Charles University

Faculty of Science

Department of Botany

Mgr. Dora Čertnerová

Genome size variation in microalgae and its evolutionary consequences

Variabilita velikosti genomu u mikrořas a její evoluční důsledky

Doctoral thesis

Supervised by doc. Mgr. Pavel Škaloud, Ph.D.

Prague, 2022

Author's declaration / Prohlášení autorky

I hereby declare that I made this thesis independently, using only the mentioned refereces. I did not submit this thesis nor its part for any other degree or diploma.

Prohlašuji, že jsem práci vypracovala samostatně a všechny v ní použité informační zdroje a literatura jsou řádně citovány. Tato práce ani její část nebyly použity k získání jiného nebo stejného akademického titulu.

Prague, 2022

Dora Čertnerová

Table of Contents:

Acknowledgements

I am very grateful to my supervisor **Pavel Škaloud** for his support and for everything he taught me. He was also the one who inspired me to study algology in the first place. When I became fascinated by the newly discovered diversity of genome size in our model algae, he allowed me to change the topic of my PhD project and provided me a space and freedome to pursue the genome size variability in microalgae much further. I will always be grateful that he trusted me.

I would also like to thank to my husband **Martin** for his love and suport. He is a constant source of inspiration for me. I love our shared enthusiasm for science, our discussions and experiments have led to more than one co-authored publication (including one that is part of this PhD project).

My gratitude also goes to the numerous scientists for their inspiration, help and support. Namely, I would like to thank **Karin Rengefors**, **Ryan Gregory**, **Paul Kron**, **Brian Husband** and **Pavel Trávníček**. I also thank to my friends and colleagues, mainly to **Martin Pusztai**, **Martina Pichrtová**, **Iva Jadrná**, **Helena Bestová** and **Katherine Drotos**.

Finally, many thanks belongs to my beloved kid **Edíček** for his patience while I was finishing my PhD project. In fact, about half of the 'General chapters' were written on the kids playgrounds while he was exploring the kids junge gym or sand and playing with his Swiss friends.

My gratitude also goes to all family members for their support and all the babysitting. I am also gratuful to Pavel Škaloud for parents-friendly enviroment in Algal speciation and evolution lab.

Abstract

Eukaryotic organisms exhibit tremendous variability in genome size with no apparent connection to their biological complexity. Although this variation is known to correlate with numerous phenotypic traits, its evolutionary consequences remain widely unknown. This particularly applies to microalgae, where the genome size estimation is often methodologically challenging. Yet, microalgae represent a promising model group to study genome size evolution owing to their lower body complexity, short generation time and large population sizes, the latter two allowing them to quickly respond to environmental challenges. The main aim of this thesis was to enhance our understanding of genome size variation in microalgae and its evolutionary consequences. To do so, together with my co-authors, I summarized the flow cytometry (FCM) protocols used for microalgae and microorganisms possessing small genomes and addressed their limitations resulting mainly from insufficient amounts of biomass, difficulties with nuclei extraction and prominent background noise due to presence of various pigments and secondary metabolites. Further, I provided best practice recommendations that include, among others, analysing young cultures, avoiding long-term cultivation, and testing different isolation buffers and nuclei isolation techniques. Second, I introduced two new easy to use nuclei isolation protocols implementing razor blade chopping of desiccated biomass, suitable for nuclei isolation of filamentous microalgae, and bead beating for nuclei isolation of solitarily living algae. Third, to thoroughly investigate intraspecific DNA content variation and its ecophysiological consequences, we further focused on golden-brown algae (chrysophytes). We employed propidium iodide FCM to estimate nuclear DNA contents. In the most comprehensive intraspecific genome size screening conducted to date on a microalgae, we revealed a substantial genome size variability among strains of *Synura petersenii*, spanning continuously 0.97– 2.02 pg of DNA. The evolutionary mechanism generating the observed variability likely operates via gradual changes of genome size accompanied by changes in genomic GC content, such as, for example, proliferation of transposable elements. A major intraspecific DNA content variability might arose from polyploidization, as assumed in some other chrysophyte species. Alternatively, two-fold intraspecific DNA content differences might represent different life cycle stages. We revealed that the chrysophytes alternate between two ploidy stages, both of which are capable of mitotic propagation and long-term survival in cultivation. With the exception of a small increase in cell size with a higher ploidy, both life cycle stages shared the same phenotype (isomorphic) and also had highly similar genomic GC contents. Consequently, the chrysophytes have an isomorphic haploid-diploid life cycle. This was also the first report of such life cycle among unicellular algae. Interestingly, the life stage transitions in chrysophytes appear to be highly synchronized among cells, possibly due to chemical signalization. Among our investigated chrysophytes, the DNA amount was positively associated with cell size and negatively associated with growth rate. As a result, strains possessing lower DNA contents should be better colonizers or have more efficient nutrient uptake. On the other hand, strains possessing higher DNA content might better tolerate toxic environments or have higher metabolite production due to potentially increased gene dosage. Yet, these putative physiological consequences were not reflected in the geographical distribution of *S. petersenii* strains possessing various DNA contents.

Key words: genome size, flow cytometry, DNA content variation, microalgae, golden-brown algae, adaptive potential

Abstrakt

Eukaryotické organismy se nehledě na svou biologickou komplexitu velmi výrazně liší velikostí genomu. I když jsou známy souvislosti této variability s mnoha fenotypovými znaky, její evoluční důsledky zůstávají do značné míry neznámé. To platí zejména u mikrořas, u kterých je určení velikosti genomu často metodologicky náročné. Mikrořasy jsou nicméně zajímavou modelovou skupinou pro studium evoluce velikosti genomu díky jejich nižší komplexitě, krátké generační době a velkým populačním hustotám, dvě poslední jmenované vlastnosti jim totiž umožňují rychle reagovat na změny prostředí. Hlavním cílem této práce bylo přinést nové poznatky o variabilitě velikosti genomu u mikrořas a jejích evolučních důsledcích. V této práci jsem se svými spoluautory shrnula protokoly průkotové cytometrie používané pro mikrořasy a mikroorganismy s malými genomy a poukázala na jejich limity, zejména spojené s nedostatečným množství biomasy, komplikacemi s izolací jader a výrazném šumu na pozadí analýz, který je obvykle způsoben různými pigmenty a sekundárními metabolity. Dále jsme uvedli několik doporučení, např. analyzovat mladé kultury, vyhýbat se dlouhodobé kultivaci a zkoušet různé izolační pufry a různé techniky izolace jader. Vyvinula jsem také dva nové, snadno aplikovatelné protokoly izolace jader, spočívající v nasekání vysušené biomasy žiletkou vhodné pro vláknité mikrořasy a rozemletí kuličkami vhodné pro jednobuněčné řasy. Abychom důkladně prozkoumali vnitrodruhovou variabilitu v množství DNA a její ekofyziologické důsledky, zaměřili jsme se detailněji na zlativky (Chrysophyceae). Množství jaderné DNA jsme měřili pomocí průtokové cytometrie s propidium jodidem. V dosud nejrozsáhlejším vnitrodruhovém screeningu velikosti genomu u mikrořas jsme odhalili značnou variabilitu velikosti genomu mezi kmeny *Synura petersenii*, s kontinuálním rozsahem 0,97- 2,02 pg DNA. Evoluční mechanismus, který vedl k této variabilitě zřejmě spočíval v postupných změnách velikosti genomu doprovázených změnami v obsahu jederného GC, jako je například šíření transpozonů. Značná vnitrodruhová variabilita v množství DNA mohla vzniknout také polyploidizací, jak se předpokládá u některých jiných druhů zlativek. Dvojnásobné vnitrodruhové rozdíly v množství DNA mohou alternativně představovat odlišná stadia životního cyklu. Odhalili jsme, že se u zlativek střídají dvě ploidní stádia, z nichž obě jsou schopna mitotického dělení a dlouhodobého přežití v kultuře. S výjimkou malého zvětšení velikosti buněk s vyšší ploidií vypadala obě stadia životního cyklu stejně (izomorfně) a měla také velmi podobné množství genomové GC. Z toho vyplývá, že zlativky mají izomorfní haploidně-diploidní životní cyklus. Jde o první popsaný případ daného životním cyklu u jednobuněčných řas. Dále je zajímavé, že se změny životních stadií u zlativek zdají být mezi buňkami synchronizované, možná díky chemické signalizaci. U námi studovaných zlativek bylo množství jaderné DNA pozitivně spjato s velikostí buněk a negativně s rychlostí růstu. V důsledku toho by kmeny, které mají menší množství DNA, měly být lepšími kolonizátory či účinněji přijímat živiny. Naopak kmeny s větším množství DNA by mohly lépe snášet toxické prostředí nebo produkovat více metabolitů díky možnému většímu počtu genů. Tyto předpokládané fyziologické důsledky se ale neodrazily v geografickém rozšíření kmenů *S. petersenii* lišících se množstvím jaderné DNA.

Klíčová slova: velikost genomu, průtoková cytometrie, variabilita obsahu DNA, zlativky, adaptivní potenciál

List of included papers

I. **Čertnerová, D.**, & Galbraith, D. 2021. Best practices in the flow cytometry of microalgae. Cytometry Part A, 99: 359-364.

II. **Čertnerová, D.** 2022. Meet the challenges of analyzing small genomes using flow cytometry (commentary). Cytometry Part A, in press.

III. **Čertnerová, D.** 2021. Nuclei isolation protocols for flow cytometry allowing nuclear DNA content estimation in problematic microalgal groups. Journal of Applied Phycology, 33: 2057–2067.

IV. **Čertnerová, D.** & Škaloud, P. 2020. Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences. Annals of Botany, 126: 1077–1087.

V. **Čertnerová, D.**, Čertner, M., Škaloud, P. 2022. Alternating nuclear DNA content in chrysophytes provides evidence of their isomorphic haploid-diploid life cycle. Algal Research, 64: 102707.

Author contribution statement

I declare that I have substantially contributed to all papers included in the thesis. Specifically, my contributions to particular papers are as follows:

I. Čertnerová, D., & Galbraith, D. 2021: Best practices in the flow cytometry of microalgae.
Literature survey, data collection, and manuscript preparation – **total contribution 80%**

II. Čertnerová, D. 2022. Meet the challenges of analyzing small genomes using flow cytometry (commentary).
Literature survey, data collection, and manuscript preparation – **total contribution 100%**

III. Čertnerová, D. 2021. Nuclei isolation protocols for flow cytometry allowing nuclear DNA content estimation in problematic microalgal groups.
Study design, data collection and analysis, and manuscript preparation – **total contribution 100%**

IV. Čertnerová, D. & Škaloud, P. 2020. Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences.
Study conception and design, field sampling, data collection, analysis and interpretation of results, and manuscript preparation – **total contribution 90%**

V. Čertnerová, D., Čertner, M., Škaloud, P. 2022. Alternating nuclear DNA content in chrysophytes provides evidence of their isomorphic haploid-diploid life cycle.
Study conception and design, data collection, analysis, and interpretation of results and manuscript preparation – **total contribution 80%**

Part A – General chapters

1. Introduction

1.1. Genome size variability and its evolutionary sources

The nuclear genome constitutes an essential cell component. While the quality of nuclear DNA (expressed by nucleotide sequences) has been in the focus of evolutionary biologists for more than a half century, its quantity per cell has received much less attention. However, eukaryotic organisms exhibit tremendous variability in genome size, spanning a 120,000-fold range (Veldhuis *et al.* 1997; Corradi *et al.* 2010). Yet, even from the early genome size studies (Mirsky and Ris 1951), it has become obvious, that this variation does not reflect the biological complexity of organisms (Thomas 1971).

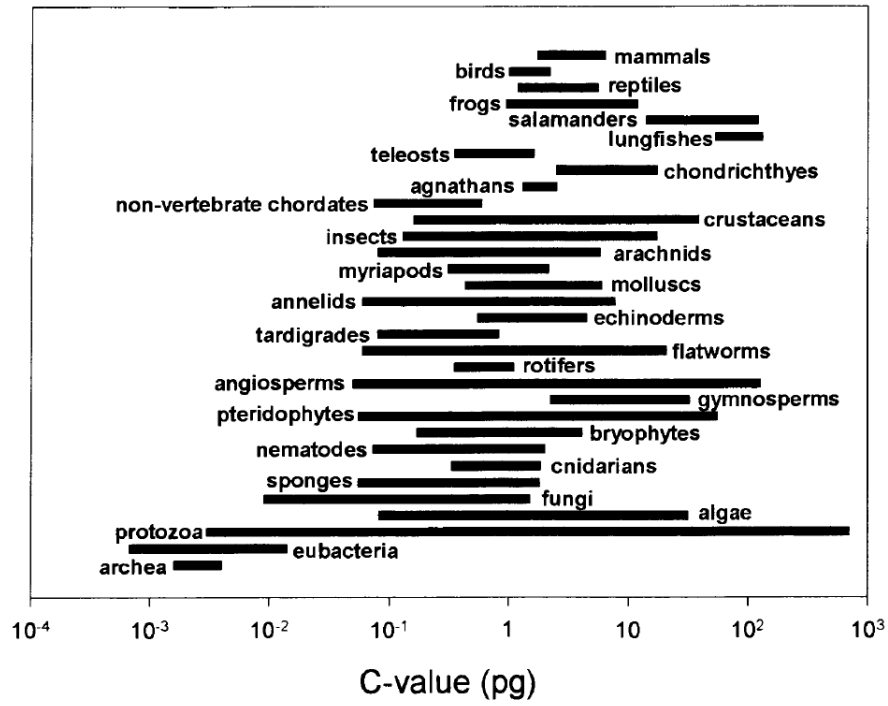


Fig. 1. The ranges in haploid genome sizes (C-values in picograms) in different groups of organisms (from Gregory 2004).

There are several evolutionary mechanisms responsible for the origin of nuclear DNA content variation. These mainly involve non-coding genomic regions, such as variation in the number and length of introns, or the amount of repetitive DNA sequences (e.g., transposable elements (TEs) or satellite DNA; Shah *et al.* 2020; Meyer *et al.* 2021). Profound shifts in genome size may be caused by whole-genome duplication (polyploidization events), chromosomal aberrations (aneuploidization) or nonhomologous meiotic recombination (Soltis and Soltis 1999; Devos *et al.* 2002; Wu *et al.* 2018; Schärer *et al.* 2020). The genome size variation might also originate

from gene duplications or hybridization events between closely related but separate species (Baack *et al.* 2005; Stelzer *et al.* 2021). The latter mechanism may be even accompanyied with the formation of reproductive barriers and ultimately lead to speciation (Coyne and Orr 2004). Despite all these mechanistic insights and progress in the elucidation of the genome structure and functions, our knowledge about the evolutionary forces driving genome size variation are still limited.

As most of the genome consists of non-coding DNA (a phenomenon called the C-value enigma; Gregory 2001a), this naturally raised the question of what is the reason for possessing such high amounts of seemingly useless DNA? There are two main directions explaining this paradox based on whether the extra DNA is beneficial or not. The 'junk DNA' theories postulate that the non-coding DNA has no advantage and it is in fact useless. Consequently, the extra DNA is randomly accumulated by genetic drift (Ohno 1972). According to the 'selfish' DNA hypothesis, the extra DNA consist of parasitic TEs (Orgel and Crick 1980; Doolittle and Sapienza 1980). Based on these neutral theories, cells are unable to control the increase of these sequences and the resulting genome size only reflects the highest tolerable maximum for a given organism. On the other hand, 'adaptive' theories propose that the genome size might have adaptive potential and could be of considerable evolutionary significance (Bennett and Leitch 2011). This is supported by the fact that the genome size correlates with the most fundamental traits. It is worth mentioning that these theories are not mutually exclusive and even neutrally evolving DNA content can be subjected to selection under certain evolutionary or environemental conditions (e.g., change to parasitic or ephemeral life strategy).

1.2. Phenotypic consequences of genome size variation

The genome size variation is often accompanied by numerous phenotypic consequences. The overall DNA amount determines nucleus size (Bennett 1972; Gregory 2001b; although some other mechanisms might be involved in controlling the nuclear size; Cantwell and Nurse 2019). This implicitly and strongly affects the cell size (a phenomenon called the karyoplasmic ratio; Wilson 1925; Cavalier-Smith 2005). The strong positive genome size – cell size correlation has been observed across the eukaryotic organisms (Gregory 2001b; Beaulieu *et al.* 2008). Furthermore, increase in cell size is actually the most common phenotypic effect and it is usually considered to be a direct consequence of the higher DNA amount (Bennett 1971; Hughes and Otto 1999).

The cell size is a particularly important trait as it inversely correlates with metabolic rate and growth rate, and directly correlates with generation time (Bennett 1972, 1987; Gregory 2002; Kozłowski *et al.* 2003; Wyngaard *et al.* 2005). As a consequence, the genome size is significantly reflected, for example, in physiology and ecology of organisms (Veselý *et al.* 2012a; Trávníček *et al.* 2019). In addition, via its effect on species' ecophysiology, the genome size may affect tolerance to stressful

environmental conditions (Nardon *et al.* 2005), ecological niche breadth (Pyšek *et al.* 2018) or even speciation and diversification rates (Igea *et al.* 2017).

Unfortunately, the current evidence of phenotypic consequences of genome size variation and its putative adaptive potential mainly come from plant and animal studies. However, unicellular eukaryotic organisms are much more suitable models. The lower complexity of their bodies provides easier disentangling of the consequences of genome size variation from other putative confounding factors (e.g., avoiding the compensatory mechanisms on cell size / number in the tissues of multicellular organisms). Further, single-celled eukaryotes have short generation time and very large population sizes, which allows them to quickly adapt to environmental changes (Lynch and Conery 2003; Foissner 2007; Ribeiro *et al.* 2013). Additionally, in cultivation, they can be kept in high densities under highly controlled conditions, which is very beneficial in ecological experiments. Despite these advantages, eukaryotic microorganisms are only rarely used in evolutionary studies on genome size.

1.3. Genome size in microalgae

1.3.1. Microalgae as a model group

Microalgae consist of photosynthetic microorganisms and their secondarily nonphotosynthetic evolutionary descendants. Even though cyanobacteria are sometimes included in microalgae, they are more and more considered as a separate group due to their prokaryotic nature (Pulz and Gross 2004). Even without cyanobacteria, microalgae are a remarkably diverse group of organisms. Although many microalgal species belong to the domain Plantae, many others are representatives of at least five distinct domains across the tree of life, including some traditionally recognized as fungi or protozoa (Burki *et al.* 2020). Because of their polyphyletic origin, microalgae differ greatly in their morphology, metabolite production, presence of specific organelles, and/or cell wall composition. Microalgae are ubiquitous across marine, freshwater and terrestrial habitats. They are crucially important primary producers and a major source of oxygen. Consequently, they are important drivers of the global ecosystem and an outstanding reservoir of biological diversity (genes, molecules, metabolic pathways and cellular processes; Falkowski *et al.* 2008). Recently, a considerable attention has been paid to microalgae as the potential source of next generation biofuels or usable metabolites (e.g., Brennan and Owende 2010; Hyka *et al.* 2013; Milano *et al.* 2016; Khan *et al.* 2018).

Microalgae were traditionally considered as simple organisms, literally 'non-vascular plants' (Copeland 1956). Yet, some of the microalgae have the most complex genomes on Earth. The nucleus of dinoflagellates, called the dinokaryon, is with constantly condensed chromosomes of liquid crystal organisation (Bouligand *et al.* 1968). Their genomes are also the largest among eukaryotes,

reaching up to 280 Gbp / pg (Veldhuis *et al.* 1997). On the other hand, evolutionary mechanisms such as extensive reduction of intergenic regions or downsizing of gene families led to the opposite extreme of genome miniaturization (Derelle *et al.* 2006). Such genome reduction was detected in green algae *Ostreococcus tauri*, *Nannochloris* spp. or red algae *Cyanidioschyzon merolae* with the smallest microalgal genome size 0.01 Gbp / pg (Maleszka 1993; Derelle *et al.* 2006; Palenik *et al.* 2007; Nelson *et al.* 2019). Apart from intracellular parasites, these are also the smallest genomes among (free living) eukaryotes. Yet, the genome size data are still available for only a fraction of microalgal diversity (Pellicer and Leitch 2020) and our knowledge in this area is thus vastly limited. However, the genome size data have recently become a prerequisite for many areas of research on microalgae. In genomics, knowing the genome size is key to designing an optimal sequencing strategy and assessing the success of resultant genome assemblies. Since the nuclear DNA content directly determines the cost of a sequencing project, low DNA contents has become a major criterion in selection of suitable algal strains (Waaland *et al.* 2004; Peters *et al.* 2004; Lin 2006). Interestingly, some pathways of DNA content increase are coupled with increased gene dosage, thus the genome size knowledge allows us selecting lineages with a potentially higher secondary metabolite production (Mason 2016; Priyadarshan 2019; Qin *et al.* 2019). Additionally, the availability of genome size data and high-level phylogenies provide the means to determine evolutionary trends in genome size variation. Such innovative studies brought new insights into microalgal nutrition modes or cell-size changes (Poulíčková *et al.* 2014; Olefeld *et al.* 2018). Further, genome size analysis allows detection of intraspecific ploidy level diversity, different cell-cycle or life-cycle stages (Vaulot *et al.* 1994; Lemaire *et al.* 1999; Gerashchenko *et al.* 2001; Houdan *et al.* 2004; Kremp and Parrow 2006; Van Dolah *et al.* 2008; Takahashi 2017; Salgado *et al.* 2017; Almeida *et al.* 2019). The latter is particularly applicable to microalgae, where the transitions between life cycle stages may be difficult to detect. In addition, it is possible to distinguish taxa and identify cryptic species on the basis of genome size differences (Figueroa *et al.* 2010).

1.3.2. Estimation of genome size in microalgae

Various methods have been used to estimate genome size in microalgae, including DNA reassociation kinetics (Rawson *et al.* 1979), pulsed-field gel electrophoresis (Yamamoto *et al.* 2001; Derelle *et al.* 2002), microdensitometry (Kapraun 2005, 2007), real-time qPCR (Créach *et al.* 2006; von Dassow *et al.* 2008) or Feulgen microdensitometry (Muravenko *et al.* 2001). Nowadays, considerable amounts of genome size data are derived from whole-genome sequencing (WGS, e.g., Armbrust 2004; Merchant *et al.* 2007; Read *et al.* 2013; Nelson *et al.* 2019). The selection of strains for WGS (or the successfully obtained WGS data) is, however, biased towards microalgae possessing small genomes and thus does not provide a realistic estimate of the scope of genome size diversity among microalgae. In

addition, the routine use of WGS as a method of choice for genome size estimation has been discouraged due to its poor quantification of genomic repeatom content (i.e., repetitive elements abundant for example in centromeres or telomeres) that may significantly underestimate the true DNA content (Doležel *et al.* 2007a; Doležel and Greilhuber 2010). On the other hand, repetitive elements may represent only a minor proportion of small genomes.

By far the most suitable method for genome size estimation is flow cytometry (FCM). This technique enables precise and rapid simultaneous analysis of thousands of fluorescent-stained nuclei in a stream of fluid (Doležel *et al.* 2007a). While FCM has found a broad spectrum of applications in genomic surveys on plants and animals (e.g., Dionisio Pires *et al.* 2004; Kron *et al.* 2007; Galbraith 2012; Chang *et al.* 2018; Sadílek *et al.* 2019), it is still only rarely applied in microalgal studies (but see Figueroa *et al.* 2010; Hyka *et al.* 2013). One of the reasons for this discrepancy is that the currently used FCM protocols are often not compatible with microalgae due to their diverse cell wall composition, various interferring pigments and secondary metabolites, or the presence of specific organelles and symbionts (Mazalová *et al.* 2011).

1.3.2.1. Current sample preparation protocols for FCM of microalgae

In the earlier studies, FCM was usually done by analysing whole intact cells. The cells can be analyzed either fresh (e.g., green algae, Pelagophyceae; Simon *et al.* 1994) or after fixation (e.g., in haptophytes; Vaulot *et al.* 1994). Various fixatives are used, such as ethanol, methanol, methanol : acetic acid mixture, formaldehyde, paraformaldehyde, or acetone (Mann and Stickle 1991; Veldhuis *et al.* 1997; LaJeunesse *et al.* 2005; Connolly *et al.* 2008; von Dassow *et al.* 2008; Figueroa *et al.* 2010; Whittaker *et al.* 2012). Following fixation, the sample is commonly washed with phosphate-buffered saline (PBS), methanol, or TE buffer (Mann and Stickle 1991; Connolly *et al.* 2008; von Dassow *et al.* 2008; Whittaker *et al.* 2012).

More recently, isolated nuclei have been preferred for FCM of microalgae. In species without a cell wall (e.g., chrysophytes), the nuclei can be released simply by osmotic bursting of cells by adding a hypotonic lysis buffer (e.g., Otto buffer; Olefeld *et al.* 2018). In microalgae that possess a cell wall, the protoplast content is very often isolated chemically using various enzymes (e.g., Mazalová *et al.* 2011; Weiss *et al.* 2011; Poulíčková *et al.* 2014). The mainly used enzymes are cellulase, macerozyme or lyticase, sometimes dissolved in a rinsing solution of PGly (composition: 27.2 mg·l^{-1} KH_2PO_4, 101 mg· l^{-1} KNO_3, 1117.6 mg· l^{-1} $CaCl_2$, 246 mg·l^{-1} $MgSO_4 \cdot 7H_2O$, 11.5 g· l^{-1} glycine, 18.016 g· l^{-1} glucose, 0.58572 g· l^{-1} MES and 65.58 g·l^{-1} mannitol). The enzymatic mixture dissolved in PGly was primarily developed for streptophyte algae, but also works for some other microalgae (Mazalová *et al.* 2011). In some cases (e.g., in *Zygnema* spp. or *Botryococcus braunii*), the enzymatic treatment needs to be

complemented with chopping the biomass using a razor blade (Mazalová *et al.* 2011; Weiss *et al.* 2011). Interestingly, the razor blade chopping alone has never been used to isolate nuclei of microalgae in order to estimate their nuclear DNA content by FCM. This is in contrast to plant FCM studies where the razor blade chopping is routinely used (Loureiro *et al.* 2021). Alternatively, phosphate buffered saline (PBS) was applied to lyse the cell walls in the marine dinoflagellate *Oxyrrhis marina* (Whiteley *et al.* 1993).

2. Aims and model group

The overall aim of my PhD project is to enhance our understanding of genome size variation in microalgae and its evolutionary consequences. Specifically, to make a thorough literature survey providing an overview of the currently used FCM protocols for microalgae and pointing to their limitations. Second, to develop new FCM protocols particularly suitable for problematic microalgae. Finally, to thoroughly investigate DNA content variation in chrysophytes and its ecophysiological consequences.

General objectives of the thesis are as follows:

Objective 1 – literature survey of the currently used FCM protocols for microalgae (paper I, paper II)
Specific questions - What are the routinely used protocols for FCM of microalgae? What are their advantages and disadvantages? What are the specific problems of FCM of microalgae compared to its application in plants and animals?

Objective 2 – developing new FCM protocols for DNA content estimation in problematic microalgal groups (paper III)
Specific questions - Are razor blade chopping and bead beating suitable methods for cell disruption and nuclei isolation in microalgae? How important is the chemical composition of a lysis buffer? Does the applicability of FCM protocols differ between filamentous and single-celled microalgae?

Objective 3 – DNA content variation in chrysophytes and its ecophysiological consequences (paper IV, paper V)
Specific questions - What is the extent of intraspecific DNA content variability in particular chrysophyte species? Do the patterns of intraspecific / intrastrain DNA content variation correspond to ploidy level shifts? Is the variability in DNA content linked to genomic GC content variation? Are there any morphological or physiological consequences of varying genome size? Is the DNA content variation among strains reflected in their ecogeographical distribution?

2.1. Model group Chrysophytes

The main model group of my PhD project are chrysophytes. They are also known as golden-brown algae due to the presence of photosyntetic pigment fucoxanthin, which gives them their brownish colour (Jeffrey *et al.* 2011). The chrysophytes are single-celled or colonial flagellates, which occur primarily in freshwater phytoplankton (but see Shi *et al.* 2011; Jeong *et al.* 2019). They are often restricted to cold waters and their blooms can cause an unpleasant fishy odour in drinking water reservoirs (Nicholls and Gerrath 1985). Although they do not possess a cell wall, silica-scaled chrysophytes (e.g., the representatives of the genera *Synura* or *Chrysosphaerella*) incorporate silicic acid and form species-specific siliceous scales on their plasmatic membrane (Kristiansen and Škaloud 2017). The silica-scaled chrysophytes are often used as bioindicators, including fossil taxa, since their scales may remain preserved in the environment. However, knowledge about the chrysophyte life cycle is limited to few publications documenting a sexual reproduction, specifically the fusion of cells that serve as gametes and the subsequent formation of a cyst (Wawrik 1972; Sandgren 1981). In recent years, the chrysophytes have drawn attention due to their remarkable DNA content diversity, ranging from 0.09 to 24.85 pg (0.09 to 24.31 Gbp; Veldhuis *et al.* 1997; Olefeld *et al.* 2018). It has been shown that this variation correlates with nutritional mode as the phototrophic species tent to have larger genomes compared to their heterotrophic counterparts (Olefeld *et al.* 2018). Further, the genome size variation is accompanied by numerous cases of major intraspecific variability (Olefeld *et al.* 2018; Majda *et al.* 2021), which was partially attributed to polyploidization events (Majda *et al.* 2021).

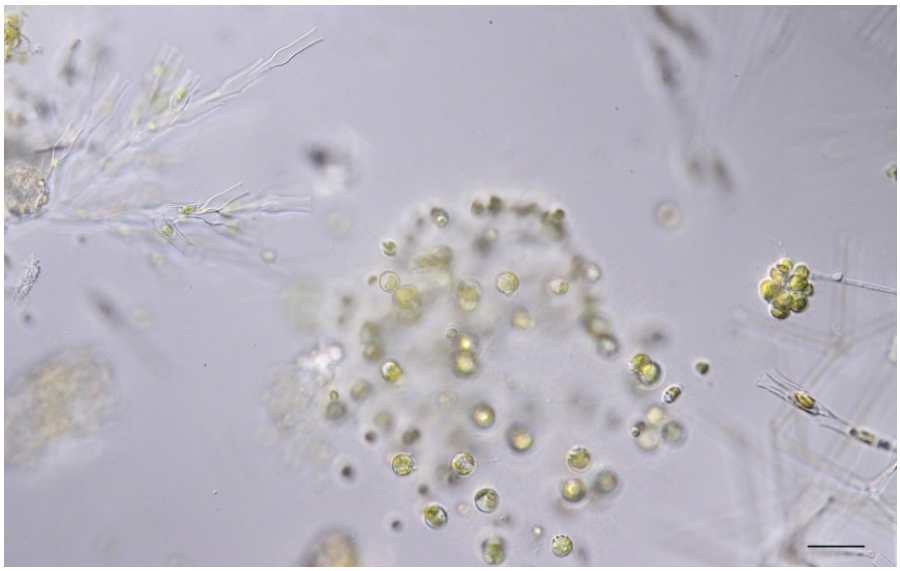


Fig. 2. Microphotographs of various chrysophytes, from the left: *Dinobryon divergens*, *Uroglena zachariasii*, and *Synura echinulata*. Scale bar represent 30 μm (courtesy of Martin Pusztai).

3. Methods

The particular methods used reflected specific objectives of each study. Hundreds of investigated microalgal samples were obtained from culture collections or sampled in natural habitats (the latter in case of most chrysophytes). Chrysophyte strains were identified by sequencing the internal transcribed spacer of nuclear ribosomal DNA (nu ITS rDNA). The key technique used in this PhD project was flow cytometry (FCM). I employed propidium iodide (PI) FCM to estimate absolute nuclear DNA contents of investigated strains. In order to obtain a suitable FCM protocol, I tested and compared several nuclei isolation techniques and lysis buffers on a set of selected microlagal strains. To ensure sufficient precision and robustness of the DNA content measurements, I implemented the following steps. I derived each genome size estimate from a mean value of (at least) three separate analyses on different days. In case of intraspecific variability, I provided a simultaneous analysis of strains differing in their DNA content to confirm the existing variation. Also, to examine the DNA content stability over generations, I re-analyzed selected strains during a long-term cultivation. To reveal a potential link between the nuclear DNA amount and a genome-wide proportion of GC bases, I analyzed the genomic CG content of particular strains by combining PI and DAPI FCM. To address potential phenotypic differences among investigated strains with different DNA contents, I examined cell size and growth rates. The cell size was analyzed either using flow imaging microscopy (FlowCAM) or optical microscopy. The growth rate was approximated from chlorophyll fluorescence yield (F_0) measured with a PAM flourometer. In order to assess putative ecogeographical trends in the distribution of genome size diversity, I examined associations between the DNA content of strains and climatic variables extracted from the Worldclim database.

4. Key results and conclusions

Particular papers included in the thesis are referred to in the following text by their corresponding Roman numerals (e.g., P–I = Paper I).

4.1. Difficulties with analyzing genome size of microalgae using FCM

Although FCM is the most suitable technique for genome size estimation, its application to microalgae is methodologically more challenging and time-consuming compared to plant and animal studies (**P-I**). Moreover, additional problems might result from general complications of analyzing small genomes using FCM (**P-II**). In papers **P-I** and **P-II** , I summarized the FCM protocols used for microalgae and microorganisms possessing small genomes, addressed their limitations and provided best practice recommendations.

The difficulties start with obtaining sufficient amounts of biomass. The cultivation is required, however, for many microalgae, it may be problematic (**P-I**). To achieve sufficient amounts of biomass is not only time-consuming but, in some cases, impossible (e.g., in some chryosphytes: *Uroglena* spp., *Uroglenopsis* spp. or *Chrysosphaerella* spp.; personal observations). Further, in case of heterotrophic microalgae (e.g., some dinoflagellates or cryptophytes), prey must be often added to the culture, and later it may be difficult to differentiate prey nuclei from those of the targeted microalgae (**P-I**). This is probably the reason to why heterotrophic microalgae have been largely avoided as a subject of FCM surveys (but see Whiteley *et al.* 1993). Similarly, it may be challenging to separate the nuclei of a studied sample from its symbionts. This is why some authors switched to specific life-cycle stages, such are Parrow and Burkholder (2002), who prioritized the analysis of zoospores in dinoflagellates.

Second, in microalgae it is often difficult to release and extract nuclei from the cells. This generally corresponds to cell wall heterogenity and complexity of its components in particular groups. Although an enzymatic treatment is often proposed as a solution (**P-I**), the enzymes may not digest the cell wall completely or only cells in specific stage might be suitable for FCM (Mazalová *et al.* 2011). Moreover, the enzymatic treatment is labour-intensive, time-consuming and often requires optimization specific to the studied group of microscopic algae. Therefore, enzymatic treatment is far from being the optimal isolation method (**P-I**). Because not many alternatives for microalgae are currently available, this has triggered an urgent need to develop and establish new nuclei isolation protocols specific to various microalgal groups. Two alternatives to enzymatic treatments were recently published in my study **P-III**. There, I introduced FCM protocols implementing bead beating and razor blade chopping after desiccating the biomass as nuclei isolation techniques. These protocols were developed with a particular focus on DNA content estimation in problematic algal groups. Both methods are easy to use and allowed

for the first time reliable DNA content estimation in green algae *Chlamydomonas noctigama*, *Microglena* sp., *Stigeoclonium* sp., and in raphidophyte *Gonyostomum semen*, among others. Chopping of desiccated biomass is particularly suitable for filamentous microalgae, specifically for Zygnematophyceae algae *Spirogyra* sp. and *Zygnema* spp., and ulvophyte algae *Trentepohlia* sp. In combination with the lysis buffer LB01, it was also the most successful nuclei isolation method of six tested protocols. This study also highlighted the importance of isolation buffer composition, since the performance of particular buffers used differed completely. The bead-beating of cells in a mixer mill is, on the other hand, convenient for single-celled species, and was especially suitable for raphidophyte *Gonyostomum semen*. Interestingly, in some cases, simple razor blade chopping of fresh biomass provided better results, as seen on the example of chlorophyte *Stigeoclonium* sp. or xanthophyte *Tribonema vulgare*, where this method resulted in more reduced background noise in FCM histograms compared to the chopping of desiccated biomass (**P-III**).

Another way how to overcome the difficulties with nuclei isolation is to analyse whole intact cells. However, this generally result in incompletely-stained nuclei and prominent cytoplasmatic and cell wall autofluorescence (**P-I**). The prominent background noise is, in fact, one of the major challenges when analyzing microscopic algae using FCM, even with properly isolated nuclei (**P-II**). Microalgae usually contain a wide variety of pigments and secondary metabolites that commonly interfere with flourescent staining and may act as staining inhibitors (e.g., phenols or tannins; Simon *et al.* 1994; Veldhuis *et al.* 1997; Loureiro *et al.* 2006; Kapraun 2007; Mazalová *et al.* 2011). In addition, the autofluorescence spectrum of photosynthetic pigments may overlap the emission spectrum of the fluorochromes used for DNA content measurement, e.g., propidium iodide (PI). The PI fluorochrome can also bind to polysaccharides from the remaining cell walls and thus contribute to the background noise and lower precision of of the measurement (i.e., increase of CV; Potter *et al.* 2016). The unwanted fluorescence from pigments and secondary metabolites is often reduced by chemical fixation of samples (**P-I**). Unfortunately, the chemical fixation brings many problems of its own and is not recommended for absolute DNA content estimation due to reduced quality of FCM analyses (Doležel *et al.* 2007b). A more suitable approach is to analyse very young cultures that have not yet accumulated as much secondary metabolites compared to cells in a stationary phase of their growth (**P-I**). Alternatively, the effect of secondary metabolites and pigments may be reduced, to some extent, by adding PVP (polyvinylpyrrolidone) and/or mercaptoethanol to the lysis buffer (Loureiro *et al.* 2021).

Further, additional difficulties are linked to a broad lack of microalgal FCM standards. This is also one of the reasons why chicken red blood cells (CRBCs) are still widely used as FCM standards for analysis of microscopic algae (e.g., Connolly *et al.* 2008; Standeren 2018). However, the nuclear DNA of chicken red blood cells has a high packing density, and red blood cells from male and female differ in genome size by 2.7 % due to the contributions of sex chromosomes, which leads to non-systematic errors in DNA content estimates (Nakamura *et al.* 1990; Hardie *et al.* 2002). There are only few established microalgal FCM standards (**P-I**). Cell-wall deficient mutant of green algae *Chlamydomonas reinhardtii* CC-400 cw15 mt+ (2C = 0.24 pg; Lemaire *et al.* 1999; Potter *et al.* 2016). The desmid *Micrasterias pinnatifida* SVCK 411 (2C = 3.4 pg; Mazalová *et al.* 2011), which, unfortunately, requires enzymatic treatment in order to release the nuclei (see above, Mazalová *et al.* 2011). More recently, the chrysophyte *Synura sphagnicola* CCAC 2959 B (2C = 0.4 pg) was introduced as a FCM standard (Olefeld *et al.* 2018). However, in the study **P-V** we revealed that the chrysophytes have a haploid-diploid life cycle and are able to alternate their life cycle stages (and thus nuclear DNA contents) during cultivation. Since their life cycle stages are indistinguishable (isomorphic), we would discourage from using the chrysophytes as FCM standards (**P-V**). As an alternative to established FCM standards, some authors use microalgae with available complete genome sequences, e.g., the diatom *Thalassiosira pseudonana* strain CCMP1335 (2C = 0.07 pg; Armbrust 2004; von Dassow *et al.* 2008). Yet, the use of genome-sequenced taxa may also not be optimal due to a potential underestimation of the total DNA content and a bias towards small genomes (see above; Doležel *et al.* 2007a; Doležel and Greilhuber 2010).

Also, additional complications might arise when the FCM outcomes are being interpreted since our knowledge of intraspecific genome size variability, the extent of polyploidization, and life cycles in microalgae remains poorly understood.

4.2. Intraspecific genome size variability in microalgae

Even though there is only a limited number of studies on genome size in microalgae, several cases of major intraspecific variation have been documented. The intraspecific variation, reaching up to 7-fold differences, have been described in desmids (*Micrasterias rotata*, *Triploceras gracile*; Poulíčková *et al.* 2014), haptophytes (*Emiliania huxleyi*; Medlin *et al.* 1996; Veldhuis *et al.* 1997; Read *et al.* 2013), diatoms (*Thalassiosira weissflogii*; von Dassow *et al.* 2008), or chrysophytes (*Synura macropora*; *S. petersenii*; **P-IV**, **P-V**). Nonetheless, the frequency of this phenomenon and its prevalence across various groups of microalgae is still poorly documented. With more than 130 analysed *S. petersenii* strains, the study **P-IV** represents to our knowledge the most comprehensive intraspecific genome size survey conducted on microscopic algae so far.

Identifying particular evolutionary mechanisms responsible for intraspecific genome size variation is often quite challenging. Smaller extents of genome size variability may arise from proliferation of TEs, unequal frequency of insertions to deletions and multiplication of larger genomic segments or entire chromosomes (aneuploidization). For example, a recent chromosomal duplication was revealed in diatom *T. pseudonana* (Armbrust 2004). Genome size shifts caused by chromosomal aberrations or increased TE activity may be accompanied by significant alterations of a genomic GC content (i.e., the relative proportion of GC base pairs; Wichman *et al.* 1993; Armbrust 2004; Derelle *et al.* 2006). Although in the study **P-IV** we were unable to identify the mechanism(s) of genome size change in chrysophyte *S. petersenii*, it seems to operate via gradual changes in genome size and lead to shifts in genomic GC content, making proliferation of TEs and/or multiplication of larger genomic segments the most likely candidates (**P-IV**). Moreover, it has been shown that even multiple-fold genome size differences can be caused by proliferation of TEs (Blommaert *et al.* 2019; Naville *et al.* 2019; Wong *et al.* 2019) or satellite DNA (Shah *et al.* 2020; Stelzer *et al.* 2021). Most cases of intraspecific genome size variability in microalgae are, however, attributed to polyploidization events (e.g., King 1960; Kapraun 2005; Bowler *et al.* 2008; Koester *et al.* 2010; Poulíčková *et al.* 2014). Polyploids were described in Zygnematophyceae algae (e.g., *Spirogyra communis*, *Netrium digitus*; King 1960; Wang *et al.* 1985) or, more recently, in chrysophytes (*Poteriospumella lacustris*, *Synura glabra*, *S. heteropora*; Majda *et al.* 2019), **P-V**). Interestingly, polyploids may be maintained in populations of unicellular microalgae by prevailing asexual reproduction (via mitotic division; **P-V**). However, karyological evidence is always needed to confirm the polyploidization, and karyotyping of small microscopic algae has proven to be challenging (**P-IV**). Alternatively, two-fold DNA content differences among analyzed microalgal samples might be attributed to the simultaneous presence of different life cycle stages.

4.3. Life cycles in microalgae

A diversity of life cycles has been observed across microalgae. However, because of the microscopic size of these organisms and a frequent lack of pronounced morphological features, it is sometimes hard to distinguish particular life-cycle stages or detect their transitions. As a consequence, our understanding of life cycles is still rather fragmentary, especially in certain microalgal groups.

Some green microalgae (e.g., *Ulothrix* spp.) dispose of haploid life cycle (Mable and Otto 1998). As microalgae were seen as simple organisms in the past (Copeland 1956), it was expected for a long time that the haploid life cycle predominantes in

microalgae. However, a diploid life cycle, mostly known from animals, occurs also among diatoms or raphidophytes (Figueroa and Rengefors 2006; Montresor *et al.* 2016; Figueroa *et al.* 2018). Finally, a haploid-diploid life cycle evolved in haptophytes, foraminifera, some dinoflagellates, and some cryptophytes and in many groups of green algae (Richerd *et al.* 1993; Mable and Otto 1998; Rousseau *et al.* 2007; Speijer *et al.* 2015; Figueroa *et al.* 2018). In most of these groups, different life cycle stages have distinct morphology. On the other hand, in an isomorphic haploid-diploid life cycle, the haploid and diploid phases are morphologically indistinguishable. Such peculiar life cycle evolved in some green and red multicellular algae, as in sea lettuce (*Ulva lactuca*; Destombe *et al.* 1989; Wichard *et al.* 2015). The first case of the isomorphic haploid-diploid life cycle among unicellular algae was reported from chrysophytes by our study **P-V**. With the exception of a small increase in cell size with higher ploidy, both life cycle stages in chrysophytes were morphologically indistinguishable. Interestingly, even the life stage transitions were captured by FCM during a long-term cultivation of strains (specifically, in *Chrysosphaerella brevispina*, *Ochromonas tuberculata* and *Synura* spp.). Similarly, the sexual reproduction in cultivation was documented in diatoms (Quijano-scheggia *et al.* 2009). In chrysophytes, both haploid and diploid stages are capable of mitotic growth and long-term survival in cultivation (**P-V**). The observed life cycle transitions between haploid and diploid ploidy level indicate that diploid strains regularly undergo meiosis in cultivation and thus the haploids must also play the role of gametes (**P-V**). Although the sexual reproduction has been previously documented in chrysophytes (Sandgren and Flanagin 1986), we have never directly observed it (**P-IV**, **P-V**). Given that each FCM measurement in the study **P-V** was conducted on thousands of cells, the life stage transitions appears to be synchronized among cells in cultivation and the same can be expected for natural populations of chrysophytes. This might ensure that they enter particular life cycle stages or produce their gametes synchronously. The life stage transitions might be triggered by chemical signalization. Production of pheromones, allowing synchronization of individuals prior to the mating process, was already described in green algae (e.g., *Volvox carteri*; Al-Hasani and Jaenicke 1992). Similarly, chemical signalling is involved in sexual induction of some diatoms (Moeys *et al.* 2016). Also, the life cycle transitions may occur very rapidly (**P-V**). For example, in chrysophytes *Ochromonas tuberculata* and *Synura soroconopea*, two ploidy level shifts were observed within two weeks. On the other hand, most other chrysophyte species seemed to alternate in ploidy considerably more slowly (**P-V**).

