Harmful Algae News

An IOC Newsletter on Toxic Algae and Algal Blooms

No. 65 - September 2020 · www.ioc-unesco.org/hab

United Nations Educational, Scientific and Cultural Organization

Intergovernmental Oceanographic Commission

Genomic resources for the domoic acid-producing diatom *Pseudo-nitzschia multistriata*

Species responsible for Harmful Algal Blooms (HABs) are among the best studied unicellular microalgae. HABs pose a serious risk to human health and are responsible for considerable economic losses in the aquaculture industry which has resulted in the funding of monitoring programs to investigate their temporal and spatial distribution, as well as research projects aimed at understanding various aspects of their biology. In recent decades, understanding about HAB diversity, physiology and trophic habits, the mechanisms that regulate toxin production, and the role of environmental variables and interactions with other members of the plankton community on their growth dynamics has increased. The rapid development of genomic- and transcriptomic-based approaches, now available for an increasing number of unicellular eukaryotes including HAB species, is allowing scientists to gain a mechanistic understanding of a broad range of biological features.

The genus *Pseudo-nitzschia* includes the majority of species that produce the neurotoxin domoic acid, the causative agent of Amnesic Shellfish Poisoning. About 50 species have been described to date and half of these are known to produce domoic acid. The importance of this genus motivated the selection of one of the most toxic species, *Pseudo-nitzschia multiseries*, to be one of the few

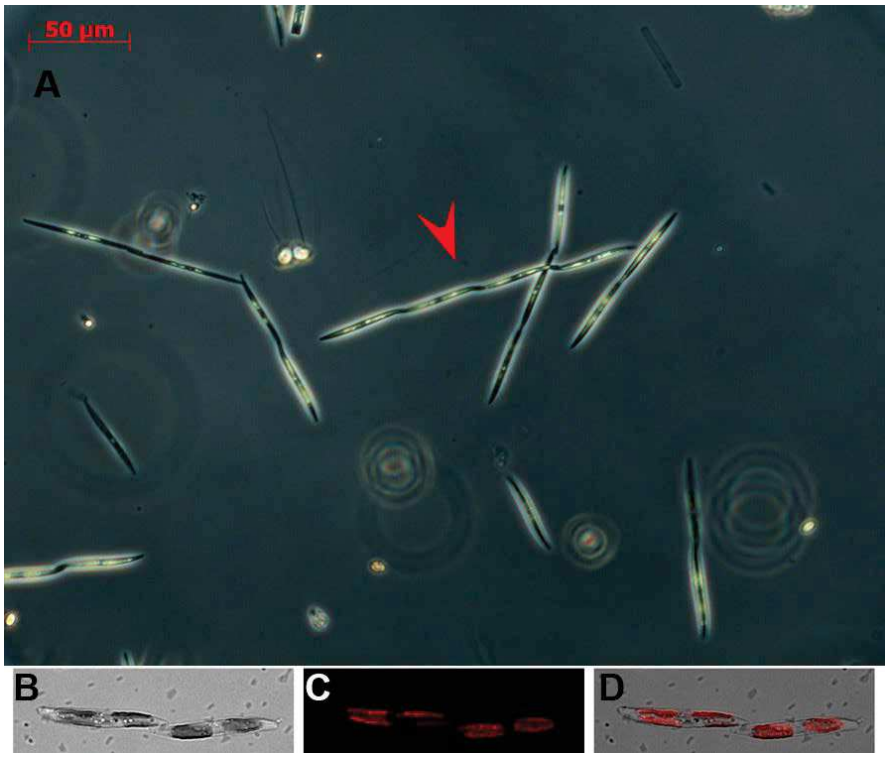

Fig. 1. A. Light micrograph of a natural phytoplankton sample collected at the Long Term Ecological Research station MareChiara (Gulf of Naples, Italy) on October 3rd 2013. Pseudo-nitzschia multistriata cells (a chain is indicated with an arrow) can be easily identified by their slightly sigmoid shape. Photo: Marina Montresor. Two Pseudo-nitzschia multistriata cells approaching cell division, B. bright field, C. chlorophyll autofluorescence in the chloroplasts, and D. merged images.

