Exploring the Impact of Climate Change and Increased CO_2 Levels on the Cyanobacteria, *Microcystis aeruginosa*

Megan Hept and Lesley Greene

Old Dominion University
Department of Chemistry and Biochemistry
Norfolk, Virginia 23529
Email: mhept@odu.edu

Abstract

Cyanobacteria are important organisms, instrumental in several key geochemical cycles and found in a wide range of environments all over the world. Some cyanobacteria produce cyanotoxins and a global health concern are cyanobacterial blooms that produce these types of toxins. *Microcystis aeruginosa* (*M. aeruginosa*) can be found in freshwater systems such as lakes and can produce microcystin, a hepatotoxin that primarily targets the liver. This cyanobacterium is predicted to rise in prevalence due to climate change, particularly increases in atmospheric CO_2. Our research had two primary objectives: (1) to complete the genome of a Lake Erie isolate of *M. aeruginosa* and conduct phylogenomic analysis, and (2) raise cultures of this isolate for five-months at present day and increased atmospheric CO_2 levels predicted for the year 2100 to identify adaptive genomic changes. The genome assembly and phylogenomic analysis was completed and provided further insight into this species. The results of the five-month study produced several key findings. One was the discovery of a novel species of gram-negative bacteria, *Lacibacter,* co-existing in culture with the *M. aeruginosa*. The other revealed that after five months there are a few discrete genetic changes seen in the cultures raised in high CO_2 for *M. aeruginosa.*

Introduction

Cyanobacterial blooms have the potential to cause harm in ecosystems worldwide, with side effects seen in blooms carrying toxin.[1] The blooms can cause problems for water quality and outcompete beneficial phytoplankton, creating issues from an environmental, human and animal health perspective.[2] One of the most well-known and problematic organisms is *Microcystis aeruginosa (M. aeruginosa)*, a freshwater cyanobacteria that is found in temperate and subtropical lakes, estuaries, and rivers.[3] This organism has been found in over 108 countries that span over six continents, and its range has been shown to be increasing in previously unaffected waterways. [4] A concern for its overgrowth is due to its ability to produce microcystin, a monocyclic heptapeptide, that is considered to be one of the most common and toxic cyanotoxins produced.[5]

While this cyanobacterium can have adverse effects on ecosystems, there is an incomplete understanding into its taxonomy and evolutionary relationships.[6] Current taxonomy typically relies on morphology, but with cyanobacteria this has been shown to be unreliable, and there is an important need to acquire genetic data and complete genomes.[7-8] One aim of this project is to acquire a completed genome of *M. aeruginosa* strain LE3, an isolate from Lake Erie in the United States, that has previously been sequenced but is in contig form.[9] *M. aeruginosa* strain LE3 has been shown to produce microcystin.[9] This cyanotoxin

primarily acts as hepatotoxin that generally targets the liver and can cause cellular damage and genotoxicity as well as other health issues.[10-11] Microcystins being one of the more common cyanotoxins that are seen worldwide are a leading problem in freshwater lakes, especially Lake Erie in North America.[12] Lake Erie is a part of the North American Laurentian Great Lakes, which represent around 20% of the available surface freshwater worldwide and 84% in the USA, and is predicted to see rises in toxic algal blooms occurrences.[13] Health in human and animal populations is not the only concern with these algal blooms as there can be serious ecological and economic impacts. Estimated costs yearly can accumulate to hundreds of millions of dollars as blooms can decrease tourism, recreation, and industry.[14]

A variety of different variables are predicted to be playing a role in domain expanse seen by *M. aeruginosa*. It is known that cyanobacteria thrive in eutrophic lakes that are dense with anthropogenic additions of nutrients, like nitrogen and phosphorous.[15] It has been estimated that 50% of the lakes in North America, Europe, and Asia are eutrophic, making cyanobacterial blooms more likely and potentially more dangerous.[16] It has also been shown that rises in temperature appear to favor *M. aeruginosa* growth and could potentially increase the level of toxin production levels.[17] Atmospheric carbon levels are predicted to rise, and *M. aeruginosa* blooms are expected to thrive in these conditions and become more competitive in their environments. [18] Specifically it has been shown that when grown at high pCO_2 in eutrophic-like conditions, *M. aeruginosa* thrives.[19] Laboratory experiments run for a short period of time have also indicated that when grown at elevated CO_2 levels, microcystin toxicity increases.[20] There are few long term experiments with elevated CO_2 levels involving *M. aeruginosa* and genomic research into the adaptions that might be seen in response to high CO_2 environments. To investigate this topic, the second aim of our study with *M. aeruginosa* LE3 is to determine if genomic adaption occurs in cultures that are grown in elevated CO_2 levels predicted for the year 2100. This will provide insight into the impact of increasing atmospheric carbon levels on the biology of important microorganisms in the future. Interestingly, this work also revealed that a second organism was co-existing with *M. aeruginosa* in the cultures. Genomic sequencing identified this organism as a novel *Lacibacter* gram-negative bacteria.