4.4. Stability of DNA content during long-term cultivation

The life cycle stage transitions are one of the reasons why we might observe changes in nuclear DNA content over long-term cultivation. Similarly, the DNA content variation in cultivated strains may be introduced by meiosis, occasionally detected in some groups of microalgae (Quijano-scheggia *et al.* 2009; **P-V**). In other cases, genome size changes in cultivation arise by different mechanisms, suggesting independent evolution of strains *in vitro*. In diatom *T. weissflogii*, DNA content differences between cultivated substrains were attributed to gene duplications, aneuploidization and polyploidization (von Dassow *et al.* 2008). Because the extent of this phenomenon is unknown and the DNA content stability during cultivation is only rarely addressed in studies (but see LaJeunesse *et al.* 2005; von Dassow *et al.* 2008; Koester *et al.* 2010; **P-IV**; **P-V**), researchers should be very cautious when interpreting intraspecific genome size variability and rather avoid a long-term cultivation before genome size analysis of microalgal species (**P-I**). In other studies, the DNA content of investigated strains remained stable during cultivation (LaJeunesse *et al.* 2005; **P-IV**), which is consistent with the general rule presuming that individuals within population share a constant nuclear DNA content (Swift 1950).

4.5. Evolutionary costs and benefits of small vs. large genomes in microalgae

An important aspect of intraspecific genome size diversity and its evolutionary maintenance is its putative adaptive potential, i.e., whether strains with a certain genome size have a fitness (dis-)advantage in some environmental or evolutionary context. The numerous associations between genome size and phenotypic traits seem to support that genome size variation may be under selection. A widely observed genome size – cell size correlation was found among countless groups of microalgal species (e.g., LaJeunesse *et al.* 2005; Connolly *et al.* 2008; von Dassow *et al.* 2008; Poulíčková *et al.* 2014; Olefeld *et al.* 2018) but also among individuals of the same species (**P-IV**; **P-V**). In unicellular algae, the cell size is particularly important since it fundamentally relates to metabolic rate, generation time or growth rate, which can further affect species temperature optima, dispersal abilities, or susceptibility to herbivores (Shuter *et al.* 1983; Cavalier-Smith 2005; **P-IV**). Specifically, as a result of the genome size – growth rate correlation, lineages possessing lower DNA contents grow (their cells divide) faster compared to lineages with higher DNA contents. This feature should be reflected in their relative fitness at least under specific environmental conditions. In aquatic microalgae, rapid population growth is a key factor for successful colonization of a new site and

effective monopolization of local resources (i.e. the monopolization hypothesis; De Meester *et al.* 2002). Once a population is well established and possibly also locally adapted, the existence of a large bank of resting propagules (e.g., chrysophyte or dinoflagellate cysts) provides a powerful buffer against newly invading lineages. Under this scenario, microalgal strains with larger genomes should be inferior colonizers of new sites, possibly sometimes outcompeted at the existing localities by other strains with smaller genomes.

Moreover, as a consequence of cell geometry, larger cells have lower surface area to volume ratio than small ones of identical shape. Thus, a lower DNA content results in cells with a relatively higher surface area which should be more efficient in nutrient uptake and at the same time, they may require lower energetic costs to maintain their life functions and have faster cell division (Lewis 1985; Hughes and Otto 1999). The "nutrient limitation hypothesis" by Lewis (1985) then predicts that cells with smaller genomes will better exploit low-nutrient environments, and this should be especially applicable to unicellular planktonic autotrophs (e.g., chrysophytes, dinoflagellates or coccolithophores). The frequency distribution of genome size categories across populations of chrysophyte *Synura petersenii* (**P-IV**) was well in line with such prediction and strains with smaller genomes were most common. However, when the distribution of *S. petersenii* strains from distinct genome size categories was explored in an ecogeographic context (**P-IV**), no signs of a large-scale environmental filtering were detected and the spatial distribution of genome size diversity seemed rather random.

It should be noted that the observed retention of substantial genome size variability within species implies there could be also some evolutionary advantages to lineages with higher DNA contents. For instance, larger cells of autophototrophs could have more efficient photosynthesis, as suggested by Finkel *et al*. (2001). Accordingly, photosynthetic chrysophytes have both larger cells and genomes compared to heterotrophic taxa (Olefeld *et al.* 2018). Also, due to a relatively smaller surface area interacting with the external environment, cells possessing larger DNA content might be preadapted to better tolerate toxic environment (Otto and Gerstein 2008). Another advantege of increased genome size is that it is often coupled with increased gene dosage (e.g. via aneuploidsation or polyploidisation), leading to higher secondary metabolite production (Mason 2016). Alternatively, the occurrence of intraspecific strains with higher DNA contents may reflect enhanced rates with which mutants with enlarged genomes originate in nature (e.g., predispositions for certain types of chromosomal mutations) and/or a low strength of selection on genome size in the surveyed populations.

Interestingly, similar associations could apply to different ploidy phases of isomorphic haploid-diploid life cycles in microalgae (Otto and Gerstein 2008). For example, the most common phenotypic effect of higher ploidy is increase in cell size (Bennett 1971; Hughes and Otto 1999), and the same was also observed between different ploidy phases in chrysophytes (**P-V**). From an ecological perspective, the existence of different life cycle stages may provide for more efficient specialization, better exploitation of the environment, and thus increasing the evolutionary success of species (Thornber 2006). In addition, microalgae with isomorphic haploid-diploid life cycle could alternate their life cycle stages depending on environmental conditions. Interestingly, the diploid stages of chrysophytes prevailed in cultivation (**P-V**). I can only speculate that nutrient rich environments, supplemented in the study with cultivation medium, may favour the diploid life cycle stage similarly to increased presence of polyploid plants on fertilized sites (Guignard *et al.* 2016).

5. Literature cited

Al-Hasani H, Jaenicke L. **1992**. Characterization of the sex-inducer glycoprotein of *Volvox carteri* f. *weismannia*. *Sexual Plant Reproduction* **5**: 8–12.

Almeida AC, Gomes T, Habuda-Stanić M, *et al.* 2019. Characterization of multiple biomarker responses using flow cytometry to improve environmental hazard assessment with the green microalgae *Raphidocelis subcapitata*. *Science of the Total Environment* **687**: 827–838.

Armbrust E V. 2004. The genome of the diatom *Thalassiosira Pseudonana*: ecology, evolution, and metabolism. *Science* **306**: 79–86.

Baack EJ, Whitney KD, Rieseberg LH. **2005**. Hybridization and genome size evolution: timing and magnitude of nuclear DNA content increases in *Helianthus* homoploid hybrid species. *New Phytologist* **167**: 623–630.

Beaulieu JM, Leitch IJ, Patel S, Pendharkar A, Knight CA. **2008**. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* **179**: 975–986.

Bennett MD. **1971**. The duration of meiosis. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **178**: 277–299.

Bennett MD. **1972**. Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society. B. Biological Sciences* **181**: 109–135.

Bennett MD. **1987**. Variation in genomic form in plants and its ecological implications. *New Phytologist* **106**: 177–200.

Bennett MD, Leitch IJ. **2011**. Nuclear DNA amounts in angiosperms: targets, trends and tomorrow. *Annals of Botany* **107**: 467–590.

Blommaert J, Riss S, Hecox-Lea B, Mark Welch DB, Stelzer CP. **2019**. Small, but surprisingly repetitive genomes: transposon expansion and not polyploidy has driven a doubling in genome size in a metazoan species complex. *BMC Genomics* **20**: 1–12.

Bouligand Y, Soyer MO, Puiseux-Dao S. **1968**. La structure fibrillaire et l'orientation des chromosomes chez les Dinoflagellés. *Chromosoma* **24**: 251–287.

Bowler C, Allen AE, Badger JH, *et al.* 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* **456**: 239–244.

Brennan L, Owende P. **2010**. Biofuels from microalgae—A review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews* **14**: 557–577.

Burki F, Roger AJ, Brown MW, Simpson AGB. **2020**. The new tree of eukaryotes. *Trends in Ecology and Evolution* **35**: 43–55.

Cantwell H, Nurse P. **2019**. Unravelling nuclear size control. *Current Genetics* **65**: 1281–1285.

Cavalier-Smith T. **2005**. Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. *Annals of Botany* **95**: 147–175.

Chang P, Tseng Y-F, Chen P-Y, Wang C-JR. **2018**. Using flow cytometry to isolate maize meiocytes for next generation sequencing: a time and labor efficient method. *Current Protocols in Plant Biology* **3**: e20068.

Connolly JA, Oliver MJ, Beaulieu JM, Knight CA, Tomanek L, Moline MA. **2008**. Correlated evolution of genome size and cell volume in diatoms (Bacillariophyceae). *Journal of Phycology* **44**: 124–131.

Copeland HF. **1956**. *The classification of lower organisms*. CA, USA: Palo Alto, California, Pacific Books.

Corradi N, Pombert J-F, Farinelli L, Didier ES, Keeling PJ. **2010**. The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. *Nature Communications* **1**: 77.

Coyne JA, Orr HA. **2004**. *Speciation*. Sunderland, MA, USA: Sinauer Associates.
Créach V, Ernst A, Sabbe K, Vanelslander B, Vyverman W, Stal LJ. **2006**. Using quantitative PCR to determine the distribution of a semicryptic benthic diatom, *Navicula phyllepta* (Bacillariophyceae). *Journal of Phycology* **42**: 1142–1154.
von Dassow P, Petersen TW, Chepurnov VA, Armbrust E. **2008**. Inter- and intraspecific relationships between nuclear DNA content and cell size in selected members of the centric diatom genus *Thalassiosira* (Bacillariophyceae). *Journal of Phycology* **44**: 335–349.
Derelle E, Ferraz C, Lagoda P, *et al.* 2002. DNA libraries for sequencing the genome of *Ostreococcus tauri* (Chlorophyta, Prasinophyceae): The smallest free-living eukaryotic cell. *Journal of Phycology* **38**: 1150–1156.
Derelle E, Ferraz C, Rombauts S, *et al.* 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences* **103**: 11647–11652.
Destombe C, Valero M, Vernet P, Couvet D. **1989**. What controls haploid—diploid ratio in the red alga, *Gracilaria verrucosa*? *Journal of Evolutionary Biology* **2**: 317–338.
Devos KM, Brown JKM, Bennetzen JL. **2002**. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome research* **12**: 1075–9.
Dionisio Pires LM, Jonker RR, Van Donk E, Laanbroek HJ. **2004**. Selective grazing by adults and larvae of the zebra mussel (*Dreissena polymorpha*): application of flow cytometry to natural seston. *Freshwater Biology* **49**: 116–126.
Van Dolah FM, Leighfield TA, Kamykowski D, Kirkpatrick GJ. **2008**. Cell cycle behavior of laboratory and field populations of the Florida red tide dinoflagellate, *Karenia brevis*. *Continental Shelf Research* **28**: 11–23.
Doležel J, Greilhuber J. **2010**. Nuclear genome size: Are we getting closer? *Cytometry Part A* **77**: 635–642.
Doležel J, Greilhuber J, Suda J. **2007a**. *Flow Cytometry with Plant Cells*. Weinheim: Wiley-VCH Verlag GmbH.
Doležel J, Greilhuber J, Suda J. **2007b**. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**: 2233–2244.
Doolittle WF, Sapienza C. **1980**. Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603.
Falkowski PG, Fenchel T, Delong EF. **2008**. The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034–9.
Figueroa RI, Estrada M, Garcés E. **2018**. Life histories of microalgal species causing harmful blooms: haploids, diploids and the relevance of benthic stages. *Harmful Algae* **73**: 44–57.
Figueroa RI, Garcés E, Bravo I. **2010**. The use of flow cytometry for species identification and life-cycle studies in dinoflagellates. *Deep-Sea Research Part II: Topical Studies in Oceanography* **57**: 301–307.
Figueroa RI, Rengefors K. **2006**. Life cycle and sexuality of the freshwater raphidophyte *Gonyostomum semen* (Raphidophyceae). *Journal of Phycology* **42**: 859–871.
Finkel ZV, Platt T, Sathyendranath S, *et al.* 2001. Light absorption and size scaling of light-limited metabolism in marine diatoms. *Limnology and Oceanography* **46**: 86–94.
Foissner W. **2007**. Protist diversity and distribution: some basic considerations In: Dordrecht: Springer, 1–8.
Galbraith DW. **2012**. Flow cytometry and fluorescence-activated cell sorting in plants: the past, present, and future. *Biomédica* **30**: 65.
Gerashchenko BI, Kosaka T, Hosoya H. **2001**. Growth kinetics of algal populations exsymbiotic from *Paramecium bursaria* by flow cytometry measurements. *Cytometry* **44**: 257–263.
Gregory TR. **2001a**. Coincidence, coevolution, or causation? DNA content, cell size, and the C-

value enigma. *Biological Reviews of the Cambridge Philosophical Society* **76**: 65–101.
Gregory TR. **2001b**. The bigger the C-value, the larger the cell: genome size and red blood cell size in Vertebrates. *Blood Cells, Molecules, and Diseases* **27**: 830–843.
Gregory TR. **2002**. A bird's-eye view of the C-value enigma: genome size, cell size, and metabolic rate in the class Aves. *Evolution* **56**: 121–130.
Gregory TR. **2004**. Macroevolution, hierarchy theory, and the C-value enigma. *Paleobiology* **30**: 179–202.
Guignard MS, Nichols RA, Knell RJ, *et al.* 2016. Genome size and ploidy influence angiosperm species' biomass under nitrogen and phosphorus limitation. *New Phytologist* **210**: 1195–1206.
Hardie DC, Gregory TR, Hebert PDN. **2002**. From pixels to picograms: A beginners' guide to genome quantification by Feulgen image analysis densitometry. *Journal of Histochemistry and Cytochemistry* **50**: 735–749.
Houdan A, Bonnard A, Fresnel J, Fouchard S, Billard C, Probert I. **2004**. Toxicity of coastal coccolithophores (Prymnesiophyceae, Haptophyta). *Journal of Plankton Research* **26**: 875–883.
Hughes JS, Otto SP. **1999**. Ecology and the evolution of biphasic life cycles. *American Naturalist* **154**: 306–320.
Hyka P, Lickova S, Přibyl P, Melzoch K, Kovar K. **2013**. Flow cytometry for the development of biotechnological processes with microalgae. *Biotechnology Advances* **31**: 2–16.
Igea J, Miller EF, Papadopulos AST, Tanentzap AJ. **2017**. Seed size and its rate of evolution correlate with species diversification across angiosperms. *PLOS Biology* **15**: e2002792.
Jeffrey SW, Wright SW, Zapata M. **2011**. Microalgal classes and their signature pigments In: Roy S, Llewellyn CA, Egeland ES, Johnsen G, eds. *Phytoplankton Pigments: Characterization, Chemotaxonomy and Applications in Oceanography*. United Kingdom: Cambridge University Press, 3–77.
Jeong M, Kim JI, Jo BY, Kim HS, Siver PA, Shin W. **2019**. Surviving the marine environment: two new species of *Mallomonas* (Synurophyceae). *Phycologia* **58**: 276–286.
Kapraun DF. **2005**. Nuclear DNA Content estimates in multicellular green, red and brown algae: phylogenetic considerations. *Annals of Botany* **95**: 7–44.
Kapraun DF. **2007**. Nuclear DNA content estimates in green algal lineages: Chlorophyta and Streptophyta. *Annals of Botany* **99**: 677–701.
Khan MI, Shin JH, Kim JD. **2018**. The promising future of microalgae: current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. *Microbial Cell Factories* **17**.
King GC. **1960**. The cytology of the desmids: the chromosomes. *New Phytologist* **59**: 65–72.
Koester JA, Swalwell JE, Von Dassow P, Armbrust EV. **2010**. Genome size differentiates co-occurring populations of the planktonic diatom *Ditylum brightwellii* (Bacillariophyta). *BMC Evolutionary Biology* **10**: 1–11.
Kozłowski J, Konarzewski M, Gawelczyk AT. **2003**. Cell size as a link between noncoding DNA and metabolic rate scaling. *Proceedings of the National Academy of Sciences* **100**: 14 080–14 085.
Kremp A, Parrow MW. **2006**. Evidence for asexual resting cysts in the life cycle of the marine peridinoid dinoflagellate, *Scrippsiella hangoei*. *Journal of Phycology* **42**: 400–409.
Kristiansen J, Škaloud P. **2017**. Chrysophyta In: Archibald JM, Simpson AGB, Slamovits CH, eds. *Handbook of the Protists*. Springer International Publishing, 1657.
Kron P, Suda J, Husband BC. **2007**. Applications of flow cytometry to evolutionary and population biology. *Annual Review of Ecology, Evolution, and Systematics* **38**: 847–876.
LaJeunesse TC, Lambert G, Andersen RA, Coffroth MA, Galbraith DW. **2005**. *Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates. *Journal of Phycology* **41**: 880–886.
Lemaire S, Hours M, Gerard-Hirne C, Trouabal A, Roche O, Jacquot J-P. **1999**. Analysis of

light/dark synchronization of cell-wall-less *Chlamydomonas reinhardtii* (Chlorophyta) cells by flow cytometry. *European Journal of Phycology* **34**: 279–286.
Lewis WMJ. **1985**. Nutrient scarcity as an evolutionary cause of haploidy. *The American Naturalist* **125**: 692–701.
Lin S. **2006**. The smallest dinoflagellate genome is yet to be found: A comment on LaJeunesse et al. "*Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates." *Journal of Phycology* **42**: 746–748.
Loureiro J, Kron P, Temsch EM, *et al.* 2021. Isolation of plant nuclei for estimation of nuclear DNA content – overview and best practices. *Cytometry Part A* **99**: 318–327.
Loureiro J, Rodriguez E, Doležel J, Santos C. **2006**. Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. *Annals of Botany* **98**: 515–527.
Lynch M, Conery JS. **2003**. The origins of genome complexity. *Science* **302**: 1401–1404.
Mable BK, Otto SP. **1998**. The evolution of life cycles. : 453–462.
Majda S, Beisser D, Boenigk J. **2021**. Nutrient-driven genome evolution revealed by comparative genomics of chrysomonad flagellates. *Communications Biology* **4**: 1–11.
Majda S, Boenigk J, Beisser D. **2019**. Intraspecific variation in protists: clues for microevolution from *Poteriospumella lacustris* (Chrysophyceae). *Genome Biology and Evolution*. **11**: 2492–2504.
Maleszka R. **1993**. Electrophoretic analysis of the nuclear and organellar genomes in the ultra-small alga *Cyanidioschyzon merolae*. *Current Genetics* **24**: 548–550.
Mann DG, Stickle AJ. **1991**. The genus *Craticula*. *Diatom Research* **6**: 79–107.
Mason AS. **2016**. *Polyploidy and Hybridization for Crop Improvement* (AS Mason, Ed.). Boca Raton: Science Publisher, USA.
Mazalová P, Šarhanová P, Ondřej V, Poulíčková A. **2011**. Quantification of DNA content in freshwater microalgae using flow cytometry: A modified protocol for selected green microalgae. *Fottea* **11**: 317–328.
Medlin LK, Barker GLA, Campbell L, *et al.* 1996. Genetic characterisation of *Emiliania huxleyi* (Haptophyta). *Journal of Marine Systems* **9**: 13–31.
De Meester L, Gómez A, Okamura B, Schwenk K. **2002**. The Monopolization Hypothesis and the dispersal–gene flow paradox in aquatic organisms. *Acta Oecologica* **23**: 121–135.
Merchant SS, Prochnik SE, Vallon O, *et al.* 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**: 245–50.
Meyer A, Schloissnig S, Franchini P, *et al.* 2021. Giant lungfish genome elucidates the conquest of land by vertebrates. *Nature* **590**: 284–289.
Milano J, Ong HC, Masjuki HH, *et al.* 2016. Microalgae biofuels as an alternative to fossil fuel for power generation. *Renewable and Sustainable Energy Reviews* **58**: 180–197.
Mirsky AE, Ris H. **1951**. The desoxyribonucleic acid content of animal cells and its evolutionary significance. *The Journal of general physiology* **34**: 451–62.
Moeys S, Frenkel J, Lembke C, *et al.* 2016. A sex-inducing pheromone triggers cell cycle arrest and mate attraction in the diatom *Seminavis robusta*. *Scientific Reports* **6**: 1–13.
Montresor M, Vitale L, D'Alelio D, Ferrante MI. **2016**. Sex in marine planktonic diatoms: insights and challenges. *Perspectives in Phycology* **3**: 61–75.
Muravenko O, Selyakh I, Kononenko N, Stadnichuk I. **2001**. Chromosome numbers and nuclear DNA contents in the red microalgae *Cyanidium caldarium* and three *Galdieria* species. *European Journal of Phycology* **36**: 227–232.
Nakamura D, Tiersch TR, Douglass M, Chandler RW. **1990**. Rapid identification of sex in birds by flow cytometry. *Cytogenetic and Genome Research* **53**: 201–205.
Nardon C, Deceliere G, Loevenbruck C, Weiss M, Vieria C, Biémont C. **2005**. Is genome size influenced by colonization of new environments in dipteran species? *Molecular Ecology* **14**: 869–

878.
Naville M, Henriet S, Warren I, *et al.* 2019. Massive changes of genome size driven by expansions of non-autonomous transposable elements. *Current Biology* **29**: 1161-1168.e6.
Nelson DR, Chaiboonchoe A, Fu W, *et al.* 2019. Potential for heightened sulfur-metabolic capacity in coastal subtropical microalgae. *iScience* **11**: 450–465.
Nicholls KH, Gerrath JF. **1985**. The taxonomy of *Synura* (Chrysophyceae) in Ontario with special reference to taste and odour in water supplies. *Canadian Journal of Botany* **63**: 1482–1493.
Ohno S. **1972**. So much 'junk' DNA in our genome In: Smith HH, ed. *Evolution of Genetic Systems.* Brookhaven Symposium in Biology, 366–370.
Olefeld JL, Majda S, Albach DC, Marks S, Boenigk J. **2018**. Genome size of chrysophytes varies with cell size and nutritional mode. *Organisms Diversity & Evolution* **18**: 163–173.
Orgel LE, Crick FHC. **1980**. Selfish DNA: the ultimate parasite. *Nature* **284**: 604–607.
Otto SP, Gerstein AC. **2008**. The evolution of haploidy and diploidy. *Current Biology* **18**: R1121–R1124.
Palenik B, Grimwood J, Aerts A, *et al.* 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 7705–10.
Pellicer J, Leitch IJ. **2020**. The Plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. *New Phytologist* **226**: 301–305.
Peters AF, Marie D, Scornet D, Kloareg B, Mark Cock J. **2004**. Proposal of *Ectocarpus siliculosus* (Ectocarpales, Phaeophyceae) as a model organism for brown algal genetics and genomics. *Journal of Phycology* **40**: 1079–1088.
Potter EE, Thornber CS, Swanson JD, McFarland M. **2016**. Ploidy distribution of the harmful bloom forming macroalgae *Ulva* spp. in Narragansett Bay, Rhode Island, USA, using flow cytometry methods. *PLoS ONE* **11**: 1–15.
Poulíčková A, Mazalová P, Vašut RJ, Šarhanová P, Neustupa J, Škaloud P. **2014**. DNA content variation and its significance in the evolution of the genus *Micrasterias* (desmidiales, streptophyta). *PLoS One* **9**: e86247.
Priyadarshan PM. **2019**. Induced mutations and polyploidy breeding In: *Plant breeding: Classical to Modern*. Singapore: Springer Singapore, 329–370.
Pulz O, Gross W. **2004**. Valuable products from biotechnology of microalgae. *Applied Microbiology and Biotechnology* **65**: 635–648.
Pyšek P, Skálová H, Čuda J, *et al.* 2018. Small genome separates native and invasive populations in an ecologically important cosmopolitan grass. *Ecology* **99**: 79–90.
Qin Y, Zhang Yuehuan, Mo R, *et al.* 2019. Influence of ploidy and environment on grow-out traits of diploid and triploid Hong Kong oysters *Crassostrea hongkongensis* in southern China. *Aquaculture* **507**: 108–118.
Quijano-scheggia S, Garce E, Andree K, Fortun JM. **2009**. Homothallic auxosporulation in *Pseudo-Nitzschia brasiliana* (Bacillariophyta). **107**: 100–107.
Rawson JRY, Eckenrode VK, Boerma CL, Curtis S. **1979**. DNA sequence organization in the alga *Euglena gracilis*. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **563**: 1–16.
Read BA, Kegel J, Klute MJ, *et al.* 2013. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature* **499**: 209–213.
Ribeiro S, Berge T, Lundholm N, Ellegaard M. **2013**. Hundred years of environmental change and phytoplankton ecophysiological variability archived in coastal sediments. *PLoS One* **8**: e61184.
Richerd S, Couvet D, Valéro M. **1993**. Evolution of the alternation of haploid and diploid phases in life cycles. II. Maintenance of the haplo-diplontic cycle. *Journal of Evolutionary Biology* **6**: 263–280.
Rousseau V, Chrétiennot-Dinet MJ, Jacobsen A, Verity P, Whipple S. **2007**. The life cycle of

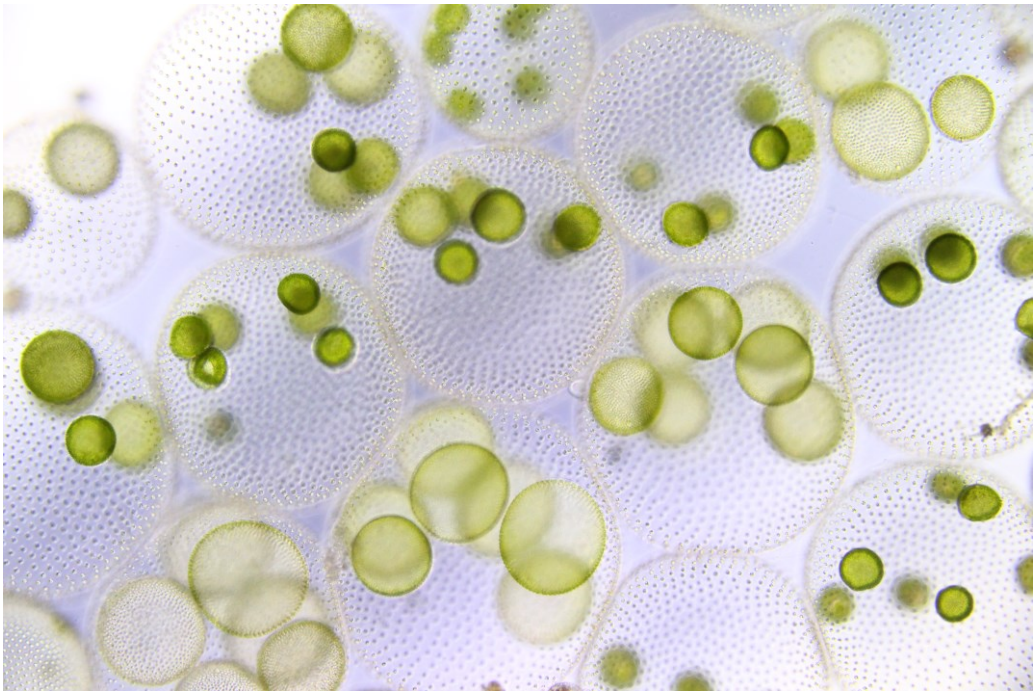
Phaeocystis: State of knowledge and presumptive role in ecology. *Biogeochemistry* **83**: 29–47.
Sadílek D, Urfus T, Vilímová J. **2019**. Genome size and sex chromosome variability of bed bugs feeding on animal hosts compared to *Cimex lectularius* Parasitizing Human (Heteroptera: Cimicidae). *Cytometry Part A* **95**: 1158–1166.
Salgado P, Figueroa RI, Ramilo I, Bravo I. **2017**. The life history of the toxic marine dinoflagellate *Protoceratium reticulatum* (Gonyaulacales) in culture. *Harmful Algae* **68**: 67–81.
Sandgren CD. **1981**. Characteristics of sexual and asexual resting cyst (statospore) formation in *Dinobryon cylindricum* Imhof (Chrysophyta). *Journal of Phycology* **17**: 199–210.
Sandgren CD, Flanagin J. **1986**. Heterothallic sexuality and density dependent encystment in the Chrysophycean alga *Synura petersenii* Korsh. *Journal of Phycology* **22**: 206–216.
Schärer L, Brand JN, Singh P, Zadesenets KS, Stelzer CP, Viktorin G. **2020**. A phylogenetically informed search for an alternative *Macrostomum* model species, with notes on taxonomy, mating behavior, karyology, and genome size. *Journal of Zoological Systematics and Evolutionary Research* **58**: 41–65.
Shah A, Hoffman JI, Schielzeth H. **2020**. Comparative analysis of genomic repeat content in gomphocerine grasshoppers reveals expansion of satellite DNA and helitrons in species with unusually large genomes. *Genome Biology and Evolution* **12**: 1180–1193.
Shi XL, Lepère C, Scanlan DJ, Vaulot D. **2011**. Plastid 16S rRNA gene diversity among eukaryotic picophytoplankton sorted by flow cytometry from the South Pacific Ocean. *PLoS ONE* **6**.
Shuter BJ, Thomas JE, Taylor WD, Zimmerman AM. **1983**. Phenotypic correlates of genomic DNA content in unicellular eukaryotes and other cells. *The American Naturalist* **122**: 26–44.
Simon N, Barlow RG, Marie D, Partensky F, Vaulot D. **1994**. Characterization of oceanic photosynthetic picoeukaryotes by flow cytometry. *Journal of Phycology* **30**: 922–935.
Soltis DE, Soltis PS. **1999**. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution* **14**: 348–352.
Speijer D, Lukeš J, Eliáš M. **2015**. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 8827–8834.
Standeren E. **2018**. Cell- and genome size responses to different temperatures in haptophytes.
Stelzer CP, Blommaert J, Waldvogel AM, Pichler M, Hecox-Lea B, Mark Welch DB. **2021**. Comparative analysis reveals within-population genome size variation in a rotifer is driven by large genomic elements with highly abundant satellite DNA repeat elements. *BMC Biology* **19**: 1–17.
Swift H. **1950**. The constancy of desoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences of the USA* **36**: 643–54.
Takahashi T. **2017**. Life cycle analysis of endosymbiotic algae in an endosymbiotic situation with *Paramecium bursaria* using capillary flow cytometry. *Energies* **10**.
Thomas CA. **1971**. The genetic organization of chromosomes. *Annual Review of Genetics* **5**: 237–256.
Thornber CS. **2006**. Functional properties of the isomorphic biphasic algal life cycle. *Integrative and Comparative Biology* **46**: 605–614.
Trávníček P, Čertner M, Ponert J, Chumová Z, Jersáková J, Suda J. **2019**. Diversity in genome size and GC content shows adaptive potential in orchids and is closely linked to partial endoreplication, plant life-history traits and climatic conditions. *New Phytologist*: 0–2.
Vaulot D, Birrien J-L, Marie D, *et al.* 1994. Morphology, ploidy, pigment composition, and genome size of cultured strains of *Phaeocystis* (Prymnesiophyceae). *Journal of Phycology* **30**: 1022–1035.
Veldhuis MJW, Cucci TL, Sieracki ME. **1997**. Cellular DNA content of marine phytoplankton using two new fluorochromes: taxonomic and ecological implications. *Journal of Phycology* **33**: 527–541.
Veselý P, Bureš P, Šmarda P, Pavlíček T. **2012**. Genome size and DNA base composition of

geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.
Waaland JR, Stiller JW, Cheney DP. **2004**. Macroalgal candidates for genomics. *Journal of Phycology* **40**: 26–33.
Wang J, Hoshaw RW, McCourt RM. **1985**. A polyploid species complex of *Spirogyra communis* (Chlorophyta) occuring in nature. *Journal of Phycology* **22**: 102–107.
Wawrik F. **1972**. Isogame Hologamie in der Gattung *Mallomonas* Perty. *Nova Hedwigia* **23**: 353–362.
Weiss TL, Johnston JS, Fujisawa K, Okada S, Devarenne TP. **2011**. Genome size and phylogenetic analysis of the A and L races of *Botryococcus braunii*. *Journal of Applied Phycology* **23**: 833–839.
Whiteley AS, Burkill PH, Sleigh MA. **1993**. Rapid method for cell cycle analysis in a predatory marine dinoflagellate. *Cytometry* **14**: 909–915.
Whittaker KA, Rignanese DR, Olson RJ, Rynearson TA. **2012**. Molecular subdivision of the marine diatom *Thalassiosira rotula* in relation to geographic distribution, genome size, and physiology. *BMC Evolutionary Biology* **12**: 209.
Wichard T, Charrier B, Mineur F, Bothwell JH, De Clerck O, Coates JC. **2015**. The green seaweed *Ulva*: A model system to study morphogenesis. *Frontiers in Plant Science* **6**: 1–8.
Wichman HA, Van Den Bussche RA, Hamilton MJ, Baker RJ. **1993**. Transposable elements and the evolution of genome organization in mammals In: Springer, Dordrecht, 149–157.
Wilson EB. **1925**. The karyoplasmic ratio In: *The Cell in Development and Heredity*. New York: The Macmillan Company, 727–733.
Wong WY, Simakov O, Bridge DM, *et al.* 2019. Expansion of a single transposable element family is associated with genome-size increase and radiation in the genus *Hydra*. *Proceedings of the National Academy of Sciences of the United States of America* **116**: 22915–22917.
Wu Y, Sun Y, Sun S, *et al.* 2018. Aneuploidization under segmental allotetraploidy in rice and its phenotypic manifestation. *Theoretical and Applied Genetics* **131**: 1273–1285.
Wyngaard GA, Rasch EM, Manning NM, Gasser K, Domangue R. **2005**. The relationship between genome size, development rate, and body size in copepods. *Hydrobiologia* **532**: 123–137.
Yamamoto M, Nozaki H, Kawano S. **2001**. Evolutionary relationships among multiple modes of cell division in the genus *Nannochloris* (Chlorophyta) revealed by genome size, actin gene multiplicity, and phylogeny. *Journal of Phycology* **37**: 106–120.

Part B – Case studies

PAPER I

Best practices in the flow cytometry of microalgae



Spherical colonies of freshwater green microalgae *Volvox aureus* (courtesy of Martin Pusztai).

Best practices in the flow cytometry of microalgae

Dora Čertnerová[1,*] and David W. Galbraith[2,3]

[1]Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 00 Prague, Czech Republic

[2]School of Plant Sciences, BIO5 Institute, Arizona Cancer Center, Department of Biomedical Engineering, University of Arizona, Tucson, Arizona, 85721

[3]Henan University, School of Life Sciences, State Key Laboratory of Crop Stress Adaptation and Improvement, State Key Laboratory of Cotton Biology, Key Laboratory of Plant Stress Biology, Jin Ming Avenue. Kaifeng 475004 China

*Author for correspondence (e-mail: dora.certnerova@gmail.com)

Running headline: Flow cytometry of microalgae

ABSTRACT

Microalgae are photosynthetic microorganisms with a major influence on global ecosystems. Further, owing to the production of various secondary metabolites, microalgae are also intensively studied for their enormous potential in biotechnology and its applications. While flow cytometry (FCM) is fast and reliable method particularly suitable for genome size estimation in plant and animal studies, its application to microalgae often comes with many methodological challenges due to specific issues (e.g. cell wall composition, and presence of various secondary metabolites). Sample preparation requires considerable amounts of biomass, chemical fixation and / or extraction of cellular components. In genome size estimation, appropriate methods for isolation of intact nuclei (using lysis buffers, razor-blade chopping, various enzymes, or bead-beating of cells) are essential for successful and high-quality analyses. Nuclear DNA amounts of microalgae diverge greatly, varying by almost 30,000-fold (0.01 to 286 pg). Even though new algal reference standards for genome size are now being introduced, animal red blood cells and nuclei from plant tissues are still predominantly used. Due to our limited knowledge of microalgal life cycles, particular caution should be taken during 1C / 2C-value (or ploidy level) assignments.

Key words: best practices, microalgae, flow cytometry, nuclear isolation, genome size, algal FCM standards

INTRODUCTION

Microalgae are an extremely diverse group of organisms, individual species of which are placed within different domains across the tree of life. These photosynthetic microorganisms occupying a wide range of habitats, from freshwater lakes to desert soils, also play key roles in the functioning of the global ecosystem. Because of their polyphyletic origin, microalgae differ greatly in their morphology, cell wall composition, protoplast content, and/or by presence of specific organelles. Analysing microalgae using flow cytometry (FCM) is, in general, methodologically more challenging and time consuming as compared to the analysis of plant or animal tissues. This is particularly due to difficulties in obtaining sufficient amounts of biomass, and in protoplast extraction (corresponding to widespread cell wall heterogeneity and variation in complexity of wall components), and due to the presence of wide variety of pigments and secondary metabolites that can interfere with fluorescent staining (Simon *et al.* 1994; Veldhuis *et al.* 1997; Kapraun 2007; Mazalová *et al.* 2011). As for other plant and animal species, FCM enables counting, sorting and/or examination of different features or physiological states of microalgae on the basis of quantification of scattered light signals and of emitted fluorescence. Here is provided a general FCM protocol and workflow for the analysis of microalgal samples. However, one should always keep in mind the enormous diversity among microalgae and the specific features of particular groups that may often require modifications of this protocol.

Obtaining biomass

The first step in FCM analysis of microalgal samples is obtaining sufficient amounts of biomass. While this task is relatively straightforward with macroalgae consisting of multicellular thalli which can be sampled in the field, analysis of microalgae usually requires further cultivation steps. In the former case, collected thalli are cleaned from epiphytes or extraneous debris and rinsed in distilled water (Gall *et al.* 1993; Reed *et al.* 1999). However, in some macroalgal groups, where the FCM analysis of thalli may be problematic, the use of unicellular life stages (e.g. zoids or spores) provides an alternative (Reed *et al.* 1999; Peters *et al.* 2004). In that case, cells are processed in the same manner as microalgal samples. When working with microscopic algae, we first need to establish a unialgal, clonal (and, if possible, axenic) culture. For heterotrophic microalgae, prey must be often added to the culture (e.g. some dinoflagellates or cryptophytes), and later it may be difficult to differentiate prey nuclei from those of the microalgal taxa. This is probably the reason as to why heterotrophic microalgae have been largely avoided as a subject of FCM (but see Whiteley *et al.* 1993). Similarly challenging is the task of separating the nuclei of a studied sample from its symbionts. To avoid this problem, it is sometimes possible to switch to specific life-cycle stages, for instance to zoospores (Parrow and Burkholder 2002; Lin *et al.* 2004). In order to increase the proportion of biological

materials in optimal condition, the culture is inoculated into fresh medium and placed under higher levels of illumination (e.g. 40 μmol m^{-2} s^{-1} or higher) a few weeks before the planned FCM analysis (usually 2 weeks, depending on the growth rate of studied algae). If possible, the use of material from different subcultures treated independently is preferable for repeated measurements. The optimal harvesting time is during the (mid-)exponential phase of culture growth. Furthermore, cells in young, exponentially-growing cultures may not have fully developed cell walls, which makes them more suitable for protoplast isolation based on enzymatic digestion (see below). This is especially crucial when working with desmids, where only young cells without fully developed cell walls are suitable for the use of enzymatic digestion (Mazalová *et al.* 2011; Poulíčková *et al.* 2014). Although the culture should have a high cell density, even more important is to aim to analyse cultures that are as young as possible, due to the potential accumulation of secondary metabolites (or even genomic changes in culture, see below) during long-term cultivation. A different approach is taken when the amount of storage compounds is the target of a study (e.g. oils or starch). In this case, microalgae are usually harvested during the stationary phase of culture growth (Shen *et al.* 2010; Přibyl *et al.* 2012).

The microalgal biomass can be harvested either using a cell scraper from agar-based cultivation media or by centrifugation of liquid cultures (ca. 1 - 100 mL of culture, depending on cell density). After removing the supernatant, the pelleted cells should be visible to the naked eye, but the cell concentration should then be precisely determined by counting. The pellet should contain at least 10^5 cells for successful analysis (e.g. (Lemaire *et al.* 1999; Parrow and Burkholder 2002; Connolly *et al.* 2008; Olefeld *et al.* 2018). When working with mucilaginous species, ultrasonication for several minutes prior to sample preparation may help subsequent release of individual cells from pellets.

Chemical fixation

Since the first flow cytometric studies on microalgae, samples have been processed by analysing the entire cells (e.g. Simon *et al.* 1994; Veldhuis *et al.* 1997; Connolly *et al.* 2008; Vaulot *et al.* 1994; von Dassow *et al.* 2008; Almeida *et al.* 2019). However, such analysis may well be affected by prominent cytoplasmic and cell wall autofluorescence due to the presence of high levels of pigments and of other classes of molecules. Sometimes the autofluorescence spectrum of cells can overlap the spectra of the fluorochromes used for DNA content measurement. To help prevent unwanted autofluorescence, chemical fixation can be employed. Various fixation protocols have been employed by different authors, with none apparently prevailing (Mann and Stickle 1991; Vaulot *et al.* 1994; Veldhuis *et al.* 1997; LaJeunesse *et al.* 2005; Connolly *et al.* 2008; von Dassow *et al.* 2008; Figueroa *et al.* 2010; Whittaker *et al.* 2012; Hong *et al.* 2016). Cell pellets are incubated in fixative, typically ethanol,

methanol, a methanol:acetic acid (3 : 1) mixture, formaldehyde, or paraformaldehyde, with incubation times ranging from tens of minutes up to 48 hours. Following centrifugation, this fixation step can be repeated up to three times and, if so, the sample is kept on ice between the washing steps. Optimal fixative concentration and incubation times need to be defined experimentally for each studied species. Glutaraldehyde fixation has also been tested as a pigment-removing fixative. However, it resulted in high background fluorescence and interfered with fluorochrome staining (Veldhuis *et al.* 1997; Vives-Rego *et al.* 2000; Parrow and Burkholder 2002; Tang and Dobbs 2007). Following fixation, the sample is washed in phosphate-buffered saline (PBS), methanol or TE buffer (Mann and Stickle 1991; Connolly *et al.* 2008; von Dassow *et al.* 2008; Whittaker *et al.* 2012; Hong *et al.* 2016; Salgado *et al.* 2017). The PBS purification was also used to separate dinoflagellate *Oxyrrhis marina* from nuclei of its prey (although this approach was successful only for unfixed cells; Whiteley *et al.* 1993). It is worth mentioning that in plant FCM studies, chemical fixation is not recommended for absolute DNA estimation (Doležel *et al.* 2007b) and the same might apply for microalgae.

Extraction of cellular components

The most common application of FCM is to detect fluorescence from stained nuclei. Despite the fixation effort, FCM analysis of entire cells may result in prominent background fluorescence and/or incompletely-stained nuclei, causing high CVs of the studied samples, or even preventing successful analysis. Hence, for total genome size estimation of microalgal samples, only the analysis performed on properly extracted nuclei is accurate enough to allow high precision of measurement. In samples without a cell wall (e.g. zoids of Ectocarpales or chrysophytes), nuclei can be extracted simply by adding lysis buffer, sometimes combined with incubation under higher temperature (e.g. 50 °C for 5 – 10 min; Peters *et al.* 2004; Olefeld *et al.* 2018; Čertnerová and Škaloud 2020). Alternatively, bead-beating of cells in a mixer mill can be employed to isolate nuclei. This method was particularly successful in the raphydophyte *Gonyostomum semen* (Rengefors *et al.* 2021). In many cases, however, enzymatic treatment to disrupt cell walls needs to be implemented prior to protoplast content extraction (Mazalová *et al.* 2011; Weiss *et al.* 2011; Poulíčková *et al.* 2014). It should be noted that enzymatic treatments are time-consuming and often require optimization specific to the studied group of algae. Enzymatic cell-wall disruption in microalgae is based on protocols adopted from plant or fungal studies (Jazwinski 1990; Doležel *et al.* 2007a). The predominantly used enzymes are cellulase, macerozyme or lyticase, sometimes dissolved in a rinsing solution of PGly (composition: 27.2 mg·l^{-1} KH_2PO_4, 101 mg· l^{-1} KNO_3, 1117.6 mg· l^{-1} $CaCl_2$, 246 mg·l^{-1} $MgSO_4 \cdot 7H_2O$, 11.5 g· l^{-1} glycine, 18.016 g· l^{-1} glucose, 0.58572 g· l^{-1} MES and 65.58 g·l^{-1} mannitol; Mazalová *et al.* 2011; Poulíčková *et al.* 2014; Weiss *et al.* 2011). The enzymatic mixture dissolved in rinsing solution

was primarily developed for streptophyte algae, but also worked for some Chlorophyta (*Chloroidium ellipsoideum*, *Tetraselmis subcordiformis*) and Ochrophyta (*Tribonema vulgare*). Nonetheless, the enzymatic mixture may not digest the cell wall completely, and, for example, only young cells of desmids with partially dissolved cell walls were suitable for FCM analysis. Sometimes the enzymatic treatment needs to be complemented with chopping algal biomass using a razor blade (e.g. in case of *Zygnema* spp.; Mazalová *et al.* 2011). It is also a common practice to check the successful enzymatic cell-wall degradation under a microscope. Grinding algal biomass in a mortar for nuclear isolation has also been tested, however without success (Mazalová *et al.* 2011).

Isolation buffers

In studies on microalgae, commonly-employed buffers are LB01 (with streptophytes, Chlorophyta, Ochrophyta, raphidophytes; Mazalová *et al.* 2011; Poulíčková *et al.* 2014; Rengefors *et al.* 2021), a MOPS-based buffer (3-(N-morpholino)propanesulfonic acid (used with dinoflagellates; LaJeunesse *et al.* 2005; Hong *et al.* 2016) or Otto buffers (with chrysophytes; Olefeld *et al.* 2018; Čertnerová and Škaloud 2020). Triton X-100 (to a final concentration of 0.05 % - 1 %) may be added to improve the sample staining, though its effect varies across different groups of algae (Veldhuis *et al.* 1997; Lemaire *et al.* 1999; Peters *et al.* 2004; Potter *et al.* 2016; Almeida *et al.* 2019). Phenols, tannins and other secondary metabolites are commonly present in microalgae and may act as staining inhibitors or lower the quality of the FCM analysis. Their adverse effect can be lowered, to some extent, by adding PVP (polyvinylpyrrolidone) and/or mercaptoethanol to the lysis buffer (Loureiro *et al.* 2021).