Content

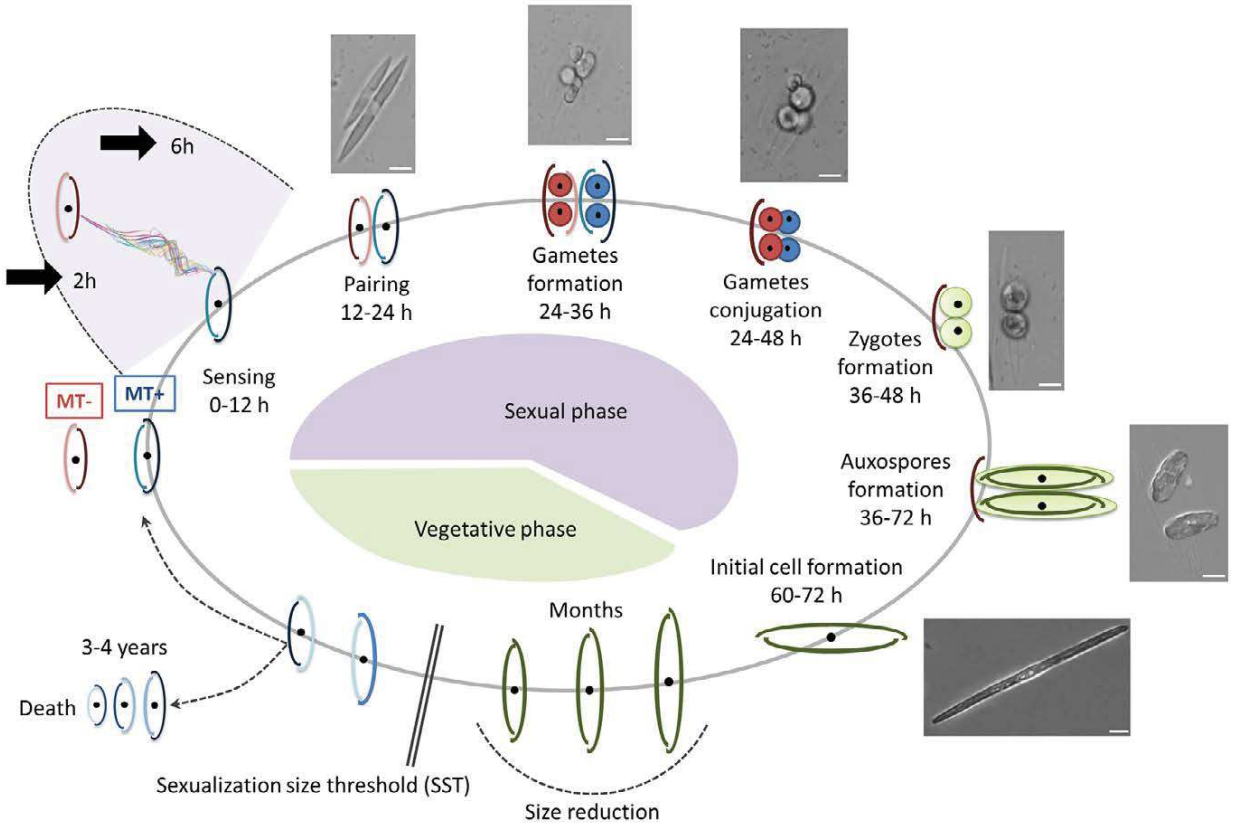

Fig. 2. Schematic drawing of the life cycle of the pennate diatom Pseudo-nitzschia multistriata. The vegetative phase is characterized by progressive cell size reduction of the population. When cells reach the sexual size threshold (SST), they can either keep decreasing in size until they die, or undergo sexual reproduction and escape the size-decreasing process producing large initial cells. The perception of chemical cues deriving from the mating partner brings cells of opposite mating type to pair and haploid gametes are produced following meiosis. Conjugation of gametes produces two expandable zygotes that develop into auxospores, within which an initial cell of maximum size is synthesized, restoring the vegetative phase of the cycle. Representative microscopic images of the different stages are shown outside the circle; bar, 10 µm. From [7].

