

## ***Interactive comment on “DNA from lake sediments reveals the long-term dynamics and diversity of *Synechococcus* assemblages” by I. Domaizon et al.***

**I. Domaizon et al.**

isabelle.domaizon@thonon.inra.fr

Received and published: 27 April 2013

We thank the reviewer for the valuable comments that improve the original manuscript. We provide thereafter a point by point answer to all the comments/suggestions made by the reviewer. We are able to provide a revised version of the manuscript that includes all these suggestions and corrections.

1. Information about the length of PCR products : Reviewer : Please provide information on the length of the different PCR products amplified (for qPCR and PCR) (maybe you could use a little illustration, showing where the primers bind and the approx. length of the different fragments).

C1294

Reply: We now provide additional information about the length of the different PCR products amplified. We prepared a table (table 1) and an associated figure (figure 1) that describe the amplicons' size, and target positions of PCR and qPCR primers and probes in the ribosomal operon of *Synechococcus* spp.

Legend of the New Table 1. (see table in Fig1 attached) will be : Size of amplicons, oligonucleotide primers and TaqMan probes used for PCR and qPCR to target total cyanobacteria and *Synechococcus* populations. F= forward, R=reverse, Tp = Taq man probe, R=reporter, Q= quencher. Target positions of PCR and qPCR primers and probes in the ribosomal operon of *Synechococcus* spp. are illustrated: the arrows indicate the 5' to 3' orientation of the oligonucleotides.

2. Suggestion : Add more details about the previous test performed to assess the level of DNA preservation : Reviewer : For the reader it would be even more convincing (I mean the authenticity of the sequences) if you could give a more detailed description of the previous tests that were performed to assess the level of DNA preservation (p. 2524, l. 17-19) - do you mean the qPCRs or any additional methods?

Reply: In the revised version, indeed, we can provide additional information about the tests that were performed to assess the level of DNA preservation. The tests were based mainly on the comparison of sequencing results (composition of cyanobacterial assemblages) obtained by targeting different size of amplicons : from qPCR amplicons (less than 100 bp length) to long fragments (~1500bp obtained from PCR reactions) by targeting 16SrRNA and ITS1 region; we also considered medium size fragments obtained from the amplification of plastidial gene (16S rRNA amplified according to the protocol described by Shi et al 2011 (~800 bp with primers PLA491F/OXY1313R which are adapted to target not only cyanobacteria but all photoautotrophs). Whatever the length of sequenced amplicons was, we obtained rather similar composition for cyanobacterial assemblages (in terms of proportion of putative genera identified from the sequencing results), and more particularly the amplification of long DNA fragments did not lead to any loss in terms of richness in comparison to short fragments.

C1295

These data showed that cyanobacterial DNA was well preserved in these sediment layers (we did perform such comparisons for the layers BF1, BF3, BF10, BF24, see table below); *Synechococcus* DNA was largely represented whatever was the length of targeted DNA fragments. Since we were particularly interested in *Synechococcus* assemblages we decided to work on the larger fragments in order to provide robust phylogenetic analyses (which were not so efficient with very short DNA fragments). The global results of these tests could be summarized in a table showing the proportion of sequences affiliated to the main cyanobacterial genera (assignment to putative taxa from BLASTN with reference sequences available in public databases and in personal databases). This table could be added as supporting information in the revised manuscript if it is useful.

In the original version, the information regarding these test appeared P2536 L 6-9 in brackets; since the reviewer thinks it could be an important point, we propose to make a full sentence with little more details to replace the initial sentence: "PCR amplification of several sizes of fragments were tested, the obtained amplicons were submitted to sequencing in order to compare the picture of taxonomic compositions provided by sequencing short (~ 150 bp) and long (up to 1500 bp) amplicons originating from the same sediment layers. The results showed that whatever the length of sequenced amplicons was, we obtained rather similar composition for cyanobacterial assemblages (in terms of proportion of putative genera), and more particularly the amplification of long DNA fragments did not led to any lost of richness (number of cyanobacterial genera detected) in comparison to short fragments. The high level of DNA preservation was thus confirmed in these sediment layers and notably for *Synechococcus* DNA that dominates all the libraries of clones." To be added in revised version P2536 L6-9.