Standardisation

For precise total DNA content estimation, it is essential to include a FCM standard. The use of internal DNA standards is highly recommended for microalgal samples, considering the frequent presence of secondary metabolites with the potential to interfere with FCM analysis. An appropriate internal standard is closely related to the studied organism with similar but not overlapping genome size. Unfortunately, due to the lack of a broad range of algal DNA standards, animal red blood cells and plant tissues are still predominantly used for this purpose (e.g. Veldhuis et al. 1997; Mazalová *et al.* 2011; Parrow and Burkholder 2002; Connolly *et al.* 2008; Olefeld *et al.* 2018; Vaulot *et al.* 1994). However, the nuclear DNA within the most commonly used standard, chicken red blood cells, has a high packing density, and red blood cells from male and female chickens differ in genome size by 2.7 % due to the contributions of sex chromosomes; these factors likely contribute to non-systematic errors in DNA content estimates (Nakamura *et al.* 1990; Hardie *et al.* 2002). Although the number of available algal standards is now rising, it is still a

negligible number in contrast to algal diversity. Examples of standards include the green microalga *Chlamydomonas reinhardtii* CC-400 cw15 mt+, which, being a cell-wall deficient mutant, is therefore easy to employ (2C = 0.24 pg; available at Chlamydomonas Resource Center, University of Minnesota; Lemaire *et al.* 1999; Potter *et al.* 2016). The chrysophyte *Synura sphagnicola* CCAC 2959 B (2C = 0.4 pg; otherwise designated LO234K-E, and available at The Central Collection of Algal Cultures (CCAC)) was recently introduced as an internal standard by Olefeld *et al.* (2018). The desmid *Micrasterias pinnatifida* SVCK 411 (2C = 3.4 pg; available at the Microalgae and Zygnematophyceae Collection Hamburg) was established as a streptophyte standard by Mazalová *et al.* (2011). It should be noted that this standard requires enzymatic treatment that degrades cell wall structure and thus enables release of nuclei (see above; Mazalová *et al.* 2011; Poulíčková *et al.* 2014). While the cultivated microalgal lineages can serve as reliable references for specific genome size classes, their limited numbers can be compensated by field-collected standards. Examples include use of the red alga *Chondrus crispus* (2C = 0.33 pg) as an internal standard (LeGall *et al.* 1993; Peters *et al.* 2004). Also, the number of microalgae with available complete genome sequences is increasing and these offer promise as new FCM standards. In von Dassow *et al.* (2008), the authors employed the diatom *Thalassiosira pseudonana* strain CCMP1335 as an internal standard (2C = 0.07 pg; available at Provasoli-Guillard National Center for Marine Algae and Microbiota). However, the use of diatoms as genome size standards may be problematic due to considerable genome flexibility during cultivation, as documented in the referenced study (see below). Even though the internal standard should be optimally treated in an identical way to the experimental microalgal sample, this is nearly impossible in most studies. As a common practice, the standard nuclei are extracted separately, and are later mixed with the microalgal sample (i.e. pseudo-internal standardisation, e.g. Parrow and Burkholder 2002; Olefeld *et al.* 2018; LaJeunesse *et al.* 2005; Hong *et al.* 2016).

Fluorescent staining

The use of different fluorescent stains allows detection a number of different content amounts and enzymatic activities of microalgal cells. Even without the addition of any stain, it is possible to determine chlorophyll content in the sample due to its autofluorescence (e.g. Almeida *et al.* 2019; Nakamura *et al.* 1990). To assess cellulose content, Calcofluor White can be used due to its ability to bind cellulose and emit blue fluorescence following UV excitation (Kwok and Wong 2003). Staining lipid globules with Nile Red fluorescent stain (9-diethylamino-5-benzo[a]phenoxazinone; Reed *et al.* 1999) allows FCM estimation of the cell lipid content. Intracellular peroxidase and reactive oxygen species (ROS) can be detected using hydroethidine or DCFH-DA (2',7'-dichlorodihydrofluorescein diacetate acetyl ester). The presence of intracellular peroxidase and ROS leads to conversion of

hydroethidine to ethidium, which is accompanied by a change of emitted fluorescence (Cid *et al.* 1996). Similarly, non-fluorescent DCFH-DA is oxidized by either intracellular peroxidase or ROS to form fluorescent DCF (7'-dichlorofluorescein; Kwok and Wong 2003). FCM can also be used to detect changes in cellular (or mitochondrial) membrane potential. Positively charged lipophilic stains, such as $DiOC_6$ (3,3'-dihexyloxacarbocyanineine) and rhodamine 123, can penetrate through organelle membranes to reach their negatively charged interiors. When an equilibrium concentration is reached, membrane depolarization or hyperpolarization causes release or uptake of the fluorescent stain, respectively (Cid *et al.* 1996; Almeida *et al.* 2019). However, by far the widest application of FCM in microalgal studies is the detection of fluorescence stained nuclei. Analysis of protoplasts following staining with a membrane-impermeable fluorochrome (e.g. propidium iodide (PI)) allows determination of cell viability in terms of plasma membrane integrity (Brussaard *et al.* 2001). Other applications are directed at analysis of DNA base composition (GC content) and genome size characteristics, allowing, for example, detection of different life-cycle stages within populations, and intraspecific ploidy level diversity or cell-cycle stages in microalgal cultures (Carre and Edmunds Jr. 1993; Vaulot *et al.* 1994; Lemaire *et al.* 1999; Gerashchenko *et al.* 2001; Houdan *et al.* 2004; Kremp and Parrow 2006; Van Dolah *et al.* 2008; Salgado *et al.* 2017; Almeida *et al.* 2019). It is also possible to distinguish taxa and identify cryptic species on the basis of differences in genome size (Figueroa *et al.* 2010; Koester *et al.* 2010). When estimating genome size in microalgae using PI, caution should be taken as the emission spectrum of PI fluorescence can overlap with the autofluorescence of photosynthetic pigments, PI can bind polysaccharides from the remaining cell walls and thus contribute to background noise and increase of CV (Potter *et al.* 2016), and its ability to fluoresce in the presence of double-stranded RNA must be considered and, if necessary, eliminated by including RNAse in the staining protocol.

Following the fluorescence staining but prior to FCM analysis, the microalgal samples are filtered (the mesh size 5 – 150 μm, depending on particular application) to prevent clogging of the fluidic system of a flow cytometer.

Special considerations for estimating genome size in microalgae

Based on the currently limited knowledge, the DNA content of microalgae varies 28,600-fold, from 0.01 detected in *Nannochloropsis* sp. (Eustigmatophyceae) to 286 pg in *Valonia* sp. (Ulvophyceae; Veldhuis *et al.* 1997; Kapraun 2005). When evaluating the outcomes of FCM analysis, authors should be very cautious in their interpretations, in light of the general lack of data on genome size variation in many groups of microalgae, a dearth of information about life cycles, as well as the possibility of rapid genome size evolution across species and within genera.

In case of two peaks being observed in a FCM uniparametric histogram of a microalgal sample, the first peak (1C) is usually considered to represent G_1-phase cells and the second (2C) belonging to G_2 cells (Figure 1; Mazalová *et al.* 2011; Poulíčková *et al.* 2014; Olefeld *et al.* 2018). Sometimes, however, only a single sample peak is observed. In this case, it is critical to determine whether the sample is in G_1 phase with no dividing cells (e.g. extremely slowly growing cultures), or whether it represents extremely synchronized cells in G_2 phase prior to mitosis. In phytoplankton species, the timing of cell division may strongly depend on the time at which cell biomass is harvested and processed (Kremp and Parrow 2006; Van Dolah *et al.* 2008). For instance, in diatoms and dinoflagellates, a peak representing the G_2 nuclei can be either prominent or completely missing (Parrow and Burkholder 2002; Connolly *et al.* 2008; Koester *et al.* 2010). On the other hand, in a highly synchronized culture of the genus *Chlamydomonas*, the G_1 peak was always present and never represented less than 29 % of analysed nuclei (Lemaire *et al.* 1999).

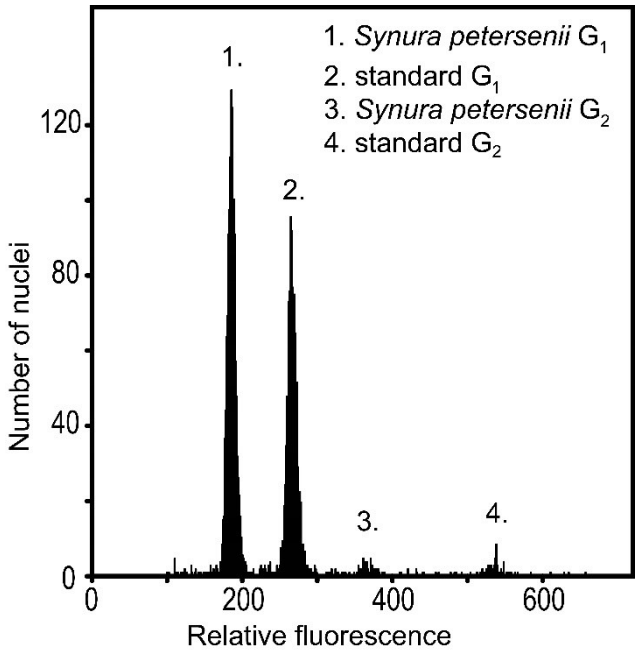


Figure 1. Flow cytometric histogram showing relative fluorescence of propidium iodide-stained nuclei of chrysophyte *Synura petersenii* and *Solanum pseudocapsicum* (reference standard) with G_1 and G_2 phase nuclei apparent for both analysed sample and standard. In this case, the CV values are 2.97 % and 2.31 %.

The G_1 sample peak is commonly referred as either *n* or 2*n* stage, usually without the knowledge of a particular life-cycle stage of the analysed sample. Unfortunately, our understanding of ploidy levels and reproduction strategies in majority of microalgae is extremely limited. Thus, vegetative cells (G_1 phase) may be

dominantly haploid (e.g. in the majority of dinoflagellates, some desmids, and presumably in chrysophytes) or diploid (e.g. in diatoms or raphidophytes; Olefeld *et al.* 2018; von Dassow *et al.* 2008; Kremp and Parrow 2006; Koester *et al.* 2010; Kapraun 2005; Temsch *et al.* 2010). Many species of microalgae are capable of vegetative growth in both sexual and asexual stages (i.e. the biphasic life cycle) and the sexual and asexual phases can be morphologically indistinguishable (as in the case of sea lettuce *Ulva* spp.; Kagami *et al.* 2005). In addition, numerous species or strains across microalgae are putative polyploids, which further complicates ploidy level assignments. Therefore, in case of any ambiguity, the DNA content of microalgal samples should be referred in pg · cell^{-1} rather than attempting to assign it to a specific 1C / 2C-value (or ploidy level).

Interestingly, genome size may differ greatly between closely related species or even between strains of the same species. Major intraspecific variation, reaching up to 7-fold differences, has been described in desmids (*Micrasterias rotata*, *Triploceras gracile*), haptophytes (*Emiliania huxleyi*), chrysophytes (*Synura petersenii*), and diatom species (*Thalassiosira punctigera*, *T. weissflogii*; Veldhuis *et al.* 1997; Mazalová *et al.* 2011; Poulíčková *et al.* 2014; Čertnerová and Škaloud 2020; Fukuda and Endoh 2006; Haig 2010). Astonishingly, genome size changes within the same strain kept in cultivation were also documented. The DNA content estimates for diatom *Thalassiosira weissflogii* CCMP 1049 differed tremendously in comparisons of three independently-conducted studies (from 0.95 to 17.25 pg; Veldhuis *et al.* 1997; Connolly *et al.* 2008, von Dassow *et al.* 2008). Von Dassow *et al.* (2008) have even reported genome size diversification of three *T. weissflogii* sub-cultures over a few years of cultivation (strains BILB2001, CCMP 1336, and CCMP 1587). A series of whole-genome duplications (polyploidy events) are the likely drivers of these genome size shifts, as previously proposed in studies on streptophytes (Hamada *et al.* 1985; Hoshaw *et al.* 1985) and dinoflagellates (Loper *et al.* 1980; Holt and Pfiester 1982). Further, the cultured microalgae may deleted or amplify specific genomic regions depending on their current environmental conditions (von Dassow *et al.* 2008). Other possible sources of genome size variation in culture include aneuploidy or meiosis introducing DNA amount variation (Quijano-scheggia *et al.* 2009). Therefore, it is highly recommended to avoid long-term cultivation, and analyse the samples as soon as possible after isolation. On the other hand, strains of the dinoflagellate genus *Symbiodinium* analysed multiple times after a varying length of cultivation, exhibited highly similar genome size suggesting its stability during the cultivation (LaJeunesse *et al.* 2005).

Due to the often challenging preparation of microalgal FCM samples and the frequently high contents of secondary metabolites interfering with the analysis, it is unusual to obtain measurements having CV values at or below 3 %, although CVs are typically below 10 % (Parrow and Burkholder 2002; LaJeunesse *et al.* 2005; Potter *et al.* 2016).

GENERAL RECOMMENDATIONS

Aside from following the general best practices in FCM as indicated in other chapters of this series, recommendations specific for microalgal studies are worth highlighting:

- Ensure a sufficient amount of input biomass by optimised cultivation (high cell abundance, absence of contamination by other organisms).
- Try to avoid analyses of strains subjected to long-term cultivation that might result in genomic or other changes.
- For DNA amount measurements, use as young cultures as possible to avoid accumulation of secondary metabolites and pigments. Also, isolate nuclei rather than attempt to analyse whole intact cells.
- In case of unsuccessful analysis, test different isolation buffers and protoplast extraction techniques (varying the intensity of razor-blade chopping or bead-beating of cells, or enzymatic treatments, as appropriate).
- In all cases, to the extent possible, and prior to FCM analyses, validate input cell populations in terms of number, purity, viability, cellular (or subcellular i.e. nuclear) integrity, and homogeneity of fluorescence staining, using light and fluorescence microscopy. Ensure the results of FCM analyses are consistent with these observations.
- When estimating genome size, use internal or pseudo-internal standardisation, as high cellular contents of secondary metabolites or pigments may lead to potential shifts in relative fluorescence.
- In case of any uncertainty regarding life cycle stage of the analysed microalgal samples, report in publications genome size estimates in absolute units (e.g. $pg \cdot cell^{-1}$) rather than attempt to assign it to 1C / 2C-value.

LITERATURE CITED

Almeida AC, Gomes T, Habuda-Stanić M, *et al.* 2019. Characterization of multiple biomarker responses using flow cytometry to improve environmental hazard assessment with the green microalgae *Raphidocelis subcapitata*. *Science of the Total Environment* **687**: 827–838.

Brussaard C, Marie D, Thyrhaug R, Bratbak G. 2001. Flow cytometric analysis of phytoplankton viability following viral infection. *Aquatic Microbial Ecology* **26**: 157–166.

Carre IA, Edmunds Jr. LN. 1993. Oscillator control of cell division in *Euglena*: cyclic AMP oscillations mediate the phasing of the cell division cycle by the circadian clock. *J Cell Sci* **104 (Pt 4**: 1163–1173.

Čertnerová D, Škaloud P. 2020. Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences. *Annals of Botany* **126**: 1077–1087.

Cid A, Fidalgo P, Herrero C, Abalde J. 1996. Toxic action of copper on the membrane system of a marine diatom measured by flow cytometry. *Cytometry* **25**: 32–36.

Coats DW. 2002. Dinoflagellate life-cycle complexities. *Journal of Phycology* **38**: 417–419.

Connolly JA, Oliver MJ, Beaulieu JM, Knight CA, Tomanek L, Moline MA. 2008. Correlated evolution of genome size and cell volume in diatoms (Bacillariophyceae). *Journal of Phycology* **44**: 124–131.

von Dassow P, Petersen TW, Chepurnov VA, Armbrust E. 2008. Inter- and intraspecific relationships between nuclear DNA content and cell size in selected membersof the centric diatom genus *Thalassiosira* (Bacillariophyceae). *Journal of Phycology* **44**: 335–349.

Van Dolah FM, Leighfield TA, Kamykowski D, Kirkpatrick GJ. 2008. Cell cycle behavior of laboratory and field populations of the Florida red tide dinoflagellate, *Karenia brevis*. *Continental Shelf Research* **28**: 11–23.

Doležel J, Greilhuber J, Suda J. 2007a. *Flow Cytometry with Plant Cells*. Weinheim: Wiley-VCH Verlag GmbH.

Doležel J, Greilhuber J, Suda J. 2007b. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**: 2233–2244.

Figueroa RI, Garcés E, Bravo I. 2010. The use of flow cytometry for species identification and life-cycle studies in dinoflagellates. *Deep-Sea Research Part II: Topical Studies in Oceanography* **57**: 301–307.

Fukuda Y, Endoh H. 2006. New details from the complete life cycle of the red-tide dinoflagellate Noctiluca scintillans (Ehrenberg) McCartney. *European Journal of Protistology* **42**: 209–219.

Gall Y Le, Brown S, Marie D, Mejjad M, Kloareg B. 1993. Quantification of nuclear DNA and G-C content in marine macroalgae by flow cytometry of isolated nuclei. *Protoplasma* **173**: 123–132.

Gerashchenko BI, Kosaka T, Hosoya H. 2001. Growth kinetics of algal populations exsymbiotic from *Paramecium bursaria* by flow cytometry measurements. *Cytometry* **44**: 257–263.

Haig D. 2010. What do we know about charophyte (Streptophyte) life cycles? *Journal of Phycology* **46**: 860–867.

Hamada J, Saito M, Ishida M. 1985. Nuclear phase in vegetative and gamete cells of *Closterium ehrenbergii*: fluorescence microspectro- photometry of DNA content. *Annual reports of the Research Reactor Institute* **18**: 56–61.

Han M, Wang R, Ding N, *et al.* 2018. Reactive oxygen species-mediated caspase-3 pathway involved in cell apoptosis of *Karenia mikimotoi* induced by linoleic acid. *Algal Research* **36**: 48–56.

Hardie DC, Gregory TR, Hebert PDN. 2002. From pixels to picograms: A beginners' guide to genome quantification by Feulgen image analysis densitometry. *Journal of Histochemistry and Cytochemistry* **50**: 735–749.

Holt JR, Pfiester LA. 1982. A technique for counting chromosomes of armored dinoflagellates,

and chromosome numbers of six freshwater dinoflagellate species. *American Journal of Botany* **69**: 1165–1168.

Hong HH, Lee HG, Jo J, *et al.* 2016. The exceptionally large genome of the harmful red tide dinoflagellate *Cochlodinium polykrikoides* margalef (Dinophyceae): Determination by flow cytometry. *Algae* **31**: 373–378.

Hoshaw RW, Wang J-C, McCourt RM, Hull HM. 1985. Ploidal Changes in Clonal Cultures of *Spirogyra communis* and Implications for Species Definition. *American Journal of Botany* **72**: 1005.

Houdan A, Bonnard A, Fresnel J, Fouchard S, Billard C, Probert I. 2004. Toxicity of coastal coccolithophores (Prymnesiophyceae, Haptophyta). *Journal of Plankton Research* **26**: 875–883.

Jazwinski MS. 1990. Preparation of extracts from yeast. *Methods in Enzymology* **182**: 154–174.

Kagami Y, Fujishita M, Matsuyama-Serisawa K, *et al.* 2005. DNA content of *Ulva compressa* (Ulvales, Chlorophyta) nuclei determined with laser scanning cytometry. *Phycological Research* **53**: 77–83.

Kapraun DF. 2005. Nuclear DNA Content Estimates in Multicellular Green, Red and Brown Algae: Phylogenetic Considerations. *Annals of Botany* **95**: 7–44.

Kapraun DF. 2007. Nuclear DNA Content Estimates in Green Algal Lineages: Chlorophyta and Streptophyta. *Annals of Botany* **99**: 677–701.

Koester JA, Swalwell JE, Von Dassow P, Armbrust EV. 2010. Genome size differentiates co-occurring populations of the planktonic diatom *Ditylum brightwellii* (Bacillariophyta). *BMC Evolutionary Biology* **10**: 1–11.

Kremp A, Parrow MW. 2006. Evidence for asexual resting cysts in the life cycle of the marine peridinoid dinoflagellate, *Scrippsiella hangoei*. *Journal of Phycology* **42**: 400–409.

Kwok ACM, Wong JTY. 2003. Cellulose synthesis is coupled to cell cycle progression at G1 in the dinoflagellate *Crypthecodinium cohnii*. *Plant physiology* **131**: 1681–91.

LaJeunesse TC, Lambert G, Andersen RA, Coffroth MA, Galbraith DW. 2005. *Symbiodinium* (Pyrrhophyta) Genome Sizes (DNA Content) Are Smallest Among Dinoflagellates. *Journal of Phycology* **41**: 880–886.

Lemaire S, Hours M, Gerard-Hirne C, Trouabal A, Roche O, Jacquot J-P. 1999. Analysis of light/dark synchronization of cell-wall-less *Chlamydomonas reinhardtii* (Chlorophyta) cells by flow cytometry. *European Journal of Phycology* **34**: 279–286.

Lin S, Mulholland MR, Zhang H, Feinstein TN, Jochem FJ, Carpenter EJ. 2004. Intense grazing and prey-dependent growth of *Pfiesteria piscicida* (Dinophyceae). *Journal of Phycology* **40**: 1062–1073.

Loper CL, Steidinger KA, Walker LM. 1980. A Simple Chromosome Spread Technique for Unarmored Dinoflagellates and Implications of Polyploidy in Algal Cultures. *Transactions of the American Microscopical Society* **99**: 343.

Loureiro J, Kron P, Temsch EM, *et al.* 2021. Isolation of plant nuclei for estimation of nuclear DNA content – overview and best practices. *Cytometry Part A* **99**: 318–327.

Mann DG, Stickle AJ. 1991. The genus *Craticula*. *Diatom Research* **6**: 79–107.

Mazalová P, Šarhanová P, Ondřej V, Poulíčková A. 2011. Quantification of DNA content in freshwater microalgae using flow cytometry: A modified protocol for selected green microalgae. *Fottea* **11**: 317–328.

Medlin LK, Barker GLA, Campbell L, *et al.* 1996. Genetic characterisation of *Emiliania huxleyi* (Haptophyta). *Journal of Marine Systems* **9**: 13–31.

Nakamura D, Tiersch TR, Douglass M, Chandler RW. 1990. Rapid identification of sex in birds by flow cytometry. *Cytogenetic and Genome Research* **53**: 201–205.

Olefeld JL, Majda S, Albach DC, Marks S, Boenigk J. 2018. Genome size of chrysophytes varies with cell size and nutritional mode. *Organisms Diversity & Evolution* **18**: 163–173.

Parrow MW, Burkholder JAM. 2002. Flow cytometric determination of zoospore DNA content

and population DNA distribution in cultured *Pfiesteria* spp. (Pyrrhophyta). *Journal of Experimental Marine Biology and Ecology* **267**: 35–51.

Peters AF, Marie D, Scornet D, Kloareg B, Mark Cock J. **2004**. Proposal of *Ectocarpus siliculosus* (Ectocarpales, Phaeophyceae) as a model organism for brown algal genetics and genomics. *Journal of Phycology* **40**: 1079–1088.

Pfiester LA, Anderson DM. **1987**. Dinoflagellate reproduction In: Taylor FJR, ed. *The Biology of Dinoflagellates*. Oxford: Blackwell, 611–48.

Potter EE, Thornber CS, Swanson JD, McFarland M. **2016**. Ploidy distribution of the harmful bloom forming macroalgae *Ulva* spp. in Narragansett Bay, Rhode Island, USA, using flow cytometry methods. *PLoS ONE* **11**: 1–15.

Poulíčková A, Mazalová P, Vašut RJ, Šarhanová P, Neustupa J, Škaloud P. **2014**. DNA content variation and its significance in the evolution of the genus *Micrasterias* (desmidiales, streptophyta). *PLoS One* **9**: e86247.

Přibyl P, Cepák V, Zachleder V. **2012**. Production of lipids in 10 strains of *Chlorella* and *Parachlorella*, and enhanced lipid productivity in *Chlorella vulgaris*. *Applied Microbiology and Biotechnology* **94**: 549–561.

Quijano-scheggia S, Garce E, Andree K, Fortun JM. **2009**. Homothallic auxosporulation in *Pseudo-nitzschia brasiliana* (Bacillariophyta). *Journal of Phycology* **107**: 100–107.

Read BA, Kegel J, Klute MJ, *et al.* 2013. Pan genome of the phytoplankton *Emiliania underpins* its global distribution. *Nature* **499**: 209–213.

Reed DC, Brzezinski MA, Coury DA, Graham WM, Petty RL. **1999**. Neutral lipids in macroalgal spores and their role in swimming. *Marine Biology* **133**: 737–744.

Rengefors K, Gollnisch R, Sassenhagen I, *et al.* 2021. Genome-wide SNP markers reveal population structure and dispersal direction of an expanding nuisance algal bloom species. *Molecular Ecology*.

Salgado P, Figueroa RI, Ramilo I, Bravo I. **2017**. The life history of the toxic marine dinoflagellate *Protoceratium reticulatum* (Gonyaulacales) in culture. *Harmful Algae* **68**: 67–81.

Shen Y, Yuan W, Pei Z, Mao E. **2010**. Heterotrophic Culture of *Chlorella protothecoides* in Various Nitrogen Sources for Lipid Production. *Applied Biochemistry and Biotechnology* **160**: 1674–1684.

Simon N, Barlow RG, Marie D, Partensky F, Vaulot D. **1994**. Characterization of oceanic photosynthetic picoeukaryotes by flow cytometry. *Journal of Phycology* **30**: 922–935.

Tang YZ, Dobbs FC. **2007**. Green autofluorescence in dinoflagellates, diatoms, and other microalgae and its implications for vital staining and morphological studies. *Applied and environmental microbiology* **73**: 2306–13.

Temsch EM, Greilhuber J, Krisai R. 2010. Genome size in liverworts. *Preslia* **82**: 63–80.

Vaulot D, Birrien J-L, Marie D, *et al.* 1994. Morphology, ploidy, pigment composition, and genome size of cultured strains of *Phaeocystis* (Prymnesiophyceae). *Journal of Phycology* **30**: 1022–1035.

Veldhuis MJW, Cucci TL, Sieracki ME. **1997**. Cellular Dna Content of Marine Phytoplankton Using Two New Fluorochromes: Taxonomic and Ecological Implications. *Journal of Phycology* **33**: 527–541.

Vives-Rego J, Lebaron P, Nebe-von Caron G. **2000**. Current and future applications of flow cytometry in aquatic microbiology. *FEMS Microbiology Reviews* **24**: 429–448.

Weiss TL, Johnston JS, Fujisawa K, Okada S, Devarenne TP. **2011**. Genome size and phylogenetic analysis of the A and L races of *Botryococcus braunii*. *Journal of Applied Phycology* **23**: 833–839.

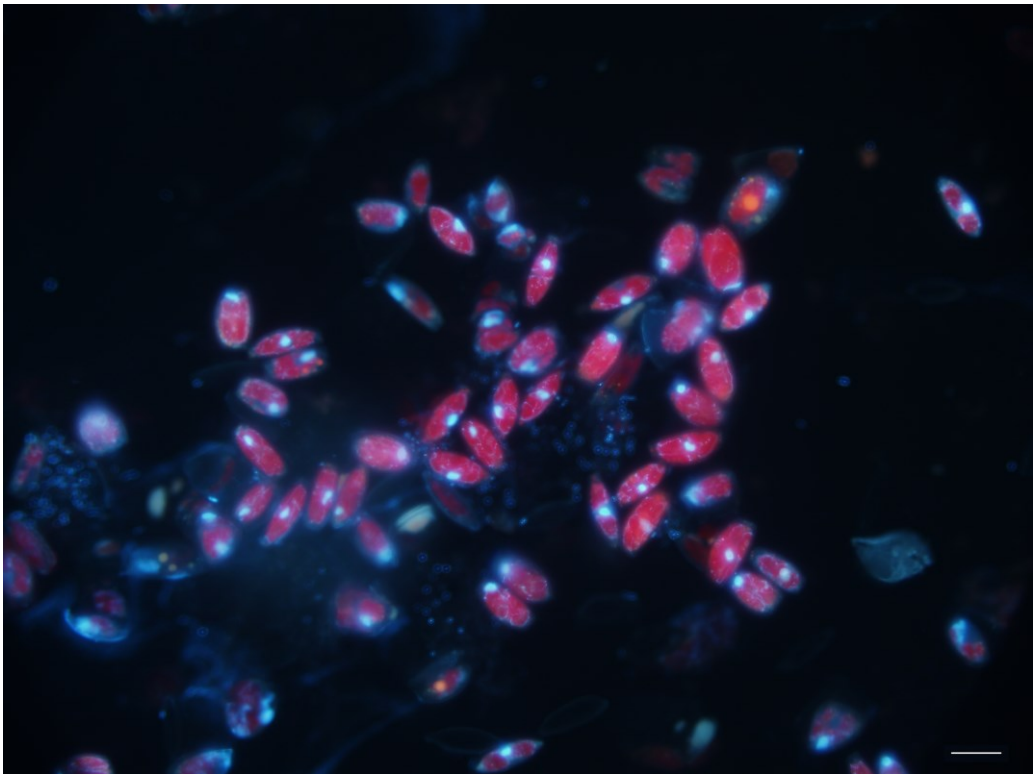
Whiteley AS, Burkill PH, Sleigh MA. **1993**. Rapid method for cell cycle analysis in a predatory marine dinoflagellate. *Cytometry* **14**: 909–915.

Whittaker KA, Rignanese DR, Olson RJ, Rynearson TA. **2012**. Molecular subdivision of the

marine diatom *Thalassiosira rotula* in relation to geographic distribution, genome size, and physiology. *BMC Evolutionary Biology* **12**: 209.

PAPER II

Meet the challenges of analysing small genomes using flow cytometry



Fluorescence microphotographs of diatom *Nitzchia* sp. possessing low nuclear DNA content stained with DAPI (blue). Scale bar represent 30 μm.

Meet the challenges of analysing small genomes using flow cytometry

Dora Čertnerová*

Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic

*Author for correspondence (e-mail: dora.certnerova@gmail.com)

Key words: flow cytometry, small genomes, DNA content estimation, microorganisms

In many fields of biodiversity research, nuclear DNA content is a crucial parameter of the study organism (individual, cellular type), allowing for example ploidy determination, cell-cycle analysis or selecting suitable organisms and optimal strategy for whole genome sequencing (WGS). Due to lower sequencing costs, small genome size represents a major advantage for WGS projects. Not surprisingly, most DNA content estimates available for small genomes have been derived from WGS data. On the other hand, the routine use of WGS as a method for genome size estimation has been discouraged due to its poor quantification of genomic content of repetitive elements (e.g. present in centromeres or telomeres) that may significantly underestimate the true DNA amount. Currently, the most suitable method for the task is flow cytometry (FCM), a rapid and easy to perform technique, using which the DNA content is estimated from the mean fluorescence intensity of nucleic acid binding dye (e.g. propidium iodide, ethidium bromide). The FCM is routinely used in immunology, cancer research or plant and animal studies, however, its application on organisms with small genomes can be highly challenging.

Even though, the complexity of organisms is not directly linked with the amount of their nuclear DNA, the small genomes are very often found among microorganisms, specifically in nano / picoplankton, unicellular parasites and most fungi, as a consequence of the positive genome size – cell size correlation (Cavalier-Smith 1985). However, even microorganisms in assumed clonal populations commonly differ in morphology, physiology or biochemistry. In fungi, the smallest measured nuclear DNA content (2.2 Mbp in *Encephalitozoon romaleae*; (Pombert *et al.* 2012) also reaches the lowest end of known DNA content among all eukaryotes. Moreover, the DNA content of other fungal species is generally not much higher (with a median value <40 Mbp; Kullman *et al.* 2005). In the study by Talhinhas *et al.* (published in the current issue of Cytometry Part A (page 343-347)), the authors nicely summarized the currently used methods for fungal genome size estimation

using FCM and addressed the potential pitfalls. Interestingly, these pitfalls are widely shared with many other groups of microorganisms with small genomes.

Until the modern sequencing techniques have been introduced, the microorganisms were largely understudied and their diversity, phylogenetic relationships, life cycles, etc. widely unexplored. Despite their major importance for the global ecosystem and common applications in biotechnology, the microorganisms' research has lagged behind plant and animal studies up to the present. However, limited research of microorganisms had consequences in low number of DNA content data, especially pronounced in contrast to their estimated diversity.

Because of the small size of their bodies, microorganisms usually need to be cultivated to obtain sufficient amounts of biomass for the FCM, which is not only time-consuming but sometimes unrealistic. For the uncultivated microorganisms in trophic interactions, another approach could be taken in simultaneous analysis of studied microorganism and its symbiont / host / prey and then to analyse these partners separately to correctly distinguish peaks of each organisms. Such approach seems especially suitable for parasites as is nicely illustrated by Talhinhas and colleagues (this issue page 343-347) for pathogenic fungi and its host plant. However, simultaneous analysis might not be suitable for organisms substantially differing in their genome size. Moreover microorganisms commonly live in microbial communities and this makes them harder to isolate or preserve in cultivation. Nonetheless, when possible, it is best to conduct the analysis on unistrain culture, ideally young and actively growing, as was also pointed out by Talhinhas and colleagues (in this issue page 343-347). Unfortunately, residua of culture media may increase background fluorescence. In fact, the background noise is one of the major challenges when analysing small genomes. For FCM analysis of plant or animals, even low sensitive flow cytometers such as CyFlow (Sysmex/Partec) are adequate, however, for FCM of microorganisms, instruments like CytoFLEX (Beckman Coulter) or FACS / LSR II (BD Biosciences), high-sensitive to small particles are more appropriate (see Fig. 1). Further, there are several ways how to reduce the background noise. Nuclei should be isolated from cells, either chemically (using enzymes) or mechanically (razor-blade chopping, bead-beating). Although razor-blade chopping is routinely used in plant FCM, it seems unsuitable for protists (i.e. single-celled eukaryotes) but useful for filamentous microorganisms as was shown by Talhinhas and colleagues (this issue page 343-347) or Čertnerová (2021). To reduce autofluorescence or adverse effect of secondary metabolites, sample can be fixed with various fixatives (ethanol, methanol, methanol:acetic acid mixture, formaldehyde, paraformaldehyde, or acetone), although the chemical fixation may not be suitable for precise genome size estimation (Doležel *et al*. 2007b). Another possibility is to test different isolation buffers. For example, the Woody Plant Buffer or Tris-$MgCl_2$ buffer seems to work with fungal samples and LB01 buffer found wide application in FCM of microalgae (Talhinhas *et al*. this issue page

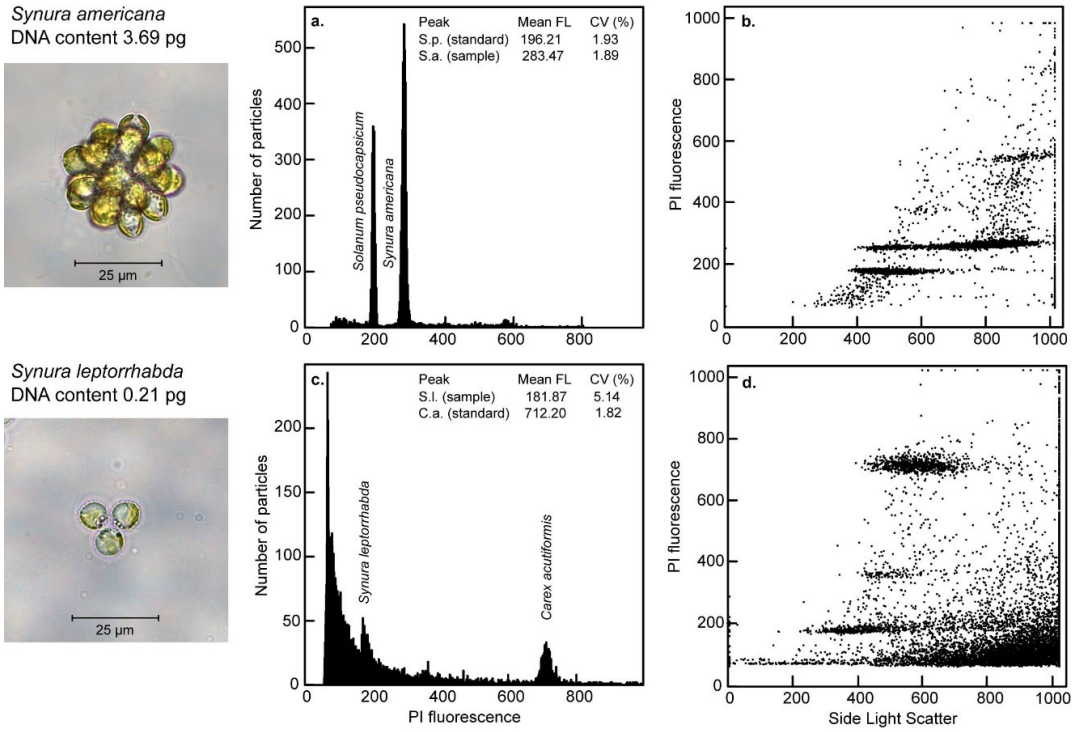


Figure 1. CyFlow FCM outputs of two chrysophyte algae of the genus *Synura* - *S. americana* with higher DNA content (3.69 pg) and *S. leptorhabda* with lower DNA content (0.21 pg) and its plant standards. Note clearly visible peaks with only minor background noise on both fluorescence histogram (a) and fluorescence vs. side scatter plot (b) in case of the first sample analysis. Conversely, higher amount of debris is present in the second analysed sample (c) with the sample DNA content approaching the limits of resolution for CyFlow instrument, yet with peaks still sufficiently separated on fluorescence vs. side scatter plot (d); unpublished data

343-347; Čertnerová 2021). However, new lysis buffers reflecting the specifics of particular groups of microorganisms still need to be developed. The lysis buffer may be further supplemented with PVP (polyvinylpyrrolidone) and / or with mercaptoethanol (Loureiro *et al.* 2021). In addition, Talhinhas and colleagues (this issue page 343-347) are suggesting to use a lower concentration of propidium iodide, however, still adequate enough to properly stain the sample nuclei. It is also convenient to visualize measurements on a side-scatter vs. fluorescence plot and apply gating to distinguish population of nuclei from a background noise if needed, as was also highlighted by the authors. In case of problematic plant or animal sample, alternative tissue / organ might help, though, this is not a possibility for most microorganisms (except few rare cases). However, despite a great effort, analysing organisms with small genomes usually leads to higher CVs and, therefore,

the criteria on acceptable precision of FCM analysis should not be generally as stringent.

Talhinhas and colleagues (this issue page 343-347) further discussed the lack of appropriate FCM standards, which is yet another important issue accompanying analysis of small genomes. In recent years, the number of newly introduced FCM standards is slowly rising up, with, for example, *Saccharomyces cerevisiae*, *Aspergillus fumigatus* or *Chlamydomonas reinhardtii* possessing very small genome sizes (1C values of 24.1, 29.2, and 0.12 pg, respectively; Veselská *et al.* 2014; Potter *et al.* 2016). In the previous work, Talhinhas *et al.* (2017) introduced additional fungal FCM standards with various genome sizes. Even so, there is still a dearth of FCM standards suitable for microorganisms, with those already introduced not easily accessible, leading to a frequent use of suboptimal standards such as chicken red blood cells or plant standards. However, these are biologically different and could be therefore influenced differently from analysed sample resulting in change of sample and standard peaks proportion.

When evaluating FCM outputs, we might have to deal with some additional challenges. The DNA content data are available for only a fraction of microorganisms and thus the range of genome size variation is widely unknown but often more diverse than expected. Fungi particularly are known for their high degree of genome size plasticity. Additionally, dearth of knowledge on life cycles of the studied organisms may lead to misinterpretation of detected fluorescence peaks in FCM histograms. Some fungal species are even heterokaryotic, i.e. possessing multiple different-sized nuclei, and hence generating several G_1 peaks (Catal *et al.* 2010). Further, variations in chromosome number and chromosome size seem to be the rule rather than the exception (Kullman 2000). Unfortunately, chromosome counts are generally problematic in microorganisms due to the small size of their cells and asynchronous cell division. This also had an impact on missing ploidy level data.

In contrast to other groups of microorganisms, fungal genome size data are listed in their own database (Kullman *et al.* 2005). Talhinhas and colleagues (this issue page 343-347) analysed these data from many different angles. They highlighted that the majority of genomes size data were obtained using WGS or static microscope-based cytometry methods, and only less than 5% were obtained with FCM. More frequent employment of FCM might thus allow researching high resolution estimates. The authors further pointed out several interesting correlations. Among others that fungal evolution towards plant mutualism or parasitism seems to be accompanied by genome size expansion and fungi interacting with plants thus possess bigger genomes when compare to saprotrophs or those interacting with animals. Similarly interesting associations with genome size were found also in different groups of microorganisms, for example, correlation of genome size with growth rate and nutritional modes in chrysophytes (Olefeld *et al.* 2018; Čertnerová and Škaloud 2020). However, much more is still waiting to be discovered with more DNA content

data available for microorganisms. This could be achieved with more routine use of FCM in microorganismal research so I fully support the authors' call for more frequent applications of FCM in fungal research (as well as in other microorganismal studies). I also believe many of these tips might find their use in other FCM applications on microorganisms, such as detecting autofluorescence or testing cell viability.

LITERATURE CITED

Catal M, King L, Tumbalam P, Wiriyajitsomboon P, Kirk WW, Adams GC. **2010**. Heterokaryotic nuclear conditions and a heterogeneous nuclear population are observed by flow cytometry in *Phytophthora infestans*. *Cytometry Part A* **77**: 769–775.

Cavalier-Smith T. **1985**. Cell volume and the evolution of genome size In: *The evolution of genome size*. Chichester: Wiley, 105–184.

Čertnerová D. **2021**. Nuclei isolation protocols for flow cytometry allowing nuclear DNA content estimation in problematic microalgal groups. *Journal of Applied Phycology* **33**: 2057–2067.

Čertnerová D, Škaloud P. **2020**. Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences. *Annals of Botany* **126**: 1077–1087.

Doležel J, Greilhuber J, Suda J. **2007**. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**: 2233–2244.

Kullman B. **2000**. Application of flow cytometry for measurement of nuclear DNA content in fungi. *Folia Cryptog. Estonica* **36**: 31–46.

Kullman B, Tamm H, Kullman K. **2005**. *Fungal Genome Size Database*. http://www.zbi.ee/fungal-genomesize/.

Loureiro J, Kron P, Temsch EM, *et al.* 2021. Isolation of plant nuclei for estimation of nuclear DNA content – overview and best practices. *Cytometry Part A* **99**: 318–327.

Olefeld JL, Majda S, Albach DC, Marks S, Boenigk J. **2018**. Genome size of chrysophytes varies with cell size and nutritional mode. *Organisms Diversity & Evolution* **18**: 163–173.

Pombert J-F, Selman M, Burki F, *et al.* 2012. Gain and loss of multiple functionally related, horizontally transferred genes in the reduced genomes of two microsporidian parasites. *Proceedings of the National Academy of Sciences* **109**: 12638–12643.

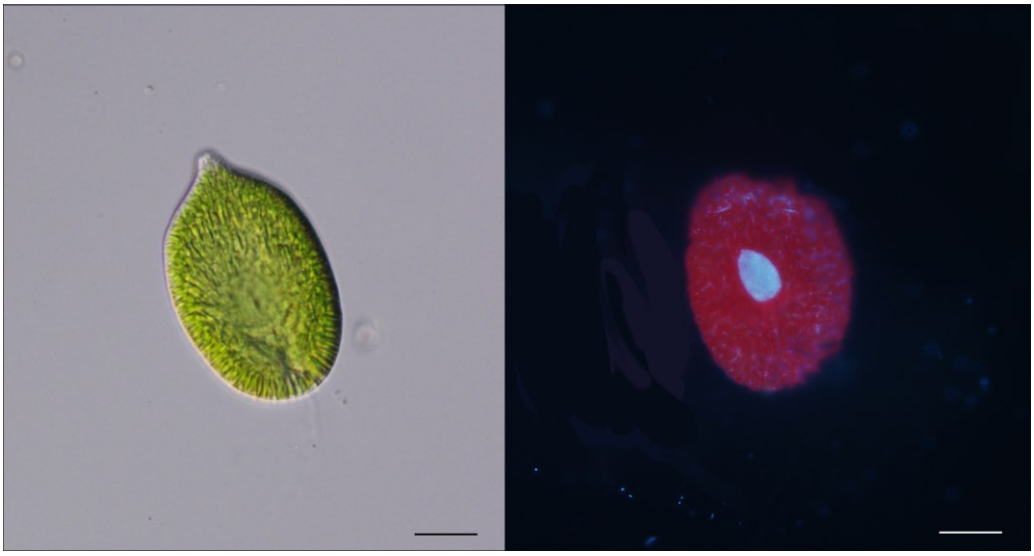
Potter EE, Thornber CS, Swanson JD, McFarland M. **2016**. Ploidy distribution of the harmful bloom forming macroalgae *Ulva* spp. in Narragansett Bay, Rhode Island, USA, using flow cytometry methods. *PLoS ONE* **11**: 1–15.

Talhinhas P, Tavares D, Ramos AP, Gonçalves S, Loureiro J. **2017**. Validation of standards suitable for genome size estimation of fungi. *Journal of Microbiological Methods* **142**: 76–78.

Veselská T, Svoboda J, Růžičková Ž, Kolařík M. **2014**. Application of flow cytometry for genome size determination in *Geosmithia fungi*: A comparison of methods. *Cytometry Part A* **85**: 854–861.

PAPER III

Nuclei isolation protocols for flow cytometry allowing nuclear DNA content estimation in problematic microalgal groups.



Microphotographs of raphidophyte *Gonyostomum semen* under light microscope (A) and DAPI fluorescence (B) showing autofluorescent chlorophyll (red) and gigantic nucleus (light blue). Scale bars represent 30 μm.

