DOI 10.18699/SBB-2022-297

Elusive recurrent bacterial contamination in a diatom culture: a case study

Morozov A.*, Galachyants Y., Marchenkov A., Petrova D., Zakharova Y.
Limnological Institute, SB RAS, Irkutsk, Russia
** morozov@lin.irk.ru*

Key words: diatoms, axenic cultures, contamination, 16S rRNA amplicon sequencing, *Sphingomonas, Fragilaria*

Motivation and Aim: Development of axenic cultures remains a crucial step for genomic studies of diatom algae. However, the genome our group has produced from supposedly axenic cultures of *Fragilaria radians* [1] was later found to contain several megabases of *Sphingomonas* sp. genetic material, suggesting contamination with this bacterium [2]. Thus, the goal of this work was to identify the source of this contamination, select non-contaminated libraries (if any) and propose ways to identify similar contamination in the future.
Methods and Algorithms: To identify the contaminated libraries, all available genomic and transcriptomic reads were mapped to the assembled genome, estimating the number of reads mapping to *Sphingomonas*-derived scaffolds. To study the contamination of the cultures themselves, the available living cultures was stained with DAPI; total DNA was also extracted from both living cultures and older frozen material, and used for 16S rRNA sequencing on Illumina MiSeq. 16S reads (both from amplicon sequencing runs and genomic libraries) were analyzed in Mothur.
Results: The original reads have come from two different *F. radians* strains and a number of sequencing platforms (454, Illumina HiSeq, PacBio). Only a subset of 454 libraries, all produced at LIN SB RAS facilities in 2011–2012, had a non-negligible number of reads mapping to *Sphingomonas* scaffolds, suggesting that contamination with bacterial genomic DNA was limited to this series of experiments. Another 454 library, also from 2011, contained a large amount of bacterial 16S rRNA but very few *Sphingomonas*-derived reads. Analysis of reads mapped to 16S reference alignment shows that they mostly come from a short region within 16S rRNA gene, suggesting that this library was independently contaminated with some metabarcoding amplicon at the sequencing facility.
Analysis of 16S rRNA amplicons from modern cultures and frozen material from 2017 suggests that the former are clean, while the latter is yet again contaminated with *Sphingomonas* sp. For the context, a number of clean libraries was produced from the same strain between 2012 and 2017; phylogenetic analysis also shows that two distinct strains were present in 2011–2012 and 2017, implying two independent contamination events. During all this time, no bacteria were detected using light microscopy with DAPI staining.
Conclusion: The produced results suggest at least two independent contaminations with *Sphingomonas* sp., most likely during the high-volume biomass production stage. Since light microscopy has failed to detect contamination in either case, we recommend supplementing it with 16S rRNA amplicon sequencing. Although some bacterial OTUs are detected even in the axenic eukaryotic cultures, they are very numerous and have

very low abundance each (typically in single reads or tens of reads at most). In contrast, contaminated samples contain a single high-abundance OTU or a number of closely related OTUs comprising the absolute majority of non-organellar 16S reads.

Acknowledgements: This research was funded by Ministry of Science and Higher Education of Russian Federation project 0279–20201–0009 (121032300191–3) "Analysis of the role of baikalian diatoms' genes and proteins via bionformatics and physico-chemical biology methods". Authors are grateful to Emmelien Vancaester and their colleagues at Ghent University for identifying the contamination in published genomic data and sharing their results.

References

1. Galachyants Y.P., Zakharova Y.R., Volokitina N.A., Morozov A.A., Likhoshway Y.V., Grachev M.A. De novo transcriptome assembly and analysis of the freshwater araphid diatom Fragilaria radians, Lake Baikal. *Sci Data*. 2019;6(1):183. doi: 10.1038/s41597-019-0191-6.
2. Vancaester E., Depuydt T., Osuna-Cruz C.M., Vandepoele K. Comprehensive and functional analysis of horizontal gene transfer events in diatoms. *Mol Biol Evol*. 2020;37(11):3243-3257. doi: 10.1093/molbev/msaa182.