Materials and Methods

Isolation of *Microcystis aeruginosa* LE3 Strain

The LE3 strain of *M. aeruginosa* was isolated in 1996 from the United States in Ohio, specifically the Western Basin of Lake Erie. The sample was collected from a freshwater algal bloom, at a depth of 1-5 meters, and at the latitude and longitude of 41.7N, 83.3 W. This was a mixed culture and was not a pure mono-isolate of just *M. aeruginosa.* This sample does have the taxonomy ID of 1194489 and the BioSample accession number of SAMN32578676. This strain has been used in genetic studies and has the ability to produce microcystin, but has remained in contig form and has never been successfully fully assembled.[9]

Culturing of the *M. aeruginosa* experimental and control samples

The *M. aeruginosa* cultures were grown in BG-11 media, a media typically used for freshwater cyanobacteria. The control cultures were grown at ambient pCO_2 levels, while the experimental cultures were grown at 1,000 ppm. All cultures were grown at 24°C and underwent a 12-hour light and dark cycle. Each group the experimental and

control, were raised in triplicates. Cultures were reinoculated approximately every two and a half weeks.

Genome Sequencing and Assembly of *Microcystis aeruginosa* strain LE3 at Time 0

DNA was extracted using a modified method for the Promega HMW Wizard Kit. DNA was sent to Novogene for both Illumina and PacBio sequencing. The PacBio and the Illumina reads were run through Unicycler with default parameters, and two genomes were produced. The culture was not a pure culture and had *M. aeruginosa* and an unknown species that were co-existing in culture. In order to assess the completeness and contamination of both genomes, two different programs were used, CheckM and Classify Microbes with GTDB-Tk - v1.7.0. CheckM was used to assess the level of completeness and level of contamination of each genome.[21] The second genome was run through CheckM and Classify Microbes with GTDB-Tk - v1.7.0 on Kbase for completeness and identification.[22]

Phylogenomic Comparison Among Completed Genomes of the Microcystis Genus

On NCBI's Genome Database, the *Microcystis* genus has a total of 13 completed genomes. Of these 13 completed genomes, 10 belong to *M. aeruginosa*, 1 belongs to *Microcystis viridis*, 1 belongs to *Microcystis panniformis*, and there is 1 *Microcystis sp*., that is an unclassified species. The genomes were analyzed using CLC Genomics Workbench, v.12. Phylogenomic analysis were also run by using The Department of Energy Systems Biology Knowledgebase (Kbase), a software and data platform that allows for biological function and comparisons to be assessed.[23] NCBI has only one other completed genome for the genus *Lacibacter* and two other genome assemblies that have full internal transcribed spacer regions (ITS), a region consisting of the 16S-23S ribosomal RNA region that has been useful in identifying bacterial species.[25] These *Lacibacter* ITS regions were isolated and run through a pairwise comparison and a phylogenetic tree using CLC Genomics Workbench, v.12.

Genome Sequencing at Five Months

At five months, the three control cultures and the three experimental cultures had DNA extracted using the same modified protocol using a Promega Wizard HMW DNA Kit. DNA was sent to NovaGene and the Illumina NovaSeq 6000 platform was used. A microbial whole genome library that was PCR-free for the experimental and culture cultures. DNA was required to meet the following expectations a total amount over 1.2 µg of genomic DNA, a volume over 20µl, and a concentration over 10 ng/µl.