Nuclei isolation protocols for flow cytometry allowing genome size estimation in problematic microalgal groups

Dora Čertnerová*

Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic

*Author for correspondence (e-mail: dora.certnerova@gmail.com)

ABSTRACT

Microalgae are fundamentally important organisms for global ecosystem functioning with high potential in biotechnology and its applications. The knowledge of their nuclear DNA content has become a prerequisite for many areas of microalgal research. Due to common presence of various pigments, secondary metabolites and complex cell walls, the nuclear DNA content estimation using flow cytometry (FCM) is, however, often laborious or even impossible with the currently used protocols. In this study the performance of six nuclei isolation protocols was compared on various problematic microalgae using FCM. The nuclei isolation methods involved osmotic bursting of cells, razor blade chopping of fresh biomass and two newly introduced protocols, razor blade chopping of desiccated biomass and bead beating. These techniques also involved the use of two different nuclei isolation solutions, Otto I + II solutions and LB01 buffer. Performance of the particular protocols differed greatly, depending on the used nuclei isolation solution and microalgal group. The most successful method was a newly adopted chopping of desiccated biomass in LB01 buffer. This method seems more appropriate for nuclei isolation in filamentous microalgae, on the other hand, bead-beating appears to be more suitable for nuclei isolation in solitarily living algae. Using the optimal protocol for a given species, their nuclear DNA content was estimated, resulting in first DNA content estimates for four investigated taxa (*Chlamydomonas noctigama*, *Gonyostomum semen*, *Microglena* sp. and *Stigeoclonium* sp.). The estimated DNA content spanned from 0.15 to 32.52 pg.

Key words: nuclei isolation, flow cytometry, microalgae, silica gel desiccation, bead beating, nuclear DNA content

INTRODUCTION

Microalgae are photosynthetic microorganisms that occur across a wide range of habitats from freshwater lakes to desert soils. Due to their polyphyletic origin across the tree of life, they are remarkably diverse group of organisms. Moreover, microalgae play a key role in the global ecosystem as primary producers and major source of oxygen. Recently, a considerable attention has been paid to microalgae as the potential source of next generation biofuels or usable metabolites (Brennan and Owende 2010; Hyka *et al.* 2013; Milano *et al.* 2016; Khan *et al.* 2018). This led to a need for microalgal DNA content data due to a number of reasons. First, this knowledge enables us to select lineages with potentially higher secondary metabolite production given that an increase in DNA content is often coupled with an increase in gene dosage (e.g. due to aneuploidsation or polyploidisation; Mason 2016; Priyadarshan 2019; Qin *et al.* 2019). Second, the recent attention drawn to microalgae accelerated the whole-genome sequencing effort and the DNA amount is the key to designing an optimal sequencing strategy. Further, the nuclear DNA content directly influences the cost of a sequencing project, hence, the low DNA content has become a major criterion in selection of appropriate algal strains (Waaland *et al.* 2004; Peters *et al.* 2004; Lin 2006). The combination of DNA content knowledge and high-level phylogeny also opens the ways to determine evolutionary trends in DNA content variation. Such innovative studies brought new insights into microalgal nutrition modes or cell-size changes (Poulíčková *et al.* 2014; Olefeld *et al.* 2018). Further, the nuclear DNA content, at least in relative units, is essential for cell cycle determination (Lemaire *et al.* 1999; Reinecke *et al.* 2018).

The most suitable method for precise and rapid nuclear DNA content estimation is flow cytometry (FCM). Using FCM, we are able to detect fluorescent-stained particles (e.g. cells, isolated nuclei) in a stream of fluid (Doležel *et al.* 2007). While FCM has found a broad spectrum of applications in genomic surveys on plants and animals (e.g. Dionisio Pires *et al.* 2004; Kron *et al.* 2007; Galbraith 2012; Chang *et al.* 2018; Sadílek *et al.* 2019), it has been only rarely applied in algal studies (but see Figueroa *et al.* 2010; Hyka *et al.* 2013).

There are several reasons causing the gap of nuclear DNA content estimates in microalgae. First, it is almost always necessary to cultivate microalgal strains from a single cell / filament to obtain sufficient amounts of biomass for FCM analysis. However, this is considerably time consuming and for some species even hard to accomplish. Because of the great diversity of microalgae, there is also a wide range of pigments and metabolites that frequently interfere with fluorescent stain and / or create pronounced background noise, prominent especially when whole - intact cells are analysed (Simon *et al.* 1994; Veldhuis *et al.* 1997; Mazalová *et al.* 2011). Although, the pronounced cytoplasmic autofluorescence as well as nonspecific background fluorescence can be lowered by chemical fixation, such approach is far from optimal due to reduced quality of FCM analyses. Instead, protoplast extraction

and preparation of nuclear suspension is much more suitable (Doležel and Bartoš 2005). To achieve this, several methods of cell-wall disruption can be implemented in a sample preparation protocol for FCM analysis. However, currently used protocols for microalgae often does not work for FCM. Commonly used nuclei isolation method is chopping the biomass by a razor blade combined with various enzymatic treatments (Mazalová *et al.* 2011; Weiss *et al.* 2011; Poulíčková *et al.* 2014). In many cases, the enzymatic treatment was applied to chemically dissolve the cell walls without the need for any further mechanical disruption (Mazalová *et al.* 2011; Poulíčková *et al.* 2014). The application of enzymatic treatment on algal samples was originally adopted from plant or fungal studies (Jazwinski 1990; Doležel *et al.* 2007) and the predominantly used enzymes for microalgal species are cellulase, macerozyme and lyticase (Mazalová *et al.* 2011; Weiss *et al.* 2011; Poulíčková *et al.* 2014). In Mazalová *et al.* (2011), the authors introduced an enzymatic treatment that was subsequently tested on a broad variety of microalgal species. The enzymatic mixture was primarily developed for Streptophyte algae (e.g. the genus *Zygnema*), but also worked with some Chlorophyta (*Chloroidium ellipsoideum*, *Tetraselmis subcordiformis*) and Ochrophyta (*Tribonema vulgare*). Despite this, the introduced protocol did not work for nearly half of the tested microalgae, among others, for the green algae *Trentepohlia* sp. or *Chlamydomonas noctigama* (referred there as *C. geitleri*).

Unfortunately, the utilization of enzymatic treatment is methodologically demanding as well as time-consuming. Moreover, due to the great algal diversity, enzymatic treatment often requires additional modifications for specific algal groups (Mazalová *et al.* 2011; Weiss *et al.* 2011; Potter *et al.* 2016). However, use of the enzymatic treatment predominates as a protoplast isolation technique in microalgal studies despite these disadvantages. To resolve the situation, new methods of nuclei isolation for FCM analysis needs to be established for microalgae. For example, the most common way of nuclei isolation in plants or seaweeds is, simple chopping tissue using a razor blade (Galbraith *et al.* 1983; Asensi *et al.* 2001; Doležel *et al.* 2007). Further, beat beating by zirconium or silica beads has been previously used to isolate nuclei of bacteria (Gryp *et al.* 2020), fungi (Griffin *et al.* 2002), plants (Roberts 2007), and animals (Harmon *et al.* 2006). Interestingly, despite its easy and rapid use, neither bead beating nor chopping by a razor blade alone were ever successfully applied to i solate microalgal nuclei for FCM.

The aim of this study is to develop new protocols of microalgal nuclei isolation and test them on a diverse set of species that were referred as problematic in the past (Mazalová *et al.* 2011; personal observation).

MATERIALS AND METHODS

Origin, cultivation and harvesting of investigated strains

Monoclonal cultures used in this study were obtained from Culture Collection of Algae of Charles University in Prague (CAUP), Culture Collection of Cryophilic Algae (CCCryo), Norwegian Culture Collection of Algae (NORCCA) and from collaborators (Table 1). The algal taxa chosen for this study were selected based on the previous difficulties with their nuclei extraction and / or FCM analysis (Mazalová *et al.* 2011; author's personal observation in pilot FCM analyses). A special focus is paid to *Zygnema* spp. strains as this species genus is the model organism in recent studies in our working group (e.g. Pichrtová *et al.* 2018; Trumhová *et al.* 2019).

The strains were cultivated either in 50 mm Petri dishes filled with Bold's Basal medium (BBM; Bischoff and Bold 1963) solidified with 1.5% agar or in 50 mL Erlenmeyer flasks filled with liquid BBM or modified WC medium (MWC; Guillard and Lorenzen 1972). The majority of cultures were maintained at 17 °C with constant light conditions under the illumination of 30 - 50 μmol photons m^{-2} s^{-1}. The *Chlamydomonas noctigama* and *Microglena* sp. strains were cultivated at 23°C with 14h light and 10h dark conditions under the illumination of 100 μmol photons m^{-2} s^{-1}. Origin details and cultivation media for particular algal strains are listed in Table 1. The cultures were transferred into a fresh medium 2 to 5 weeks before the planned FCM analyses and their biomass growth regularly checked. Afterwards, the culture biomass was harvested in their exponential phase of growth. Approximately 15-30-mg bulk of biomass was collected from cultures growing on solidified medium (BBM-agar) using an inoculation needle with a bent tip. Similarly, 2 mL of strains cultivated in liquid medium (BBM or MWC) were transferred into an Eppendorf tube, centrifuged (5,500 rpm for 5 min) and superfluous medium removed by pipetting.

Nuclei isolation and staining

In total, six nuclei isolation protocols were subsequently tested on the studied algal strains. In each protocol, either LB01 buffer (15 mM Tris, 2 mM Na2EDTA, 0.5 mM spermine tetrahydrochloride, 80 mM KCl, 20 mM NaCl, 0.1 % (v/v) Triton X-100; pH = 8.0; Doležel *et al.* 1989) or a two-step Otto protocol (Otto I solution consisting of 0.1 M citric acid, 0.5% Tween 20 with pH = 2.0 - 3.0 and Otto II solution consisting of 0.4 M $Na_2HPO_4 \cdot 12H_2O$ with pH = 8.0 – 9.0; Otto 1990) were used.

Protocol 1

Single-celled algal strains (*C. noctigama*, *Microglena* sp. and *Gonyostomum semen*) were prepared for the FCM analysis without any protoplast extraction, i.e. whole cells of each strain were mixed with 550 μL of ice-cold LB01 lysis buffer or Otto I solution, to attempt a release of nuclei by osmotic bursting of cells. The suspension

Table 1. Original collection site and cultivation media for the investigated algal strains.

Class	Species	Strain	Original collection site	Cultivation medium
Zygnematophyceae	*Spirogyra* sp.	CAUP K902	pond near Winterthur, Switzerland	BBM-agar
Zygnematophyceae	*Zygnema* sp.	13 179-4	near Pyramiden, Svalbard	BBM-agar
Zygnematophyceae	*Zygnema* sp.	15 Osor 2	puddle, near Osor, Croatia	BBM-agar
Zygnematophyceae	*Zygnema* sp.	CCCryo 171-04	Mountain creek, Poatina, Tasmania, Australia	BBM-agar
Chlorophyceae	*Chlamydomonas noctigama*	CAUP G224 (SAG 6.73 / UTEX 2289)	Hvězda pond, Northern Moravia, Czech Republic	BBM
Chlorophyceae	*Microglena* sp.	Fio 17	Lake Fiolen, Småland, Sweden	BBM
Chlorophyceae	*Stigeoclonium* sp.	CAUP J603	Žebrákovský creek, river basin of Sázava, Czech-Moravian Highlands, Czech Republic	BBM-agar
Raphidophyceae	*Gonyostomum semen*	NIVA-2/10 (BO-182)	Lake Bökesjön, Scandia, Sweden	modified WC
Ulvophyceae	*Trentepohlia* sp.	CAUP J1601	bark, Singapore	BBM-agar
Xanthophyceae	*Tribonema vulgare*	CAUP D501	Palach Pond near Lednice, Czech Republic	BBM-agar

was thoroughly mixed and filtered through a 42 μm nylon mesh into a special 3.5-mL cuvette for direct use with the flow cytometer. Following a 20-min. incubation* at room temperature, staining solution consisting of either 1 mL Otto II solution or 550 μL LB01 lysis buffer, of 50 μg · mL^{-1} propidium iodide, of 50 μg · mL^{-1} RNase IIA and of 2 μ · mL^{-1} β-mercaptoethanol was added to the sample.

Protocol 2

Harvested biomass was transferred to a plastic Petri dish and chopped by a razor blade in 550 μL of ice-cold Otto I solution. The resulting suspension was thoroughly mixed and filtered through a 42 μm nylon mesh into a special 3.5-mL cuvette for direct use with the flow cytometer. Following a 20-min. incubation* at room temperature, staining solution consisting of 1 mL of Otto II solution, of 50 μg · mL^{-1} propidium iodide, of 50 μg · mL^{-1} RNase IIA and of 2 μl μL · mL^{-1} β-mercaptoethanol was added to the sample.

Protocol 3

Harvested biomass was transferred into a plastic Petri dish and chopped by a razor blade in 550 μL of ice-cold lysis buffer LB01. The resulting suspension was thoroughly mixed and filtered through a 42 μm nylon mesh into a special 3.5-mL cuvette for direct use with the flow cytometer. Following a 20-min. incubation* at room temperature, staining solution consisting of 550 μL of LB01 lysis buffer, of 50 μg · mL^{-1} propidium iodide, of 50 μg · mL^{-1} RNase IIA and of 2 μL · mL^{-1} β-mercaptoethanol was added to the sample.

Protocol 4

Harvested biomass was desiccated by transferring into 2-mL Eppendorf tube and placed with an open lid into a zip-lock bag filled with silica gel for 2 to 5 days. The dry algal biomass was then transferred in a plastic Petri dish and chopped by a razor blade in 550 μL of ice-cold lysis buffer LB01. The sample preparation was further completed according to Protocol no. 3.

Protocol 5

Approximately 10 glass beads of 1.5 mm diameter (Sigma-Aldrich) were added into 2-mL Eppendorf tube containing 550 μL of ice-cold Otto I solution and a biomass pellet. The cells were disrupted for 3 min at 25 Hz using Retsch MM200 mixer mill (Retsch, Inc., Haan, Germany). The nuclei suspension was then filtered through a 42 μm nylon mesh into a special 3.5-mL cuvette for direct use with the flow cytometer. Following a 20-min. incubation* at room temperature, staining

*If visible sediment was present after 20-min. incubation, an upper layer of nuclei suspension was transferred into a new cuvette and used as a material for analysis.

solution consisting of 1 mL Otto II solution, 50 μg · mL^{-1} propidium iodide, 50 μg · mL^{-1} RNase IIA and 2 μL · mL^{-1} β-mercaptoethanol was added to the sample.

Protocol 6

Approximately 10 glass beads of 1.5 mm diameter (Sigma-Aldrich) were added into 2-mL Eppendorf tube containing 550 μL of ice-cold lysis buffer LB01 and the pellet of biomass. The cells were disrupted for 3 min at 25 Hz using Retsch MM200 mixer mill (Retsch, Inc., Haan, Germany). The sample was filtered through a 42 μm nylon mesh into a special 3.5-mL cuvette for direct use with the flow cytometer. Following a 20-min. incubation* at room temperature, staining solution consisting of 550 μL of LB01 lysis buffer, of 50 μg · mL^{-1} propidium iodide, of 50 μg · mL^{-1} RNase IIA and of 2 μL · mL^{-1} β-mercaptoethanol was added to the sample.

Standardization

Initially, nuclei suspensions for FCM analysis were prepared without a standard. When a suitable nuclei extraction protocol was found for a given species, an internal standard was included into following analyses. Four different plants were used as standards in this study – wild clone of *Carex acutiformis* (2C = 0.82 pg; Veselý *et al.* 2012), commercial clone of *Solanum pseudocapsicum* (2C = 2.59 pg; Temsch *et al.* 2010), wild clone of *Bellis perennis* (2C = 3.38 pg; Schönswetter *et al.* 2007) and *Vicia faba* cv. Inovec (2C = 26.90 pg; Doležel *et al.* 1992). To release the standard nuclei, ca. 20- mg piece of fresh leaf tissue was chopped with a razor blade either together with an algal sample (Protocols 2, 3 and 4) or separately, in a fraction of used nuclei isolation solution and later mixed with the protoplast suspension containing the remaining solution (Protocol 5 and 6). When razor chopping was used to isolate nuclei of both algal sample and plant standard, the algal biomass was chopped slightly less than the plant standard. The resulting nuclei suspension was filtered and stained as described in Protocols 1 – 6.

DNA content estimation

The stained samples were immediately analysed using a Partec CyFlow SL cytometer (Partec GmbH, Münster, Germany) equipped with a green solid-state laser (Cobolt Samba, 532 nm, 100 mW) and aside from PI fluorescence intensity optical parameters Forward Scatter (FSC) and Side Scatter (SSC) were recorded. Each sample measurement was taken for up to 5,000 particles. The success rate of particular protocol was evaluated as follows: 1) No peak – sample peak undistinguishable from the background noise or not detected; 2) Poor analysis – sample peak visible but its position hardly recognizable from the background noise (yet apparent on a relative fluorescence vs. side scatter plots); 3) Good result – sample peak clearly visible with reduced background noise.

To properly analyse DNA content of the studied algal strains, at least three measurements were done on separate days to obtain precise value and to minimize

the effect of random instrumental shift. The resulting FCM histograms were analysed using FloMax ver. 2.4d (Partec). The lowest fluorescence intensity sample peaks were identified as G_1 (vegetative cells) and additional peaks with double fluorescence intensity (if observed) as G_2. Gating of sample nuclei in fluorescence vs. side-scatter plots was necessary to remove the background noise connected to the populations of interest in order to obtain more accurate results (with an exception of *G. semen*). The absolute nuclear DNA content was calculated as the sample G_1 peak mean fluorescence / standard G_1 peak mean fluorescence × standard 2C DNA content (according to Doležel and Bartoš 2005). Since the ploidy level or life cycle stage of studied organisms is generally unknown, the DNA content results are given in pg · cell^{-1}, i.e. the absolute nuclear DNA content measured per cell (1 pg ≈ 978 Mbp; Doležel *et al.* 2003). The quality and accuracy of resulting DNA content estimates was expressed by averaged coefficient of variation (CV) for individual sample peaks and standard deviation (SD) for measurements error averaged from the three independent measurements.

RESULTS

Comparison of isolation protocols

Altogether, six nuclei isolation protocols were tested and compared on a set of ten problematic algal taxa (Table 1). Protocol success rate was evaluated using a three quantitative scale (see Materials and methods). The results differed greatly according to the used protocol and algal sample tested (Table 2).

Except analysing cells only mixed with nuclei isolation solution (Protocol 1), broadly used technique of razorblade sample chopping (Protocol 2 and 3) was the least successful method in this study. When Otto I + II solutions were used (Protocol 2), none of the tested strains resulted in a visible sample peak. The razor- blade chopping technique was successful only in combination with LB01 isolation buffer (Protocol 3), resulting in a clearly visible and well separated peaks for microalga *Stigeoclonium* sp. and *Tribonema vulgare*. A partial success of Protocol 3 was also achieved for all *Zygnema* strains, however, still leading to a high background noise and often hardly distinguishable sample peak (Fig. 1a, d).

In contrast, Protocol 4, combining sample desiccation with razor blade chopping in LB01 isolation buffer, was the most successful from all the tested methods. This protocol resulted in clearly visible and well separated peaks for microalgae *Spirogyra* sp., *Trentepohlia* sp. (Fig. 1h, k) and all three analysed strains of *Zygnema* sp. (Fig. 1g, j). Interestingly, the same quality of analysis with *Trentepohlia* was observed when Otto I + II solutions were used instead of LB01 buffer, however, this was not examined for any other microalgal strain. The Protocol 4 was further partially successful for the species *Stigeoclonium* sp. and *T. vulgare*, however, leading to a more pronounced background noise compared to the same method without the desiccation step (Protocol 3). On the other hand, this method failed to result in any

Table 2. Comparison of six different nuclei isolation protocols applied across the studied algal strains, including the type of biomass, used nuclei isolation solution and cell disruption technique. The outcomes are categorized as No peak (sample peak undistinguishable from the background noise or not detected), Poor analysis (sample peak visible but its position hardly recognizable from background noise; yet apparent on a Side Scatter), Good result (sample peak clearly visible with minimum background noise) and Not tested.

	Biomass	Nuclei isolation solution	Cell disruption	*Spirogyra* sp.	*Zygnema* spp.	*Chlamydomonas noctigama*	*Microglena* sp.	*Stigeoclonium* sp.	*Gonyostomum semen*	*Trentepohlia* sp.	*Tribonema vulgare*
Protocol 1	fresh	LB01 / Otto I + II	osmotic	Not tested	Not tested	No peak	No peak	Not tested	No peak	Not tested	Not tested
Protocol 2	fresh	Otto I + II	razor chopping	No peak	No peak	No peak	No peak	No peak	No peak	No peak	No peak
Protocol 3	fresh	LB01	razor chopping	No peak	Poor analysis	No peak	No peak	Good result	No peak	No peak	Good result
Protocol 4	desiccated	LB01	razor chopping	Good result	Good result	No peak	No peak	Poor analysis	No peak	Good result[1]	Poor analysis
Protocol 5	fresh	Otto I + II	bead beating	Not tested	Not tested	No peak	No peak	Not tested	No peak	Not tested	Not tested
Protocol 6	fresh	LB01	bead beating	Not tested	No peak	Poor analysis	Poor analysis	No peak	Good result	No peak	No peak

[1]The same quality of FCM analysis in *Trentepohlia* strain was observed when Otto I + II solutions instead of LB01 buffer were used in Protocol 3.

sample peaks for the species *Chlamydomonas noctigama*, *Microglena* sp. and *Gonyostomum semen* (Fig. 1c, f). The only successful method for analysing these microalgal species was the Protocol 6. In this protocol, the nuclei were extracted by bead-beating cells in LB01 isolation buffer. This method was particularly suitable for *G. semen*, where it led to a high-quality analysis with nearly no visible background noise (Fig. 1i, l). Contrarily, the analyses of *C. noctigama* and *Microglena* sp. were of very low quality (pronounced background noise and poor peak delimitation). However, the Protocol 6 was the only protocol leading to any sample peak for these species. Interestingly, the same method of nuclei isolation by bead beating successful for species *C. noctigama*, *Microglena* sp. and *G. semen* did not work when Otto I + II solutions were used (Protocol 5) instead of LB01 isolation buffer (Protocol 6). Therefore, the Protocol 5 was not further examined for the remaining strains.

Nuclear DNA content estimation

When the most suitable protocol for particular species was found, their absolute nuclear DNA content per cell was thoroughly investigated (Table 3). The nuclear DNA content of studied microalgal strains is given in pg of absolute nuclear DNA per cell with equivalent values in Gbp (1 pg ≈ 0.978 Gbp; Doležel *et al.* 2003). The DNA content differed greatly, spanning from 0.15 (0.14) to 32.52 pg (31.81 Gbp). The smallest DNA content belonged to the representatives of the class Chlorophyceae with 0.15 pg (0.14 Gbp) for *Stigeoclonium* sp., 0.33 pg (0.33 Gbp) for *C. noctigama*, and 0.44 pg (0.43 Gbp) for *Microglena* sp. and to the representative of the class Xanthophyceae with 0.34 pg (0.34 Gbp) for *T. vulgare*. In contrast, the largest measured DNA content of 32.52 pg (31.81 Gbp) belonged to *G. semen* from the class Raphidophyceae. The three analysed strains of the genus *Zygnema* varied in their DNA content (1.11 – 2.86 pg ≈ 1.09 – 2.73 Gbp). The highest quality of DNA content estimates was observed with in *G. semen* and one of *Zygnema* strains, with coefficients of variation (CVs) <2 % (1.14 and 1.75 %, respectively). On the other hand, the lowest quality of DNA content estimates was documented in *Stigeoclonium* sp., *Microglena* sp. and *Spirogyra*, exceeding 13 % (13.51, 13.54, and 13.65 %, respectively).

DISCUSSION

Nuclei isolation protocols.

In this study, two new nuclei isolation protocols for FCM are proposed and applied on various samples of microalgae. The newly introduced methods involve either sample desiccation before razor- blade chopping or bead beating of the sample biomass. Both methods are easy to use and bring satisfactory results of DNA content estimation in microalgae, even for problematic taxa. These new methods were compared with more broadly used techniques for microalgae, i.e. analysis of

Table 3. Absolute nuclear DNA content per cell estimated for the studied algal strains. The average DNA content estimates based on three independent measurements are provided in pg of DNA (with and in equivalent values in Gbp), along with average coefficient of variation for analyses (CV) and details on the used internal reference standard.

Species	Strain	Average genome size		Mean CV (%)	Internal reference standard
		[pg] ± SD	[Gbp]		
Zygnema sp.	13 179-4	1.112 ± 0.05	1.087	4.50	*Bellis perennis* (2C = 3.38 pg)
Zygnema sp.	15 Osor 2	2.394 ± 0.10	2.342	4.18	*Bellis perennis* (2C = 3.38 pg)
Zygnema sp.	CCCryo 171-04	2.856 ± 0.05	2.793	1.75	*Bellis perennis* (2C = 3.38 pg)
Stigeoclonium sp.	CAUP J603	0.148 ± 0.02	0.144	13.51	*Carex acutiformis* (2C = 0.82 pg)
Tribonema vulgare	CAUP D501	0.342 ± 0.01	0.335	2.92	*Carex acutiformis* (2C = 0.82 pg)
Trentepohlia sp.	CAUP J1601	1.167 ± 0.02	1.141	1.56	*Solanum pseudocapsicum* (2C = 2.59 pg)
Gonyostomum semen	NIVA-2/10 (BO-182)	32.523 ± 0.37	31.807	1.14	*Vicia faba* cv. Inovec (2C = 26.9 pg)
Microglena sp.	Fio19	0.443 ± 0.06	0.434	13.54	*Carex acutiformis* (2C = 0.82 pg)
Chlamydomonas noctigama	CAUP G224 (SAG 6.73 / UTEX 2289)	0.333 ± 0.01	0.326	3.00	*Carex acutiformis* (2C = 0.82 pg)
Spirogyra sp.	CAUP K902	1.026 ± 0.14	1.003	13.65	*Solanum pseudocapsicum* (2C = 2.59 pg)

osmotic bursting of cells (applied on unicellular algae) and razor blade chopping of fresh biomass (all tested algae).

Moreover, these new techniques allowed for the first time DNA content estimation in *Chlamydomonas noctigama*, *Gonyostomum semen*, *Microglena* sp. and *Stigeoclonium* sp. Further, identical strains of *C. noctigama* (strain CAUP G224) and *Stigeoclonium* sp. (strain CAUP J603) were already examined in the study Mazalová *et al.* (2011) using enzymatic mixture for protoplast extraction, however, without any success. In this work, *C. noctigama* was successfully analysed by applying bead beating of the biomass in LB01 isolation buffer (Protocol 6). Interestingly, a suitable method for analysing *Stigeoclonium* sp. was simple razor blade chopping of the biomass in LB01 isolation buffer (Protocol 3), not a protoplast extraction using enzymatic mixture (Mazalová *et al.* 2011) or any other method used in this study (except Protocol 4, see later). Moreover, the Protocol 3 was also the best method to analyse *T. vulgare*. Although this taxa taxon was already successfully analysed with the use of enzymatic mixture in the study Mazalová *et al.* (2011), the enzymatic treatment is methodologically demanding as well as time consuming. In contrast, razor blade chopping of a fresh sample is very simple and rapid method and sometimes, as seen on the example of *Stigeoclonium* sp. and *T. vulgare*, also the optimal method for FCM without the need for further optimization. Therefore, this simple method is still worth a try when conducting pilot FCM measurements on other microalgal species. Both *Stigeoclonium* sp. and *T. vulgare* were also successfully analysed using desiccation step followed by razor blade chopping (Protocol 4), however, resulting in a reduced quality of the FCM analysis.

In general, razor blade chopping of biomass in LB01 isolation buffer either preceded by the desiccation step (Protocol 4) or without it (Protocol 3) appears to be more efficient way of nuclei isolation in filamentous microalgae. The success of the desiccation using silica gel is especially interesting since this led to a decrease of quality in FCM analysis of vascular plants (Kolář *et al.* 2012). However, desiccating the biomass of microalgae might have reduced the negative effect of secondary metabolites interfering with DNA staining. For example, high amounts of secondary metabolites such as phenols were documented in *Trentepohlia* sp., *Spirogyra* sp. and *Zygnema* spp. (Simić *et al.* 2012; Pichrtová *et al.* 2013; Mridha *et al.* 2017). Phenolic compounds can significantly decrease the quality of FCM analyses (Loureiro *et al.* 2006a), and the desiccation might reduce their negative effect (along with possibly other metabolites) on FCM analysis. On the other hand, optimal algal material for FCM analysis are young cultures approximately 3 to 5 weeks after their inoculation into a fresh medium, yet, young *Zygnema* cells are known to contain high amounts of phenolic compounds (Holzinger *et al.* 2018), contradicting the benefits of using young cultures. However, analysis of *Zygnema* spp. cultures older than 5 weeks resulted only in a background noise (data not shown). Another explanation could be the putative role of desiccation in disturbing layers of polysaccharide present on

Zygnema and *Spirogyra* filaments (Palacio-López *et al.* 2019), facilitating the release of their nuclei.

Bead-beating of biomass in LB01 buffer (Protocol 6) seems to be more suitable for solitarily living algae. Even though the cell disruption by bead-beating was previously used to isolate DNA of algae (e.g. Countway and Caron 2006), to my knowledge, it has never used as a method for nuclei isolation in algal FCM. This technique was particularly suitable for *G. semen*, where it resulted in clear FCM histograms with very limited background noise (Fig. 1l, i). Bead-beating of cells in LB01 buffer (Protocol 6) is also the only method that gained any DNA content estimates for *C. noctigama* and *Microglena* sp. not only in this study but also including unsuccessful attempts in Mazalová *et al.* (2011). However, the outcomes were of very poor quality and further optimizat ion is needed to obtain more precise results.

It is worth emphasizing the importance of selecting optimal nuclei isolation solution when employing FCM on algal samples. In this study, only two nuclei isolation solution were used (LB01 buffer and Otto I + II solutions), however, their performance was completely different. When using Otto I + II solutions (Protocol 1, Protocol 2 and Protocol 5), the analyses led to no visible sample peaks (with the exception of *Trentepohlia* sp. with equally good results under the use of both buffers). Vast majority of the successful analyses were done using LB01 isolation buffer. The differences between LB01 buffer and Otto I + II solutions are in their different chemical composition but also in strikingly distinct pH level (2-3 and 8, respectively; Loureiro *et al.* 2006b). This stresses the importance of selecting an optimal isolation solution and comparing to others might be a next step in further optimization.

DNA content estimates of the studied algae.

The nuclear DNA content of four algal taxa was successfully estimated with the smallest measured DNA content represented by *Stigeoclonium* sp. with 0.15 pg (0.14 Gbp). To my knowledge, this also represents the first DNA content estimate for whole order Chaetophorales. On the other hand, the largest DNA content measured in this study belongs to raphidophyte *G. semen* with 32.52 pg (31.81 Gbp). The only representative of the class Raphidophyceae that has been analysed for DNA content so far was marine *Heterosigma carterae* possessing a genome five times smaller (5.43 - 6.12 pg / 5.31 – 5.98 Gbp; Veldhuis *et al.* 1997). In contrast, more DNA content data are available for the genus *Chlamydomonas* with estimates ranging from 0.08 – 0.40 pg (0.08 – 0.39 Gbp; Chiang and Sueoka 1967; Kates *et al.* 1968; Cattolico and Gibbs 1975; Spring *et al.* 1978; Veldhuis *et al.* 1997; Merchant *et al.* 2007; Reinecke *et al.* 2018; Nelson *et al.* 2019). However, only a few of these estimates were acquired using FCM. The DNA content of *C. noctigama* estimated in this study (0.33 pg / Gbp) is rather large but still falling within the previously published range.

The identical strain of *T. vulgare* (CAUP D 501) was previously analysed by Mazalová *et al.* (2011), however, leading to a slightly different result of 0.41 pg (0.40

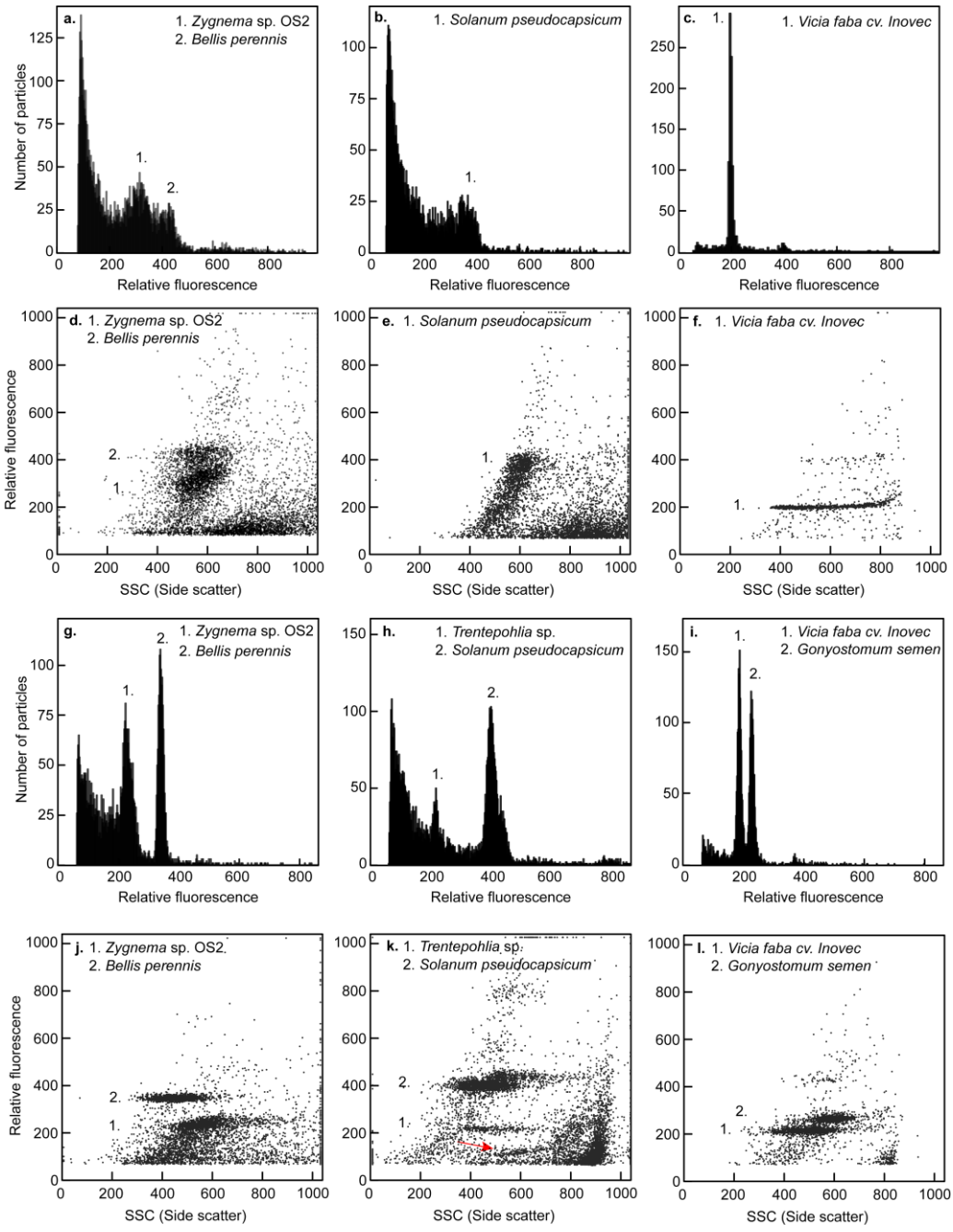


Figure 1. Flow cytometric fluorescence histograms (a-c, g-i) and fluorescence vs. side-scatter plots (d-f, j-l) summarizing the results of poor-quality (a-f) and suitable (g-l) nuclei isolation protocols of *Zygnema* sp. OS2, *Trentepohlia* sp. and *G. semen* strains with internal reference standards. The nuclei of *Zygnema* sp. OS2 isolated by a razor blade chopping of fresh biomass in LB01 buffer (Protocol 3) resulted in visible sample and standard peaks with pronounced background noise (a, d), while using razor blade chopping of desiccated biomass (Protocol 4) led to prominent sample and standard peaks (g, j). Isolation of *Trentepohlia* sp. nuclei using the

Protocol 3 did not result in visible sample peak (b, e), contrary to nuclei isolation with the Protocol 4, where the sample peak is clearly visible and well separated (h, k). Note that in fluorescence vs. side-scatter plot the peak of presumed haploid zoospores can be identified (k, indicated by the arrow). The Protocol 4 did not result in any sample peak for *G. semen* (c, f), however, when nuclei were isolated by cell bead-beating in LB01 buffer (Protocol 6), it led to a clear sample peak with nearly no background noise (i, l).

Gbp) compared to 0.34 pg / Gbp estimated in this study. This variance might be induced by use of a different FCM standard. In this study, the plant *Carex acutiformis* was used in opposite to *Raphanus sativus* cv. Saxa used in the study Mazalová *et al.* (2011). However, the latter FCM standard displays many difficulties like high CVs, polyploidy, higher presence of secondary metabolites, reported genome size of different values, therefore its use was repeatedly discouraged (Doležel *et al.* 1992; Praça-Fontes *et al.* 2011; Park *et al.* 2016; Šmarda *et al.* 2019). The only available DNA content data for the genera *Trentepohlia* and *Spirogyra* originate from DAPI microdensitometry (Kapraun 2005, 2007; López-Bautista *et al.* 2006). However, DAPI fluorescent stain binds to adenine–thymine-rich regions and therefore may brings the erroneous estimates of AT:GC ratio of the sample and the reference standard (Doležel *et al.* 1992). For the both genera, the DNA content estimates in this study were the first ones acquired using FCM. The previously published estimates for *Trentepohlia* sp. span from 1.08 to 4.01 pg (1.10 – 4.10 Gbp; López-Bautista *et al.* 2006; Kapraun 2007) and thus the estimate measured in this study (1.17 pg / 1.14 Gbp) falls within the published range. Interestingly, in relative fluorescence vs. side scatter plots of some *Trentepohlia* sp. FCM analyses is possible to identify three sample peaks that differ in their ploidy level (Fig. 1k). An abundant population of nuclei belonging to the intermediate ploidy was identified as nuclei of vegetative filament (G_1). The peak of highest *Trentepohlia* sp. ploidy was determined as dividing nuclei of the vegetative filament (G_2), unfortunately partially overlapping with the standard nuclei. The peak of the lowest ploidy with smallest population of nuclei may represent haploid zoospores. The presence of sporangia containing zoospores was subsequently confirmed by observation using light microscopy. The previous DNA content estimates of *Spirogyra* sp. (3.91 – 4.01 pg / 4.00 – 4.10 Gbp; Kapraun 2005) were four times higher than in this study (1.03 pg / 1.00 Gbp). Despite the fact that the measurements were conducted on different *Spirogyra* strains and by different techniques, these results probably reflect high DNA content variability within the *Spirogyra* genus. Similarly, DNA content variability within the genus *Zygnema* will likely be much higher than documented to date. The known DNA content range is from 0.49 to 1.5 pg (0.50 – 1.54 Gbp; Kapraun 2005; Mazalová *et al.* 2011). However, analyses of three *Zygnema* strains displayed DNA content between 1.11 to 2.86 pg (1.09 - 2.80 Gbp) and thus the previous DNA content range for the genus was nearly doubled.

I believe the presented new nuclei isolation protocols will provide alternative ways of microalgal FCM and apply to a broad range of various species of microalgae. Hopefully, the newly introduced protocols will help to extend yet very limited DNA content data of microalgae and these data will subsequently serve to various microalgae applications.

FUNDING

No funding was received for conducting this study.

ACKNOWLEDGEMENTS

I thank Martina Pichrtová (Department of Botany, Charles University) and Karin Rengefors (Department of Biology, Lund University) for providing several algal cultures used in this study. I also thank Paul Kron (Department of Integrative Biology, Guelph University) for his suggestions at the initial stage of this study.

LITERATURE CITED

Asensi A, Gall EA, Marie D, Billot C, Dion P, Kloareg B. **2001**. Clonal propagation of *Laminaria digitata* (Phaeophyceae) sporophytes through a diploid cell-filament suspension. *Journal of Phycology* **37**: 411–417.

Bischoff H, Bold H. **1963**. *Phycological Studies IV. Some Soil Algae from Enchanted Rock and Related Algal Species*. University Austin, Texas, USA.

Brennan L, Owende P. **2010**. Biofuels from microalgae—A review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews* **14**: 557–577.

Cattolico RA, Gibbs SP. **1975**. Rapid filter method for the microfluorometric analysis of DNA. *Analytical Biochemistry* **69**: 572–582.

Chang P, Tseng Y-F, Chen P-Y, Wang C-JR. **2018**. Using Flow Cytometry to Isolate Maize Meiocytes for Next Generation Sequencing: A Time and Labor Efficient Method. *Current Protocols in Plant Biology* **3**: e20068.

Chiang KS, Sueoka N. **1967**. Replication of chloroplast DNA in *Chlamydomonas reinhardi* during vegetative cell cycle: its mode and regulation. *Proceedings of the National Academy of Sciences* **57**: 1506–1513.

Countway PD, Caron DA. **2006**. Abundance and distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as revealed by quantitative PCR. *Applied and environmental microbiology* **72**: 2496–506.

Dionisio Pires LM, Jonker RR, Van Donk E, Laanbroek HJ. **2004**. Selective grazing by adults and larvae of the zebra mussel (*Dreissena polymorpha*): application of flow cytometry to natural seston. *Freshwater Biology* **49**: 116–126.

Doležel J, Bartoš J. **2005**. Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany* **95**: 99–110.

Doležel J, Bartoš J, Voglmayr H, Greilhuber J. **2003**. Nuclear DNA content and genome size of trout and human. *Cytometry Part A : the journal of the International Society for Analytical Cytology* **51A**: 127–128.

Doležel J, Binarová P, Lucretti S. **1989**. Analysis of Nuclear DNA content in plant cells by Flow cytometry. *Biologia Plantarum* **31**: 113–120.

Doležel J, Greilhuber J, Suda J. **2007**. *Flow Cytometry with Plant Cells*. Weinheim: Wiley-VCH

Verlag GmbH.

Doležel J, Sgorbati S, Lucretti S. **1992**. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* **85**: 625–631.

Figueroa RI, Garcés E, Bravo I. **2010**. The use of flow cytometry for species identification and life-cycle studies in dinoflagellates. *Deep-Sea Research Part II: Topical Studies in Oceanography* **57**: 301–307.

Galbraith DW. **2012**. Flow cytometry and fluorescence-activated cell sorting in plants: the past, present, and future. *Biomédica* **30**: 65.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. **1983**. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science (New York, N.Y.)* **220**: 1049–51.

Griffin DW, Kellogg CA, Peak KK, Shinn EA. **2002**. A rapid and efficient assay for extracting DNA from fungi. *Letters in Applied Microbiology* **34**: 210–214.

Gryp T, Glorieux G, Joossens M, Vaneechoutte M. **2020**. Comparison of five assays for DNA extraction from bacterial cells in human faecal samples. *Journal of Applied Microbiology* **129**: 378–388.

Guillard RRL, Lorenzen CJ. **1972**. Yellow-green algae with chlorophyllide C. *Journal of Phycology* **8**: 10–14.

Harmon AF, Zarlenga DS, Hildreth MB. **2006**. Improved methods for isolating DNA from *Ostertagia ostertagi* eggs in cattle feces. *Veterinary Parasitology* **135**: 297–302.

Holzinger A, Albert A, Aigner S, *et al.* 2018. Arctic, Antarctic, and temperate green algae *Zygnema* spp. under UV-B stress: vegetative cells perform better than pre-akinetes. *Protoplasma* **255**: 1239–1252.

Hyka P, Lickova S, Přibyl P, Melzoch K, Kovar K. **2013**. Flow cytometry for the development of biotechnological processes with microalgae. *Biotechnology Advances* **31**: 2–16.

Jazwinski MS. **1990**. Preparation of extracts from yeast. *Methods in Enzymology* **182**: 154–174.

Kapraun DF. **2005**. Nuclear DNA Content Estimates in Multicellular Green, Red and Brown Algae: Phylogenetic Considerations. *Annals of Botany* **95**: 7–44.

Kapraun DF. **2007**. Nuclear DNA Content Estimates in Green Algal Lineages: Chlorophyta and Streptophyta. *Annals of Botany* **99**: 677–701.

Kates JR, Chiang KS, Jones RF. **1968**. Studies on DNA replication during synchronized vegetative growth and gametic differentiation in *Chlamydomonas reinhardtii*. *Experimental Cell Research* **49**: 121–135.

Khan MI, Shin JH, Kim JD. **2018**. The promising future of microalgae: current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. *Microbial Cell Factories* **17**.

Kolář F, Lučanová M, Tešitel J, Loureiro J, Suda J. **2012**. Glycerol-treated nuclear suspensions\-an efficient preservation method for flow cytometric analysis of plant samples. *Chromosome Research* **20**: 303–315.

Kron P, Suda J, Husband BC. **2007**. Applications of Flow Cytometry to Evolutionary and Population Biology. *Annual Review of Ecology, Evolution, and Systematics* **38**: 847–876.

Lemaire S, Hours M, Gerard-Hirne C, Trouabal A, Roche O, Jacquot J-P. **1999**. Analysis of light/dark synchronization of cell-wall-less *Chlamydomonas reinhardtii* (Chlorophyta) cells by flow cytometry. *European Journal of Phycology* **34**: 279–286.

Lin S. **2006**. The smallest dinoflagellate genome is yet to be found: A comment on LaJeunesse et al. "*Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates." *Journal of Phycology* **42**: 746–748.

López-Bautista JM, Kapraun DF, Chapman RL. **2006**. Nuclear DNA content estimates in the Trentepohliales (Chlorophyta): Phylogenetic considerations. *Algological Studies/Archiv für*

Hydrobiologie, Supplement Volumes **120**: 41–50.

Loureiro J, Rodriguez E, Doležel J, Santos C. 2006a. Comparison of four nuclear isolation buffers for plant DNA flow cytometry. *Annals of Botany* **98**: 679–689.

Loureiro J, Rodriguez E, Doležel J, Santos C. 2006b. Flow Cytometric and Microscopic Analysis of the Effect of Tannic Acid on Plant Nuclei and Estimation of DNA Content. *Annals of Botany* **98**: 515–527.

Mason AS. 2016. *Polyploidy and Hybridization for Crop Improvement* (AS Mason, Ed.). Boca Raton: Science Publisher, USA.

Mazalová P, Šarhanová P, Ondřej V, Poulíčková A. 2011. Quantification of DNA content in freshwater microalgae using flow cytometry: A modified protocol for selected green microalgae. *Fottea* **11**: 317–328.