In supplemental information we propose to add the results for qPCR and PCR assays that are presented in the MS : The legend for these Sup data (Data S1. see Figure 2 attached) will be : Comparison of cyanobacterial assemblage' pictures obtained from sequencing of different amplicons that varied in length (from ~100bp, to ~1500bp from

C1296

16s gene+ITS1 region). Results are expressed as proportion of sequences affiliated to the main genera (identified from sedimentary DNA originating from 3 sediment layers (BF1, BF10, BF24), by assignment by BLASTN with identity >95%)

3.Possibility of using FST values to look for genetic differentiation between the different sediment layers: Reviewer : To estimate the levels of genetic variation within and between the defined 16SrRNA OTUs obtained from the different sediment layers - I would suggest calculating pairwise FST values to look for genetic differentiation between the different sediment layers you analysed (eg using Arlequin). Although you are not looking at real populations of *Syn.*, Fst values and the net nucleotide divergence (comparing all pairs of sedimentlayers analysed) would give a good estimate of the genetic differentiation between the different sediment layers and these results could be compared to the UniFrac analysis. Although you have identified the highest diversity in the ITS Fragment, I miss some phylogenetically/statistical analysis for ITS, like the ones which you applied to the 16Sr-RNA data set. This genetic marker might offer the possibility to show higher genetic differentiation between your genetic lineages. The results of pairwise Fst values for this data set is possibly also very interesting. Could you give an explanation why you did not perform any tests with ITS sequences?

Reply: We initially thought that the discrimination of ITS types (in addition to phylogenetic analyses of 16S OTUs) was sufficient to highlight the genetic variation that could be detected within *Synechococcus* assemblages retrieved for different periods of the lake history. Additional analyses on ITS and 16S are of course possible. As suggested by the reviewer, we performed pairwise genetic differentiation analyses for both 16S and ITS data. It allows providing additional analyses/information on ITS region.

Consequently this additional information could be added in the revised version. We propose to add a sentence at the end of material & methods section P2526 L29 (end of §entitled : "Post sequencing and phylogenetic analyses"): "To estimate the genetic differentiation of *Synechococcus* (from 16SrRNA OTUs and from ITS types) between different sediment layers we calculated Nei's Gst coefficient of differentiation (equiv-

C1297

alent to Wright's  $F_{st}$ ; Nei, 1987) using Mega5 software (Tamura et al., 2011). We calculated each pairwise  $G_{st}$  values to look for genetic differentiation between the 8 different analyzed layers (BF1, BF 3, BF10,12, BF 15, BF 20, BF 24, BF 27). Values obtained were then used in XLSTAT (Addinsoft 2008) to configure coordinates for a multidimensional scaling (MDS) approach."

A Figure could be added in the main body of the manuscript or in the supplemental data (Data S3) (see figure 3 attached). Comments regarding this figure will be included in the results' section P2534 from L23: Values obtained for genetic differentiation were higher for ITS region (from 0.004 (BF3-BF1) to 0.334 (BF15-BF24)) than they were for 16s rRNA (from 0.001 (BF20-BF1) to 0.089 (BF24-BF20)). The results obtained from these genetic variation estimates (both for 16s and ITS) globally tend to confirm the results obtained from UNIFRAC analysis. The periods of strong eutrophication (1981-1983 and 1972-1973) are clearly discriminated from other grouped periods (i.e. 2001-2003, 1951-1952). The main difference with Unifrac results is the good similarity between the layer BF1 (2008-2009) and BF3 (2001-2003) for which low and non significant GST values are estimated for both 16sgene and ITS region.

4.Suggestion regarding genetic differentiation linked to other environmental factors than temperature: Reviewer : In your study you describe summer air temperature as the driving factor for the abundance of Cyanobacteria and especially Syn., but the analysis of Syn. strains (Unifrac analysis) indicates also that possibly phylogenetically related groups of Syn. Occurred at different trophic periods during the lakes's history, which would mean that although the trophic level of a lake has a minor influence on the presence of Syn. , that the genetic diversity suggests a correlation between similar genotypes at a certain trophic state of the lakes. Have you thought of something like this? I think this result is very interesting as it shows that beyond the pure proof of Cyanobacteria /Syn.DNA through time (and its changes correlated to temperature) also genetic differentiation within specific groups of Syn. might be linked to other environmental factors which have not taken into consideration before. If you agree, maybe

C1298

you could state on this in the manuscript.