diatoms selected for genome sequencing when this was still a demanding and time consuming effort. The *P. multiseries* genome has been publicly available since 2011 (*https://mycocosm.jgi.doe.gov/Psemu1/Psemu1.home.html*) and despite being quite fragmented, possibly due to the high content of repetitive sequences it has represented a valuable resource used to explore diatom biology in general and in many comparative genomic analyses. Subsequently, *Pseudo-nitzschia* transcriptomes became available through the Marine Microbial Eukaryotic Transcriptome Sequencing Project (MMETSP), enlarging the repertoire of sequence resources for the genus (*https://www.imicrobe.us/#/search/mmetsp*). Currently, thanks to this initiative, transcriptomics data are available for *P. arenysensis*, *P. delicatissima*, *P. fraudulenta*, *P. australis*, *P. heimii* and *P. pungens*. Transcriptomics data have also been generated in dedicated studies for other species, namely *P. multiseries, P. granii, P. seriata* and *P. obtusa* [1-3].

Here we illustrate genomic resources that are available to the research community for another toxic *Pseudo-nitzschia* species, *P. multistriata.* We also provide a few examples of questions that can be addressed by genomic approaches.

Pseudo-nitzschia multistriata was described in 1993 by Hideaki Takano from coastal waters in southern Japan. The species produces domoic acid [4] and has been reported from various regions world-wide (Mediterranean Sea, Gulf of Mexico, Malaysia, Singapore, New Zealand, the Pacific coast of Mexico, see [5]). Cells of *P. multistriata* can easily be identified in light microscopy due to their distinct 'sigmoid' shape (Figure 1). This species began to be recorded at the Long Term Ecological Research station Mare Chiara in the Gulf of Naples in 1993 and since then is regularly recorded in summer-early autumn. The relative ease of identification made it a model for a long term study of population structure using microsatellites [6]. This study required the isolation of a few hundred strains to obtain their DNA. Availability of all of these strains represented a great resource facilitating the use of *P. multistriata* as a genetic model, allowing genetic crosses and production of new generations in the laboratory.

We would like to avail of this opportunity to publicize our request for *P. multistriata* strains to all of our colleagues working with phytoplankton samples who may find the species, easily distinguishable from other *Pseudo-nitzschia* due to the curved tips and sigmoid shape. Strains from different geographic locations will greatly enrich ongoing population genomics analyses and could reveal more about core and dispensable regions of the genome, selection, and ultimately about the evolution of this toxic species.

Fig. 3. The genome browser available on the SZN BioInforma platform.

The genome could be sequenced exploiting inbred strains, obtained from the cross of a first generation of sibling strains [7]. Because of the lower polymorphism of inbreds, it was possible to reconstruct long fragments of DNA from reads obtained with the Illumina technology, overcoming the limits of short reads assembly. Indeed, the assembly of high polymorphic sequence reads make it harder for common assembly software to bind small sequences together to produce the long fragments called scaffolds.

The 59 Mb genome obtained enhanced the range of possible approaches to explore the genetic basis of many of the species features, including the capability to undergo sexual reproduction. *P. multistriata*, like the majority of pennate diatoms, has a heterothallic mating system. This implies that sexual reproduction, the process in diatoms that counteracts progressive cell size reduction to produce large-sized F1 cells, only occurs when cells of opposite mating type (MT+ and MT-) make contact (Figure 2) [7]. Gene expression studies based on RNA-seq have focused on the life cycle and on the sexual reproduction phase, revealing differences in the transcriptomic behavior between the two mating types, even if they are morphologically identical. One of the most exciting findings enabled by the genome sequence availability was the discovery of the first mating type determining gene for diatoms. How a cell becomes MT+ or MT- was unknown; differential expression of a specific gene between groups of MT+ and MT- strains was linked to structural differences in a specific genomic region, providing the first clue on how the MT is specified [8].