Merchant SS, Prochnik SE, Vallon O, *et al.* 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science (New York, N.Y.)* **318**: 245–50.

Milano J, Ong HC, Masjuki HH, *et al.* 2016. Microalgae biofuels as an alternative to fossil fuel for power generation. *Renewable and Sustainable Energy Reviews* **58**: 180–197.

Mridha A, Nandi C, Pal R, Paul S. 2017. Studies on few fresh water green algal species reveals *Spirogyra triplicata* as the repository of high phenolic and flavonoid content exhibiting enhanced anti-oxidant property. *Journal of Pharmacognosy and Phytochemistry* **6**: 1291–1297.

Nelson DR, Chaiboonchoe A, Fu W, *et al.* 2019. Potential for Heightened Sulfur-Metabolic Capacity in Coastal Subtropical Microalgae. *iScience* **11**: 450–465.

Olefeld JL, Majda S, Albach DC, Marks S, Boenigk J. 2018. Genome size of chrysophytes varies with cell size and nutritional mode. *Organisms Diversity & Evolution* **18**: 163–173.

Otto F. 1990. DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. *Methods in Cell Biology* **33**: 105–110.

Palacio-López K, Tinaz B, Holzinger A, Domozych DS. 2019. Arabinogalactan proteins and the extracellular matrix of charophytes: A sticky business. *Frontiers in Plant Science* **10**: 447.

Park CH, Baskar TB, Park SY, *et al.* 2016. Metabolic profiling and antioxidant assay of metabolites from three radish cultivars (*Raphanus sativus*). *Molecules* **21**.

Peters AF, Marie D, Scornet D, Kloareg B, Mark Cock J. 2004. Proposal of *Ectocarpus siliculosus* (Ectocarpales, Phaeophyceae) as a model organism for brown algal genetics and genomics. *Journal of Phycology* **40**: 1079–1088.

Pichrtová M, Holzinger A, Kulichová J, *et al.* 2018. Molecular and morphological diversity of *Zygnema* and *Zygnemopsis* (Zygnematophyceae, Streptophyta) from Svalbard (High Arctic). *European Journal of Phycology* **53**: 492–508.

Pichrtová M, Remias D, Lewis LA, Holzinger A. 2013. Changes in Phenolic Compounds and Cellular Ultrastructure of Arctic and Antarctic Strains of *Zygnema* (Zygnematophyceae, Streptophyta) after Exposure to Experimentally Enhanced UV to PAR Ratio. *Microbial Ecology* **65**: 68–83.

Potter EE, Thornber CS, Swanson JD, McFarland M. 2016. Ploidy distribution of the harmful bloom forming macroalgae *Ulva* spp. in Narragansett Bay, Rhode Island, USA, using flow cytometry methods. *PLoS ONE* **11**: 1–15.

Poulíčková A, Mazalová P, Vašut RJ, Šarhanová P, Neustupa J, Škaloud P. 2014. DNA content variation and its significance in the evolution of the genus *Micrasterias* (desmidiales, streptophyta). *PLoS One* **9**: e86247.

Praça-Fontes MM, Carvalho CR, Clarindo WR, Cruz CD. 2011. Revisiting the DNA C-values of the genome size-standards used in plant flow cytometry to choose the "best primary standards." *Plant Cell Reports* **30**: 1183–1191.

Priyadarshan PM. 2019. Induced Mutations and Polyploidy Breeding In: *Plant breeding: Classical to Modern*. Singapore: Springer Singapore, 329–370.

Qin Y, Zhang Yuehuan, Mo R, *et al.* 2019. Influence of ploidy and environment on grow-out traits of diploid and triploid Hong Kong oysters *Crassostrea hongkongensis* in southern China. *Aquaculture* **507**: 108–118.

Reinecke DL, Castillo-Flores A, Boussiba S, Zarka A. **2018**. Polyploid polynuclear consecutive cell-cycle enables large genome-size in *Haematococcus pluvialis*. *Algal Research* **33**: 456–461.

Roberts A V. 2007. The use of bead beating to prepare suspensions of nuclei for flow cytometry from fresh leaves, herbarium leaves, petals and pollen. *Cytometry Part A* **71A**: 1039–1044.

Sadílek D, Urfus T, Vilímová J. **2019**. Genome Size and Sex Chromosome Variability of Bed Bugs Feeding on Animal Hosts Compared to *Cimex lectularius* Parasitizing Human (Heteroptera: Cimicidae). *Cytometry Part A* **95**: 1158–1166.

Schönswetter P, Suda J, Popp M, Weiss-Schneeweiss H, Brochmann C. **2007**. Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. *Molecular Phylogenetics and Evolution* **42**: 92–103.

Simić S, Kosanić M, Ranković B. **2012**. Evaluation of in vitro antioxidant and antimicrobial activities of green microalgae *Trentepohlia umbrina*. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* **40**: 86–91.

Simon N, Barlow RG, Marie D, Partensky F, Vaulot D. **1994**. Characterization of oceanic photosynthetic picoeukaryotes by flow cytometry. *Journal of Phycology* **30**: 922–935.

Šmarda P, Knápek O, Březinová A, *et al.* 2019. Genome sizes and genomic guanine+cytosine (GC) contents of the Czech vascular flora with new estimates for 1700 species. *Preslia* **91**: 117–142.

Spring H, Grierson D, Hemleben V, *et al.* 1978. DNA contents and numbers of nucleoli and pre-rRNA-genes in nuclei of gametes and vegetative cells of *Acetabularia mediterranea*. *Experimental Cell Research* **114**: 203–215.

Temsch EM, Greilhuber J, Krisai R. **2010**. Genome size in liverworts. *PRESLIA* **82**: 63–80.

Trumhová K, Holzinger A, Obwegeser S, Neuner G, Pichrtová M. **2019**. The conjugating green alga *Zygnema* sp. (Zygnematophyceae) from the Arctic shows high frost tolerance in mature cells (pre-akinetes). *Protoplasma* **256**: 1681–1694.

Veldhuis MJW, Cucci TL, Sieracki ME. **1997**. Cellular Dna Content of Marine Phytoplankton Using Two New Fluorochromes: Taxonomic and Ecological Implications. *Journal of Phycology* **33**: 527–541.

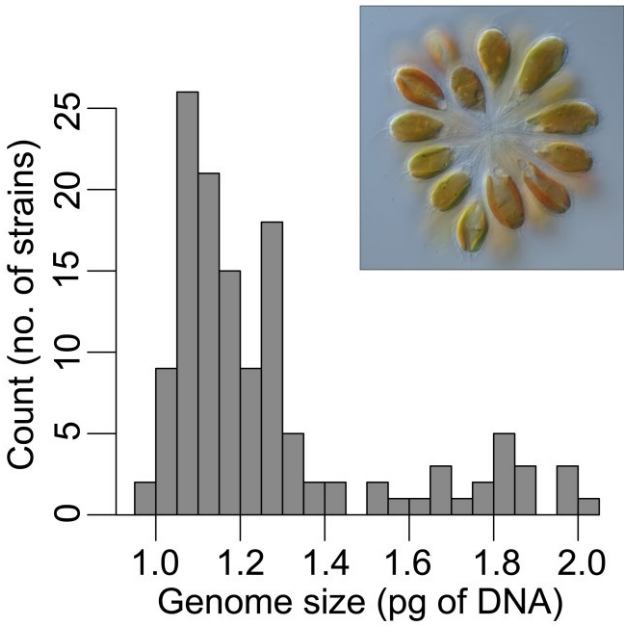
Veselý P, Bureš P, Šmarda P, Pavlíček T. **2012**. Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.

Waaland JR, Stiller JW, Cheney DP. **2004**. Macroalgal candidates for genomics. *Journal of Phycology* **40**: 26–33.

Weiss TL, Johnston JS, Fujisawa K, Okada S, Devarenne TP. **2011**. Genome size and phylogenetic analysis of the A and L races of *Botryococcus braunii*. *Journal of Applied Phycology* **23**: 833–839.

PAPER IV

Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences.



Intraspecific genome size variation among strains of *Synura petersenii*.

Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences

Dora Čertnerová*, Pavel Škaloud

Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, CZ-128 01 Prague, Czech Republic

*Author for correspondence (e-mail: dora.certnerova@gmail.com)

ABSTRACT

Background and Aims: While the nuclear DNA content variation and its phenotypic consequences have been well described in plants and animals, much less of this topic is known from unicellular algae and protists in general. The dearth of data is especially pronounced when it comes to intraspecific genome size variation. This study attempts to investigate the extent of intraspecific variability in genome size and its adaptive consequences in a microalgae species.

Methods: Propidium iodide flow cytometry was used to estimate the absolute genome size of 131 strains (isolates) of the golden-brown algae *Synura petersenii* (Chrysophyceae, Stramenopiles), identified by identical ITS rDNA barcodes. Cell size, growth rate and genomic GC content were further assessed on a subset of strains. Geographic location of 67 sampling sites across the Northern hemisphere was used to extract climatic database data and to evaluate ecogeographical distribution of genome size diversity.

Key Results: Genome size ranged continuously from 0.97 pg to 2.02 pg of DNA across the investigated strains. The genome size was positively associated with cell size and negatively associated with growth rate. Climatic conditions did not have a significant effect on genome size variation. No clear trends in geographic distribution of strains of particular genome size were detected and strains of different genome size occasionally coexisted at the same locality. Genomic GC content was significantly associated only with genome size via a quadratic relationship.

Conclusions: This study presents the most comprehensive intraspecific genome size screening conducted on a protist to date. Genome size variability in *Synura petersenii* was likely triggered by evolutionary mechanism operating via gradual changes in genome size accompanied by changes in genomic GC content, such as e.g. proliferation of transposable elements. The variation was reflected in cell size and relative growth rate, possibly with adaptive consequences.

Key words: intraspecific DNA content variation, genome size, flow cytometry, golden-brown algae, *Synura petersenii*, GC content, biovolume, growth rate, environmental conditions, ITS

INTRODUCTION

The nuclear genome constitutes an essential cell component. While the quality of nuclear DNA (expressed by nucleotide sequences or presence of specific alleles) has been in the focus of biologists for more than a half of century, its quantity per cell has gathered far less attention. Nonetheless, genomes contain orders of magnitude more DNA than required to sustain the basic cell functioning (a phenomenon called "C-value enigma"; Gregory 2001a) and this may suggest that even the overall quantity of nuclear DNA has an adaptive potential (Mirsky and Ris, 1951; Thomas, 1971; Cavalier-Smith, 2005). Our knowledge of the genome size variation and its evolutionary consequences mainly comes from plant and animal studies (e.g. Leitch *et al.* 1998; Gregory 2005; Beaulieu *et al.* 2008; Liedtke *et al.* 2018; Trávníček *et al.* 2019). Much less of this topic is known from single-celled eukaryotes, despite the fact that protist genome size ranges more than 28,600-fold compared to 6,600-fold variation among plants and animals (Veldhuis *et al.* 1997; Gregory 2005; Keeling and Slamovits 2005; Pellicer and Leitch 2019).

There are several evolutionary mechanisms responsible for the nuclear DNA amount variation. The genome size can either increase or decrease by chromosomal aberrations (aneuploidy), non-homologous recombination, and changes in the relative genome-wide frequency of insertions to deletions (Devos *et al.* 2002; Roux *et al.* 2003; Lynch and Conery 2003; Wu *et al.* 2018). Conversely, only increase in genome size is possible via higher activity of transposable elements and, more abruptly, by whole-genome doubling (polyploidization; Soltis and Soltis 1999; Kidwell 2002; Cavalier-Smith 2005; Sun *et al.* 2012). Recent or past hybridization events between closely related but separate species can also contribute to nuclear DNA amount differentiation (Baack *et al.* 2005).

Variation in nuclear DNA content is usually accompanied with phenotypic consequences. The genome size directly affects nucleus size (Sparrow and Evans 1961; Baetcke *et al.* 1967; Bennett 1972; Gregory 2001b; Zubáčová *et al.* 2008) and through it also fundamentally relates to the cell size (Cavalier-Smith and Beaton 1999; Gregory 2001a; Gregory 2001b; Cavalier-Smith 2005). The genome size – cell size correlation, also known as karyoplasmic ratio, has been observed across the eukaryotic tree of life, including many protist lineages (Wilson 1925; Bennett 1972; Suzuki *et al.* 1982; Shuter *et al.* 1983; Veldhuis *et al.* 1997; LaJeunesse *et al.* 2005; Connolly *et al.* 2008; von Dassow *et al.* 2008). Although the exact cause of this relationship is still debated, evolutionary changes in cell size may either cause or be caused by changes in genome size (Gregory 2001a). The cell size is a particularly important trait in single-celled organisms as it inversely correlates with metabolic rate, growth rate and directly correlates with generation time (Van't Hof and Sparrow 1963; Bennett 1972; Shuter *et al.* 1983; Gregory 2001a; Cavalier-Smith 2005). Changes in cell size may also be reflected in protist ecology, for example by altering the grazing pressure, efficiency of nutrient acquisition and/or light harvesting

(Garcia-Pichel 1994; Finkel *et al.* 2001; Smetacek *et al.* 2004; Irwin *et al.* 2006). Therefore, genome size variation could be (via cell size) subject to natural selection in populations of protists (Cavalier-Smith 2005).

Another understudied genomic parameter with possibly adaptive nature is the relative genome-wide frequency of AT to GC base pairs, often expressed as a %GC content (Bennett and Leitch 2011; Šmarda and Bureš 2012). The higher genomic GC content is sometimes associated with extreme climatic conditions such as pronounced cold, drought or temperature fluctuations, probably due to the increased thermal stability of the DNA double helix (Šmarda and Bureš 2012; Šmarda *et al.* 2014; Trávníček *et al.* 2019). Additionally, the GC content variation may be linked to changes in genome size, since the genomic nucleobase composition might be altered by high activity of transposable elements (TEs) or chromosomal aberrations (Wichman *et al.*, 1993; Armbrust, 2004; Derelle *et al.*, 2006). However, there is only little known about the evolutionary impact of GC content variation, and what is known, comes almost exclusively from studies on prokaryotes, plants and animals (Goodsell and Dickerson 1994; Hildebrand *et al.* 2010; Šmarda *et al.* 2014; Mugal *et al.* 2015; Veleba *et al.* 2017; Trávníček *et al.* 2019).

Both genome size and GC content may be estimated using flow cytometry (FCM). This technique is based on measuring the properties of fluorescent-stained particles (e.g. cells, isolated nuclei) in a stream of fluid and allows rapid and precise nuclear DNA analysis (Doležel *et al.* 2007). While the FCM has found a wide spectrum of applications in genomic surveys on plants and animals (Dionisio Pires *et al.* 2004; Kron *et al.* 2007; Galbraith 2012; Pellicer and Leitch 2014), its potential was not yet much explored in protist studies (but see Figueroa *et al.* 2010) and robust methodological protocols allowing work with diverse protist material are missing.

The dearth of data is especially pronounced when it comes to intraspecific genome size variation in unicellular eukaryots. This can be attributed to analysing only single strain per species in most studies (Veldhuis *et al.* 1997; Mazalová *et al.* 2011) and, possibly, to the fact that contrary to e.g. plant studies, best-practice protocols preventing false reports based on methodological and instrumental errors (Greilhuber 2005) are not routinely applied.

To study the evolution of genome size and its phenotypic and physiological consequences, we chose golden-brown algae *Synura petersenii* (Chrysophyceae, Stramenopiles) as our model system. In general, protist species have short generation time and huge population size, which allows them to quickly respond to environmental change (Lynch and Conery 2003; Foissner 2007; Ribeiro *et al.* 2013). *Synura petersenii* is an autotrophic flagellate with assumed worldwide distribution. It creates colonies of cells covered by siliceous scales with species-specific ornamentation and a characteristic pronounced central keel (Kristiansen and Preisig 2007). The species has recently undergone thorough taxonomic revision supported by molecular markers and morphometric analysis of its siliceous scales, which

revealed 15 separate species in the formerly recognized *S. petersenii* species complex (Wee *et al.* 2001; Boo *et al.* 2010; Kynčlova *et al.* 2010; Škaloud *et al.* 2012, 2014; Jo *et al.* 2016). One of these taxonomically revised species, *S. petersenii* s. str. is used as a model species in this study. During our pilot FCM analysis, we detected intraspecific genome size variation among the strains of *S. petersenii*. The general aims of the study are to prove existence of intraspecific genome size variation in *S. petersenii* and investigate in detail, for the first time, the extent of intraspecific variability in genome size in a microalgae species. Additionally, we ask the following questions: i) Is the variability in DNA content linked to GC content variation? ii) Are there any phenotypic and physiological consequences of varying genome size? iii) Is genome size variation among strains reflected in their ecogeographical distribution?

MATERIALS AND METHODS

Origin, cultivation and identification of the investigated strains

Altogether, 131 isolates of the species *Synura petersenii* were obtained from various fresh-water localities across the Northern hemisphere. Sampling details are listed in Supplementary Data Table S1. To establish new cultures, water samples were taken using a 25 μm mesh plankton net and single *Synura* colonies were captured by micro-pipetting and transferred into separate culture wells filled with WC medium (Guillard and Lorenzen 1972). All cultures were maintained at 17 °C (cooling box Pol-Eko Aparatura Sp.J., model ST 1, Wodzisław Śląski, Poland) with 24-h light mode under the illumination of 30 μmol $m^{-2}s^{-1}$ (TLD 18W/33 fluorescent lamps, Philips, Amsterdam, Netherlands).

All strains were identified based on their Internal Transcribed Spacer sequence of nuclear ribosomal DNA (ITS1, 5.8S and ITS2 rDNA) since this is the most variable of commonly used molecular markers in this group (Jo *et al.* 2016). For this purpose, genomic DNA was extracted from a centrifuged pellet of cells by InstaGene Matrix (Bio-Rad, USA) and the resulting supernatant was directly used as a PCR template. The amplifications were performed using a universal primer ITS4 (White 1990) and a lineage-specific primer Kn1.1 (Wee *et al.* 2001). The PCR reactions were carried out in a total volume of 20 μl with a PCR mix containing 0.2 μL of MyTaqHS DNA polymerase (Bioline), 4 μL of MyTaqHS Buffer (Bioline), 0.4 μL of each primer, 14 μL of double-distilled water and 1 μL of template DNA (not quantified). The amplifications were performed in Eppendorf Mastercycler ep Gradient 5341 (Eppendorf GmbH, Hamburg, Germany) using the following program: 1 min of denaturation at 95 °C; followed by 35 cycles of denaturation at 95 °C (15 s), annealing at 52 °C (30 s) and elongation at 72 °C (40 s), concluded with the final extension at 72 °C (7 min) and held at 10 °C. The PCR products were sized on a 1% agarose gel and then purified using AMPure XP magnetic beads (Agencourt). The

purified DNA templates were sequenced by Sanger Sequencing method at Macrogen, Inc. (Seoul, Korea, http://dna. macrogen.com). Finally, the obtained sequences were identified using the BLAST in National Center for Biotechnology Information Search database (NCBI) and personal ITS database created during previous studies (Škaloud *et al.* 2012, 2014). The strains with identical ITS rDNA sequence to *Synura petersenii* were transferred into Erlenmeyer flasks with 30 ml of WC medium and kept in longer cultivation. The collection was further supplemented with five strains from previous studies (Kynčlová *et al.* 2010; Kim *et al.* 2019).

DNA content estimation

To estimate nuclear genome size of our strains, we employed propidium iodide flow cytometry (FCM). Approximately two weeks before the planned FCM analyses, cultures were inoculated into fresh medium. For sample preparation, 1 ml of well-grown culture was centrifuged (5500 rpm for 5 min) and the superfluous medium was removed by pipetting. Consequently, 350 µl of ice-cold nuclei isolation buffer Otto I (0.1 M citric acid, 0.5% Tween 20; Otto 1990) was added to algal pellet, causing the release of sample nuclei. The resulting suspension was thoroughly shaken and kept on ice. Plant *Solanum pseudocapsicum* (2C = 2.59 pg; Temsch *et al.* 2010) was used as an internal standard. To release nuclei of the standard, ca. 20-mg piece of fresh leaf tissue was chopped with a razor blade in a plastic Petri dish with 250 µl of ice-cold Otto I buffer. Both suspensions (with algal and standard nuclei) were thoroughly mixed and filtered through 42-µm nylon mesh into a special 3.5-ml cuvette for direct use with the flow cytometer. Following a 20-min. incubation at room temperature, the sample was mixed with 1 ml of staining solution consisting of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12H_2O$; Otto 1990), 50 µg ml^{-1} of propidium iodide, 50 µg ml^{-1} of RNase IIA and 2 µl ml^{-1} β-mercaptoethanol. The stained sample was immediately analysed using a Partec CyFlow SL cytometer (Partec GmbH, Münster, Germany) equipped with green solid-state laser (Cobolt Samba, 532 nm, 100 mW). Measurements on each sample were done up to 5000 particles and the resulting FCM histograms analysed using FloMax ver. 2.4d (Partec, Münster, Germany). Since there is no knowledge of ploidy level in the genus *Synura* (Olefeld *et al.*, 2018), we identified the first sample peak on the FCM histogram as G_1 (vegetative cells) and the second peak as G_2 (dividing cells). The absolute nuclear DNA amount (C-value) was calculated as sample G_1 peak mean fluorescence / standard G_1 peak mean fluorescence × standard 2C DNA content (according to Doležel 2005).

In case of low quality measurement, i.e. G_1 sample coefficient of variation (CV) above 5%, the sample preparation and analysis was repeated. To minimize the effect of random instrumental shift, each *S. petersenii* strain was analysed at least three times on separate days. Whenever the three independent genome size estimates differed by more than 3%, the most outlying measurement was discarded and a new

measurement conducted, until this condition was fulfilled. In order to corroborate genome size differences among strains, simultaneous analysis of multiple selected strains (i.e. A64, D55, G61) was performed. We also tested for strain genome size stability during its cultivation. Following inoculation into a fresh medium, three strains exhibiting highest variation among repeated measurements (i.e. F19, G16, H11) were analysed once a week for the period of six to seven weeks. Two other strains (961, S63.B3) were then re-analysed after two years following the first measurements.

GC content estimation

To assess variation in genomic GC content, we analysed 38 strains of *S. petersenii* using FCM with AT-selective dye DAPI (4',6-diamidino-2-phenylindole) and compared the results with propidium iodide FCM. The strains for GC content estimation, cell size measurements and growth rate analysis (see below) were selected representatively across the whole range of genome size diversity, however, a different set of strains was used for each assessment (Supplementary Data Table S2). This was due to unavailability of some strains at the time of particular assessments (e.g. the cultured strains did not survive, provided limited biomass or a contamination occurred), replacement strains with similar genome size were then randomly selected. We employed the same sample preparation as for propidium iodide FCM, except the staining solution consisted of 1 ml of Otto II buffer, 4 μg ml^{-1} of DAPI and 2 μl ml^{-1} of β-mercaptoethanol. Samples were immediately analysed using Partec PA II flow cytometer (Partec GmbH, Münster, Germany) equipped with a 488 nm UV LED as a source of excitation light. Analyses were run up to 5000 particles and the resulting FCM histograms were analysed using FloMax ver. 2.4d (Partec, Münster, Germany). Computation of the base content was done according to Šmarda *et al.* (2008) via publicly available excel spreadsheet (http://sci.muni.cz/botany/systemgr/download/Festuca/ATGCFlow.xls).

Cell size measurements

After two weeks of cultivation in a fresh medium, the cell size of 39 selected *S. petersenii* strains was analysed by imaging flow cytometry using Benchtop B3 Series FlowCAM (Fluid Imaging Technologies, Yarmouth, Maine, USA). The FlowCAM settings were AutoImage mode, 50 μm flow cell, 20x objective and flow Rate 0.020 ml/min. The mean biovolume of 100 cells per strain was calculated on manually selected images using VisualSpreadsheet® Particle Analysis Software ver. 4.11.12 as the volume of a sphere of the area-based radius measured automatically by the FlowCAM for each cell.

Growth rate

Test of growth rate was performed on eight replicate cultures of each of 31 selected *S. petersenii* strains. The chlorophyll fluorescence, i.e. the effective quantum yield of photochemical energy conversion in PSII (Φ_{PSII}) of cultures starting with the same initial concentration was measured daily at the same hour over 15 days using PAM 2500 fluorometer (Heinz Walz GmbH, Effeltrich, Germany). The variable Φ_{PSII} is a relative parameter calculated as $(F_M' - F)/F_M'$, where F is steady-state fluorescence in the light-adapted state and F_M' is the maximum fluorescence in the light-adapted state measured after the application of a saturation pulse (Roháček and Barták 1999). The growth rate was subsequently derived as inverse value of median time at which the population density reaches ½ the carrying capacity, i.e. the inflection point (t-mid value^{-1}) in R software ver. 3.4.3 (R Development Core Team, 2017) using the package growthcurver ver. 0.3.0 (Sprouffske and Wagner 2016).

Ecogeographical patterns

Geographical distribution of the genome size diversity in *Synura* was visualised on a map using ArcGIS 10.0 (ESRI, Redlands, California, USA). In order to assess putative ecogeographical trends in the distribution of genome size diversity well beyond any obvious spatial patterns, we also tested for associations between genome size of *Synura* and database-derived climatic variables. We used ArcGIS to extract climate data from 19 Bioclim variables of the WorldClim database ver. 2 (http://worldclim.org/version2; Fick and Hijmans 2017) downloaded in the highest available resolution (30 arc seconds). Climatic conditions may affect aquatic microalgae e.g. via temperature regulated onset and duration of their seasonal blooms or precipitation frequently being associated with input of nutrients into aquatic ecosystems (Baek *et al.* 2009). Each of 67 sampling sites were assigned values of the climatic variables and geographic latitude, included as an additional ecogeographically relevant parameter. When multiple strains were collected at a site, we only retained those with genome size estimates differing by at least 5% (i.e. arbitrarily selected threshold to prevent pseudoreplication of data) and such strains were then treated as independent observations. If two or more strains with similar genome size (<5% difference) originated from the same site, all but one randomly selected strain were excluded from the dataset. The resulting dataset consisting of 82 strains was analysed using a redundancy analysis (RDA) implemented in Canoco 5 (Lepš and Šmilauer 2014), genome size of *Synura* was used as an explanatory variable. All response variables were standardized prior the RDA and statistical significance was tested using a Monte Carlo test with 999 permutations. The RDA was then also used to test the effect of GC content.

Statistical analysis

Unless stated otherwise, statistical data analysis was conducted in R. Separate linear regression models were applied to test whether the variation in cell size and growth rate can be explained by genome size (log transformed). These analyses were conducted on a subset of 39 and 31 strains, respectively, for which data were available. Both models were also re-run using the GC content as an explanatory variable.

We then employed a regression model to assess the relation between GC content (response variable) and genome size (explanatory variable, log transformed). Due to previous reports of a quadratic relationship between the two genomic parameters in plants (e.g. Šmarda *et al.* 2014), we used manual AIC-based forward selection with the function *addterm* from the R package MASS ver. 7.3-50 (Venables and Ripley 2002) to test whether incorporating the logarithm of genome size either in linear (~ log.GS) or quadratic form (~ log.GS + I[log.GS^2]) will significantly improve the model performance.

RESULTS

Intraspecific variability in genome size

We successfully determined absolute nuclear DNA amount in 131 strains of *Synura petersenii* with identical ITS rDNA sequence from 67 localities across the Northern hemisphere (Fig. 1). The sampled *S. petersenii* strains exhibited a 2.1-fold variation in their genome size, ranging from 0.971 to 2.022 pg of DNA (Supplementary Data Table S2). Frequency distribution of the genome size values was conspicuously positively skewed (median = 1.170 pg, mean = 1.296 pg; Fig. 2). Sufficient precision of flow cytometric measurements was ensured by relatively low coefficients of variation (CVs) for both sample and standard G_1 nuclei peaks (mean CV = 3.29% and 2.25%, respectively), see Fig. 3A for a representative analysis. The intraspecific variability in genome size was also verified in a simultaneous analysis of three *S. petersenii* strains with contrasting nuclear DNA amounts that resulted in three clearly differentiated peaks in the flow cytometric histogram (Fig. 3B).

To corroborate the longer term stability of genome size differences among strains in cultivation, two selected strains were re-analysed two years following the first measurements: strain 961 (1.071 and 1.077 pg of DNA, respectively) and strain S63.B3 (1.516 and 1.492 pg of DNA, respectively). In both cases the deviation between repeated estimates fell within the limits of instrumental precision (i.e. mean CV). Similarly, three other strains investigated for genome size stability via regular weekly measurements (F19, G16, H11) did not show any substantial deviations from the original estimates (Fig. 4).

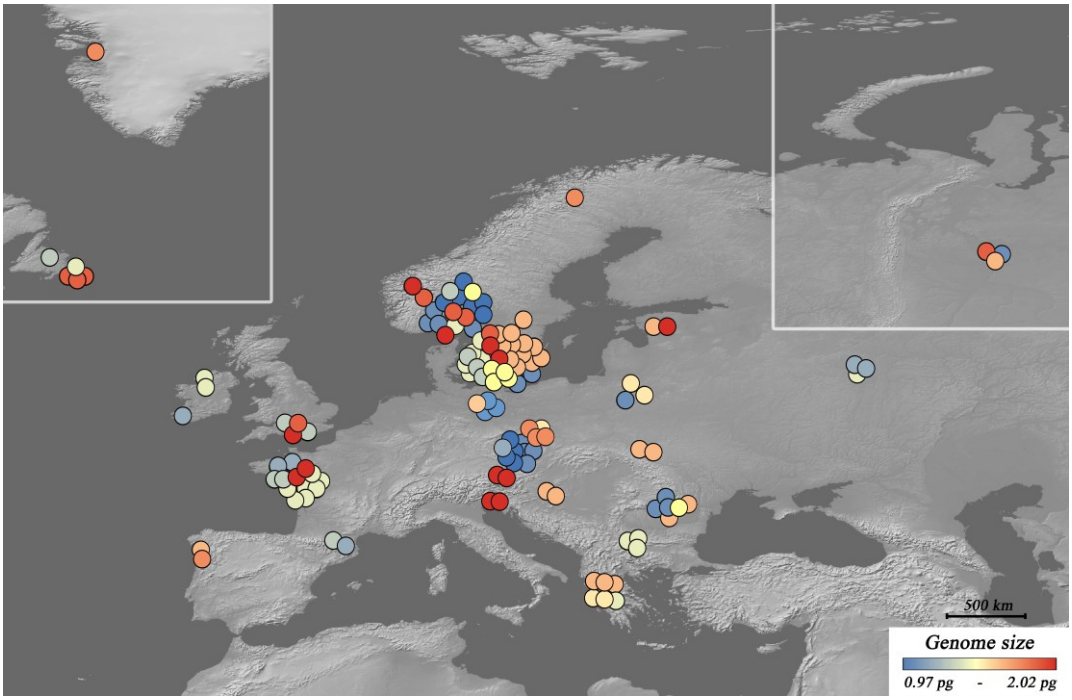


Figure 1. Distribution of 131 strains of *Synura petersenii* under study and their estimated genome size. Symbol colour refers to different genome size categories.

Phenotypic consequences and ecogeographical distribution of genome size diversity

Genome size was significantly associated with both cell size ($F_{1,37} = 6.25$, $P = 0.017$, $R^2 = 0.145$; Fig. 5A) and growth rate ($F_{1,29} = 5.03$, $P = 0.033$, $R^2 = 0.148$; Fig. 5B), increase in genome size led to increase in cell size and decrease in growth rate. However, the two observed relationships were not affected by a putative strong correlation between the cell size and growth rate ($t_{23} = -1.32$, $P = 0.201$, $r = -0.265$).

Synura strains with smaller or larger genome size did not display any apparent spatial trends in their geographic distribution (Fig. 1). This was supported by the lack of significant association between spatial distribution of genome size diversity and climatic variables or latitude in the redundancy analysis (RDA; $P = 0.505$, 999 permutations). The first, constrained RDA axis explained only 1.0% of overall variation, whereas the second, unconstrained axis explained 41.6% of variation (Supplementary Data Figure S1).

Diversity in genomic GC content

The other genomic parameter, GC content, varied from 37.1% to 41.2%, with mean value of 39.5% (Supplementary Data Table S2). The GC content had no significant effect on either cell size ($F_{1,15} = 0.29$, $P = 0.598$, $R^2 = 0.019$) or growth rate

($F_{1,11}$ = 1.96, P = 0.189, R^2 = 0.151). No significant association between GC content diversity and climatic variables or latitude was detected in RDA (P = 0.890, 999 permutations). The first, constrained RDA axis explained only 0.8% of overall variation, whereas the second, unconstrained axis explained 42.2% of variation (Supplementary Data Figure S1). On the other hand, a significant quadratic relationship was detected between GC content and genome size ($F_{2,35}$ = 6.95, P = 0.003, R^2 = 0.284; Fig. 5C). The appropriateness of including the predictor in a quadratic form was corroborated using manual forward selection, as such model significantly outperformed both the linear relationship ($F_{1,36}$ = 0.11, P = 0.743, R^2 = 0.003) and the null model without any predictors.

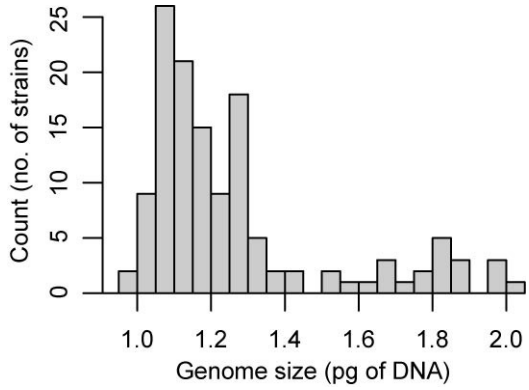


Figure 2. Frequency distribution of genome size categories among investigated *Synura petersenii* strains.

DISCUSSION

Genome size variability a its evolutionary sources

Currently, there is an apparent dearth of genome size estimates from protists, especially regarding the degree of intraspecific genome size variation. This might be due to the fact that the use of flow cytometry (FCM), efficient and widely applied technique of nuclear DNA content estimation (Doležel *et al.* 2007), is often methodologically challenging in protists as a result of difficulties in obtaining sufficient amounts of biomass, protoplast extraction or the presence of wide variety of pigments and secondary metabolites interfering with fluorescent staining (Veldhuis *et al.* 1997; Kapraun 2007; Mazalová *et al.* 2011; Poulíčková *et al.* 2014). Considering the above, we adopted a FCM protocol with several steps improving the robustness of our estimates that included inferring each genome size estimate from mean value of three analyses on different days, simultaneous analysis of strains with different genome size to confirm the existing differences

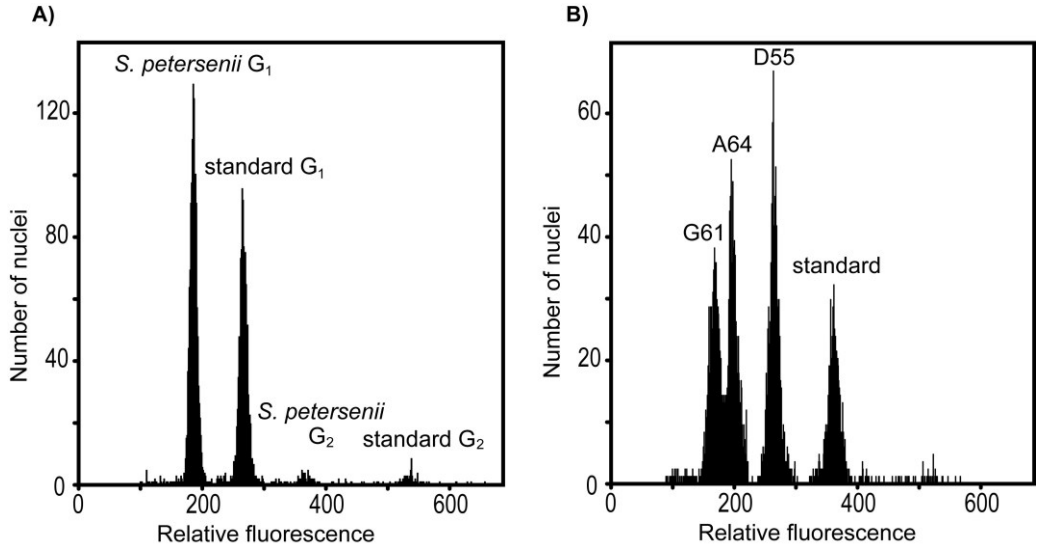


Figure 3. Flow-cytometric histograms showing relative fluorescence of propidium iodide-stained nuclei of *Synura petersenii* samples and Solanum pseudocapsicum (standard). In (A), a representative analysis of single S. petersenii strain with G_1 and G_2 phase nuclei apparent for both the analysed sample and standard. In (B), a simultaneous analysis of three strains with distinct genome size (G61, A64, D55) to confirm the existing differences.

(multiple peaks in a FCM histogram) and re-analysing strains after some period of time to account for genome size stability under cultivation. In our study, we successfully estimated genome size of more than 130 strains belonging to a single microalgae species and, to our knowledge, this is the most comprehensive intraspecific genome size screening conducted on protists so far. We revealed considerable variability in genome size of *Synura petersenii*, ranging two-fold across the analysed strains, from 0.97 to 2.02 pg of DNA. Our estimates (median value = 1.17 pg) are not consistent with an earlier estimate of *S. petersenii* genome revision

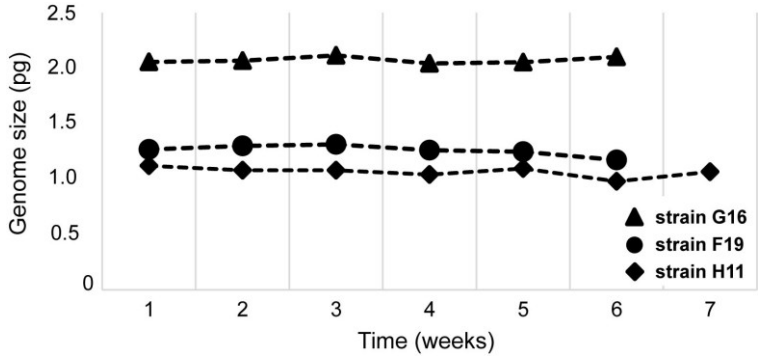


Figure 4. Temporal stability of genome size of three selected Synura petersenii strains.

size of 0.78 pg made by Olefeld *et al.* (2018). However, the published data belonged to the strain WA18K-A (with other designation CCMP 2892) that was in taxonomic

of *S. petersenii* species complex assigned to a different species, *S. heteropora* (Škaloud *et al.* 2014). Therefore, we present the first genome size data for *S. petersenii* sensu stricto.

There are several possible scenarios of what could be the source of genome size diversity observed among *S. petersenii* strains. First, owing to the robust FCM protocol, consistent methodology and generally high precision of our analyses, we are convinced that the error of measurement have not substantially contributed to the genome size variation. Taking into account the two-fold difference between the lowest and highest genome size estimates, alternating life cycle stages or whole genome doubling (polyploidization) events would seem as likely explanations.

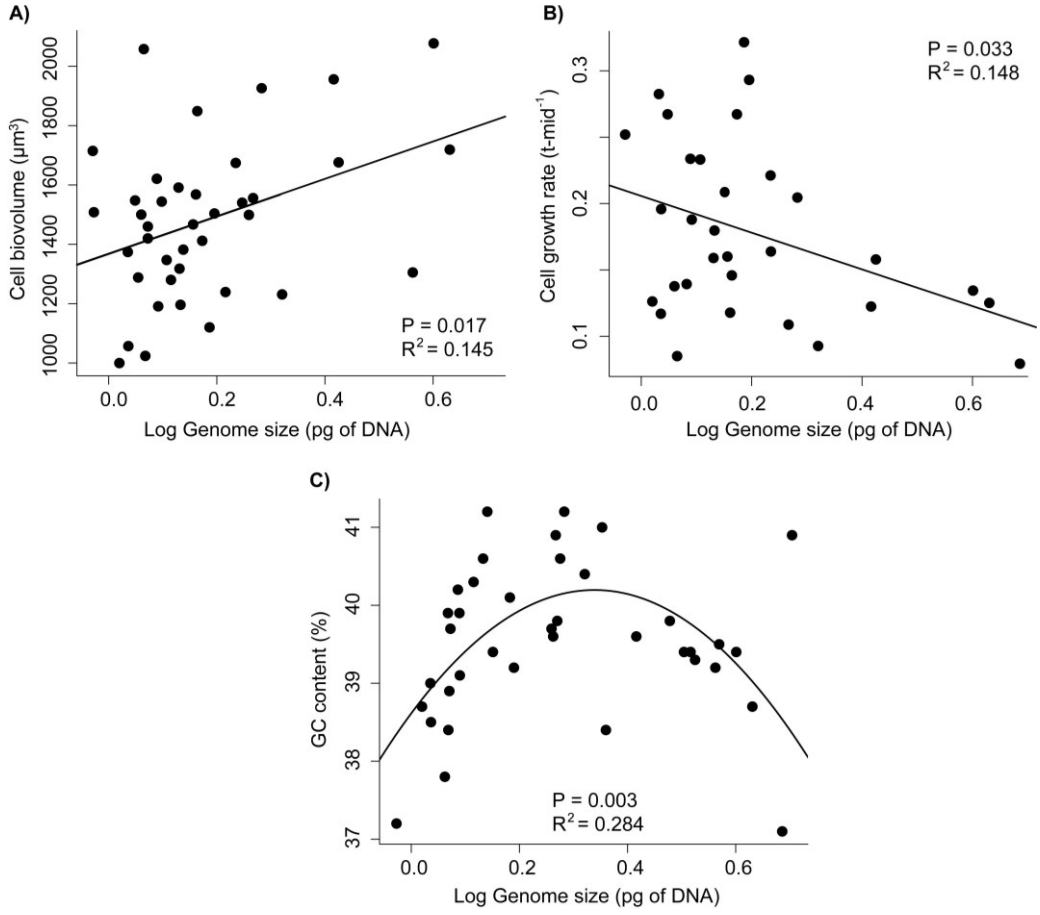


Figure 5. Three genomic and phenotypic traits associated with intraspecific genome size diversity in Synura petersenii: cell size (A), relative growth rate (B) and genomic GC content (C). Model predictions are depicted using lines in (A-B) and the curve of quadratic function in (C).

However, none of these mechanisms can be the sole source of the diversity observed in *Synura* as there were no discrete genome size categories that would

reflect the inherent ploidy shifts (Fig. 2). Another argument against the alternating life cycle stages is that strains re-analysed after weeks (or two years) exhibited more or less stable genome size estimates (Fig. 4). This contrasts with the genome size differentiation that emerged within a long-term cultivated strain of *Thalassiosira weissflogii* belonging to diatoms, a more intensively studied group of Stramenopiles, implying the ability to rapidly change the DNA amount (von Dassow *et al.* 2008), possibly in context of sexual reproduction or (theoretically) local adaptation. On the other hand, genome size appeared to be stable in cultivation of another diatom *Ditylum brightwellii* where intraspecific variation among strains was also previously detected (Koester *et al.* 2010). While we are unable to exclude the possibility that strains at both extreme ends of the observed genome size continuum are in fact different ploidy cytotypes or distinct stages of *S. petersenii* life cycle, other evolutionary mechanisms operating with more gradual increases or decreases in genome size were most probably involved.

Alternative explanations may be provided by proliferation of transposable elements (TEs), unequal frequency of insertions to deletions, and multiplication of larger genomic segments or even whole chromosomes (supernumerary chromosomes or aneuploidy; Jones *et al.* 2008; Šmarda and Bureš 2010; Ruiz-Ruano *et al.* 2011; Stelzer *et al.* 2019). For example, genome size diversity in the diatom *Thalassiosira weissflogii* was attributed to polyploidization, aneuploidization and gene duplications (von Dassow *et al.* 2008). A recent chromosome doubling was also detected in diatom *T. pseudonana* (Armbrust 2004). Unfortunately, despite a considerable effort, we did not succeed with karyotyping of *S. petersenii* strains and neither chromosome counts nor complete genomic sequences are available for any representative of the class Chrysophyceae (incl. the genus *Synura*). It is thus unclear whether prompt karyotype evolution or chromosomal aberrations could be responsible for the observed intraspecific genome size variation. Genome size changes caused by chromosomal aberrations or increased TE activity may be accompanied by significant alterations of genomic GC content (i.e. the relative proportion of GC base pairs; Wichman *et al.* 1993; Armbrust 2004; Derelle *et al.* 2006). Interestingly, we found a significant relationship between the genome size of *Synura* strains and their genomic GC content, which had quadratic nature and predicted highest GC content in medium-sized genomes (Fig. 5C). A similar quadratic relationship between the two variables was previously documented in monocot plants (Veselý *et al.* 2012; Šmarda *et al.* 2014), where it was explained by involvement of GC-rich LTR retrotransposons in genome size expansion in combination with a mechanism responsible for decreasing GC content in large genomes (e.g. lower energetic cost of dATPs and dTTPs synthesis leading to their mis-incorporation into the newly synthesized DNA as a mutational bias toward AT-rich genome; Rocha and Danchin 2002; Grover and Wendel 2010). It is worth emphasizing that in

contrast to a 207-fold genome size variation among monocot plants in the dataset analysed by Šmarda *et al.* (2014), we were able to detect the significant quadratic relationship with GC content on a very fine scale of two-fold genome size variation. This might suggest that the observed GC content variation is a mere by-product of the mechanism governing genome size evolution in *S. petersenii*, which is in line with the absence of any biological or environmental correlates of GC content diversity in our dataset.

Intraspecific variability vs. cryptic diversity

As a general rule in multicellular organisms, individuals belonging to the same species share a constant nuclear DNA content (Swift 1950). Nonetheless, intraspecific genome size variation manifested either via multiple ploidy cytotypes or at a homoploid level is occasionally observed among both plants and animals (Jeffery *et al.* 2016; Kolář *et al.* 2017; Stelzer *et al.* 2019). There is also some evidence of intraspecific genome size variation among microalgae coming from diatoms (von Dassow *et al.* 2008; Koester *et al.* 2010), desmids (Poulíčková *et al.* 2014) and haptophytes (Medlin *et al.* 1996; Veldhuis *et al.* 1997; Read *et al.* 2013). Nonetheless, the frequency of this phenomenon and its prevalence across various groups of protists is still poorly documented and the evolutionary mechanisms involved are only exceptionally addressed. Theoretically, there are two mutually non-exclusive evolutionary scenarios that would result in intraspecific genome size variation. Firstly, the mechanisms of genome size change might act recurrently with high enough frequency to compensate for only transient character of induced changes (e.g. via aneuploidy, presence of supernumerary chromosomes). Secondly, intraspecific genome size variation could be maintained in populations when it is coupled with a reproductive barrier that prevents crosses between conspecific individuals with different genome sizes. The reproductive barrier in the latter scenario could either directly result from the mechanism inducing genome size differences (e.g. polyploidization) or arise independently (e.g. spatial or temporal isolation, specific mate recognition systems).