Reply: We agree and we could add in the revised version a sentence to introduce this idea/comment, moreover the results obtained for genetic differentiation contribute to support this idea. "Beyond the changes in Synechococcus proportion that are correlated to temperature, our results suggest a genetic differentiation within Synechococcus assemblage that might be linked to trophic status (similar genotypes at a certain trophic state) and possibly to other environmental factors which have not taken into consideration." this sentence could be included in the revised version P2541 L15 (the initial sentence is consequently separated in two parts).

5.Indicate on fig. 1 the major environmental events that influenced Lake Bourget over the last 100 years.

Reply:In the revised version we indicate on the figure 1 (the initial figure 1 will be separated in two figures according to reviewer 2 request) the major environmental events that influenced Lake Bourget over the last 100 years (i.e. period of eutrophication, re\_oligotrophication, etc).

6.Suggestion for Figure 3 (figure 3 of the online MS): Reviewer: This figure has very much information, but this information is partly not explained (e.g. the groups which sometimes have numbers or letters and the description of clades which is not really visible anymore). I would suggest to try giving all the clusters /clades informative names like, e.g. subalpine (if that is possible) and to give a short explanation in the caption (although you mention some of the formerly defined groups in the text, it would be helpful to see them in the figure caption once more).

Reply:For the Figure 3, we added description in the caption, and gave all the clusters /clades informative names like, as well as a short explanation in the caption. The revised version is prepared with this modification.

7.Remark regarding the single sequence types : Reviewer: Some OTUs (e.g. OTUb

C1299

16,17,18) only occur as single sequence types in just one sediment layer and are not unambiguously associated to another reference, how do you explain their authenticity?

Reply: Some OTUs occur as single sequence types in just one sediment layer and are not unambiguously associated to another reference (i.e. OTU5, OTU6). We have indeed to be very careful in considering these sequences, we cannot be absolutely sure that the sequences actually corresponded to a past specific taxa, even though the global preservation of DNA is obviously very good, some degradation of DNA due to diagenesis process exists and also some bias in sequencing. Thus, it's hard to be sure of the reliability of these single sequences, consequently we added a sentence in the results and in the caption of figure 3 (now figure 4) to highlight this point => sentence added in revised version: "Some of these OTUs (i.e. OTUs 5, 6, 16, 17, 18) might be considered with caution since they occur as single sequence types in just one sediment layer and are not unambiguously associated to another reference. " (to be added at the end of §3.4 p2533 L23).

8. Suggestion for Figure 3 (figure 3 of the online MS) : remove the diversity estimates and presented them in another table  
Reviewer: I would also remove the diversity estimates and would present them in another table, which should also indicate the ranges of these estimations.

Reply: In the revised version of the MS, we could easily remove the diversity estimates and present them in another table => new table=table 3

9. Suggestion related to Figure S1 (online MS) Reviewer: the rarefaction curves (except number BF24) do not really show that the diversity was fully covered, as all the curves do not reach saturation. You should indicate this in the text and consider that the real diversity is probably higher but the genetic approach (different DNA preservation, specificity of the primers, cloning etc.) influences the obtained species diversity in your data set.

Reply: As suggested by the reviewer, we indicated in the text that the rarefaction curves

C1300

showed that the diversity was not fully covered, and consequently that the real diversity is probably higher but the genetic approach (different DNA preservation, specificity of the primers, cloning etc.) influences the obtained species diversity in our data set. A new sentence : "Rarefaction curve analyses performed for *Synechococcus* sequences showed a rather good coverage of the richness especially for the layer BF24, however, the richness was not fully covered and we assume that the real diversity is probably higher (Supplementary material, Figure S1)" replace the initial sentence p2531 L26-27).

Minor comments

10.p. 2522, l. 24, and p. 2523, l. 7, we indicated the approx. length of the PCR product, by referring to the table 1 (newly produced in this revised version) "Table 1 describes the amplicons' size and target positions of qPCR primers and probes used in this study to target cyanobacteria or specifically *Synechococcus* spp." was added for the revised version p2531 L21).

11.p. 2528, l. 26, the reviewer asked us to give another sentence to explain the UniFrac method, since our explanation seemed incomplete; accordingly we corrected the sentence and we provided a more accurate definition of UniFrac method: "Unweighted UniFrac measures the distance between two communities by calculating the fraction of the branch length in a phylogenetic tree that leads to descendants in either, but not both, of the two communities." this sentence replace the initial sentence in §2.5.2., p2528 L23).

