Besides, the range of extracted DNA is not provided anywhere (not even in the supplementary material), I would be curious to see how much you managed you extract.

Thanks for your comments. The protocol of FastDNA SPIN Kit has a bead beating step in lysis buffer. We agree with the comments that the DNA extractions for the filters are not sufficient. That's the reason why we didn't make the comparison based on the absolute abundance of DNA and lipids. However, since the DNA extraction efficiencies for different groups of Archaea are hypothetically identical, the ratio of MGII/Archaea could reflect the relative variation of MGII along the Pearl River and its estuary.

The total abundance of DNA is listed in Table 1.

Line 18 (page 8): Where the qPCR conditions tested by the authors or previously tested? If these primers have not been tested before the authors should demonstrate with supplementary data how specific these qPCR reactions are (especially the one for the MGII). Besides, no efficiency nor R2 values of the qPCR assays are provided.

The primers targeting on archaeal 16S (Bano et al., 2004), MGII 16S (Massana et al., 1997; Teira et al., 2004) and archaeal amoA gene (Francis et al., 2005) are proven to be specific in published research.

We add the following paragraph into the manuscript section 2.3.1 –

PCR targeting the different genes were conducted before the qPCR. The PCR bands were recovered by Gel Extraction Kit (omega) and sequenced on the 3730 sequencing platform. The sequences were annotated as the corresponding target genes, which demonstrated the specificity of those qPCR primers. The melting curve analysis was performed to demonstrate that the fluorescent signal obtained in a given reaction was consistent with the expected profile for specific PCR products on the basis of comparison with standards. The R^2 values for standard curve were above 0.99. The efficiency of each qPCR was between 87% and 99%.

Line 21 (page 8): 454 sequencing (as mentioned above would be better to say 16S rRNA gene pyrosequencing), was only done in n=3 SPM samples, no replicates. Dangerous to make such assumptions based in such limited dataset. Also the authors don't report the number of sequences that were recovered from each sample. These should be comparable to make proper comparisons between the samples as seen in Figure 2.

The 454 sequencing data was to show the distribution of the archaeal community composition. It is unfortunate that we did not have replicates in this study. However, the result showed in this study (Fig. 2) was similar to the distribution of archaeal community composition in Wang et al. (2015). The two data sets were collected in the same study area, but different times (samples in this study were collected in 2011; samples in Wang et al., 2015 were from 2010). However, the qPCR data were based on samples from three water columns, i.e. fresh water samples (n=2), mixing water samples (n = 6) and seawater samples (n = 4), and sediments (n =3). The major observation in this study is based upon the regression analysis between lipid data and qPCR data (rather than pyrosequencing data). Therefore, it is reasonable to make the conclusion based on our sample set.

As for the number of sequences, a total of 9,343 effective sequences with an average length of 531 bp were generated. Sequence numbers are 2,751, 2,987 and 3,695 for river water, mixing water and seawater samples, respectively. This information was added into the caption notes of Figure 2.

Line 5-8 (page 9): The taxonomy assignation of archaeal 16S rRNA gene reads can be problematic depending on the classifier used. It is recommended that the authors provide further prove of the identity of the archaeal sequences (such a phylogenetic tree of representative sequences). Results and Discussion.

Thanks for your suggestion. The primary Figure 2 showing the distribution of archaeal community composition was replaced by the phylogenetic tree based on the dominated OTUs (Fig. 2). Although the taxonomic classify showed that they had similar archaeal composition, we further found that the dominated OTUs in the river water, mixing zone and seawater were significantly different based on the phylogenetic tree. Comparing with the MG-I (only one dominated OTU), the MG-II are more diverse in both mixing zone (5 dominated OTUs) and sea water (4 dominated OTUs). This suggests that the archaeal composition may account for the variation of GDGTs in different environments.

Lines 1 (page 14): ". . .were produced in situ in the PR estuary by the source microorganisms": Which microorganisms? According to Figure 2 only 17% of the sequences are affiliated to MGII and more than 30% to Thaumarchaeota so the GDGT in situ production could also well be attributed to MGI, right?

We deleted the term "by the source microorganisms". According to the qPCR data in Table 1, the abundance of MG-II is significantly higher than MG-I in the mixing water (avg. 23% MG-II vs. <1% MG-I) and seawater (avg. 8.7% MG-II vs. 0.4 % MG-I) station. On the other hand, compared with the linear relationship between %GDGTs and the ratio of MGII/Archaea, there is no relationship exhibited between %GDGTs and the

ratio of MGI/Archaea. These results suggest that 1) MG-I may not be a significant source of the in situ produced lipids in the studying area, 2) the source organism is more likely to be MG-II.

Lines 12-24 (page 15): The increased ratio of GDGT-2/3 ratio in deep water column responsible to the warm bias of TEX_{86}-derived temperature has been recently suggested to be related to differences in the GDGT produced by deep water Thaumarchaeota MGI (Villanueva et al., Environmental Microbiology in press doi: 10.1111/1462-2920.12508). As this paragraph is phrased it seems that the authors suggest that the GDGT-2/3 ratio variation in deep waters could be attributed to MGII as suggested for the authors in this study. Rewrite to make this part clearer.

The change was made.

We greatly appreciate the valuable and insightful comments made by this reviewer.