A unique insight into mechanisms maintaining intraspecific genome size variation was recently documented in one species of rotifer (Stelzer *et al.* 2019). The variation, apparent already at the within-population level, was possible due to independently segregating large genomic elements present in males. Regardless of the genome size difference, individuals were able to interbreed and produce viable offspring, stressing their identity to one species (Stelzer *et al.* 2019). However, under natural conditions the species relies predominantly on asexual reproduction mediated by parthenogenetic females. Similar to rotifers, *Synura petersenii* also reproduces mainly clonally (by a cell division), though sexual reproduction has been documented (Sandgren and Flanagin 1986). Despite our great effort, we were unable to experimentally interbreed *S. petersenii* strains. Neither crosses between

contrasting genome size categories nor those performed between strains with similar-sized genomes were successful, possibly suggesting inadequate conditions for mating. It thus remains unclear whether the different genome size categories of *S. petersenii* strains are coupled with a reproductive barrier or not. Another similarity between our studied *S. petersenii* populations and the case study on rotifers is that many strains of different genome size categories co-occurred contemporarily at the same locality. We detected common presence of two or more strains differing in their genome size (up to 1.8-fold difference) on 14 out of 67 localities (21%). However, the actual rate may be even higher as our sampling strategy was primarily focused at between-locality comparisons and there seem to be no clear trends in geographical distribution of genome size categories. It is likely that the prevalence of clonal reproduction contributes to the maintenance of strains of different genome size categories and their coexistence in *S. petersenii* populations. *Synura petersenii* is a colonial species and it is generally unknown whether the colonies are composed of genetically identical cells or may combine multiple genotypes (strains). Since the cultures for this study were established from one colony of cells each and always had uniform genome size, we hypothesize that strains of different genome size coexist at a locality in well-separated colonies.

Our results cannot rule out the scenario that various genome size categories in *S. petersenii* are coupled with reproductive barriers and thus reflect cryptic diversity within the taxon. Were this the case, it would mean that the nuclear ITS rDNA region is not always a sufficient molecular marker to separate microalgae species, even though this marker (and ribosomal DNA in general) is widely used as a barcode for species identification in many algal studies (e.g. Helms *et al.* 2001; Connell 2002; von Dassow *et al.* 2008; Whittaker *et al.* 2012; Jo *et al.* 2016). Genome size estimation using flow cytometry might then serve as a useful tool for identifying potential cryptic diversity in protists. Such approach was already applied in some diatoms and harmful dinoflagellates (Figueroa *et al.* 2010; Koester *et al.* 2010).

Adaptive role of genome size variation

An important aspect of intraspecific genome size diversity and its evolutionary maintenance is its putative adaptive potential, i.e. whether strains with certain genome size have a fitness (dis-)advantage in some environmental or evolutionary context. Among *S. petersenii* strains, increase in genome size resulted in significant increase in cell size and significant decrease in relative growth rate. The genome size – cell size correlation has been previously documented in many other protists (LaJeunesse *et al.* 2005; Connolly *et al.* 2008; von Dassow *et al.* 2008; Poulíčková *et al.* 2014; Olefeld *et al.* 2018), though in our study the relationship was not as tight as presumed, explaining 14.5% of the overall variation (Fig. 5A). There are two likely explanations for this discrepancy. Firstly, in the other studies the relationship was

tested across different species, thus with a much broader range of both genome sizes and cell sizes, increasing the chance of finding a general trend. Secondly, the cells of *Synura* lack a cell wall and flexibly adjust their volume under varying temperature, nutrient composition etc. (Němcová *et al.* 2010; Řezáčová-Škaloudová *et al.* 2010; Pichrtová and Němcová 2011), which could have contributed to residual model variance.

The genome size was further associated with relative growth rate of particular strains (14.8% of overall variation explained), leading to up to three-fold difference in growth rates between strains from lowest and highest genome size categories (Fig. 5B). As a result, strains with larger genomes could not grow and divide as quickly as their counterparts with smaller genome size, a feature that should be reflected in their relative fitness at least under specific environmental conditions. In water microorganisms, rapid population growth is a key factor for successful colonization of a new site and effective monopolization of local resources (i.e. the monopolization hypothesis; De Meester *et al.* 2002). Once a population is well established and possibly also locally adapted, the existence of a large bank of resting propagules (in this case *Synura* cysts) provides a powerful buffer against newly invading genotypes. Under this scenario, *S. petersenii* strains with larger genomes should be inferior colonizers of new sites, possibly sometimes outcompeted at the existing localities by other strains with smaller genomes. This is in line with the frequency distribution of genome size categories across *S. petersenii* populations, which was strongly skewed towards smaller genomes (Fig. 2). The strains with larger genomes could then be maintained in populations either due to their recurrent in situ origin from smaller-genome progenitors or because of other compensatory adaptive traits that were not included in our study, possibly stemming from their larger cell size – e.g. more efficient nutrient uptake and/or photosynthesis (Finkel *et al.* 2001; Irwin *et al.* 2006).

The question is whether the identified phenotypic consequences of genome size variation could also have translated into contrasting ecogeographical distributions of *Synura* strains.

Based on our results, this does not seem to be the case. It was already suggested by the lack of any clear geographical trends in distribution of strains from particular genome size categories (Fig. 1) and further strengthened by the occasional co-occurrence of strains with different genome size at the sampled localities. Redundancy analysis (RDA) on a dataset consisting of 19 database-derived climatic variables and geographical latitude characterising the *Synura* sampling sites provided a more comprehensive assessment (Supplementary Data Figure S1). The non-significant effect of genome size in RDA indicated that current spatial distribution of different genome size categories in *Synura* is not a result of large-scale environmental filtering. To our knowledge, our study was the first attempt to evaluate intraspecific genome size variation in protists in an ecogeographical context.

CONCLUSIONS

Genome size variation and its evolutionary consequences are highly understudied among protists, particularly on the intraspecific level with nearly no data available. We present the most comprehensive intraspecific genome size screening conducted to date, revealing a gradient of continuous genome size variation among *S. petersenii* strains. Even though we were unable to identify the main evolutionary mechanism responsible for genome size variation in this species, it likely operates via gradual changes in genome size which are accompanied with changes in genomic GC content. We hypothesize that proliferation of transposable elements and multiplication of larger genomic segments or even whole chromosomes are the most likely scenarios. Interestingly, the genome size variability was reflected in cell size and relative growth rate but not in distinct ecogeographical distribution of strains. Occasionally, we even detected strains with different genome size coexisting at the same locality. Whether these strains are associated with reproductive barriers (suggesting cryptic diversity within *S. petersenii*) remained unresolved, though prevailing clonal reproduction of the species could substantially contribute to the maintenance of local genome size diversity even in their absence.

SUPPLEMENTARY DATA

Supplementary data are available online at https://academic.oup.com/aob and consist of the following. Table S1. Collection details for *Synura petersenii* strains used in this study. Table S2. Genomic and physiological parameters recorded on investigated strains of *Synura petersenii*. Figure S1. Redundancy analyses testing associations between genomic parameters and climatic conditions across the collection sites.

FUNDING

This work was supported by the Charles University [GAUK, project no. 1304317]. Part of the work was carried out with the support of RECETOX Research Infrastructure [ID LM2015051, MEYS CR, 2016–2019].

ACKNOWLEDGEMENTS

We would like to thank to E. Gusev (Papanin's Institute for Biology of Inland Waters, Russian Academy of Sciences) for providing us with cultures.

LITERATURE CITED

Armbrust E V. 2004. The Genome of the Diatom *Thalassiosira Pseudonana*: Ecology, Evolution, and Metabolism. *Science* **306**: 79–86.

Baack EJ, Whitney KD, Rieseberg LH. 2005. Hybridization and genome size evolution: timing and magnitude of nuclear DNA content increases in *Helianthus* homoploid hybrid species. *New Phytologist* **167**: 623–630.

Baek SH, Shimode S, Kim H, Han M-S, Kikuchi T. 2009. Strong bottom–up effects on phytoplankton community caused by a rainfall during spring and summer in Sagami Bay,

Japan. *Journal of Marine Systems* **75**: 253–264.

Baetcke KP, Sparrow AH, Nauman CH, Schwemmer SS. **1967**. The relationship of DNA content to nuclear and chromosome volumes and to radiosensitivity (LD50). *Proceedings of the National Academy of Sciences of the United States of America* **58**: 533–40.

Beaulieu JM, Leitch IJ, Patel S, Pendharkar A, Knight CA. **2008**. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* **179**: 975–986.

Bennett MD. **1972**. Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **181**: 109–135.

Bennett MD, Leitch IJ. **2011**. Nuclear DNA amounts in angiosperms: targets, trends and tomorrow. *Annals of Botany* **107**: 467–590.

Boo SM, Kim HS, Shin W, *et al.* 2010. Complex phylogeographic patterns in the freshwater alga *Synura* provide new insights into ubiquity vs. endemism in microbial eukaryotes. *Molecular Ecology* **19**: 4328–4338.

Cavalier-Smith T. **2005**. Economy, Speed and Size Matter: Evolutionary Forces Driving Nuclear Genome Miniaturization and Expansion. *Annals of Botany* **95**: 147–175.

Cavalier-Smith T, Beaton MJ. **1999**. The skeletal function of non-genic nuclear DNA: new evidence from ancient cell chimaeras. *Genetica* **106**: 3–13.

Connell L. **2002**. Rapid identificationof marine algae (Raphidophyceae) using three-primer PCR amplification of nuclear internal transcribed spacer (ITS) regions from fresh and archived material. *Phycologia* **41**: 15–21.

Connolly JA, Oliver MJ, Beaulieu JM, Knight CA, Tomanek L, Moline MA. **2008**. Correlated evolution of genome size and cell volume in diatoms (Bacillariophyceae). *Journal of Phycology* **44**: 124–131.

von Dassow P, Petersen TW, Chepurnov VA, Virginia Armbrust E. **2008**. Inter- and intraspecific relationships between nuclear DNA content and cell size in selected membe rsof the centric diatom genus *Thalassiosira* (Bacillariophyceae). *Journal of Phycology* **44**: 335–349.

Derelle E, Ferraz C, Rombauts S, *et al.* 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences* **103**: 11647–11652.

Devos KM, Brown JKM, Bennetzen JL. **2002**. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome research* **12**: 1075–9.

Dionisio Pires LM, Jonker RR, Van Donk E, Laanbroek HJ. **2004**. Selective grazing by adults and larvae of the zebra mussel (*Dreissena polymorpha*): application of flow cytometry to natural seston. *Freshwater Biology* **49**: 116–126.

Doležel J. **2005**. Plant DNA Flow Cytometry and Estimation of Nuclear Genome Size. *Annals of Botany* **95**: 99–110.

Doležel J, Greilhuber J, Suda J. **2007**. *Flow Cytometry with Plant Cells*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Fick SE, Hijmans RJ. **2017**. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**: 4302–4315.

Figueroa RI, Garcés E, Bravo I. **2010**. The use of flow cytometry for species identification and life-cycle studies in dinoflagellates. *Deep-Sea Research Part II: Topical Studies in Oceanography* **57**: 301–307.

Finkel ZV, Platt T, Sathyendranath S, *et al.* 2001. Light absorption and size scaling of light-limited metabolism in marine diatoms. *Limnology and Oceanography* **46**: 86–94.

Foissner W. 2007. Protist diversity and distribution: some basic considerations In: Springer, Dordrecht, 1–8.

Galbraith DW. **2012**. Flow cytometry and fluorescence-activated cell sorting in plants: the past, present, and future. *Biomédica* **30**: 65.

Garcia-Pichel F. **1994**. A model for internal self-shading in planktonic organisms and its implications for the usefulness of ultraviolet sunscreens. *Limnology and Oceanography* **39**: 1704–1717.

Goodsell DS, Dickerson RE. **1994**. Bending and curvature calculations in B-DNA. *Nucleic acids research* **22**: 5497–503.

Gregory TR. **2001a**. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological Reviews of the Cambridge Philosophical Society* **76**: 65–101.

Gregory TR. **2001b**. The Bigger the C-Value, the Larger the Cell: Genome Size and Red Blood Cell Size in Vertebrates. *Blood Cells, Molecules, and Diseases* **27**: 830–843.

Gregory TR. **2005**. Genome Size Evolution in Animals. *The Evolution of the Genome*: 3–87.

Greilhuber J. **2005**. Intraspecific Variation in Genome Size in Angiosperms: Identifying its Existence. *Annals of Botany* **95**: 91–98.

Grover CE, Wendel JF. **2010**. Recent Insights into Mechanisms of Genome Size Change in Plants. *Journal of Botany* **2010**: 1–8.

Guillard RRL, Lorenzen CJ. **1972**. Yellow-green algae with chlorophyllide C. *Journal of Phycology* **8**: 10–14.

Helms G, Friedl T, Rambold G, Mayrhofer H. **2001**. Identification of Photobionts from the lichen family Physciaceae using algal-specific ITS rDNA sequencing. *The Lichenologist* **33**: 73–86.

Hildebrand F, Meyer A, Eyre-Walker A. **2010**. Evidence of Selection upon Genomic GC-Content in Bacteria (MW Nachman, Ed.). *PLoS Genetics* **6**: e1001107.

Irwin AJ, Finkel Z V., Schofield OME, Falkowski PG. **2006**. Scaling-up from nutrient physiology to the size-structure of phytoplankton communities. *Journal of Plankton Research* **28**: 459–471.

Jeffery NW, Hultgren K, Chak STC, Gregory TR, Rubenstein DR. **2016**. Patterns of genome size variation in snapping shrimp (V Katju, Ed.). *Genome* **59**: 393–402.

Jo BY, Kim JI, Škaloud P, Siver PA, Shin W. **2016**. Multigene phylogeny of *Synura* (Synurophyceae) and descriptions of four new species based on morphological and DNA evidence. *European Journal of Phycology* **51**: 413–430.

Jones RN, Viegas W, Houben A. **2008**. A Century of B Chromosomes in Plants: So What? *Annals of Botany* **101**: 767–775.

Kapraun DF. **2007**. Nuclear DNA Content Estimates in Green Algal Lineages: Chlorophyta and Streptophyta. *Annals of Botany* **99**: 677–701.

Keeling PJ, Slamovits CH. **2005**. Causes and effects of nuclear genome reduction. *Current Opinion in Genetics & Development* **15**: 601–608.

Kidwell MG. **2002**. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **115**: 49–63.

Kim JI, Shin H, Škaloud P, *et al.* 2019. Comparative plastid genomics of Synurophyceae: inverted repeat dynamics and gene content variation. *BMC Evolutionary Biology* **19**: 20.

Koester JA, Swalwell JE, Von Dassow P, Armbrust EV. **2010**. Genome size differentiates co-occurring populations of the planktonic diatom *Ditylum brightwellii* (Bacillariophyta). *BMC Evolutionary Biology* **10**: 1–11.

Kolář F, Čertner M, Suda J, Schönswetter P, Husband BC. **2017**. Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. *Trends in Plant Science* **22**: 1041–1055.

Kristiansen J, Preisig HR. **2007**. *Süßwasserflora von Mitteleuropa Freswater Flora of Central Europe: Chrysophyte and Haptophyte Algae* (B Büdel, G Gärtner, L Krienitz, H R.Preisig, and M Schagerl, Eds.). Berlin: Springer.

Kron P, Suda J, Husband BC. **2007**. Applications of Flow Cytometry to Evolutionary and Population Biology. *Annual Review of Ecology, Evolution, and Systematics* **38**: 847–876.

Kynčlova A, Škaloud P, Škaloudova M. **2010**. Unveiling hidden diversity in the *Synura petersenii* species complex (Synurophyceae, Heterokontophyta). *Nova Hedwigia*: 283–298.

LaJeunesse TC, Lambert G, Andersen RA, Coffroth MA, Galbraith DW. **2005**. *Symbiodinium* (Pyrrhophyta) Genome Sizes (DNA Content) Are Smallest Among Dinoflagellates1. *Journal of Phycology* **41**: 880–886.

Leitch I, Chase MW, Bennett MD. **1998**. Phylogenetic Analysis of DNA C-values provides Evidence for a Small Ancestral Genome Size in Flowering Plants. *Annals of Botany* **82**: 85–94.

Lepš J, Šmilauer P. **2014**. *Multivariate analysis of ecological data using Canoco 5* (2nd ed.). Cambridge: Cambridge University Press.

Liedtke HC, Gower DJ, Wilkinson M, Gomez-Mestre I. **2018**. Macroevolutionary shift in the size of amphibian genomes and the role of life history and climate. *Nature Ecology & Evolution* **2**: 1792–1799.

Lynch M, Conery JS. **2003**. The Origins of Genome Complexity. *Science* **302**: 1401–1404.

Mazalová P, Šarhanová P, Ondřej V, Poulíčková A. **2011**. Quantification of DNA content in freshwater microalgae using flow cytometry: a modified protocol for selected green microalgae. *Fottea* **11**: 317–328.

Medlin LK, Barker GLA, Campbell L, *et al.* 1996. Genetic characterisation of *Emiliania huxleyi* (Haptophyta). *Journal of Marine Systems* **9**: 13–31.

De Meester L, Gómez A, Okamura B, Schwenk K. **2002**. The Monopolization Hypothesis and the dispersal–gene flow paradox in aquatic organisms. *Acta Oecologica* **23**: 121–135.

Mugal CF, Arndt PF, Holm L, Ellegren H. **2015**. Evolutionary consequences of DNA methylation on the GC content in vertebrate genomes. *G3* **5**: 441–7.

Němcová Y, Neustupa J, Kviderová J, Řezáčová-Škaloudová M. **2010**. Morphological plasticity of silica scales of *Synura echinulata* (Synurophyceae) in crossed gradients of light and temperature - a geometric morphometric approach. *Nova Hedwigia* **136**: 21–32.

Olefeld JL, Majda S, Albach DC, Marks S, Boenigk J. **2018**. Genome size of chrysophytes varies with cell size and nutritional mode. *Organisms Diversity & Evolution* **18**: 163–173.

Otto F. 1990. DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. In: Crissman HA, Darzynkiewicz Z., eds. *Methods in cell biology*. New York: Academic Press, 105–110.

Pellicer J, Leitch IJ. **2014**. The Application of Flow Cytometry for Estimating Genome Size and Ploidy Level in Plants In: Humana Press, Totowa, NJ, 279–307.

Pellicer J, Leitch IJ. **2019**. The Plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. *New Phytologist*. PMID: 31608445

Pichrtová M, Němcová Y. **2011**. Effect of temperature on size and shape of silica scales in *Synura petersenii* and *Mallomonas tonsurata* (Stramenopiles). *Hydrobiologia* **673**: 1–11.

Poulíčková A, Mazalová P, Vašut RJ, Šarhanová P, Neustupa J, Škaloud P. **2014**. DNA content variation and its significance in the evolution of the genus *Micrasterias* (desmidiales, streptophyta). *PLoS ONE* **9**. e86247

Read BA, Kegel J, Klute MJ, *et al.* 2013. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature* **499**: 209–213.

Řezáčová-Škaloudová M, Neustupa J, Němcová Y. **2010**. Effect of temperature on the variability of silicate structures in *Mallomonas kalinae* and *Synura curtispina* (Synurophyceae). *Nova Hedwigia* **136**: 55–69.

Ribeiro S, Berge T, Lundholm N, Ellegaard M. 2013. Hundred Years of Environmental Change and Phytoplankton Ecophysiological Variability Archived in Coastal Sediments. PLoS ONE **8**: 1–8.

Rice A, Šmarda P, Novosolov M, *et al.* 2019. The global biogeography of polyploid plants. *Nature Ecology & Evolution* **3**: 265–273.

Rocha EPC, Danchin A. **2002**. Base composition bias might result from competition for metabolic resources. *Trends in Genetics* **18**: 291–294.

Roháček K, Barták M. 1999. Technique of the modulated chlorophyll fluorescence: basic concepts, useful parameters, and some applications. *Photosynthetica* **37**: 339–363.

Roux N, Toloza A, Radecki Z, Zapata-Arias FJ, Doležel J. 2003. Rapid detection of aneuploidy in *Musa* using flow cytometry. *Plant Cell Reports* **21**: 483–490.

Ruiz-Ruano FJ, Ruiz-Estévez M, Rodríguez-Pérez J, López-Pino JL, Cabrero J, Camacho JPM. 2011. DNA Amount of X and B Chromosomes in the Grasshoppers *Eyprepocnemis plorans* and *Locusta migratoria*. *Cytogenetic and Genome Research* **134**: 120–126.

Sandgren CD, Flanagin J. 1986. Heterothallic sexuality and density dependent encystment in the Chrysophycean alga *Synura petersenii* Korsh. *Journal of Phycology* **22**: 206–216.

Shuter BJ, Thomas JE, Taylor WD, Zimmerman AM. 1983. Phenotypic Correlates of Genomic DNA Content in Unicellular Eukaryotes and Other Cells. *The American Naturalist* **122**: 26–44.

Škaloud P, Kynčlová A, Benada O, Kofroňová O, Škaloudová M. 2012. Toward a revision of the genus *Synura*, section Petersenianae (Synurophyceae, Heterokontophyta): morphological characterization of six pseudo-cryptic species. *Phycologia* **51**: 303–329.

Škaloud P, Škaloudová M, Procházková A, Němcová Y. 2014. Morphological delineation and distribution patterns of four newly described species within the *Synura petersenii* species complex (Chrysophyceae, Stramenopiles). *European Journal of Phycology* **49**: 213–229.

Šmarda P, Bureš P. 2010. Understanding intraspecific variation in genome size in plants. *Preslia* **82**: 41–61.

Šmarda P, Bureš P. 2012. The Variation of Base Composition in Plant Genomes In: *Plant Genome Diversity Volume 1*. Vienna: Springer Vienna, 209–235.

Šmarda P, Bureš P, Horová L, *et al.* 2014. Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proceedings of the National Academy of Sciences of the United States of America* **111**: E4096-E4102.

Šmarda P, Bureš P, Horová L, Foggi B, Rossi G. 2008. Genome Size and GC Content Evolution of *Festuca*: Ancestral Expansion and Subsequent Reduction. *Annals of Botany* **101**: 421–433.

Smetacek V, Assmy P, Henjes J. 2004. The role of grazing in structuring Southern Ocean pelagic ecosystems and biogeochemical cycles. *Antarctic Science* **16**: 541–558.

Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution* **14**: 348–352.

Sparrow AH, Evans HJ. 1961. Nuclear factors affecting radiosensitivity. I. The influence of nuclear size and structure, chromosome complement, and DNA content. *Brookhaven symposia in biology* **14**: 76–100.

Sprouffske K, Wagner A. 2016. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**: 172.

Stelzer C-P, Pichler M, Stadler P, Hatheuer A, Riss S. 2019. Within-population genome size variation is mediated by multiple genomic elements that segregate independently during meiosis. *Genome Biology and Evolution* **11(12)**: 3424–3435.

Sun C, López Arriaza JR, Mueller RL. 2012. Slow DNA Loss in the Gigantic Genomes of Salamanders. *Genome Biology and Evolution* **4**: 1340–1348.

Suzuki T, Nishibayashi S, Kuroiwa T, Kanbe T, Tanaka K. 1982. Variance of ploidy in *Candida albicans*. *Journal of bacteriology* **152**: 893–6.

Swift H. 1950. The constancy of desoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences of the United States of America* **36**: 643–54.

Temsch EM, Greilhuber J, Krisai R. 2010. Genome size in liverworts. *Preslia* **82**: 63–80.

Trávníček P, Čertner M, Ponert J, Chumová Z, Jersáková J, Suda J. 2019. Diversity in genome size and GC content shows adaptive potential in orchids and is closely linked to partial endoreplication, plant life-history traits and climatic conditions. *New Phytologist* **224**: 1642–1656.

Van't Hof J, Sparrow AH. **1963**. A relationship between DNA content, nuclear volume, and minimum mitotic cycle time. *Proceedings of the National Academy of Sciences of the United States of America* **49**: 897–902.

Veldhuis MJW, Cucci TL, Sieracki ME. **1997**. Cellular DNA Content of Marine Phytoplankton Using Two New Fluorochromes: Taxonomic and Ecological Implications. *Journal of Phycology* **33**: 527–541.

Veleba A, Šmarda P, Zedek F, Horová L, Šmerda J, Bureš P. **2017**. Evolution of genome size and genomic GC content in carnivorous holokinetics (Droseraceae). *Annals of Botany* **119**: 409–416.

Venables WN, Ripley BD. **2002**. *Modern Applied Statistics with S*. New York: Springer.

Veselý P, Bureš P, Šmarda P, Pavlíček T. **2012**. Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.

Wee JL, Fasone LD, Sattler A, Starks WW, Hurley DL. **2001**. ITS/5.8S DNA sequence variation in 15 isolates of *Synura petersenii* Korshikov (Synurophyceae). *Nova Hedwigia* **122**: 245–258.

White TJ. **1990**. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics In: Innis MA, ed. *PCR Protocols: A Guide to Methods and Applications*. San Diego: Academic Press, 315–322.

Whittaker KA, Rignanese DR, Olson RJ, Rynearson TA. **2012**. Molecular subdivision of the marine diatom *Thalassiosira rotula* in relation to geographic distribution, genome size, and physiology. *BMC Evolutionary Biology* **12**.

Wichman HA, Van Den Bussche RA, Hamilton MJ, Baker RJ. **1993**. Transposable elements and the evolution of genome organization in mammals In: Springer, Dordrecht, 149–157.

Wilson EB. **1925**. The karyoplasmic ration In: *The Cell in Development and Heredity*. New York: The Macmillan Company, 727–733.

Wu Y, Sun Y, Sun S, *et al.* 2018. Aneuploidization under segmental allotetraploidy in rice and its phenotypic manifestation. *Theoretical and Applied Genetics* **131**: 1273–1285.

Zubáčová Z, Cimbůrek Z, Tachezy J. **2008**. Comparative analysis of trichomonad genome sizes and karyotypes. *Molecular and Biochemical Parasitology* **161**: 49–54.

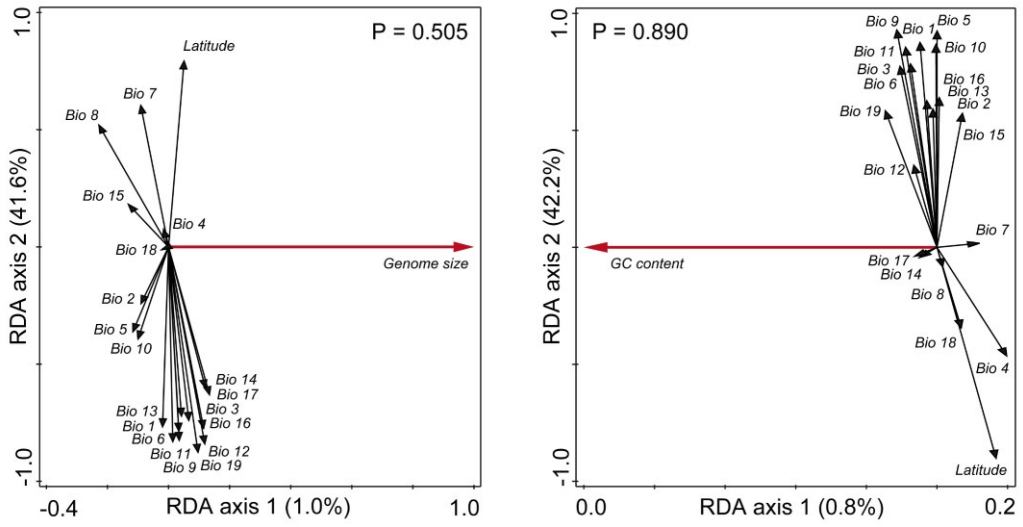


Figure S1. Redundancy analysis (RDA) on a dataset of 19 Bioclim variables and geographical latitude characterizing the collection sites. The effect of genome size (on the left) and of GC content (on the right) of residing *Synura* strains was tested using a Monte Carlo test with 999 permutations.

Table S1. Collection details for *Synura petersenii* strains used in this study.

Strain	Collection site	GPS coordinates	Sampling date	Environmental variables		
				Temp. (°C)	pH	Conductivity ($\mu S\ cm^{-1}$)
952	Djupsvatnet lake, Hovet, Norway	60°42'59.220"N, 8°6'0.144"E	10.09.2015	9.3	7.3	11
958	Djupsvatnet lake, Hovet, Norway	60°42'59.220"N, 8°6'0.144"E	10.09.2015	9.3	7.3	11
961	Großer Cramolsee pond, Zechlinerhütte, Rheinsberg, Germany	53°9'2.916"N, 12°50'24.612"E	12.11.2015	11.4	6.9	84
A21	Großer Cramolsee pond, Zechlinerhütte, Rheinsberg, Germany	53°9'2.916"N, 12°50'24.612"E	12.11.2015	11.4	6.9	84
A22	Großer Cramolsee pond, Zechlinerhütte, Rheinsberg, Germany	53°9'2.916"N, 12°50'24.612"E	12.11.2015	11.4	6.9	84
A23	Großer Cramolsee pond, Zechlinerhütte, Rheinsberg, Germany	53°9'2.916"N, 12°50'24.612"E	12.11.2015	11.4	6.9	84
A55	A small pond near Labe River, Jaroměř, Czech Republic	50°21'57.495"N, 15°55'33.460"E	26.02.2016	5.5	8.0	467
A60	A small pond near Labe River, Jaroměř, Czech Republic	50°21'57.495"N, 15°55'33.460"E	26.02.2016	5.5	8.0	467
A64	A small pond near Labe River, Jaroměř, Czech Republic	50°21'57.495"N, 15°55'33.460"E	26.02.2016	5.5	8.0	467

A66	Podhradská tůň, Bakov nad Jizerou, Czech Republic	50°27'40.000"N, 14°54'42.152"E	19.03.2016	NA	NA	NA
A76-2	Hinnerydssjöarna lake, Strömsnäsbruk, Sweden	56°36'57.456"N, 13°36'39.996"E	29.04.2016	9.6	6.9	66
A79	Hinnerydssjöarna lake, Strömsnäsbruk, Sweden	56°36'57.456"N, 13°36'39.996"E	29.04.2016	9.6	6.9	66
A80	Hinnerydssjöarna lake, Strömsnäsbruk, Sweden	56°36'57.456"N, 13°36'39.996"E	29.04.2016	9.6	6.9	66
A95	Salen lake, Grimslöv, Sweden	56°46'49.404"N, 14°33'42.012"E	1.5.2016	8.6	6.9	99
A96	Möckeln lake, Älmhult N, Sweden	56°34'8.259"N, 14°7'53.049"E	30.4.2016	8.0	6.6	91
A97	Salen lake, Grimslöv, Sweden	56°46'49.404"N, 14°33'42.012"E	1.5.2016	8.6	6.9	99
B3	Hinnerydssjöarna lake, Strömsnäsbruk, Sweden	56°36'57.456"N, 13°36'39.996"E	29.04.2016	9.6	6.9	66
B6	Unnamed pond, Ljungby, Sweden	56°39'12.420"N, 13°26'31.596"E	29.4.2016	9.0	6.0	37
B9	Unnamed pond, Ljungby, Sweden	56°39'12.420"N, 13°26'31.596"E	29.4.2016	9.0	6.0	37
B15	Vättern lake, Motala, Sweden	58°44'4.128"N, 14°59'1.608"E	30.4.2016	6.9	7.0	144
B24	Salen lake, Grimslöv, Sweden	56°46'49.404"N, 14°33'42.012"E	1.5.2016	8.6	6.9	99

B41	A small pond near Labe River, Jaroměř, Czech Republic	50°21'57.495"N, 15°55'33.460"E	26.02.2016	5.5	8.0	467
B50	Vizír lake, Hamr, Czech Republic	48°57'52.620"N, 14°53'18.840"E	9.10.2016	7.9	6.6	89
B53	Vizír lake, Hamr, Czech Republic	48°57'52.620"N, 14°53'18.840"E	9.10.2016	7.9	6.6	89
B54	Vizír lake, Hamr, Czech Republic	48°57'52.620"N, 14°53'18.840"E	9.10.2016	7.9	6.6	89
B87	Pískovna Cep I, Czech Republic	48°55'2.338"N, 14°53'1.630"E	9.10.2016	14.6	7.5	185
B88	Pískovna Cep I, Czech Republic	48°55'2.338"N, 14°53'1.630"E	9.10.2016	14.6	7.5	185
B98	Pískovna Cep I, Czech Republic	48°55'2.338"N, 14°53'1.630"E	9.10.2016	14.6	7.5	185
C52	St John's Lough, Co. Leitrim, Ireland	54°3'4.572"N, 7°52'35.724"W	3.2.2017	6.0	8.4	112
C77	St John's Lough, Co. Leitrim, Ireland	54°3'4.572"N, 7°52'35.724"W	3.2.2017	6.0	8.4	112
D40	Hatchet Moor Inclosure, East Boldre, Brockenhurst, Great Britain	50°48'30.060"N, 1°29'0.384"W	19.03.2017	12.0	8.7	143
D45	Hatchet Moor Inclosure, East Boldre, Brockenhurst, Great Britain	50°48'30.060"N, 1°29'0.384"W	19.03.2017	12.0	8.7	143
D55	Unnamed pond, Saint-Philbert-de-Grand-Lieu, France	47°2'48.192"N, 1°39'7.128"W	12.03.2017	12.0	6.9	603

D57	Unnamed pond, Saint-Philbert-de-Grand-Lieu, France	47°2'48.192"N, 1°39'7.128"W	12.03.2017	12.0	6.9	603
D62	Unnamed pond, Saint-Joachim, France	47°25'38.964"N, 2°16'15.240"W	14.03.2017	11.0	7.6	370
D65	Unnamed pond, Saint-Joachim, France	47°25'38.964"N, 2°16'15.240"W	14.03.2017	11.0	7.6	370
D67	Étang de la Grippé, Varades, France	47°23'9.276"N, 0°59'40.920"W	14.03.2017	13.0	9.3	607
D83	Unnamed pond next to Cadnam River, Great Britain	50°56'51.144"N, 1°32'58.020"W	19.03.2017	12.0	8.7	300
D91	Le Grand Étang, Sains, France	48°33'17.244"N, 1°34'50.664"W	16.03.2017	13.0	7.8	351
E19	Unnamed pond, Brockenhurst, Great Britain	50°50'20.616"N, 1°25'35.796"W	19.03.2017	12.0	8.5	408
E28	Unnamed pond, Saint-André-des-Eaux, France	47°19'27.300"N, 2°16'32.160"W	14.03.2017	16.0	8.7	1081
E34	Unnamed pond, Saint-Joachim, France	47°25'38.964"N, 2°16'15.240"W	14.03.2017	11.0	7.6	370
E42	Le Grand Étang, Sains, France	48°33'17.244"N, 1°34'50.664"W	16.03.2017	13.0	7.8	351
E51	Unnamed pond, Sucé-sur-Erdre, France	47°20'41.532"N, 1°30'10.800"W	13.03.2017	13.0	6.5	298
E53	Unnamed pond, Sucé-sur-Erdre, France	47°20'41.532"N, 1°30'10.800"W	13.03.2017	13.0	6.5	298

E54	Unnamed pond, Sucé-sur-Erdre, France	47°20'41.532"N, 1°30'10.800"W	13.03.2017	13.0	6.5	298
F8	Unnamed pond, Genarp, Sweden	55°38'6.168"N, 13°25'3.430"E	27.03.2017	NA	8.0	152
F9	Unnamed pond, Genarp, Sweden	55°38'6.168"N, 13°25'3.430"E	27.03.2017	NA	8.0	152
F10	Unnamed pond, Genarp, Sweden	55°38'6.168"N, 13°25'3.430"E	27.03.2017	NA	8.0	152
F19	Krankesjön lake, Lund, Sweden	55°41'30.987"N, 13°29'34.054"E	27.03.2017	NA	7.8	355
F22	Krankesjön lake, Lund, Sweden	55°41'30.987"N, 13°29'34.054"E	27.03.2017	NA	7.8	355
F23	Krankesjön lake, Lund, Sweden	55°41'30.987"N, 13°29'34.054"E	27.03.2017	NA	7.8	355
F26	Krankesjön lake, Lund, Sweden	55°41'30.987"N, 13°29'34.054"E	27.03.2017	NA	7.8	355
F42	Vederslövssjön lake, Vederslöv, Sweden	56°47'4.524"N, 14°44'15.911"E	01.04.2017	NA	6.9	112
F46	Unnamed pond, Klagshamn, Sweden	55°31'39.810"N, 12°55'37.182"E	03.04.2017	NA	8.8	1094
F49	Boire de Champtocé, Maine-et-Loire, France	47°24'32.508"N, 0°51'44.172"W	03.04.2017	13	7.7	424
F50	Unnamed pond, Tygelsjö, Sweden	55°31'11.125"N, 12°59'25.393"E	03.04.2017	NA	7.8	892

F51	Unnamed pond, Tygelsjö, Sweden	55°31'11.125"N, 12°59'25.393"E	03.04.2017	NA	7.8	892
F52	Unnamed pond, Tygelsjö, Sweden	55°31'11.125"N, 12°59'25.393"E	03.04.2017	NA	7.8	892
F54	Unnamed pond, Tygelsjö, Sweden	55°31'11.125"N, 12°59'25.393"E	03.04.2017	NA	7.8	892
F56	Unnamed pond, Tygelsjö, Sweden	55°31'11.125"N, 12°59'25.393"E	03.04.2017	NA	7.8	892
F59	Unnamed pond, Tygelsjö, Sweden	55°31'11.125"N, 12°59'25.393"E	03.04.2017	NA	7.8	892
F61	Unnamed pond, Tygelsjö, Sweden	55°31'11.125"N, 12°59'25.393"E	03.04.2017	NA	7.8	892
F68	Svaneholmssjön lake, Skurup, Sweden	55°30'2.394"N, 13°28'46.938"E	03.04.2017	NA	7.8	369
F71	Svaneholmssjön lake, Skurup, Sweden	55°30'2.394"N, 13°28'46.938"E	03.04.2017	NA	7.8	369
F76	Svaneholmssjön lake, Skurup, Sweden	55°30'2.394"N, 13°28'46.938"E	03.04.2017	NA	7.8	369
F85	Häckebergasjön lake, Genarp, Sweden	55°34'45.545"N, 13°25'54.869"E	03.04.2017	NA	8.3	392
F87	Häckebergasjön lake, Genarp, Sweden	55°34'45.545"N, 13°25'54.869"E	03.04.2017	NA	8.3	392
G16	Åsrumvannet lake, Larvik Municipality, Norway	59°9'24.599"N, 10°2'50.492"E	28.04.2017	8.2	7.9	85

G17	Gjennestadvannet lake, Stokke, Norway	59°14'6.778"N, 10°14'26.812"E	28.04.2017	9.0	7.4	104
G23	Gjennestadvannet lake, Stokke, Norway	59°14'6.778"N, 10°14'26.812"E	28.04.2017	9.0	7.4	104
G34	Rodbyvatnet lake, Hurum, Norway	59°35'35.912"N, 10°29'17.045"E	28.04.2017	5.1	7.5	101
G44	Rodbyvatnet lake, Hurum, Norway	59°35'35.912"N, 10°29'17.045"E	28.04.2017	5.1	7.5	101
G50	Fiskumvannet, Eikeren, Norway	59°42'13.183"N, 9°49'47.402"E	28.04.2017	7.0	7.0	75
G61	Fiskumvannet, Eikeren, Norway	59°42'13.183"N, 9°49'47.402"E	28.04.2017	7.0	7.0	75
H11	Gjennestadvannet lake, Stokke, Norway	59°14'6.778"N, 10°14'26.812"E	28.04.2017	9.0	7.4	104
H14	Fabrikkdammen lake, Røyken, Norway	59°45'29.851"N, 10°26'31.049"E	28.04.2017	5.1	8.1	206
H16	Fabrikkdammen lake, Røyken, Norway	59°45'29.851"N, 10°26'31.049"E	28.04.2017	5.1	8.1	206
H18	Åshildrødtjernet lake, Sandefjord, Norway	59°10'32.394"N, 10°7'8.602"E	28.04.2017	10.5	7.4	63
H27	Gjersjøen lake, Oppegård, Norway	59°45'30.175"N, 10°46'42.704"E	28.04.2017	5.5	7.5	238
H29	Gjersjøen lake, Oppegård, Norway	59°45'30.175"N, 10°46'42.704"E	28.04.2017	5.5	7.5	238

H31	Gjersjøen lake, Oppegård, Norway	59°45'30.175"N, 10°46'42.704"E	28.04.2017	5.5	7.5	238
H48	Unnamed pond, Larvik, Norway	59°4'51.618"N, 10°2'9.024"E	28.04.2017	7.7	7.1	190
H50	Unnamed pond, Larvik, Norway	59°4'51.618"N, 10°2'9.024"E	28.04.2017	7.7	7.1	190
H97	Second Pond, Goulds, St. John's, Newfoundland, Canada	47°27'27.576"N, 52°43'46.272"W	25.05.2017	8.0	7.4	77
I2	Second Pond, Goulds, St. John's, Newfoundland, Canada	47°27'27.576"N, 52°43'46.272"W	25.05.2017	8.0	7.4	77
I18	Quidi Vidi Lake, St. John's, Newfoundland, Canada	47°34'39.972"N, 52°41'48.408"W	25.05.2017	9.0	7.4	775
I21	Quidi Vidi Lake, St. John's, Newfoundland, Canada	47°34'39.972"N, 52°41'48.408"W	25.05.2017	9.0	7.4	775
Ir 38B	Upper Lake, County Kerry, Ireland	51°59'50.388"N, 9°33'2.603"W	1.4.2010	NA	NA	NA
J35	Unnamed pond, Bishop's Falls, Newfoundland, Canada	48°57'8.928"N, 55°30'26.208"W	25.05.2017	11.0	7.7	50
L30	Lacul Cocor, Câmpina, Romania	45°9'2.700"N, 25°44'54.024"E	16.11.2017	7.7	6.7	235
L37	Lacul Cocor, Câmpina, Romania	45°9'2.700"N, 25°44'54.024"E	16.11.2017	7.7	6.7	235
L43	Lacul Cocor, Câmpina, Romania	45°9'2.700"N, 25°44'54.024"E	16.11.2017	7.7	6.7	235

L45	Unnamed pond, Gorgota, Romania	44°47'32.856"N, 26°5'6.036"E	17.11.2017	8.2	8.4	720
L46	Unnamed pond, Gorgota, Romania	44°47'32.856"N, 26°5'6.036"E	17.11.2017	8.2	8.4	720
L55	Unnamed pond, Botevgrad, Bulgaria	42°54'20.772"N, 23°48'55.584"E	17.11.2017	8.6	8.0	387
L59	Unnamed pond, Botevgrad, Bulgaria	42°54'20.772"N, 23°48'55.584"E	17.11.2017	8.6	8.0	387
L61	Unnamed pond, Botevgrad, Bulgaria	42°54'20.772"N, 23°48'55.584"E	17.11.2017	8.6	8.0	387
L71	Rivio Lake, Greece	38°43'35.976"N, 21°11'51.576"E	19.11.2017	16.1	8.3	543
L89	Rivio Lake, Greece	38°43'35.976"N, 21°11'51.576"E	19.11.2017	16.1	8.3	543
L90	Rivio Lake, Greece	38°43'35.976"N, 21°11'51.576"E	19.11.2017	16.1	8.3	543
M16	Unnamed pond, Metsovo, Greece	39°47'58.524"N, 21°9'36.684"E	20.11.2017	6.5	7.6	178
M18	Unnamed pond, Metsovo, Greece	39°47'58.524"N, 21°9'36.684"E	20.11.2017	6.5	7.6	178
M22	Unnamed pond, Metsovo, Greece	39°47'58.524"N, 21°9'36.684"E	20.11.2017	6.5	7.6	178
N19	Kis-Balaton, Balatonmagyaród, Hungary	46°36'55.387"N, 17°10'4.520"E	22.11.2017	7.4	7.7	794

N22	Kis-Balaton, Balatonmagyaród, Hungary	46°36'55.387"N, 17°10'4.520"E	22.11.2017	7.4	7.7	794
O28	Rio Rabagão, Montalegre, Portugal	41°40'38.316"N, 7°59'1.716"W	07.02.2018	9.0	6.0	22
O39	Encoro das Conchas, Ourense, Spain	41°56'21.948"N, 8°2'3.192"W	07.02.2018	7.1	6.0	56
P64	Unnamed pond, Laghetti delle Mucille, Ronchi dei Legionari, Italy	45°49'16.498"N, 13°31'31.141"E	14.03.2018	11.8	8.5	460
P68	Unnamed pond, Laghetti delle Mucille, Ronchi dei Legionari, Italy	45°49'16.498"N, 13°31'31.141"E	14.03.2018	11.8	8.5	460
Q88	Jezioro Studzieniczne, Augustov, Poland	53°52'1.236"N, 23°5'32.640"E	15.04.2018	13.1	8.6	273
R021	Unnamed pond, Gus-Khrustalny District, Vladimir Oblast, Russia	55°35'2.040"N, 40°26'14.640"E	12.10.2013	NA	NA	NA
R022	Unnamed pond, Gus-Khrustalny District, Vladimir Oblast, Russia	55°35'2.040"N, 40°26'14.640"E	12.10.2013	NA	NA	NA
R023	Unnamed pond, Gus-Khrustalny District, Vladimir Oblast, Russia	55°35'2.040"N, 40°26'14.640"E	12.10.2013	NA	NA	NA
R10	Unnamed pond, Szafranki, Poland	53°15'36.324"N, 22°36'12.852"E	16.04.2018	16.9	7	190
S102.C2	Medenice lake, Staňkov, Czech Republic	48°56'21.232"N, 14°57'21.049"E	14.11.2012	4.8	7.6	122
S114.C7	Unnamed pond near Vistasälven River, Kiruna, Sweden	67°53'1.194"N, 18°59'57.041"E	21.08.2013	13.00	6.3	134

S31	Raudtee karjäär lake, Selja, Pärnumaa, Estonia	58°30'31.824"N, 24°48'18.144"E	20.04.2018	12.2	7.5	289
S51	Väikejärv lake, Parika LKA, Viljandi maakond, Estonia	58°29'24.684"N, 25°45'26.496"E	20.04.2018	12.4	6.9	89
S63.B3	Tasersuaq lake, Ilulissat, Greenland, Denmark	69°13'19.100"N, 51°5'3.050"W	02.08.2011	13.4	6.0	50
S7.7	Babín pond, Matějov, Žďárské vrchy, Czech Republic	49°32'31.679"N, 15°53'48.638"E	2007	NA	NA	NA
S84	Jezioro Głuche, Gmina Krasnopol, Poland	54°3'43.164"N, 23°14'33.864"E	23.05.2018	15.2	8.1	443
T4	Sukhodolka River, Borynychi, Ukraine	49°29'50.712"N, 24°13'1.020"E	30.04.2018	13.3	8.0	646
T13	Sukhodolka River, Borynychi, Ukraine	49°29'50.712"N, 24°13'1.020"E	30.04.2018	13.3	8.0	646
T53	Unnamed pond, Les Estanyols, Bolquère, France	42°30'53.244"N, 2°4'35.292"E	08.06.2018	14.6	6.2	227
T54	Unnamed pond, Les Estanyols, Bolquère, France	42°30'53.244"N, 2°4'35.292"E	08.06.2018	14.6	6.2	227
W15	Unnamed stream, Khanty-Mansiysky District, Russia	60°56'53.448"N, 68°20'27.096"E	28.06.2018	9.8	7.7	69
W36	Unnamed stream, Batovo, Khanty-Mansiysky District, Russia	60°24'21.456"N, 69°49'21.252"E	28.06.2018	11.4	7.9	244
W81	Gornopravdinsk pond near Irtyš River, Khanty-Mansiysky District, Russia	60°3'45.468"N, 69°54'27.612"E	28.06.2018	13.3	8.7	328

W92	Unnamed pond, Pfanndlbrunn, Bad Mitterndorf, Sonnenalm, Austria	47°33'30.594"N, 13°54'34.855"E	14.06.2018	15.0	7.3	693
X1	Unnamed pond, Pfanndlbrunn, Bad Mitterndorf, Sonnenalm, Austria	47°33'30.594"N, 13°54'34.855"E	14.06.2018	15.0	7.3	693

Table S2. Genomic and physiological parameters recorded on investigated strains of *Synura petersenii*.