Diatom genomes, when compared to estimates for other toxic phytoplankton such as dinoflagellate species, are generally smaller and less "intimidating", and genome sequencing is approachable for even small teams, without the requirement to establish large consortia. The *P. multistriata* genome sequencing project was a small project funded by a Marie Curie Career integration grant, it mainly involved two teams, the SZN in Naples (Italy) and the TGAC in Norwich (UK) (now Earlham Institute), cost around 10,000 euros and was initially released on a TGAC browser which was password protected in line with institute policies. To make it more easily accessible, the genome browser is now being hosted by the bioinformatics service at the SZN, (*http://bioinfo.szn.it/pmultistriata/*), embedded in the service platform (*http://bioinfo.szn.it/*; Figure 3) where it is freely accessible together with all the related RNA-seq data produced so far. Beyond gene tags, researchers have the possibility to browse the genome information visualizing repetitive elements, conserved regions in common with other species and other features, with links provided to the data files and to the relevant literature. Finally, within the European project EMBRIC and with the assistance of European Bioinformatics Institute teams, additional work on the sequencing data has been done to have it released in Ensembl (*https://protists.ensembl.org/Pseudonitzschia_multistriata/Info/Index*), one of the major genome browsers, originally created for vertebrate genomes but with an expanding section for protists. Availability in Ensembl makes the data readily discoverable by the HAB community, by scientists working on very distantly related organisms, and facilitates its use in large scale comparative studies. Unfortunately due to specific requirements of the Ensembl browser which did not support the original files generated for *P. multistriata*, the SZN and the Ensembl versions of the genome are not uniform; the scaffold names are different and, in Ensembl, all *P. multistriata* proteins are reported as "unknown proteins" while a detailed functional annotation is provided by the platform available from SZN.

While these discrepancies highlight the existence of difficulties in crosslinking –omics data from different resources, especially for non-model systems as well as the need for constant and supported interaction between wet lab scientists and bioinformaticians, the availability of *P. multistriata* genome resources was a big opportunity for comparative analyses. It has enabled and strengthened important discoveries, such as the conservation of a cluster of four genes involved in domoic acid synthesis, originally identified in the *P. multiseries* genome [9].

Continued on page 5

Fig. 4. Some Pseudo-nitzschia fans at Stazione Zoologica Anton Dohrn: (standing from left to right) Svenja Mager, Maria Immacolata Ferrante , Francesco Manfellotto, Anna Santin, Maria Valeria Ruggiero, Monia Russo, Viviana Di Tuccio, (sitting from right to left) Rossella Annunziata, Antonella Ruggiero, Pina Marotta, Marina Montresor

When tides collide: Harmful cyanobacterial and microalgal blooms in Florida and implications for risk assessment

Cyanobacterial blooms are a regular occurrence in southern Florida. Water releases from Lake Okeechobee to maintain the water level in this large lake regularly occur along the St. Lucie Canal to the eastern seaboard or along the Caloosahatchee to Fort Myers and the Gulf of Mexico (Fig. 1). Due to increased nutrient loading in Lake Okeechobee, cyanobacterial blooms, frequently composed of *Microcystis aeruginosa* are released from the lake and contaminate both these waterways. This results in large, microcystin-containing cyanobacterial blooms that can affect tourism and fisheries (Fig. 2). Press reports cataloguing these economic and health impacts in Florida are common.

In summer 2018, the situation on the west coast of Florida was exacerbated by the appearance of a bloom of *Karenia brevis* in the Gulf of Mexico. This brevetoxin-producing algal bloom resulted in the deaths of marine mammals and fish with numbers in the thousands reported on the beaches around southwest Florida. Due to the large discharge of water from Lake Okeechobee, pulses of freshwater ended up in the estuarine and coastal environments.

Sampling of the Caloosahatchee in 2018 showed high concentrations of *Microcystis aeruginosa* and microcystin-LR according to microscopy, UPLC-PDA and UPLC-MS [1]. BMAA was also detected by triple quadrupole mass spectrometry at much lower concentration. Although such blooms dominated the freshwater environment of the river, visually intact *Microcystis* colonies were observed in brackish water in the sound and estuary past the tide line. Conversely, although diatoms and dinoflagellates could not be observed in freshwater samples, ELISA analysis for brevetoxins showed low positive concentrations in the Caloosahatchee and large concentrations in the brackish and marine environment.