12.p. 2527, l. 4, we gave the full name for TCC i.e. Thonon Culture Collection

13.p. 2529, l. 7, we added a hyphen is missing (Giguët-Covex)

14. Table 2 (which is now Table 4): we indicate at the bottom of the table the total number of clones identified per time interval (=sediment layer) as suggested by the reviewer.

C1301

15. Figure 1: we titled the first column of the figure with e.g. years AD as suggested by the reviewer

16. Question from the reviewer regarding Figure 3 (online MS): Could it be that you indicated the wrong clade with "Cyanobium gracile", because this clade does not include any Cyanobium sequences? If I am wrong, please indicate why that cluster is named like this? Reply. Thanks for highlighting this error, indeed, this group A does not include any Cyanobium sequences, we corrected on the phylogenetic tree.

17. Figure 4 (which is now Figure 5): We indicated the 'BF abbreviations' for the different time intervals in the caption as suggested by the reviewer.

18. Table S1. Here, reviewer 1 proposed to include the Fst values. We have prepared a figure in order to show the results obtained for these values both for 16S and ITS (see comment/reply 3) thus we suggest not to modify the table S1.

Interactive comment on Biogeosciences Discuss., 10, 2515, 2013.

C1302

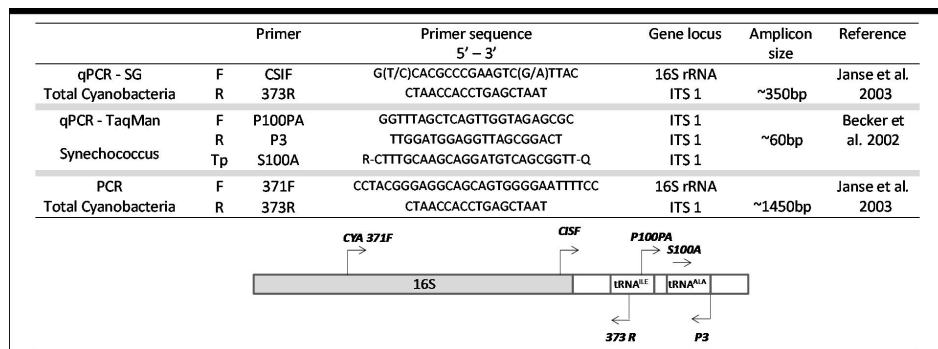


Fig. 1.

C1303

Sediment layer	origin of amplicons	Chroococcales				Nostocales		Oscillatoriales	
		Synechococcus	Microcystis	Chroococcus	others	Anabaena	Nostoc	Planktothrix	Leptolyngbia
BF1	16S ITS (< 100bp)	50	6	6	13			19	6
	16S ITS (~1500 bp)	72	10	4	1			10	1
BF10	16S ITS (< 100bp)	65	6		6	12	6		6
	16S ITS (~1500 bp)	83	7	4	3	3			
BF24	16S ITS (< 100bp)	73	7				13		7
	16S ITS (~1500 bp)	93	2	2	1		1		2

Fig. 2.

C1304

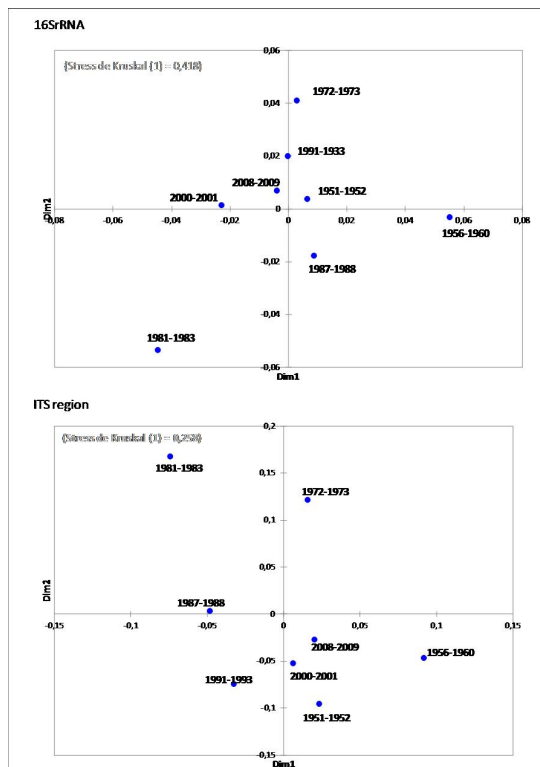


Fig. 3.

C1305