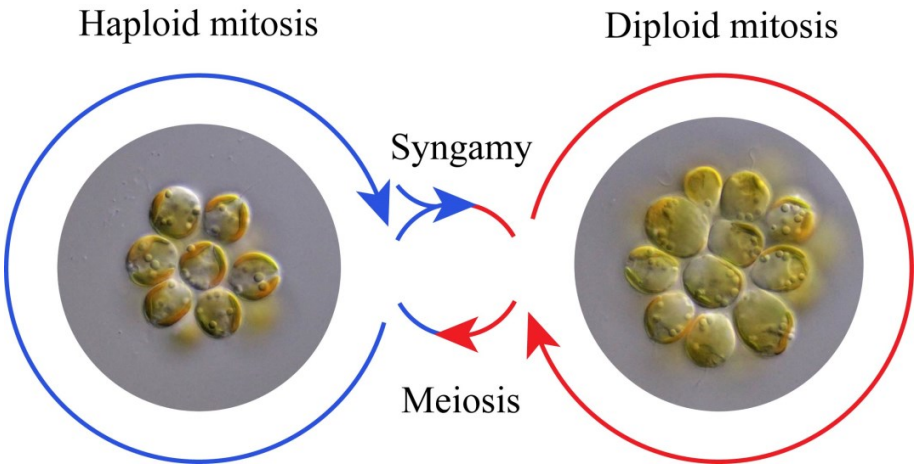
Strain	Genome size (pg)	Genome size (Gbp)	GC content (%)	Biovolume (μm^3)	Growth rate (t-mid)
952	1.690	1.653	39.3	-	-
958	1.767	1.728	39.5	-	-
961	1.071	1.048	38.4	-	-
A21	1.067	1.043	-	2058	11.75
A22	1.070	1.046	39.9	1024	-
A23	1.276	1.248	-	-	-
A55	1.217	1.190	-	-	-
A60	1.433	1.402	38.4	-	-
A64	1.423	1.391	41.0	-	-
A66	1.075	1.051	-	1460	-
A76-2	1.189	1.162	-	1412	3.74
A79	1.148	1.123	-	1382	-
A80	1.103	1.078	-	1544	-
A95	1.187	1.161	-	-	-
A96	1.205	1.178	-	1120	3.11
A97	1.138	1.113	-	1591	-
B3	1.142	1.117	40.6	1196	5.56
B6	1.050	1.027	-	1548	-
B9	1.755	1.716	39.2	1305	-
B15	1.075	1.052	39.7	1420	-
B24	1.175	1.150	-	1568	8.48
B41	1.378	1.348	40.4	1231	10.79
B50	1.049	1.026	-	-	3.74
B53	1.020	0.997	38.7	1000	7.92
B54	1.037	1.015	38.5	1057	5.11
B87	1.065	1.042	-	-	-
B88	1.035	1.013	-	-	-
B98	1.056	1.033	-	1288	-
C52	1.216	1.189	-	1504	3.41
C77	1.169	1.143	-	1467	6.25
D40	1.113	1.088	-	-	-
D45	1.132	1.108	-	-	-
D55	1.824	1.784	39.4	2077	7.43
D57	1.138	1.113	-	-	-
D62	1.182	1.156	-	-	-
D65	1.162	1.137	-	-	-
D67	1.165	1.139	-	-	-
D83	1.973	1.929	-	-	-
D91	1.163	1.138	39.4	-	4.79
E19	1.738	1.700	-	-	-
E28	1.178	1.152	-	1849	6.86
E34	1.123	1.098	-	-	-

E42	1.139	1.114	-	-	-
E51	1.095	1.071	-	-	-
E53	1.088	1.064	-	-	-
E54	1.113	1.088	-	-	-
F8	1.113	1.088	-	1347	4.29
F9	1.096	1.071	-	1191	5.32
F10	1.122	1.097	40.3	1280	-
F19	1.253	1.225	-	-	-
F22	2.022	1.978	40.9	-	-
F23	1.104	1.079	-	-	-
F26	1.268	1.240	-	-	-
F42	1.140	1.115	-	1318	6.29
F46	1.879	1.837	38.7	1719	7.98
F49	1.967	1.923	-	-	-
F50	1.306	1.277	40.9	1556	9.19
F51	1.264	1.236	-	-	4.52
F52	1.265	1.237	-	1674	6.10
F54	1.327	1.298	41.2	1926	4.89
F56	1.296	1.268	39.7	1499	-
F59	1.241	1.214	-	1239	-
F61	1.280	1.252	-	1540	-
F68	1.317	1.288	40.6	-	-
F71	1.271	1.243	-	-	-
F76	1.300	1.271	39.6	-	-
F85	1.110	1.086	-	-	-
F87	1.296	1.267	-	-	-
G16	1.986	1.942	37.1	-	12.58
G17	1.062	1.038	-	1500	7.26
G23	1.031	1.008	-	-	3.54
G34	1.037	1.014	-	-	-
G44	1.088	1.064	-	-	-
G50	1.093	1.069	39.9	1621	4.28
G61	1.064	1.040	37.8	-	-
H11	1.073	1.050	38.9	-	-
H14	1.120	1.096	-	-	-
H16	1.062	1.039	-	-	-
H18	1.811	1.771	-	-	-
H27	1.094	1.070	-	-	-
H29	0.973	0.952	37.2	1508	-
H31	0.971	0.950	-	1715	3.97
H48	1.161	1.135	-	-	-
H50	1.094	1.070	39.1	-	-
H97	1.656	1.619	39.4	-	-
I2	1.676	1.639	39.4	-	-
I18	1.613	1.577	39.8	-	-
I21	1.151	1.126	41.2	-	-
Ir 38B	1.080	1.056	-	-	-

J35	1.106	1.081	-	-	-
L30	1.051	1.028	-	-	-
L37	1.076	1.052	-	-	-
L43	1.073	1.050	-	-	-
L45	1.230	1.203	-	-	-
L46	1.268	1.241	-	-	-
L55	1.143	1.118	-	-	-
L59	1.152	1.127	-	-	-
L61	1.170	1.145	-	-	-
L71	1.209	1.182	39.2	-	-
L89	1.181	1.155	-	-	-
L90	1.200	1.173	40.1	-	-
M16	1.265	1.237	-	-	-
M18	1.310	1.281	39.8	-	-
M22	1.302	1.273	-	-	-
N19	1.276	1.248	-	-	-
N22	1.277	1.249	-	-	-
O28	1.366	1.336	-	-	-
O39	1.270	1.242	-	-	-
P64	1.845	1.804	-	-	-
P68	1.863	1.822	-	-	-
Q88	1.228	1.201	-	-	-
R021	1.101	1.076	-	-	-
R022	1.086	1.062	-	-	7.17
R023	1.143	1.118	-	-	-
R10	1.051	1.028	-	-	-
S102.C2	1.090	1.066	40.2	-	-
S114.C7	1.530	1.496	-	1676	6.33
S31	1.286	1.257	-	-	-
S51	1.880	1.839	-	-	-
S63.B3	1.516	1.482	39.6	1956	8.16
S7.7	1.036	1.014	39.0	1374	8.55
S84	1.203	1.177	-	-	-
T4	1.265	1.237	-	-	-
T13	1.250	1.223	-	-	-
T53	1.120	1.095	-	-	-
T54	1.095	1.071	-	-	-
W15	1.585	1.550	-	-	-
W36	1.044	1.021	-	-	-
W81	1.275	1.247	-	-	-
W92	1.839	1.799	-	-	-
X1	1.841	1.800	-	-	-

PAPER V

Alternating nuclear DNA content in chrysophytes provides evidence of their isomorphic haploid-diploid life cycle.



An illustrative scheme of the isomorphic haploid-diploid life cycle of chrysophytes on the example of *S. petersenii*.

Alternating nuclear DNA content in chrysophytes provides evidence of their isomorphic haploid-diploid life cycle

Dora Čertnerová[1]*, Martin Čertner[1,2] & Pavel Škaloud[1]

[1]*Department of Botany, Faculty of Science, Charles University, Benátská 2, CZ-128 00 Prague, Czech Republic*
[2]*Institute of Botany, The Czech Academy of Sciences, Zámek 1, CZ-252 43 Průhonice, Czech Republic*

*Author for correspondence (e-mail: dora.certnerova@gmail.com)

ABSTRACT

Across eukaryotic organisms there is a great diversity of life cycles. This particularly applies to unicellular eukaryotes (protists), where the life cycles are still largely unexplored, although this knowledge is key to understanding their biology.

To detect the often inconspicuous transitions among life cycle stages, we focused at shifts in ploidy levels within strains of unicellular chrysophyte alga. Representatives of three genera (*Chrysosphaerella*, *Ochromonas*, and *Synura*) were analysed for nuclear DNA contents using a propidium iodide flow cytometry. Selected strains exhibiting ploidy level variation were also surveyed for DNA base composition (GC content) and cell size. Additionally, we tracked ploidy level changes in seven strains under long-term cultivation.

An alternation of two ploidy levels was revealed in the life cycle of chrysophytes with both life cycle stages capable of mitotic growth and long-term survival in cultivation. With the exception of a small increase in cell size with higher ploidy, both life cycle stages shared the same phenotype and also had highly similar genomic GC content. Further, we detected three ploidy levels in two *Synura* species (*S. glabra*, *S. heteropora*), where the highest ploidy (putatively 4x) most likely resulted from a polyploidization event.

Consequently, chrysophytes have a haploid-diploid life cycle with isomorphic life cycle stages. As far as we know, this is the first report of such life cycle strategy in unicellular algae. Life cycle stages and life stage transitions seem to be synchronized among all cells coexisting within a culture, possibly due to chemical signals. Particular life stages may be more successful under certain environmental conditions, for our studied strains the diploid stage prevailed in cultivation.

Key words: isomorphic haploid-diploid life cycle, chrysophytes, alternation of life stages, nuclear DNA content, flow cytometry, *Synura*

INTRODUCTION

Across eukaryotic organisms there is a considerable diversity of life cycles. Individuals at particular life cycle stages may differ, for example, by their overall morphology, environmental requirements, or by the number of chromosome sets in cell nuclei (ploidy level). Many organisms alternate between two stages, a haploid phase with one set of chromosomes reduced by meiosis and a duplicated diploid phase following the fusion of gametes (Otto and Gerstein 2008; Beukeboom and Nicolas Perrin 2014). Depending on whether the both stages are more-or-less equally represented in the life cycle or one of them largely predominates, various life cycles can be recognized. In a diploid life cycle, organisms switch between a short haploid phase (usually restricted to unicellular gametes) and the prevailing diploid phase, only which is capable of mitotic growth. Such a life cycle occurs among diatoms, raphidophytes or budding yeasts; however, it is best known from animals, including humans (Figueroa and Rengefors 2006; Montresor *et al.* 2016; Figueroa *et al.* 2018; Fischer *et al.* 2021). A haploid life cycle is characterized by mitosis restricted to the haploid phase, which also lasts for most of the organism's lifespan, as the only diploid stage is a unicellular zygote. The haploid life cycle evolved in stoneworts (charophytes) and in some other green algae (Mable and Otto 1998). However, probably the most widespread among organisms is a haploid-diploid life cycle. Here, both the haploid and diploid stages are capable of mitotic growth. This life strategy dominates in land plants, red and brown algae, basidiomycete fungi, but also occurs in many groups of green algae and various groups of unicellular algae (Richerd *et al.* 1993; Mable and Otto 1998; Rousseau *et al.* 2007; Speijer *et al.* 2015; Figueroa *et al.* 2018). A peculiar form of this life cycle evolved in some green and red algae, where the haploid and diploid phases are morphologically indistinguishable (Destombe *et al.* 1989; Wichard *et al.* 2015). This isomorphic haploid-diploid life cycle can be found, for example, in sea lettuce (*Ulva lactuca*; Wichard *et al.* 2015).

Knowledge of the life cycle and identification of particular life cycle stages are key not only to understand the basic biology of studied organisms, but also to correctly assess their genome size (1C vs. 2C value), and hence to design an optimal sequencing strategy, and properly interpret population genetic or genomic data (Pirrello *et al.* 2018; Otto and Rosales 2020; Čertnerová and Galbraith 2021). In algae, understanding the life cycles is also essential to predict formation of blooms and toxins production (Figueroa *et al.* 2018). However, the knowledge of algal life cycles remains still largely fragmented, particularly for most unicellular algae. This could be attributed to their microscopic size and a frequent lack of pronounced morphological features, which makes life cycle stage transitions harder to detect.

Here, we attempt to overcome the problem by measuring nuclear DNA contents and looking for shifts in a ploidy level that should be associated with the life cycle transitions.

To broaden our knowledge of algal life cycles, we chose chrysophytes as a model group. The chrysophytes, also known as golden-brown algae, are single-celled or colonial flagellates, which occur primarily in freshwater phytoplankton and their blooms can cause an unpleasant fishy odour in drinking water reservoirs (Nicholls and Gerrath 1985). In some taxa (e.g. among the representatives of the genera *Synura* and *Chrysosphaerella*), the cells are covered by species-specific silica scales (Kristiansen and Škaloud 2017). However, not much is known about the chrysophyte life cycle. Undifferentiated cells may serve as gametes (Wawrik 1972). Fusion of the gametes was observed in several cases, specifically the apical fusion in *Kephyrion*, *Stenocalyx*, *Chrysolykos* and *Dinobryon* or the posterior fusion in *Synura* and *Mallomonas* (Fott 1959; Wawrik 1972). The fusion of gametes is followed by cyst formation (Kristiansen and Škaloud 2017). Some colonial species even produce separate male and female colonies (Sandgren 1981). According to Sandgren and Flanagin (Sandgren and Flanagin 1986), the genus *Synura* is heterothallic and its sexuality might be induced at high cell densities. In recent years, the chrysophytes have drawn attention due to their remarkable DNA content diversity, ranging from 0.09 to 24.85 pg (0.09 to 24.31 Gbp), accompanied by numerous cases of major intraspecific variability (Olefeld *et al.* 2018; Čertnerová and Škaloud 2020; Majda *et al.* 2021). This variation was either attributed to polyploidization (i.e., whole-genome doubling) or its source remained unresolved (Čertnerová and Škaloud 2020; Majda *et al.* 2021).

The present study was stimulated by our repeated detection of intraspecific DNA content variation arising in cultures of some of our investigated taxa that opened the question whether this could be attributed to unprecedented rates of certain evolutionary processes (e.g. polyploidization, aneuploidization, proliferation of transposable elements; De Storme and Mason 2014) or whether it constitutes an inherent part of organisms' life cycles. By employing flow cytometry on selected, DNA content variable taxa we are asking the following specific questions: 1) What are the patterns of DNA content variation among and within strains; and do these correspond to ploidy level shifts (i.e., two-fold DNA content differences)? 2) Of what character are the temporal changes in the DNA content of strains over time in cultivation? 3) Is intraspecific / intra-strain DNA content variation linked with differences in genomic base composition (i.e. GC content; no differences expected under the scenario of whole genome duplication)? 4) Are there any apparent phenotypic differences between intraspecific strains with different DNA contents?

MATERIALS AND METHODS

Origin and cultivation of the investigated strains

For this study, we selected chrysophyte taxa where intraspecific DNA content variation was detected during our previous unpublished work. Altogether 61 chrysophyte strains were obtained from 49 various freshwater localities across the Northern hemisphere, comprising 59 isolates of the genus *Synura*, one isolate of *Ochromonas tuberculata* and one isolate of *Chrysosphaerella brevispina*. The sampling details are listed in Supplementary data Table S1. To establish new cultures, water samples were taken using a 25 μm mesh plankton net and single cells or colonies were captured by micro-pipetting and transferred into separate culture wells filled either with MES buffered DY-IV (in case of *S. sphagnicola* and *Ochromonas tuberculata*; Andersen *et al.* 1997) or with WC medium (Guillard and Lorenzen 1972). The culture collection was supplemented with seven previously established cultures (Korshikov 1929; Škaloud *et al.* 2014, 2019; Čertnerová and Škaloud 2020). All cultures were maintained at 17 °C (cooling box Pol-Eko Aparatura Sp.J., model ST 1, Wodzisław Śląski, Poland) with a 24-h light mode under illumination of 30 μmol m^{-2} s^{-1} (TLD 18W/33 fluorescent lamps, Philips, Amsterdam, Netherlands). The generation time of *Synura* cells under these cultivation conditions can be approximated as 2 days (based on Kim, Jin *et al.* 2008). Subsequently, the strains were transferred into Erlenmeyer flasks filled with 30 mL of growth medium and kept for longer cultivation with re-inoculations into a fresh medium every three months.

Phylogenetic analyses

To genetically identify *Synura* strains, the internal transcribed spacer of nuclear ribosomal DNA (ITS1, 5.8S and ITS2 rDNA; nu ITS rDNA) of individual isolates was sequenced. For this purpose, genomic DNA was extracted from a centrifuged pellet of cells by InstaGene Matrix (Bio-Rad, Hercules, CA, USA), and the resulting supernatant directly used as a PCR template. Amplifications were performed using the universal primer ITS4 (White 1990) and a genus-specific primer Kn1.1 (Wee *et al.* 2001). PCRs were carried out in a total volume of 20 μL with a PCR mix containing 0.2 μL of MyTaqHS DNA polymerase (Bioline, Memphis, TN, USA), 4 μL of MyTaqHS buffer (Bioline), 0.4 μL of each primer, 14 μL of double distilled water and 1 μL of template DNA (not quantified). Amplifications were performed in Eppendorf Mastercycler ep Gradient 5341 (Eppendorf GmbH, Hamburg, Germany) using the following program: 1 min of denaturation at 95 °C; followed by 35 cycles of denaturation at 95 °C (15 s), annealing at 52 °C (30 s) and elongation at 72 °C (40 s), concluded with a final extension at 72 °C (7 min) and held at 10 °C. The PCR products were sized on a 1% agarose gel and then purified using AMPure XP

magnetic beads (Agencourt, Beckman Coulter, Brea, CA, USA). The purified DNA templates were sequenced using the Sanger sequencing method at Macrogen, Inc. (Amsterdam, Netherlands, https://dna. macrogen.com). Finally, the obtained sequences were identified using BLAST in the National Center for Biotechnology Information (NCBI) Search database and our own ITS database built up during previous studies (Škaloud *et al.* 2012, 2014, 2019, 2020; Čertnerová and Škaloud 2020).

The phylogenetic tree was inferred by the maximum likelihood (ML) analysis using RAxML 8.1.20 (Stamatakis 2014), applying the GTR+Γ evolutionary model. Bootstrap analysis was performed with the rapid bootstrapping procedure, using 100 pseudoreplicates. Bayesian posterior probabilities were computed by MrBayes 3.2.6 (Ronquist *et al.* 2012). Two parallel Monte Carlo Markov chains runs were carried out for 3 million generations each with one cold and three heated chains. Trees and parameters were sampled every 100th generation. Convergence of the two runs was assessed during the run by calculating the average standard deviation of split frequencies. The "burn-in" was specified at the value 1,000 using the "sump" command. All analyses were run at the Cyberinfrastructure for Phylogenetic Research (CIPRES) Portal (http://www.phylo.org/sub_sections/portal; Miller *et al.* 2010)

DNA content estimation and ploidy level assignment

To estimate nuclear DNA contents of the obtained strains, we employed propidium iodide flow cytometry (PI FCM). Approximately two weeks before the planned FCM analyses, cultures were inoculated into fresh medium. For sample preparation, 1 mL of well-grown culture was centrifuged (5,500 rpm for 5 min) and the superfluous medium was removed by pipetting. Consequently, 350 μL of ice-cold nuclei isolation buffer Otto I (0.1 M citric acid, 0.5% Tween 20; Otto 1990) was added to the algal pellet, causing an osmotic rupture of cells and release of the sample nuclei. The resulting suspension was thoroughly shaken and kept on ice. Plants *Solanum pseudocapsicum* (2C = 2.59 pg; Temsch *et al.* 2010) or *Carex acutiformis* (2C = 0.82 pg; Veselý *et al.* 2012) were used as a (pseudo-)internal standard, depending on the sample DNA content. To release nuclei of the standard, ca. 20-mg piece of fresh leaf tissue was chopped with a razor blade in a plastic Petri dish with 250 μL of ice-cold Otto I buffer. Both suspensions (with algal and standard nuclei) were thoroughly mixed and filtered through a 42 μm nylon mesh into a special 3.5-mL cuvette for direct use with the flow cytometer. Following a 20-min. incubation at room temperature, the sample was mixed with 1 mL of staining solution consisting of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12H_2O$; Otto 1990), 50 $\mu g \cdot mL^{-1}$ PI, 50 $\mu g \cdot mL^{-1}$ RNase IIA and 2 $\mu L \cdot mL^{-1}$ β-mercaptoethanol. The stained sample was

immediately analysed using a Partec CyFlow SL cytometer (Partec GmbH, Münster, Germany) equipped with a green solid-state laser (Cobolt Samba, 532 nm, 100 mW). In each sample, 5,000 particles were measured and the resulting FCM histograms were analysed using FloMax ver. 2.4d (Partec). The first sample peak in the FCM histogram was identified as G_1 (vegetative cells) and a second peak with twice the relative fluorescence as G_2 (dividing cells). The absolute nuclear DNA content (C-value) was calculated as sample G_1 peak mean fluorescence / standard G_1 peak mean fluorescence × standard 2C DNA content (according to Doležel 2005). In case of low-quality measurements (i.e. G_1 sample peaks with coefficient of variation (CV) >5%), both sample preparation and analysis were repeated. To minimize the effect of random instrumental shift, each strain was analysed at least three times on separate days and the estimates averaged. Each time the three independent DNA content estimates differed by >3%, the most outlying measurement was discarded and a new measurement was carried out; however, the DNA content results of six strains (968, S20.45, S71.B4, V29, X40, K8) were averaged after five consecutive analyses with DNA content differences >3% (not exceeding 6%). Nuclear DNA content is reported in absolute units per cell (pg of DNA and equivalent values in Gbp). Since the chrysophytes are presumed to be haploids (Sandgren 1991; Olefeld *et al.* 2018), when the DNA content variation within a species corresponded to two-fold differences, we referred to the lowest value as haploid (1x) and to its multiples as diploid (2x) or tetraploid (4x), with full awareness that the base ploidy level still needs to be verified. Despite our previous considerable effort, karyotyping and chromosome counts of various chrysophyte strains were not successful. To additionally corroborate the observed intraspecific ploidy level variation, a simultaneous analysis of three *S. glabra* strains (G11, F45 and L13), each representing a different ploidy level, was performed. Since eight strains indicated ploidy level change during cultivation, if possible, their cultures were repeatedly re-analysed (up to 8-times) within three consecutive years. In one strain of *S. petersenii* (C87), it appeared that three G_1 sample peaks differing in their ploidy level were present. Consequently, individual cells of this strain were inoculated into new subcultures (C87-1 – C87-6) and later (repeatedly) analysed for their DNA content. To avoid misinterpretation of G_2 peaks for a G_1 peak during the ploidy assignment, the presence of both G_1 and G_2 peaks was thoroughly checked and confirmed (usually clearly apparent in the relative fluorescence vs. side scatter plot) in all the analyses performed.

GC content estimation

To reveal a potential association between the nuclear DNA amount and a genome-wide proportion of GC bases, we analysed the genomic GC content of nine

strains belonging to four *Synura* species (*S. americana*, *S. glabra*, *S. macropora*, *S. sphagnicola*). The strains were analysed using FCM with the AT-selective dye DAPI (4',6-diamidino-2-phenylindole) and the results were directly compared with the PI FCM outputs for particular strains. We employed the same sample preparation as for PI FCM, except that the staining solution consisted of 1 mL of Otto II buffer, 4 $\mu g \cdot mL^{-1}$ DAPI and 2 $\mu L \cdot mL^{-1}$ β-mercaptoethanol. The stained samples were immediately analysed using a Partec PA II flow cytometer (Partec GmbH, Münster, Germany) equipped with a 488-nm UV LED as a source of excitation light. In each sample, 5,000 particles were measured and the resulting FCM histograms were analysed using FloMax. Computation of the GC base content was done according to (Šmarda *et al.* 2008) via a publicly available Excel spreadsheet (http://sci.muni.cz/botany/systemgr/download/Festuca/ATGCFlow.xls). Each strain was analysed at least three times on separate days and the final estimate was averaged from the individual measurements.

Cell size measurements

In our search for phenotypic differences between intraspecific strains with different DNA contents we chose one trait – cell size. This trait is particularly important for unicellular organisms and a tight relationship between the cell size and nuclear DNA content has already been demonstrated (i.e. the nucleotypic effect; Bennett 1971). We selected 12 strains belonging to five *Synura* species (*S. americana*, *S. glabra*, *S. macropora*, *S. petersenii,* and S. *sphagnicola*) exhibiting intraspecific ploidy level variation. Before the analyses, 50 μL of each strain at the exponential phase of growth was inoculated into 4 mL of fresh medium and cultivated for 2 weeks. After this period, microphotographs of individual cells were taken using a Leica DM2500 LED optical microscope with 40× magnification. The cell size was later estimated for each strain using ImageJ ver. 1.45s (Schneider *et al.* 2012) as object area on the microphotograph. The final estimates were based on a median value of 30 cells measured per each strain.

RESULTS

Nuclear DNA content variation in chrysophytes

Altogether, this study was performed on 68 strains representing three chrysophyte genera, *Synura*, *Chrysosphaerella*, and *Ochromonas* (Supplementary data Table S1), where we previously detected intraspecific DNA content variation. Using the nuclear ITS rDNA (nu ITS rDNA) molecular barcode, we identified nine species of *Synura*: *S. americana*, *S. glabra*, *S. heteropora*, *S. hibernica*, *S. lanceolata*, *S. macropora*, *S. petersenii*, *S. soroconopea*, and *S. sphagnicola* (Fig. 1). Although the nu ITS rDNA showed minor sequence variations in several strains, this variability was not associated with DNA content differences. We successfully estimated absolute nuclear DNA contents for all strains. Intraspecific DNA content variation largely corresponded to the presence of different ploidy levels and was detected either in among-strain comparisons (Table 1; Fig. 2; for more detailed data, see Supplementary Table S2) or during repeated measurements on a single strain in cultivation (Fig. 3; Fig. 4; for more detailed data, see Supplementary Table S3). In each species, the lowest nuclear DNA content detected (for simplicity arbitrarily assigned as haploid; 1x) also had a corresponding diploid value (2x) with twice the DNA amount. A minor DNA content variation among strains in some species of *Synura* (Table 1, Fig. 1) did not compromise the overall pattern. The only exception was *S. macropora*, where higher DNA contents ranging 3.46 - 3.70 pg were not multiples of the lowest value (1.48 pg).

Table 1. Nuclear DNA content and ploidy level variation among chrysophyte species. Ploidy level assignment is arbitrary, assuming the lowest DNA content category in each species corresponds to a haploid. The predominant ploidy level in each species is marked with an asterisk.

Species		No. of strains	Nuclear DNA content [pg; median ± SD] 1x ploidy level	2x ploidy level	4x ploidy level
Chrysosphaerella brevispina		1	0.2	0.5*	
Ochromonas tuberculata		1	0.4	0.8*	
Synura	*americana*	3	1.1 ± 0.02*	2.2	
	glabra	27	1.0 ± 0.02	2.0 ± 0.09 *	3.8
	heteropora	15	0.8 ± 0.06	1.4 ± 0.07*	3.0
	hibernica	1	1.8*	3.6	
	lanceolata	1	0.7	1.4*	
	macropora	6	1.5	3.6 ± 0.09*	
	petersenii	1	1.0	2.1*	
	soroconopea	1	1.5	3.2*	
	sphagnicola	9	0.4 ± 0.03*	0.8	

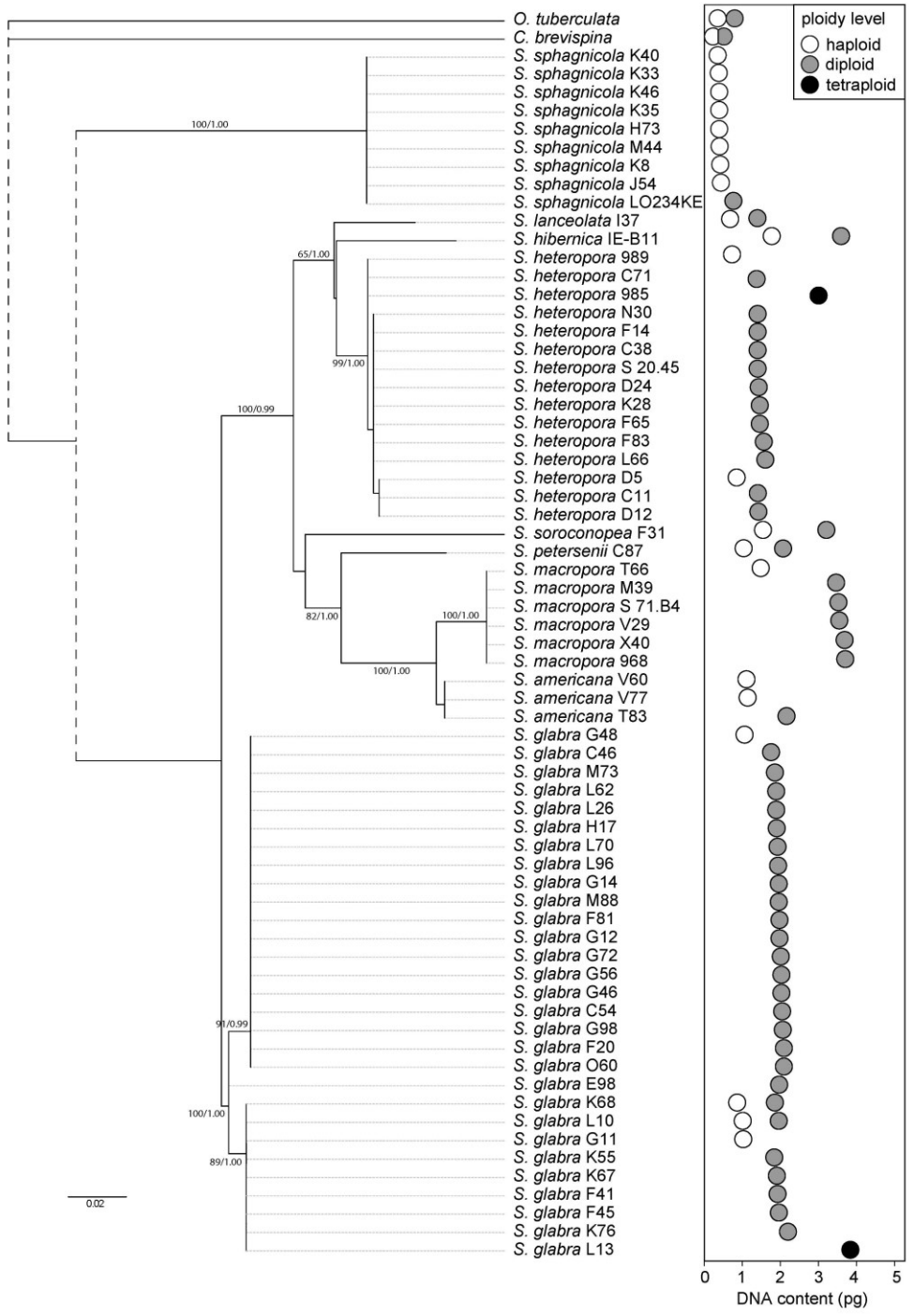


Figure 1. Phylogenetic relationships and DNA contents of investigated chrysophyte strains. Maximum likelihood (ML) phylogeny of nine *Synura* species is based on ITS rDNA sequences. Values at the nodes indicate statistical support estimated by – ML bootstrap (left) and MrBayes posterior node probability (right). Only statistical supports higher than 60/0.95 are shown. Scale

bar – estimated number of substitutions per site. The phylogenetic relationships of the species *Ochromonas tuberculata* and *Chrysosphaerella brevispina* are illustrated by dashed line following phylogenetic analysis in Kristiansen and Škaloud (Kristiansen and Škaloud 2017). The strains possess various DNA contents from 0.23 to 3.83 pg and haploid (white), diploid (grey) or tetraploid (black) ploidy level.

Ploidy level transitions during long-term cultivation were captured in seven strains of six chrysophyte taxa (*Chrysosphaerella* sp., *Ochromonas tuberculata*, *S. glabra*, *S. hibernica*, *S. lanceolata*, and *S. soroconopea*; Fig. 3 and Fig. 4; Supplementary Table S3). These transitions occurred in both directions, from haploid to diploid level and vice versa; multiple transitions suggesting alternation of the two ploidies over the course of time were observed in some strains. Moreover, the presence of three different ploidy levels (1x, 2x, 4x) was detected in two species (*S. glabra*, *S. heteropora*; Fig. 2; Table 1). Interestingly, different ploidy cytotypes in these two species even coexisted at one locality, within the same algal bloom. An initial indication of the presence of a tetraploid cytotype also in *S. petersenii* strain C87 was not corroborated and most likely represented G_2 of the diploid cells. When we re-inoculated individual cells from the original colonies and established six subcultures, they became fixed for either haploid or diploid ploidy level. Interestingly, repeated analyses of *S. petersenii* strain C87 subcultures did not show any further ploidy transitions over the period of two years.

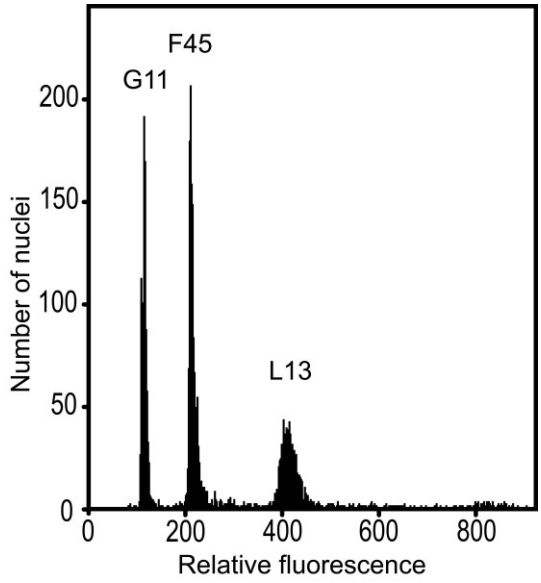


Figure 2. A simultaneous flow cytometric analysis of three strains of *Synura glabra* that differ in their ploidy level (G11 - haploid, F45 - diploid, and L13 - tetraploid). Nuclei were stained with propidium iodide.

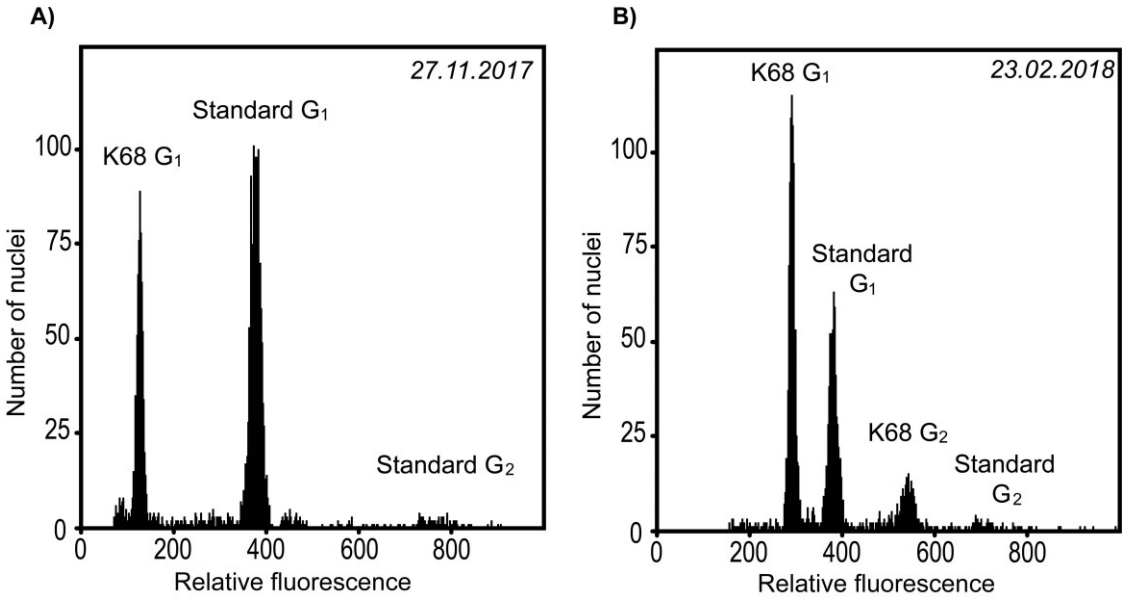


Figure 3. A life cycle stage transition and the inherent ploidy shift from haploid (A) to diploid (B) level documented in *Synura glabra* strain K68 by two consecutive flow cytometric measurements. Plant *Solanum pseudocapsicum* was used as a reference standard.

Genomic GC content

The genomic GC content ranged from 38.0 to 47.7% among nine strains representing different ploidy levels of four *Synura* species (*S. americana*, *S. glabra*, *S. macropora, and S. sphagnicola*). In general, the GC content was highly similar for intraspecific ploidy cytotypes, despite the multiple-fold difference in their DNA content (Table 2). Specifically, the GC content values of strains ranged from 39.6 to 40.8% in *S. americana*, from 38.0 to 38.6% in *S. glabra*, and from 47.3 to 47.7% in *S. sphagnicola*. A marked exception were *S. macropora* strains exhibiting increased GC content variation 38.3 - 42.2%.

Phenotypic consequences of ploidy level variation

Ploidy level increase (i.e., doubling of nuclear DNA content) was associated with greater cell size in the investigated chrysophyte taxa (Table 2). In intraspecific comparisons, shifts from haploid to diploid ploidy level resulted on average in the cell size increase by 34% (range 13 – 55%). This trend was observed both among different strains of the same species and among haploid and diploid subcultures of the same strain (in the case of *S. petersenii* strain C87). Interestingly, except for the change in cell size, the overall phenotypes of different ploidy levels appeared to be the same (see Figs. 5 and 6).

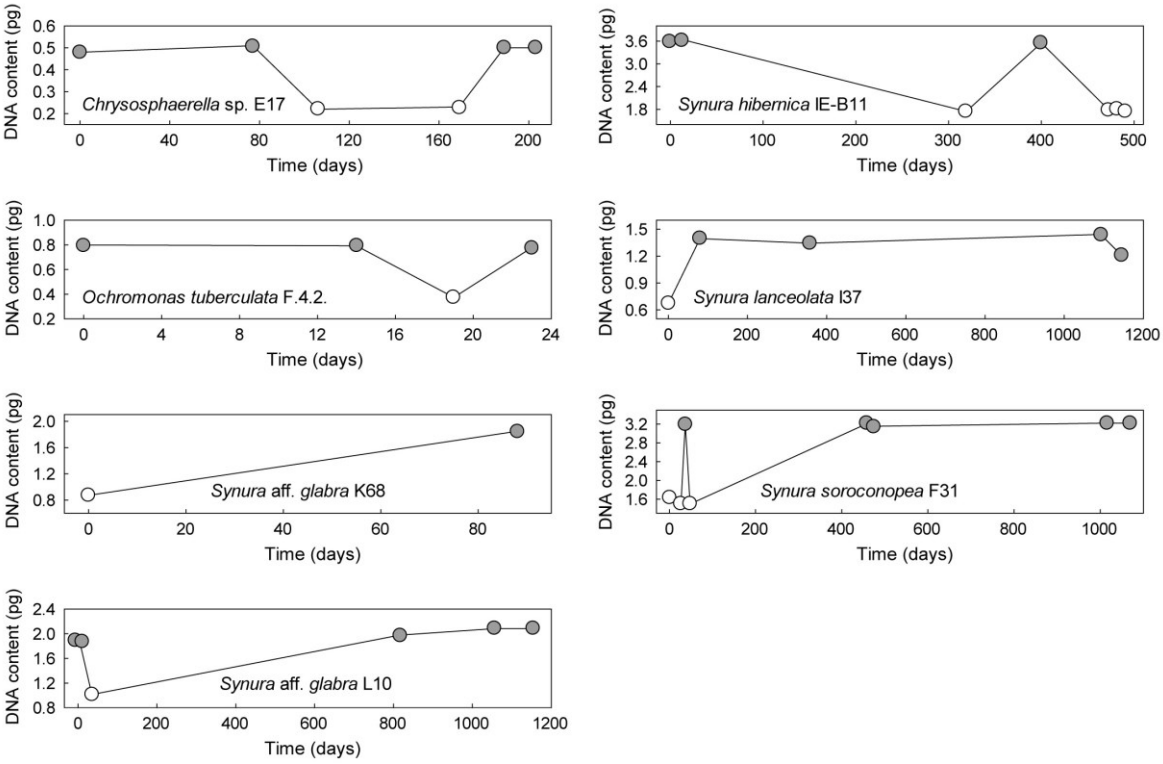


Figure 4. Two-fold differences in nuclear DNA content captured by a time series of flow cytometric measurements on seven strains of chrysophyte taxa alternating between the haploid (white) and diploid (grey) life cycle stage.

Table 2. Genomic GC content and cell size estimates for selected strains of *Synura* exhibiting intraspecific ploidy level variation.

Species	Strain	Ploidy level	DNA content	GC content [mean ± SD]	Cell size [mean ± SD]
S. americana	V77	1x	1.13 pg	39.6 ± 0.1 %	101 ± 23 μm²
	T83	2x	2.16 pg	40.8 ± 0.2 %	114 ± 10 μm²
S. glabra	G11	1x	1.02 pg	38.6 ± 0.3 %	80 ± 12 μm²
	K67	2x	1.91 pg	38.4 ± 0.2 %	124 ± 21 μm²
	L13	4x	3.83 pg	38.0 ± 0.2 %	137 ± 23 μm²
S. macropora	T66	1x	1.48 pg	38.3 ± 0.2 %	139 ± 25 μm²
	S71.B4	2x	3.53 pg	42.2 ± 0.2 %	99 ± 13 μm²
S. petersenii	C87-6	1x	1.08 pg	-	104 ± 11 μm²
	C87-2	2x	2.07 pg	-	139 ± 11 μm²
	C87-5	2x	2.12 pg	-	157 ± 17 μm²
S. sphagnicola	K35	1x	0.39 pg	47.3 ± 0.0 %	112 ± 21 μm²
	LO234KE	2x	0.76 pg	47.7 ± 0.1 %	132 ± 24 μm²

DISCUSSION

Intraspecific ploidy level variation in chrysophytes

With the exception of a single species (*S. macropora*, see below), DNA content variation within and among conspecific strains was not random but corresponded to different ploidy levels. Two to three ploidy levels were detected within the particular chrysophyte species investigated, here for simplicity referred to as haploid (1x), diploid (2x) and tetraploid (4x).

Aside from the (almost) exactly two-fold differences in nuclear DNA contents, the intraspecific ploidy level variation was also supported by the highly similar genomic base composition (GC content), a trait that otherwise spanned rather broadly among the species (38.0 - 47.7%). Moreover, we are convinced that our records of the higher ploidy levels (2x, 4x) are not simply artefacts caused by the occurrence of dividing cells in G_2 phase of the life cycle (i.e., after duplication of genomic DNA but not yet entering mitosis). Before assigning a ploidy, we always checked for the presence of at least a small population of nuclei with twice the fluorescence intensity (~ DNA content) than our peaks of interest in flow cytometric histograms, that correspond to G_2-phase cells (Sliwinska *et al.* 2021). Key insights into the evolutionary processes responsible for the observed DNA content variation were provided by long-term cultivation of strains, when we detected alternation between the lower and higher ploidy states over time (Fig. 4). This suggests that ploidy level shifts are an inherent part of the chrysophytes' life cycles. Additionally, our records of three different ploidy levels in two species of *Synura* (*S. glabra*, *S. heteropora*) suggest that part of the DNA content diversity is also contributed by polyploidization.