Assembly and Alignment of Experimental and Control Cultures at Five Months

Sequencing data was acquired from Novogene in the form of fasta files and uploaded onto CLC workbench program. Each sample's sequencing data was run through the CLC for map reads to reference, with the reference genome. The read mapping produced was run through the extract consensus sequence and a consensus sequence was produced, with areas of low coverage chosen to be represented by N's. Each culture had a genome produced for a total of three control genomes and three experimental genomes at five months and were all aligned to the reference genome. This allowed for a visualization to compare the genomes for regions and points of difference. Coding regions that had no differences amongst the controls but showed differences in the experimental culture's genomes were assessed. Once a coding region had been identified, the region for each genome was extracted out and put into a

nucleotide alignment. The coding region for all was also translated into a protein sequence to see if the differences seen produced a change in the amino acid sequence. Amino acid changes were noted and analyzed for what type of change occurred.

Results and Discussion

Complete Genome of *M. aeruginosa* strain LE3 and a novel *Lacibacter sp.* Strain and Characterization

The genome of the *M. aeruginosa* LE3 strain were submitted to NCBI Genomes and annotated using the NCBI's prokaryotic genome annotation pipeline.[26] The genome was accepted and will be published on the database. The genome of *M. aeruginosa* strain LE3 is 4,984,962 base pairs in length and has 4,710 protein coding regions. The CheckM results for the first genome was a completeness of 99.89% and a 0.07% contamination and it was assigned to the phylum Cyanobacteria.

The *Lacibacter* genome had a score of 100% and a contamination score of 0.25% and was assigned to the phylum *Bacteroidetes*. The results of the GTDB analysis showed this organism belonged to the genus *Lacibacter* but that its species was novel. The genome was accepted into NCBI's data and was given the strain name of *Lacibacter sp.* strain MH-610. The genome is 3,984,569 base pairs in length and has 3,582 protein coding regions. The genomes have been deposited in NCBI Genomes database awaiting publication.

Phylogenetic relationships of *M. aeruginosa* LE3 and of *Lacibacter sp.* strain MH-610

In the NCBI database to date (2023), there are a number of *M. aeruginosa* sequences, but there are only 10 completed genomes for this species. Of these 10 complete genomes, the vast majority are from Asia, and only one from North America. This North American *M. aeruginosa* strain is from a algal bloom in Florida but was found to be non-toxic.[27] This makes the completed genome of LE3 of high importance, due to its prevalence in Lake Erie, a key freshwater reservoir in North America, and it being the first completed toxic genome of *M. aeruginosa* for North America. Out of all the completed genomes to date, the LE3 isolate most closely resembles *M. aeruginosa* strain NIES-298, with the genomes being on the same branch in the phylogenomic tree, with every replication placing them together (Figure 1). To compare all the annotated *Microcystis* genomes, a compute Pangenome/Pangenome circle plot were run. The LE3 isolate when compared to all annotated genomes showed there is a high number of core genes, genes that are found in every *Microcystis* genome (Figure 2). The pangenome analysis also showed that LE3 has a small number of non-core genes, genes that are found exclusively in the LE3 strain (Figure 2). The genus *Lacibacter* did not have a large variety of genomes to work with but of the available sequences available, key phylogenetic analysis were able to be made. The novel strain found in the *M. aeruginosa* culture was further confirmed to be a novel species of *Lacibacter* (Figure 3).