The findings of the "crossing-over" of blooms raise questions concerning the risk assessment of cyanobacterial and dinoflagellate toxins in estuarine environments. First, how far can toxins produced by freshwater cyanobacterial blooms be found in saline environments? Conversely, can marine diatom and dinoflagellate toxins be routinely found in freshwater environments, most likely through navigation and tidal flows? Our data indicate that both are possible, although a greater understanding of the effect of salinity on the growth of *Microcystis* is required, building upon previous work that has shown that e.g. microcystin concentrations can be increased after salt shock [2]. Furthermore, in other estuarine environments a wide range of marine and freshwater toxins have been identified [3].

If the particular toxin(s) are known within freshwater or marine blooms, then adequate risk assessment is possible to protect human and animal health. However, when multiple blooms occur at the same location, then toxicity assessment and risk assessment are necessary to determine whether there are interactions between the toxicants that may change the toxicological outcome. Communication is also essential in order to alert the public to potential issues that may arise from exposure to one or more classes of aquatic toxin(s), whether from harmful cyanobacterial and/or algal blooms. Further communication may be necessary as, depending on the country, marine and freshwaters may be under the jurisdiction of different agencies with different accountabilities and responsibilities. The presence of such multiple blooms may also have effects upstream as, in the case of Florida, changes to discharges from Lake Okeechobee may have to occur when marine blooms are present. When tides collide, procedures need to be in place to determine the potential risk, communicate with relevant parties and perform toxicity assessments when necessary. In addition, collection of autopsy materials can also determine what toxicants may be causing adverse health effects, whether to wild or domestic animals or humans. Ultimately, legislation and good stewardship of waters and relevant contingencies can minimise the potential risk of adverse health effects from occurring.

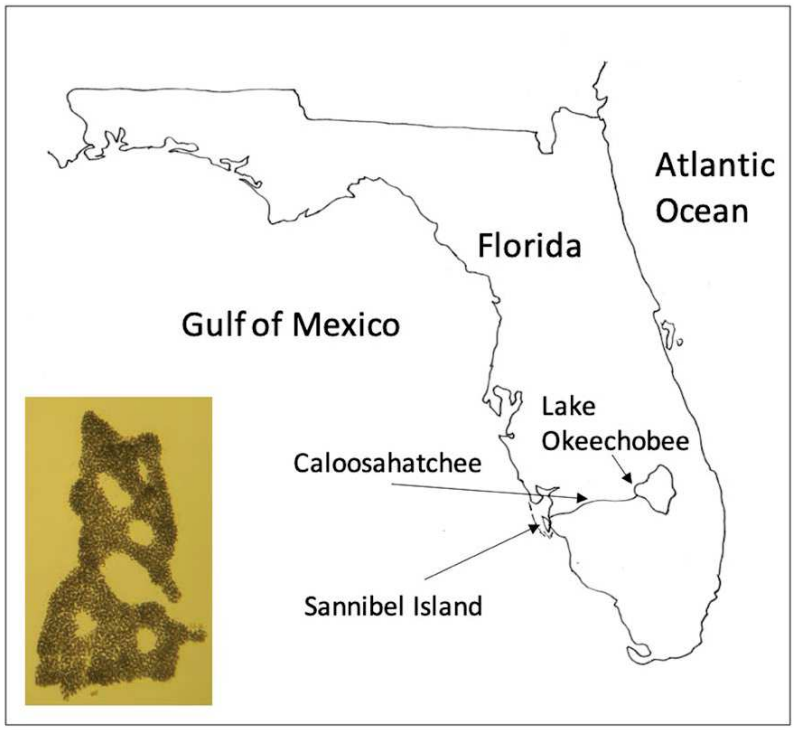

Fig. 1. Location of Lake Okeechobee and Caloosahatchee River in Southern Florida, USA where a bloom of Microcystis aeruginosa released from the lake collided with a red tide of brevetoxin-producing Karenia brevis near Sannibel Island. Bottom left Insert, a colony of Microcystis aeruginosa (approx. 500 m width) from Caloosahatchee River

Acknowledgements

We are grateful to Dr. John Cassini, Jason

Fig. 2. Bloom of Microcystis aeruginosa at Lake Okeechobee, Florida, summer 2018

Pim and Calusa Waterkeepers for assistance in sampling.