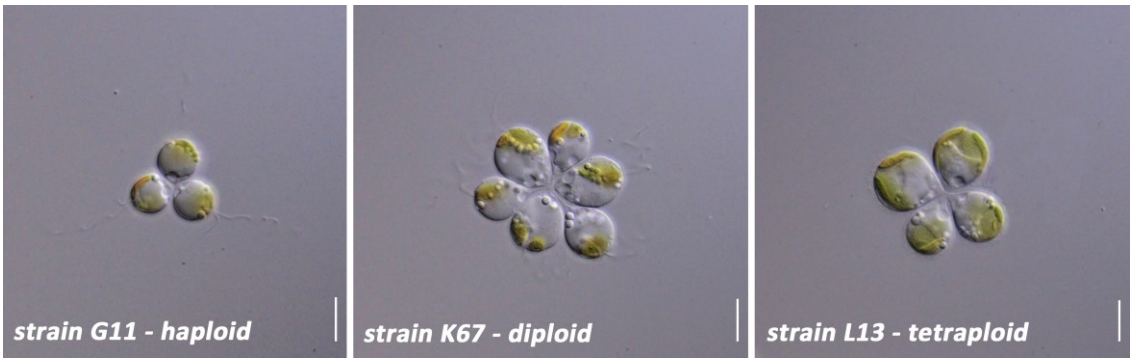


Figure 5. Phenotypes of haploid, diploid and tetraploid strains of *Synura glabra*, mainly differing by the size of cells in their colonies. Scales = 10 μm.

While most (minor) deviations from the two-fold differences in nuclear DNA contents can be attributed to the error of measurement or to the higher content of secondary metabolites interfering with DNA staining, some might have arisen from genetic differentiation of strains during evolution. A prominent case of the latter are the strains of *S. macropora* with the DNA contents of either 1.5 pg or 3.6 pg, which also have distinct genomic GC contents (Table 2).

These strains could represent different ploidy levels of distinct lineages (cryptic taxa) within *S. macropora*, though their evolutionary differentiation must have occurred relatively recently given that the nu ITS rDNA region has not yet diversified (see Fig. 1). Similarly, pronounced shifts in nuclear DNA content not followed by the nu ITS rDNA diversification were already described among *S. petersenii* strains (Čertnerová and Škaloud 2020).

Life cycle of chrysophytes

Even though the chrysophytes with over 1,100 described species (Guiry and Guiry 2021) constitute a relatively large group of microscopic algae, the knowledge of their life cycles is very fragmentary. There are few instances in the older literature describing the formation of gametes and their fusion (Wawrik 1972; Sandgren 1981; Sandgren and Flanagin 1986), but neither the sexual reproduction in chrysophytes, nor other key aspects of their life cycle have been since then a subject of targeted study. Our flow cytometric data has clearly shown that the life cycle of chrysophytes involves alternation between a lower and a higher ploidy state. As a consequence, ploidy level shifts may occur within a single strain maintained in cultivation, and independently collected strains of a single species may exhibit ploidy level variation. Moreover, based on repeated DNA content measurements conducted on strains under long-term cultivation, we are convinced that the lower ploidy state is not restricted to short-lived gametes and the higher ploidy state is not equivalent to a zygote. The chrysophyte cells at either ploidy state are able of mitotic propagation which allows them longer-term persistence in cultivation. Our findings are thus consistent with chrysophytes having a haploid-diploid life cycle. This life cycle was already proposed for several groups of unicellular algae, namely for haptophytes, foraminifera, some dinoflagellates and some cryptophytes (Rousseau *et al.* 2007; Speijer *et al.* 2015; Figueroa *et al.* 2018).

As a general rule, different life cycle stages in these groups are morphologically well distinguishable. For example, in the haptophyte species *Emiliania huxleyi*, diploids are covered by prominent calcified scales and lack flagella, while flagellated haploids are covered by inconspicuous organic scales (von Dassow *et*

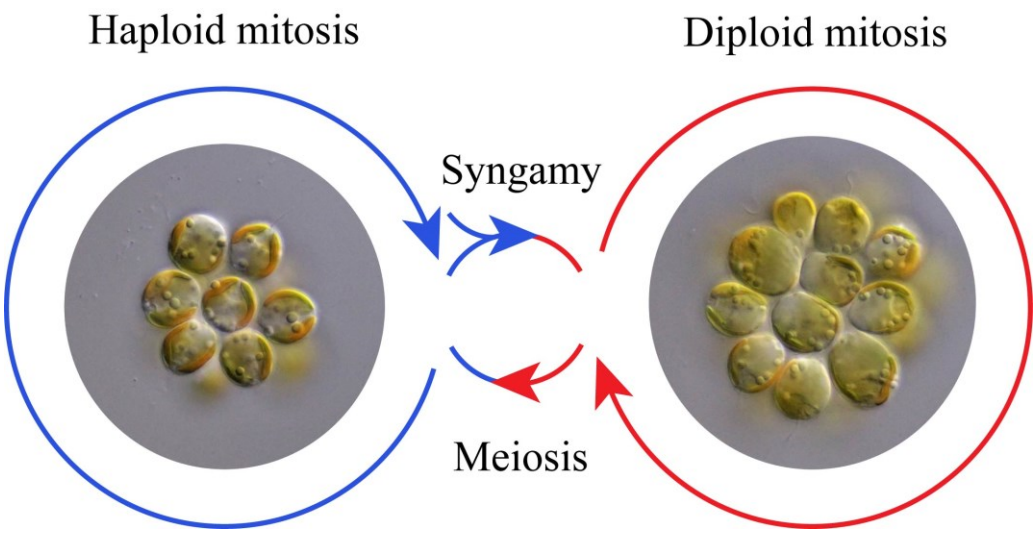


Figure 6. An illustrative scheme of the isomorphic haploid-diploid life cycle of chrysophytes. Both haploid and diploid stages are capable of vegetative reproduction and have a more-or-less uniform phenotype (shown here on example of *S. petersenii* strains C87-6 and C87-2). Syngamy refers to the fusion of haploid gametes.

al. 2009). In another haptophyte species, *Phaeocystis globosa*, haploids and diploids differ by the presence of organic scales (Rousseau *et al.* 2007). Different life cycle stages of the cryptophyte *Cryptomonas* were even previously described as two separate genera, *Cryptomonas* and *Campylomonas* (Hoef-Emden and Melkonian 2003). In chrysophytes, this does not seem to be the case and different life cycle stages are morphologically indistinguishable. Apart from a small increase in cell size with the higher ploidy level, we have not observed any apparent phenotypic differences between the two stages. Hence, the chrysophytes have an isomorphic haploid-diploid life cycle (see Fig. 6 for a scheme). Interestingly, it is the first record of this life strategy among unicellular algae. There are two consequences that can be derived from the life cycle of chrysophytes. First, even though capable of long-term persistence via mitotic propagation, the haploids must also play the role of gametes. Second, the observed ploidy level shifts indicate that diploid strains regularly undergo meiosis in cultivation. In spite of the latter, we never directly observed the sexual reproduction nor have identified preferred conditions under which it occurs.

Questions arise as to how often life stage transitions occur, what triggers them, and whether chrysophyte species spend more time as haploids or diploids. Our results suggest that, at least *ex situ*, under cultivation, life cycle transitions may occur very rapidly. For example, in *Ochromonas tuberculata* and *Synura soroconopea*, two ploidy level shifts were observed within two weeks. On the other hand, most other

species seemed to alternate in ploidy considerably more slowly (Fig. 4). It should be noted though, that intervals in our time series of flow cytometric measurements were highly irregular as ploidy level variation within strains was usually first detected only by coincidence. Consequently, it is very likely that some life stage transitions have remained undetected in our long-term cultivation experiment. Also, the strains and taxa used in this study were specifically selected from a much broader collection based on the intraspecific DNA content variation that was detected during our ongoing research. It seems the probability of detecting the life stage transitions is rather low, though it does not necessarily mean such transitions are rare in natural populations of chrysophytes.

Interestingly, since each flow cytometric measurement was conducted on thousands of cells, our results also prove that all cells in each culture had synchronous life cycles. The only exception was an initial coexistence of haploids and diploids in *S. petersenii* strain C87, but soon after their re-inoculation, the life cycles synchronized and six resulting subcultures became fixed for either the haploid or the diploid stage. Possibly, the first measurement might have caught the cells in the middle of a life stage transition. It can be expected that such synchronous life stage transitions are also applicable to individuals within natural populations. In general, chemical signalization among cells could be associated with the life stage transitions in chrysophytes, ensuring they enter particular life cycle stages or produce their gametes synchronously. Production of pheromones, allowing synchronization of individuals prior to the mating process was already described in green algae (e.g. *Volvox carteri*; Al-Hasani and Jaenicke 1992). Similarly, chemical signalling is involved in sexual induction of some diatoms (Moeys *et al.* 2016). When we summarize our DNA content measurements across the studied chrysophyte species (Table 1, Fig. 4), it seems that the diploid life cycle stage predominates in most, though there are also exceptions where the haploids are more common (i.e., *Synura americana*, *S. hibernica* and *S. sphagnicola*). This suggests that the relative duration of these two stages in the life cycles of chrysophytes could differ from taxa to taxa, possibly reflecting their environmental preferences and/or population genetic processes.

Evolutionary benefits of a haploid-diploid life cycle

It is widely believed that the differentiation of life cycle on two alternating ploidy phases comes with many evolutionary advantages. From a genetic point of view, the haploid stage allows more efficient selection of beneficial alleles and immediate elimination of deleterious mutations, whereas the diploid stage has the ability to accumulate mutations at a higher rate and mask deleterious mutations, resulting in increased genetic diversity (Lewis and Wolpert 1979; Otto and Gerstein 2008). From

an ecological perspective, the existence of two or more life cycle stages may provide for more efficient specialization, better exploitation of the environment, and thus increasing the evolutionary success of species (Thornber 2006). The evolution and maintenance of haploid-diploid life cycles was nicely demonstrated by Hughes and Otto (Hughes and Otto 1999). Using a genetic model parametrized with demographic data of red algae *Gracilaria gracilis*, the authors conclude that haploid-diploid life cycles may be evolutionarily stable as long as resource competition between haploids and diploids is sufficiently weak. As an example from microscopic algae, the populations of haptophyte *Emiliania huxleyi* often suffer from infections by phycodnaviruses. However, the transition from heavily calcified susceptible diploids to resistant haploids lacking calcified covering serves as an escape mechanism (Frada *et al.* 2008). Interestingly, even subtle phenotypic differences between the two stages in isomorphic species may have significant ecological consequences (Destombe *et al.* 1993; Dyck and DeWreede 1995; Hughes and Otto 1999).

The only apparent difference we have observed between the phenotypes of conspecific haploids and diploids was a small increase in cell size. A positive association between the ploidy level and cell size was consistently observed in all chrysophyte species (see Table 2), in spite of a considerable variation in cell shapes due to the lack of a cell wall (Fig. 5). Increase in cell size is actually the most common phenotypic effect of higher ploidy and it is usually considered to be a direct consequence of the doubling of nuclear DNA content (Bennett 1971; Hughes and Otto 1999). Particularly in unicellular organisms, the cell size is one of the most important traits because it fundamentally relates to metabolic rate, growth rate or generation time, but can also affect species temperature optima, dispersal abilities, or susceptibility to herbivores (Van't Hof and Sparrow 1963; Shuter *et al.* 1983; Cavalier-Smith 2005). As a consequence of their geometry, larger cells have lower surface area to volume ratio than small ones of identical shape. Thus, smaller haploid cells with a relatively higher surface area should be more efficient in nutrient uptake and at the same time, they may require lower energetic costs to maintain their life functions and have faster cell division (Lewis 1985; Hughes and Otto 1999). The "nutrient limitation hypothesis" by Lewis (Lewis 1985) then predicts that haploids will better exploit low-nutrient environments, and this should be especially applicable to unicellular planktonic autotrophs (e.g., chrysophytes). On the other hand, diploids might be preadapted to better tolerate toxic environments due to their relatively smaller surface area interacting with the external environment (Otto and Gerstein 2008).

As was already mentioned above, even though chrysophytes can persist in either haploid or diploid stage in cultivation, the diploids seem to prevail among strains

and taxa. We can only speculate that nutrient-rich environments, supplemented in our study with cultivation medium, may favour the diploid life cycle stage in line with Lewis's hypothesis (Lewis 1985). Since both stages are part of a one life cycle and principally should alternate in populations, any putative associations between the occurrence of haploids / diploids and environmental conditions will likely reflect temporal (e.g. seasonal) changes in the suitability of particular habitats for chrysophyte taxa. From this perspective, it would be interesting to observe life stage transitions in chrysophyte cultures being induced by the input of a fresh medium (although our preliminary observations do not support it, data not shown). On the other hand, the cultures reaching a certain "population density" could also serve as a trigger, and a particularly important one in the terms of potentially forecasting a bloom formation.

We believe that due to their isomorphic haploid-diploid life cycle, unicellular chrysophytes may serve as a unique model to answer various questions related to the evolutionary importance of ploidy level variation among unicellular algae, and the potential benefits and costs of having different ploidy phases in a life cycle or maintaining diploid and polyploid cytotypes. These will be the subject of our further investigation.

DNA contents and polyploidy in chrysophytes

The available data shows that there is a considerable variation in nuclear DNA contents among chrysophytes, equalling to a 276-fold difference between *Segregatospumella dracosaxi* with the smallest reported DNA content and *Mallomonas caudata* with the largest DNA content (0.09 pg and 24.85 pg, respectively; Veldhuis *et al.* 1997; Olefeld *et al.* 2018; Čertnerová and Škaloud 2020). While the nuclear DNA content variation might have played an important role in the chrysophyte evolution, DNA content estimates are still available for only a small fraction of the overall species richness in Chrysophyceae (32 species ~ 3%). In addition to the insights into the chrysophytes' life cycles, this study provides a large number of DNA content estimates. Here, we provide the first DNA content estimates for the genus *Chrysosphaerella* and the species *Ochromonas tuberculata*, *Synura glabra*, *S. hibernica*, *S. lanceolata*, *S. macropora*, and *S. soroconopea*. Our estimates for the species *S. petersenii* (1.04 / 2.09 pg) fall well within the previously described wide range of intraspecific DNA contents (Čertnerová and Škaloud 2020). The previously published estimates for two strains of *S. heteropora* (S 20.45, WA18K-A) are 1.51 pg and 1.55 pg, respectively (Olefeld *et al.* 2018), which is in agreement with our estimates for 12 strains (mean = 1.44 pg, range = 1.38-1.61 pg) and also confirms the overall prevalence of the diploid life cycle stage. Similarly, with the exception of a single strain (L0234KE), our DNA content estimates for eight *S. sphagnicola* strains

corroborate the previously published data (0.40 pg; Olefeld *et al.* 2018). Note that Olefeld *et al.* (2018) considered that chrysophytes are presumably haploid, but when providing a 1C value, the authors divided the estimated DNA contents by two as if they were diploids. On the other hand, the DNA content of *S. sphagnicola* strain L0234KE, introduced as a flow cytometric (FCM) standard in Olefeld *et al.* (2018), was referred to as 0.40 pg, while our estimate for the same strain is 0.76 pg. This inconsistency most likely results from a life stage transition of the strain between the two measurements. Even though the best practice is to apply closely related species as FCM standards (Temsch *et al.* 2021), our findings show that the use of chrysophyte FCM standards should be highly discouraged due to their isomorphic haploid-diploid life cycle and the resulting DNA content instability.

In addition to the haploid – diploid life stage transitions, the presence of a third, higher ploidy detected in two chrysophyte species (*S. glabra* and *S. heteropora*) suggests the involvement of polyploidy. The tetraploid cytotype was quite rare, comprising just 3% of the investigated strains. Both the precisely two-fold DNA content difference from respective diploids and ITS rDNA homogeneity favour their autopolyploid origin (i.e., intraspecific polyploidy). At one locality of *S. heteropora*, the tetraploids even coexisted with haploids in a mixed-ploidy population (strains no. 985 and 989; Vltava river, Prague, Czech Republic). A similar coexistence of *S. glabra* diploids and tetraploids in Inari lake (strains K76 and L13; Inari, Finland) could either represent the lower- and higher-ploidy life stages of a single clone, or belong to two different clones, both at the higher-ploidy stage. We can hypothesize that such coexistence of different ploidy cytotypes in a population can be maintained by their prevalent asexual reproduction via mitotic division. The polyploidy was recently observed in another chrysophyte, *Poteriospumella lacustris* (Majda *et al.* 2019), and is likely a recurrent phenomenon in chrysophytes.

CONCLUSIONS

In this study, we revealed that chrysophytes have a haploid-diploid life cycle. Chrysophyte taxa alternate between two ploidy states, both of which are capable of mitotic propagation and long-term survival in cultivation. The two life cycle stages are morphologically undistinguishable, apart from a small increase in cell size with the higher ploidy level. This is the first report of an isomorphic haploid-diploid life cycle among unicellular algae. Interestingly, our flow cytometric measurements also revealed that life cycles are well synchronized among cells coexisting within a culture and the same can be expected for natural populations of chrysophytes. The relative duration of the life cycle stages may differ from taxa to taxa. While the

diploid stage prevailed in most of our studied chrysophyte representatives, in others the haploid stage seemed to dominate. Collectively, our results provide new unique insights into the chrysophyte life cycle. The chrysophytes may serve as a suitable model for studying the benefits and costs of having different ploidy stages in a life cycle.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online at
https://doi.org/10.1016/j.algal.2022.102707

FUNDING

This work was supported by the Charles University [GAUK, project no. 1304317]. Additional support was provided by the Czech Academy of Sciences (long-term research development project no. RVO 67985939) and by the Charles University Research Centre program No. 204069.

ACKNOWLEDGEMENTS

We are indebted to Martin Pusztai, Helena Bestová and Iva Jadrná (Department of Botany, Charles University) for their help with sample collection. We also thank to Yvonne Němcová (Department of Botany, Charles University) and Jens Boenigk (Department of Biodiversity, University of Duisburg-Essen) for providing us with the cultures of *Ochromonas tuberculata* and *S. sphagnicola* (LO234KE), respectively.

LITERATURE CITED

Al-Hasani H, Jaenicke L. 1992. haracterization of the sex-inducer glycoprotein of *Volvox carteri* f. *weismannia*. *Sexual Plant Reproduction* **5**: 8–12.

Andersen RA, Morton SL, Sexton JP. 1997. Provasoli-Guillard National Center for culture of marine phytolankton 1997 list of strains. *Journal of Phycology* **33**: 1–75.

Bennett MD. 1971. The duration of meiosis. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **178**: 277–299.

Beukeboom LW, Nicolas Perrin. 2014. *The evolution of sex determination*. Oxford University Press, USA.

Cavalier-Smith T. 2005. Economy, Speed and Size Matter: Evolutionary Forces Driving Nuclear Genome Miniaturization and Expansion. *Annals of Botany* **95**: 147–175.

Čertnerová D, Galbraith DW. 2021. Best practices in the flow cytometry of microalgae. *Cytometry Part A* **99**: 359–364.

Čertnerová D, Škaloud P. 2020. Substantial intraspecific genome size variation in golden-brown algae and its phenotypic consequences. *Annals of Botany* **126**: 1077–1087.

von Dassow P, Ogata H, Probert I, *et al.* 2009. Transcriptome analysis of functional differentiation between haploid and diploid cells of *Emiliania huxleyi*, a globally significant photosynthetic calcifying cell. *Genome Biology* **10**.

Destombe C, Godin J, Nocher M, Richerd S, Valero M. 1993. Differences in response between haploid and diploid isomorphic phases of *Gracilaria verrucosa* (Rhodophyta: Gigartinales) exposed to artificial environmental conditions. *Hydrobiologia* **260/261**: 131–137.

Destombe C, Valero M, Vernet P, Couvet D. 1989. What controls haploid—diploid ratio in the red alga, *Gracilaria verrucosa*? *Journal of Evolutionary Biology* **2**: 317–338.

Doležel J. **2005**. Plant DNA Flow Cytometry and Estimation of Nuclear Genome Size. *Annals of Botany* **95**: 99–110.

Dyck LJ, DeWreede RE. **1995**. Patterns of seasonal demographic change in the alternate isomorphic stages of. *Phycologia* **34**: 390–395.

Figueroa RI, Estrada M, Garcés E. **2018**. Life histories of microalgal species causing harmful blooms: Haploids, diploids and the relevance of benthic stages. *Harmful Algae* **73**: 44–57.

Figueroa RI, Rengefors K. **2006**. Life cycle and sexuality of the freshwater raphidophyte *Gonyostomum semen* (Raphidophyceae). *Journal of Phycology* **42**: 859–871.

Fischer G, Liti G, Llorente B. **2021**. The budding yeast life cycle: More complex than anticipated? *Yeast* **38**: 5–11.

Fott B. **1959**. Zur Frage der Sexualität bei den Chrysomonaden. *Nova Hedwigia* **1**: 115–129.

Frada M, Probert I, Allen MJ, Wilson WH, De Vargas C. **2008**. The "Cheshire Cat" escape strategy of the coccolithophore *Emiliania huxleyi* in response to viral infection. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 15944–15949.

Guillard RRL, Lorenzen CJ. **1972**. Yellow-green algae with chlorophyllide C. *Journal of Phycology* **8**: 10–14.

Guiry MD, Guiry GM. **2021**. *AlgaeBase. World-wide electronic publication, National University of Ireland, Galway.*

Hoef-Emden K, Melkonian M. **2003**. Revision of the Genus *Cryptomonas* (Cryptophyceae): A Combination of Molecular Phylogeny and Morphology Provides Insights into a Long-Hidden Dimorphism. *Protist* **154**: 371–409.

Hughes JS, Otto SP. **1999**. Ecology and the evolution of biphasic life cycles. *American Naturalist* **154**: 306–320.

Kim, Jin H, Shin, Mi O, Lee, Kyung L, Kim, Han S. **2008**. Effect of environmental conditions on the growth of *Synura petersenii* (Synurophyceae) in vitro and two eutrophic water bodies in Korea. *Nova Hedwigia* **86**: 529–544.

Korshikov AA. **1929**. Studies on the Chrysomonads I In: *Archiv für Protistenkunde*.253–290.

Kristiansen J, Škaloud P. **2017**. Chrysophyta In: Archibald JM, Simpson AGB, Slamovits CH, eds. *Handbook of the Protists*. Springer International Publishing, 1657.

Lewis WMJ. **1985**. Nutrient Scarcity as an Evolutionary Cause of Haploidy. *The American Naturalist* **125**: 692–701.

Lewis J, Wolpert L. **1979**. Diploidy, evolution and sex. *Journal of Theoretical Biology* **78**: 425–438.

Mable BK, Otto SP. **1998**. The evolution of life cycles. : 453–462.

Majda S, Beisser D, Boenigk J. **2021**. Nutrient-driven genome evolution revealed by comparative genomics of chrysomonad flagellates. *Communications Biology* **4**: 1–11.

Majda S, Boenigk J, Beisser D. **2019**. Intraspecific variation in protists: clues for microevolution from *Poteriospumella lacustris* (Chrysophyceae) (J Archibald, Ed.). *Genome Biology and Evolution*.

Miller MA, Pfeiffer W, Pfeiffer W. **2010**. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Gateway Computing Environments Workshop*.1–8.

Moeys S, Frenkel J, Lembke C, *et al.* 2016. A sex-inducing pheromone triggers cell cycle arrest and mate attraction in the diatom *Seminavis robusta*. *Scientific Reports* **6**: 1–13.

Montresor M, Vitale L, D'Alelio D, Ferrante MI. **2016**. Sex in marine planktonic diatoms: insights and challenges. *Perspectives in Phycology* **3**: 61–75.

Nicholls KH, Gerrath JF. **1985**. The taxonomy of *Synura* (Chrysophyceae) in Ontario with special reference to taste and odour in water supplies. *Canadian Journal of Botany* **63**: 1482–1493.

Olefeld JL, Majda S, Albach DC, Marks S, Boenigk J. **2018**. Genome size of chrysophytes varies with cell size and nutritional mode. *Organisms Diversity & Evolution* **18**: 163–173.

Otto F. **1990**. DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. *Methods in Cell Biology* **33**: 105–110.

Otto SP, Gerstein AC. **2008**. The evolution of haploidy and diploidy. *Current Biology* **18**: R1121–R1124.

Otto SP, Rosales A. **2020**. Theory in service of narratives in evolution and ecology. *American Naturalist* **195**: 290–299.

Pirrello J, Deluche C, Frangne N, *et al.* 2018. Transcriptome profiling of sorted endoreduplicated nuclei from tomato fruits: how the global shift in expression ascribed to DNA ploidy influences RNA-Seq data normalization and interpretation. *Plant Journal* **93**: 387–398.

Richerd S, Couvet D, Valéro M. **1993**. Evolution of the alternation of haploid and diploid phases in life cycles. II. Maintenance of the haplo-diplontic cycle. *Journal of Evolutionary Biology* **6**: 263–280.

Ronquist F, Teslenko M, van der Mark P, *et al.* 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology* **61**: 539–42.

Rousseau V, Chrétiennot-Dinet MJ, Jacobsen A, Verity P, Whipple S. **2007**. The life cycle of *Phaeocystis*: State of knowledge and presumptive role in ecology. *Biogeochemistry* **83**: 29–47.

Sandgren CD. **1981**. Characteristics of sexual and asexual resting cyst (statospore) formation in *Dinobryon cylindricum* Imhof (Chrysophyta). *Journal of Phycology* **17**: 199–210.

Sandgren CD. **1991**. Chrysophyte reproduction and resting cysts: a paleolimnologist´s primer. *Journal of Paleolimnology* **5**: 1–9.

Sandgren CD, Flanagin J. **1986**. Heterothallic sexuality and density dependent encystment in the Chrysophycean alga *Synura petersenii* Korsh. *Journal of Phycology* **22**: 206–216.

Schneider CA, Rasband WS, Eliceiri KW. **2012**. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675.

Shuter BJ, Thomas JE, Taylor WD, Zimmerman AM. **1983**. Phenotypic Correlates of Genomic DNA Content in Unicellular Eukaryotes and Other Cells. *The American Naturalist* **122**: 26–44.

Škaloud P, Kynčlová A, Benada O, Kofroňová O, Škaloudová M. **2012**. Toward a revision of the genus *Synura*, section Petersenianae (Synurophyceae, Heterokontophyta): morphological characterization of six pseudo-cryptic species. *Phycologia* **51**: 303–329.

Škaloud P, Škaloudová M, Doskočilová P, Kim JI, Shin W, Dvořák P. **2019**. Speciation in protists: Spatial and ecological divergence processes cause rapid species diversification in a freshwater chrysophyte. *Molecular Ecology* **28**: 1084–1095.

Škaloud P, Škaloudová M, Jadrná I, *et al.* 2020. Comparing Morphological and Molecular Estimates of Species Diversity in the Freshwater Genus *Synura* (Stramenopiles): A Model for Understanding Diversity of Eukaryotic Microorganisms. *Journal of Phycology* **56**: 574–591.

Škaloud P, Škaloudová M, Procházková A, Němcová Y. **2014**. Morphological delineation and distribution patterns of four newly described species within the *Synura petersenii* species complex (Chrysophyceae, Stramenopiles). *European Journal of Phycology* **49**: 213–229.

Sliwinska E, Loureiro J, Leitch IJ, *et al.* 2021. Application-based guidelines for best practices in plant flow cytometry. *Cytometry Part A*: 1–33.

Šmarda P, Bureš P, Horová L, Foggi B, Rossi G. **2008**. Genome Size and GC Content Evolution of *Festuca*: Ancestral Expansion and Subsequent Reduction. *Annals of Botany* **101**: 421–433.

Speijer D, Lukeš J, Eliáš M. **2015**. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 8827–8834.

Stamatakis A. **2014**. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

De Storme N, Mason A. **2014**. Plant speciation through chromosome instability and ploidy change: Cellular mechanisms, molecular factors and evolutionary relevance. *Current Plant Biology* **1**: 10–33.

Temsch EM, Greilhuber J, Krisai R. **2010**. Genome size in liverworts. *PRESLIA* **82**: 63–80.

Temsch EM, Koutecký P, Urfus T, Šmarda P, Doležel J. **2021**. Reference standards for flow

cytometric estimation of absolute nuclear DNA content in plants. *Cytometry Part A*: 1–15.

Thornber CS. **2006**. Functional properties of the isomorphic biphasic algal life cycle. *Integrative and Comparative Biology* **46**: 605–614.

Van't Hof J, Sparrow AH. **1963**. A relationship between DNA content, nuclear volume, and minimum mitotic cycle time. *Proceedings of the National Academy of Sciences of the USA* **49**: 897–902.

Veldhuis MJW, Cucci TL, Sieracki ME. **1997**. Cellular Dna Content of Marine Phytoplankton Using Two New Fluorochromes: Taxonomic and Ecological Implications. *Journal of Phycology* **33**: 527–541.

Veselý P, Bureš P, Šmarda P, Pavlíček T. **2012**. Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.

Wawrik F. **1972**. Isogame Hologamie in der Gattung *Mallomonas* Perty. *Nova Hedwigia* **23**: 353–362.

Wee JL, Fasone LD, Sattler A, Starks WW, Hurley DL. **2001**. ITS/5.8S DNA sequence variation in 15 isolates of *Synura petersenii* Korshikov (Synurophyceae). *Nova Hedwigia* **122**: 245–258.

White TJ. **1990**. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics In: Innis MA, ed. *PCR Protocols: A Guide to Methods and Applications*. San Diego: Academic Press, 315–322.

Wichard T, Charrier B, Mineur F, Bothwell JH, De Clerck O, Coates JC. **2015**. The green seaweed *Ulva*: A model system to study morphogenesis. *Frontiers in Plant Science* **6**: 1–8.

Table S1. Collection details for chrysophyte strains used in this study and GenBank accession numbers for nuclear ITS sequences.

Species	Strain	Collection site	GPS coordinates	Sampling date	nr ITS GenBank accession number*
Chrysosphaerella brevispina	E17	Unnamed pond next to Beaulieu Road, Hampshire, United Kingdom	50°50'20.6"N, 1°25'35.8"W	19.03.2017	-
Ochromonas tuberculata	F.4.2.	Břehyňský stream, Doksy, Czech Republic	50°34'34.1"N, 14°41'33.5"E	07.07.2006	-
Synura S. americana	T83	Unnamed lake, Surgutsky District, Russia	61°08'48.3"N, 73°38'36.6"E	28.05.2018	**OL803930**
S. americana	V60	Unnamed lake, Khanty-Mansiysky District, Russia	61°02'05.0"N, 67°56'35.5"E	01.06.2018	**OL803931**
S. americana	V77	Unnamed lake, Khanty-Mansiysky District, Russia	60°56'55.5"N, 68°45'07.4"E	01.06.2018	**OL803932**
S. glabra	C46	Češík pond, Hradec Králové, Czech Republic	50°11'14.9"N, 15°51'55.6"E	26.11.2016	**OL803933**
S. glabra	C54	Ballinamore canal, Co. Leitrim, Ireland	54°00'40.8"N, 8°00'08.5"W	03.02.2017	**OL803934**
S. glabra	E98	Erken lake, Norrtälje, Sweden	59°50'39.9"N, 18°34'50.3"E	13.03.2017	**OL803935**
S. glabra	F20	Kranksjon lake, Lund, Sweden	55°41'31.0"N, 13°29'34.1"E	27.03.2017	**OL803936**
S. glabra	F41	Vederslövssjön lake, Växjö, Sweden	56°47'04.5"N, 14°44'15.9"E	01.04.2017	**OL803937**
S. glabra	F45	Vederslövssjön lake, Växjö, Sweden	56°47'04.5"N, 14°44'15.9"E	01.04.2017	**OL803938**
S. glabra	F81	Häckebergasjön lake, Genarp, Sweden	55°34'45.5"N, 13°25'54.9"E	03.04.2017	**OL803939**
S. glabra	G11	Vederslövssjön lake, Växjö, Sweden	56°47'04.5"N, 14°44'15.9"E	01.04.2017	**OL803940**

S. glabra	G12	Mosseelva lake, Moss, Norway	59°26'10.3"N, 10°40'59.0"E	28.04.2017	**OL803941**
S. glabra	G14	Årsumsvannet lake, Larvik, Norway	59°09'24.6"N, 10°02'50.5"E	28.04.2017	**OL803942**
S. glabra	G46	Nautjønna lake, Sandefjord, Norway	59°09'58.3"N, 10°05'42.0"E	28.04.2017	**OL803943**
S. glabra	G48	Eikeren lake, Øvre Eiker Municipality, Norway	59°42'13.2"N, 9°49'47.4"E	28.04.2017	**OL803944**
S. glabra	G56	Eikeren lake, Øvre Eiker Municipality, Norway	59°42'13.2"N, 9°49'47.4"E	28.04.2017	**OL803945**
S. glabra	G72	Goksjø lake, Sandefjord, Norway	59°10'24.3"N, 10°08'07.2"E	28.04.2017	**OL803946**
S. glabra	G98	Gjennestadvatnet lake, Sandefjord, Norway	59°14'06.8"N, 10°14'26.8"E	28.04.2017	**OL803947**
S. glabra	H17	Åshildrødtjernet lake, Sandefjord, Norway	59°10'32.4"N, 10°07'08.6"E	28.04.2017	**OL803948**
S. glabra	K55	Ore Mountains, Czech Republic	NA	28.09.2017	**OL803949**
S. glabra	K67	Ore Mountains, Czech Republic	NA	28.09.2017	**OL803950**
S. glabra	K68	Ore Mountains, Czech Republic	NA	28.09.2017	**OL803951**
S. glabra	K76	Inari lake, Nellim, Finland	68°50'54.7"N, 28°19'15.5"E	02.10.2017	**OL803952**
S. glabra	L10	Inari lake, Nellim, Finland	68°50'54.7"N, 28°19'15.5"E	02.10.2017	**OL803953**
S. glabra	L13	Inari lake, Nellim, Finland	68°50'54.7"N, 28°19'15.5"E	02.10.2017	**OL803954**
S. glabra	L26	Mioriţa pond, Brašov, Romania	45°35'33.9"N, 25°33'02.3"E	15.11.2017	**OL803955**
S. glabra	L62	Unnamed pond, Botevgrad, Bulgaria	42°54'20.8"N, 23°48'55.6"E	17.11.2017	**OL803956**
S. glabra	L70	Limni Amvrakia lake, Fities, Greece	38°43'36.0"N, 21°11'51.6"E	19.11.2017	**OL803957**
S. glabra	L96	Limni Lisimachia lake, Aggelokastro, Greece	38°33'07.9"N, 21°22'48.3"E	19.11.2017	**OL803958**

S. glabra	M73	Wiel van Collee pond, Leerdam, Netherlands	51°53'25.4"N, 5°06'14.0"E	08.11.2017	**OL803959**
S. glabra	M88	Merwedekanaal, Arkel, Netherlands	51°52'17.2"N, 4°59'50.6"E	08.11.2017	**OL803960**
S. glabra	O60	Unnamed pond, Sandiás, Spain	42°04'48.4"N, 7°46'17.4"W	07.02.2018	**OL803961**
S. heteropora	985	Vltava river, Prague, Czech Republic	50°00'08.3"N, 14°24'09.4"E	30.11.2015	**OL803962**
S. heteropora	989	Vltava river, Prague, Czech Republic	50°00'08.3"N, 14°24'09.4"E	30.11.2015	**OL803963**
S. heteropora	C11	Pískovna Cep I, Suchdol nad Lužnicí, Czech Republic	48°55'02.3"N, 14°53'01.6"E	8.10.2016	**OL803964**
S. heteropora	C38	Jáma pond, Hradec Králové, Czech Republic	50°11'10.8"N, 15°51'25.1"E	26.11.2016	**OL803965**
S. heteropora	C71	Lough Nafurnace, Derryvrisk, Ireland	53°22'11.7"N, 9°32'35.4"W	4.2.2017	**OL803966**
S. heteropora	D5	Věžák pond, Hrubá Skála, Czech Republic	50°30'58.3"N, 15°11'32.0"E	NA	**OL803967**
S. heteropora	D12	Komořany lagoons, Prague, Czech Republic	49°59'24.8"N, 14°24'04.2"E	NA	**OL803968**
S. heteropora	D24	Modřany lagoons, Prague, Czech Republic	49°59'52.7"N, 14°24'12.8"E	NA	**OL803969**
S. heteropora	F14	Kranksjon lake, Lund, Sweden	55°41'31.0"N, 13°29'34.1"E	27.03.2017	**OL803970**
S. heteropora	F65	Svaneholmssjön, Skurup, Sweden	55°30'02.4"N, 13°28'46.9"E	03.04.2017	**OL803971**
S. heteropora	F83	Häckebergasjön lake, Genarp, Sweden	55°34'45.5"N, 13°25'54.9"E	03.04.2017	**OL803972**
S. heteropora	K28	Harasov lake, Kokořín, Czech Republic	50°24'39.8"N, 14°34'07.3"E	18.04.2017	**OL803973**
S. heteropora	L66	Unnamed pond, near Parakalamos, Greece	39°52'28.6"N, 20°34'59.1"E	18.11.2017	**OL803974**
S. heteropora	N30	Pateira de Fermentelos dam, Oiã, Portugal	40°33'21.5"N, 8°30'31.5"W	05.02.2018	**OL803975**
S. heteropora	S20.45 / CAUP B 709	Crinan Canal, Lochgilphead, UK	56°03'37.8"N, 5°28'37.7"W	2008	**OL803976**

S. hibernica	Irsko IE-B11	Glendulagh Lake, Co. Galway, Ireland	53°27'59.0"N, 9°44'19.4"W	2008	HG514206
S. lanceolata	I37	Quidi Vidi Lake, Newfoundland, Canada	47°34'40.0"N, 52°41'48.3"W	24.05.2017	**OL803978**
S. macropora	968	Unnamed pond, near Mažice, Czech Republic	49°13'15.8"N, 14°36'40.1"E	30.11.2015	**OL803979**
S. macropora	M39	Lang Nieuwlandsche Achterwetering canal, Arkel, Netherlands	51°52'52.1"N, 5°01'33.2"E	8.11.2017	**OL803980**
S. macropora	S71.B4	Podhradská tůň, Bakov nad Jizerou, Czech Republic	50°27'40.1"N, 14°54'42.3"E	30.11.2015	**OL803981**
S. macropora	T66	Lac du Ticou, Bolquère, France	42°30'42.1"N, 2°04'12.4"E	20.05.2018	HG514228
S. macropora	V29	Unnamed pond, Khanty-Mansiysky District, Russia	61°01'10.1"N, 68°05'02.4"E	01.06.2018	**OL803983**
S. macropora	X40	Unnamed pond, Khanty-Mansiysky District, Russia	61°15'40.5"N, 73°37'07.8"E	29.05.2018	**OL803984**
S. petersenii	C87	Lough Ramor, Co. Cavan, Ireland	53°49'50.7"N, 7°05'03.0"W	2.2.2017	**OL803985**
S. soroconopea	F31	Trummen lake, Växjö, Sweden	56°51'38.0"N, 14°49'26.6"E	01.04.2017	**OL803986**
S. sphagnicola	H73	Máchovo lake, Doksy, Czech Republic	50°34'45.2"N, 14°40'01.1"E	28.04.2017	**OL803987**
S. sphagnicola	J54 / S150.D10	Great Rattling Brook, Newfoundland, Canada	48°48'24.7"N, 55°31'54.8"W	28.05.2017	**OL803988**
S. sphagnicola	K8 / S152.E5	Unnamed pond, Newfoundland, Canada	48°56'37.1"N, 55°49'23.8"W	29.05.2017	MK322792
S. sphagnicola	K33	Peat-bog, Jizera Mountains, Czech Republic	50°49'52.6"N, 15°14'39.9"E	07.06.2017	MK322795
S. sphagnicola	K35	Klečové louky peat-bog, Jizera Mountains, Czech Republic	50°50'14.2"N, 15°14'45.8"E	07.06.2017	**OL803991**
S. sphagnicola	K40	Peat-bog, Jizerka, Jizera Mountains, Czech Republic	50°49'40.7"N, 15°19'41.5"E	07.06.2017	**OL803992**
S. sphagnicola	K46	Unnamed pond, Cep, Czech Republic	48°55'23.3"N, 14°50'23.2"E	28.05.2017	**OL803993**
S. sphagnicola	LO234KE	Teichwiesen lake, Rohrbach bei Mattersburg, Austria	47°43'00.3"N, 16°27'26.4"E	NA	MK322785

S. sphagnicola	M44	Conne River Pond, Newfoundland, Canada	48°10'43.5"N, 55°29'02.1"W	28.05.2017	**OL803995**

* New sequences are indicated in bold type.

Table S2. Nuclear DNA contents for 60 strains of the genus *Synura* that exhibited ploidy-level variation among intraspecific strains during our investigation. Plant *Solanum pseudocapsicum* was used as a (pseudo)internal standard in most flow-cytometric analyses with the exception of measurements on *S. sphagnicola*, where it was supplemented with *Carex acutiformis*.

Species	Strain	DNA content	
		[pg]	[~Gbp]
S. americana	T83	2.155 ± 0.03	2.107
	V60	1.099 ± 0.01	1.075
	V77	1.131 ± 0.02	1.106
S. glabra	C46	1.745 ± 0.02	1.707
	C54	2.049 ± 0.01	2.004
	E98	1.952 ± 0.00	1.909
	F20	2.088 ± 0.02	2.042
	F41	1.920 ± 0.01	1.878
	F45	1.956 ± 0.03	1.913
	F81	1.976 ± 0.02	1.933
	G11	1.022 ± 0.01	1.000
	G12	1.977 ± 0.01	1.933
	G14	1.959 ± 0.03	1.916
	G46	2.029 ± 0.02	1.984
	G48	1.060 ± 0.01	1.036
	G56	2.018 ± 0.03	1.973
	G72	2.008 ± 0.03	1.964
	G98	2.057 ± 0.02	2.012
	H17	1.908 ± 0.02	1.866
	K55	1.832 ± 0.02	1.792
	K67	1.908 ± 0.01	1.866
	K76	2.202 ± 0.01	2.154
	L13	3.832 ± 0.03	3.747
	L26	1.895 ± 0.02	1.854
	L62	1.881 ± 0.01	1.840
	L70	1.917 ± 0.02	1.875
	L96	1.933 ± 0.02	1.891
	M73	1.861 ± 0.03	1.820
	M88	1.964 ± 0.03	1.921
	O60	2.090 ± 0.01	2.044
S. heteropora	985	3.004 ± 0.02	2.938
	989	0.727 ± 0.02	0.711
	C11	1.390 ± 0.02	1.359
	C38	1.410 ± 0.02	1.379
	C71	1.377 ± 0.03	1.346
	D12	1.413 ± 0.02	1.382
	D24	1.433 ± 0.03	1.401
	D5	0.840 ± 0.03	0.821
	F14	1.406 ± 0.02	1.375

	F65	1.465 ± 0.02	1.433
	F83	1.559 ± 0.02	1.525
	K28	1.445 ± 0.02	1.414
	L66	1.607 ± 0.03	1.572
	N30	1.398 ± 0.01	1.367
	S20.45 / CAUP B 709	1.418 ± 0.04	1.387
S. macropora	968	3.704 ± 0.05	3.623
	M39	3.464 ± 0.03	3.388
	S71.B4	3.529 ± 0.05	3.451
	T66	1.477 ± 0.02	1.444
	V29	3.549 ± 0.06	3.471
	X40	3.688 ± 0.04	3.607
S. sphagnicola	H73	0.391 ± 0.02	0.382
	J54	0.440 ± 0.00	0.430
	K33	0.378 ± 0.00	0.370
	K35	0.388 ± 0.01	0.379
	K40	0.352 ± 0.01	0.344
	K46	0.384 ± 0.01	0.376
	K8	0.428 ± 0.04	0.419
	LO234KE	0.764 ± 0.02	0.748
	M44	0.411 ± 0.01	0.402

Table S3. Ploidy level shifts over time detected within strains of seven chrysophyte species maintained in long-term cultivation. Plant *Solanum pseudocapsicum* was used as a (pseudo)internal standard in most flow-cytometric analyses with the exception of measurements on *Chrysosphaerella brevispina*, where it was supplemented with *Carex acutiformis*.

Species	Strain	Date of analysis	DNA content [pg]
Chrysosphaerella brevispina	E17	24.04.2019	0.48
		10.07.2019	0.51
		08.08.2019	0.22
		10.10.2019	0.23
		30.10.2019	0.50
		13.11.2019	0.50
Ochromonas tuberculata	F.4.2.	20.12.2018	0.80
		03.01.2019	0.79
		08.01.2019	0.37
		12.01.2019	0.78
Synura			
S. aff. *glabra*	K68	27.11.2017	0.87
		23.02.2018	1.85
	L10	23.11.2017	1.89
		27.11.2017	1.87
		28.12.2017	1.01
		17.02.2020	1.98
		13.10.2020	2.08
S. hibernica	Ir IE-B11	19.12.2017	3.59
		01.01.2018	3.64
		03.11.2018	1.75
		23.01.2019	3.55
		05.04.2019	1.80
		15.04.2019	1.81
		24.04.2019	1.75
S. lanceolata	I37	17.10.2017	0.67
		05.01.2018	1.39
		10.10.2018	1.34
		13.10.2020	1.44
S. petersenii	C87	29.08.2017	1.05 + 2.11
		06.09.2017	1.03 + 2.06
	C87-1	08.01.2019	2.11
		13.10.2020	2.16
		04.12.2020	2.07
	C87-2	03.01.2019	2.07
	C87-3	10.10.2018	2.01
		08.01.2019	2.08
	C87-4	10.10.2018	1.07

		08.01.2019	1.04
	C87-5	08.01.2019	2.12
		13.10.2020	2.14
		04.12.2020	2.10
	C87-6	10.10.2018	1.08
		08.01.2019	1.08
		13.10.2020	1.12
		04.12.2020	1.12
S. soroconopea	F31	01.01.2018	1.62
		02.02.2018	1.51
		07.02.2018	3.21
		13.02.2018	1.50
		05.04.2019	3.21
		15.04.2019	3.15
		13.10.2020	3.22