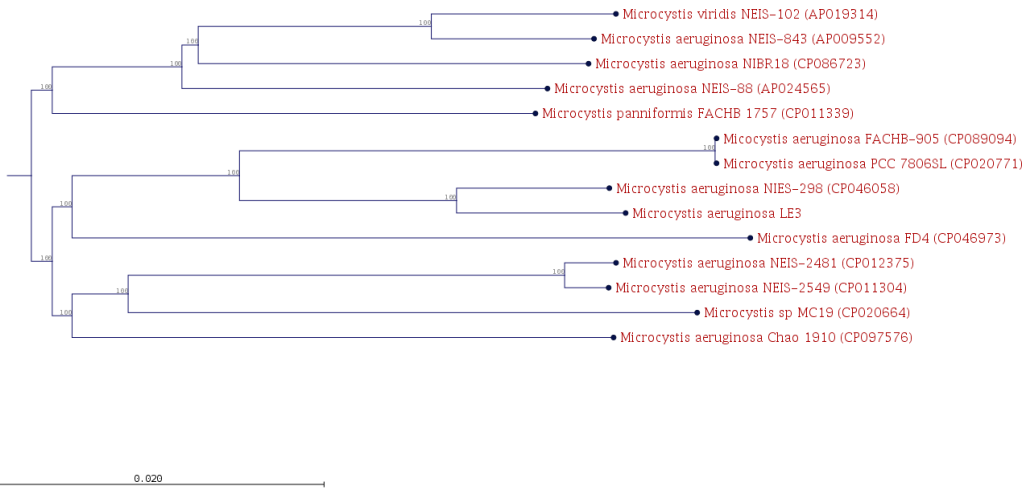

Fig. 1. Phylogenomic tree of all completed genomes in the *Microcystis* genus. The tree represents all the completed genomes of the *Microcystis* genus on NCBI's database. The phylogenomic tree was produced using the Neighbor-Joining algorithm, with distance being measured by the Jukes-Canthor method, and a 1,000 boot strap replicates were performed. Organisms are listed by their full Latin name, followed by their strain name and then their NCBI accession number.

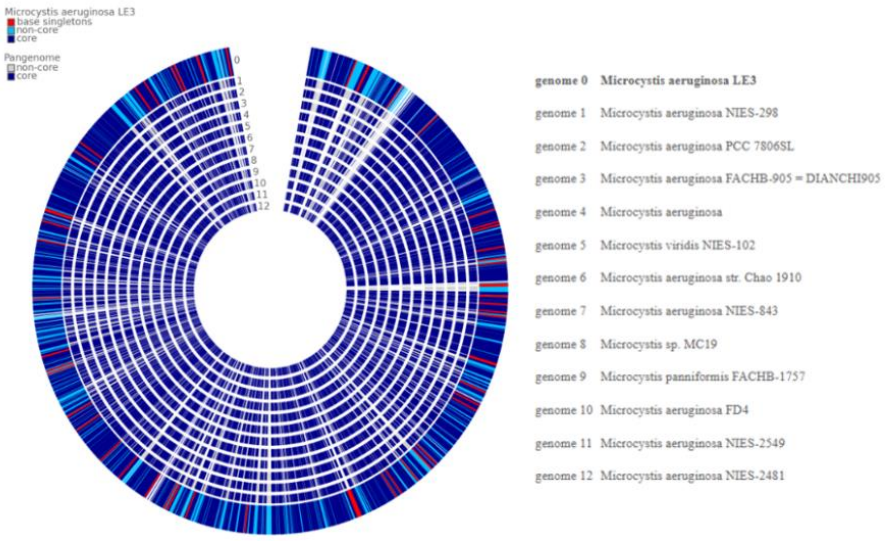

Fig. 2. Pangenome circle plot of all complete, annotated genomes of the *Microcystis* genus. The results of the pangenome circle plot performed in Kbase. The LE3 strain serves as the base genome for this circle, and the right half of the figure shows what number correlates to a genome. Core genes are those that are found in every one of the genomes that is being analyzed and is represented by dark blue. Non-core genes are those that are found in multiple genomes but not all genomes analyzed and is represented with a light grey in non-base genomes and light blue in the base genome. Singletons are genes that are exclusively found in the base genome and are represented by red.

Fig. 3. Phylogenetic Tree of the genus *Lacibacter* and the available ITS regions. The phylogenomic tree was produced using the Neighbor-Joining algorithm, with distance being measured by the Jukes-Canthor method, and a 1,000 boot strap replicates were performed. Organisms are listed by their full Latin name, followed by their strain name and then their NCBI accession number.

Mutations after 5 months at high CO_2 for the mix culture

Random mutation and variability among genomes are expected, however mutations were only reported in protein coding regions in which only the high carbon genomes were mutated. A total of 3 protein coding regions had the same mutations in 2 out of the 3 high carbon-raised genomes of the *Microcystis* (Table 1). Most of the *Microcystis* mutations appeared to be nonsynonymous and changed the protein sequence. The *Microcystis* results indicate that while there appears to be a few genes that adaptations could potentially be seen, the genome appears to be stable, and few adaptations seem to have occurred for the increased carbon levels (Table 1). The calcium-binding protein seems to be the most effective of the protein coding regions, with four separate mutations shared by 2 of the genomes (Table 1). Of note is that chloroplasts are involved in intracellular calcium homeostasis, which may indicate an adaptation to a higher carbon environment.[28] The other gene of note is the IS630 family transposase, which belongs to a category of mobile genome elements, miniature inverted repeat transposases (MITEs).[29]

Table 1. Mutations in *Microcystis* cultures raised in an increased carbon environment.