References

1. *Metcalf JS et al (in press). Neurotox Res*
2. *Matthiensen et al 2000. In: de Koe WJ et al (eds), Mycotoxins and Phycotoxins in Perspective at the Turn of the Millenium. Proc Xth Int IUPAC symposium on Mycotoxins and Phycotoxins, Guaruja,Brazil, 21-25 May 2000, pp 527-536*
3. *Peacock MB et al 2018. Harmful Algae 73: 138-147*

Authors

James S Metcalf, Nara Souza, Sandra Anne Banack & Paul Alan Cox, Brain Chemistry Labs, Jackson, WY 83001, USA

Email corresponding author:
james@ethnomedicine.org

Continued from page 3

The raw and processed sequencing data generated can be beneficial for many applications. Some are already routine methods, such as the use of the genome as a reference to map RNA-seq reads and interpret transcriptomics studies, comparative analyses with other genomes to define structural and/or phylogenetic relationships, estimation of the extent of genomic variations and rearrangements, and the identification of regions under selection, contributing to evolutionary insights in populations and species. Other analyses that can benefit from the existence of a reference genome include support in taxonomic assignation and in the interpretation of results from metagenomics and metatranscriptomics projects. Combining information from such approaches can lead to innovative ways to monitor phytoplankton; for example, genes found highly induced under specific experimental conditions in the laboratory can become markers to follow specific processes in the natural environment by qPCR or metatranscriptomics. Examples include genes involved in the synthesis of toxins, genes expressed during sexual reproduction, genes activated during specific stress conditions. We are now familiar with the fact that community composition can be described with barcode sequences. In the future we will be able to identify organisms by retrieving their genomes, as is now the case for bacteria. Moreover, species-specific genes allow specific functions from individual species to be understood, e.g. to obtain a picture of their metabolic state, or the phase of their life history. Finally, for many traditional or more innovative methods for the study of gene function, the presence of a reference genome is mandatory. Sophisticated manipulations and genome editing methods, including the CRISPR/Cas9 technology, are now within reach for this *Pseudo-nitzschia* species.

In view of the potential for *P. multistriata* to be used as a model for planktonic diatom life cycles, more studies are ongoing, and a further leap forward is expected thanks to the support of the Marine Microbial Initiative of the Gordon and Betty Moore Foundation, which is funding a project (*http://www.szn.it/index.php/en/research/integrative-marine-ecology/research-projects-emi/disco*) to dissect the mechanisms underlying sex determination and controlling transitions between life cycle phases, to discover the genes and genomic regions that are under selection in natural populations, and to assess the effects of sex on genome evolution. More *P. multistriata* genomes are being re-sequenced thanks to this project, and technological advances and new technologies gradually becoming available will be exploited to improve the current genome assembly, to explore epigenetic control and to bring our knowledge of *Pseudo-nitzschia* to a higher functional level.

References

1. *Bender et al 2014. Front Mar Sci 1: 3*
2. *Cohen et al 2018. Environ Microbiol 20: 3109-26*
3. *Harðardóttir et al 2019. BMC Mol Biol 20: 7*
4. *Orsini et al 2002. Eur J Phycol 37: 247-257*
5. *Bates et al 2018. Harmful Algae 79: 3-43*
6. *Ruggiero et al 2018. ISME J 12: 463–72*
7. *Basu et al 2017. New Phytol 215: 140–156*
8. *Russo et al 2018. Nat Comm 9: 5050*
9. *Brunson et al 2018. Science 361: 1356-1358*

Authors

Maria Immacolata Ferrante & Marina Montresor, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Napoli (Italy)

Email corresponding author:
mariella.ferrante@szn.it