Gene	Number of Genomes Mutated	Change in DNA sequence	Synonymous or nonsynonymous	Change in Protein sequence
calcium-binding protein	2	C » T C » A A » T T » C G » C T » C	Synonymous Nonsynonymous Nonsynonymous Synonymous Nonsynonymous Nonsynonymous	No change Q » N Q »N/H No change V » L V » L
IS630 family transposase	2	T » C	Nonsynonymous	L » P
Hypothetical protein	2	C » G	Nonsynonymous	R » G

Conclusions

The project consisted of two main aims. The first aim was to complete the genome of the *Microcystis aeruginosa* strain LE3, a species that was isolated from Lake Erie and conduct a phylogenomic analysis. The second was to grow this organism at present day and future atmospheric CO_2 levels and assess the genomes for adaption. The genome of *M. aeruginosa* was completed and is pending future release, as well as a novel species of the genus *Lacibacter* which was discovered alongside it. In addition, phylogenetic studies were conducted to assess evolutionary relationships among the completed *Microcystis* genomes. The same was performed with the *Lacibacter* genome and it was confirmed to be a novel species that will undergo further physiological characterization. The genomes of *Microcystis* were sequenced at five months and genomes assembled and compared. For *Microcystis* a total 3 different protein coding regions had the same mutations for two out of the three high-CO_2 genomes. This analysis indicates that at five months *M. aeruginosa* did not experience a high degree of adaption to the high CO_2 environment. It appears that there may be select sequence regions that will adapt or are currently mutating, but longer study is needed to see if these are true changes and if any more occur. Analysis of microcystin production at present day and high CO_2 conditions is still in the progress.

Acknowledgements

We thank Peter Bernhardt and Margaret Mulholland (Old Dominion University) for growing the cyanobacterial/bacterial cultures and very helpful discussions and guidance.

References

1. Huisman J, Codd GA, Paerl HW, et al. Cyanobacterial blooms. Nat Rev Microbiol 2018;16(8):471–483; doi: 10.1038/s41579-018-0040-1.
2. Paerl HW, Otten TG. Harmful cyanobacterial blooms: causes, consequences, and controls. Microb Ecol 2013;65(4):995–1010; doi: 10.1007/s00248-012-0159-y.

3. Harke MJ, Steffen MM, Gobler CJ, et al. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, Microcystis spp. Harmful Algae 2016;54:4–20; doi: https://doi.org/10.1016/j.hal.2015.12.007.
4. Kramer BJ, Davis TW, Meyer KA, et al. Nitrogen limitation, toxin synthesis potential, and toxicity of cyanobacterial populations in Lake Okeechobee and the St. Lucie River Estuary, Florida, during the 2016 state of emergency event. PLoS One 2018;13(5):e0196278.
5. Massey IY, Yang F. A Mini Review on Microcystins and Bacterial Degradation. Toxins (Basel) 2020;12(4); doi: 10.3390/toxins12040268.
6. Struck TH, Feder JL, Bendiksby M, et al. Finding Evolutionary Processes Hidden in Cryptic Species. Trends Ecol Evol 2018;33(3):153–163; doi: 10.1016/j.tree.2017.11.007.
7. Otsuka S, Suda S, Li R, et al. Morphological variability of colonies of Microcystis morphospecies in culture. J Gen Appl Microbiol 2000;46(1):39–50; doi: 10.2323/jgam.46.39.
8. Cai H, McLimans CJ, Beyer JE, et al. Microcystis pangenome reveals cryptic diversity within and across morphospecies. Sci Adv 2023;9(2):eadd3783; doi: 10.1126/sciadv.add3783.
9. Meyer KA, Davis TW, Watson SB, et al. Genome sequences of lower Great Lakes Microcystis sp. reveal strain-specific genes that are present and expressed in western Lake Erie blooms. PLoS One 2017;12(10).
10. McLellan NL, Manderville RA. Toxic mechanisms of microcystins in mammals. Toxicol Res (Camb) 2017;6(4):391–405; doi: 10.1039/c7tx00043j.
11. Arman T, Clarke JD. Microcystin Toxicokinetics, Molecular Toxicology, and Pathophysiology in Preclinical Rodent Models and Humans. Toxins (Basel) 2021;13(8); doi: 10.3390/toxins13080537.
12. Steffen MM, Belisle BS, Watson SB, et al. Status, causes and controls of cyanobacterial blooms in Lake Erie. J Great Lakes Res 2014;40(2):215–225; doi: https://doi.org/10.1016/j.jglr.2013.12.012.
13. Watson SB, Miller C, Arhonditsis G, et al. The re-eutrophication of Lake Erie: Harmful algal blooms and hypoxia. Harmful Algae 2016;56:44–66; doi: 10.1016/j.hal.2016.04.010.
14. Smith RB, Bass B, Sawyer D, et al. Estimating the economic costs of algal blooms in the Canadian Lake Erie Basin. Harmful Algae 2019;87:101624; doi: 10.1016/j.hal.2019.101624.
15. Malone TC, Newton A. The Globalization of Cultural Eutrophication in the Coastal Ocean: Causes and Consequences. Front Mar Sci 2020;7.
16. Bhagowati B, Ahamad KU. A review on lake eutrophication dynamics and recent developments in lake modeling. Ecohydrol Hydrobiol 2019;19(1):155–166; doi: https://doi.org/10.1016/j.ecohyd.2018.03.002.
17. Bui T, Dao T-S, Vo T-G, et al. Warming Affects Growth Rates and Microcystin Production in Tropical Bloom-Forming Microcystis Strains. Toxins (Basel) 2018;10(3); doi: 10.3390/toxins10030123.
18. Sandrini G, Ji X, Verspagen JMH, et al. Rapid adaptation of harmful cyanobacteria to rising $CO_{(2)}$. Proc Natl Acad Sci U S A 2016;113(33):9315–9320; doi: 10.1073/pnas.1602435113.
19. Ma J, Wang P, Wang X, et al. Cyanobacteria in eutrophic waters benefit from rising atmospheric $CO_{(2)}$ concentrations. Sci Total Environ 2019;691:1144–1154; doi: 10.1016/j.scitotenv.2019.07.056.
20. Sandrini G, Cunsolo S, Schuurmans JM, et al. Changes in Gene Expression, Cell Physiology and Toxicity of the Harmful Cyanobacterium Microcystis Aeruginosa at Elevated CO_2. Front Microbiol 2015;6.
21. Parks DH, Imelfort M, Skennerton CT, et al. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 2015;25(7):1043–1055; doi: 10.1101/gr.186072.114.
22. Chaumeil P-A, Mussig AJ, Hugenholtz P, et al. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 2020;36(6):1925–1927; doi: 10.1093/bioinformatics/btz848.
23. Arkin AP, Cottingham RW, Henry CS, et al. KBase: The United States Department of Energy

Systems Biology Knowledgebase. Nat Biotechnol 2018;36(7):566–569; doi: 10.1038/nbt.4163.
24. Li L, Stoeckert CJJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 2003;13(9):2178–2189; doi: 10.1101/gr.1224503.
25. Huo D, Chen Y, Zheng T, et al. Characterization of Microcystis (Cyanobacteria) Genotypes Based on the Internal Transcribed Spacer Region of rRNA by Next-Generation Sequencing. Front Microbiol 2018;9:971; doi: 10.3389/fmicb.2018.00971.
26. Tatusova T, DiCuccio M, Badretdin A, et al. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 2016;44(14):6614–6624; doi: 10.1093/nar/gkw569.
27. Hidetoshi U, L. HT, H. SJ, et al. Complete Genome Sequence of Microcystis aeruginosa FD4, Isolated from a Subtropical River in Southwest Florida. Microbiol Resour Announc 2020;9(38):e00813-20; doi: 10.1128/MRA.00813-20.
28. Navazio L, Formentin E, Cendron L, et al. Chloroplast Calcium Signaling in the Spotlight . Front Plant Sci 2020;11.
29. Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol Rev 2014;38(5):865–891; doi: 10.1111/1574-6976.12067.