BMC Methods

METHODOLOGY

Open Access

High molecular weight DNA extraction from mucilage rich zygnematophycean green algae for long read DNA sequencing

Xuehuan Feng[1], Charlotte Permann[2], Andreas Holzinger[2] and Yanbin Yin[1*]

Abstract

Background Zygnematophyceae green algae represent the closest living relatives of land plants. Adaptions to hydro-terrestrial environments are evident through the production of mucilage carbohydrates, which are secreted outside algal cell walls to retain water. However, the mucilage poses significant challenges for the extraction of high molecular weight (HMW) DNA.

Methods To address this, we have developed an efficient protocol optimized for algae nuclei isolation and HMW DNA extraction with modified CTAB method, facilitating the use of third-generation long read sequencing technologies, e.g., PacBio or Oxford Nanopore. Furthermore, we have benchmarked the performance of our method against eight established DNA extraction methods or commercial kits.

Results Of the eight existing DNA extraction methods assessed, the PowerPlant DNeasy Kit, prominent for its use in plant DNA extraction, was the only protocol to successfully isolate DNA from *Zygnema circumcarinatum* algae. However, the DNA quality was insufficient for applications in PacBio or Oxford Nanopore sequencing. In contrast, our novel method not only yielded a high DNA concentration but also high purity with optimal A260/A230 and A260/A280 ratios suitable for long read DNA sequencing. Notably, the integrity of algal DNA obtained via our method surpassed that from commercial kits, demonstrating a significant increase in the length of extracted DNA, with a peak at 55.7 kb compared to 17.6 kb for the PowerPlant DNeasy Kit. Additionally, our method substantially reduced the organellar DNA, lowering it from 72.5% in the commercial kits to 9.6%, thereby enhancing the yield of nuclear genomic DNA.

Discussion Our method represents a significant advancement in the extraction of DNA from challenging plants and algae characterized by high extracellular mucilage contents. Our protocol removes most organellar DNA, and thus substantially increases the proportion of nuclear DNA reads in the sequencing data to lower the cost for nuclear genome sequencing. Our cost-efficient method will facilitate the whole genome sequencing of mucilage-rich algae and plants.

Keywords High molecular weight DNA, Mucilage, Zygnematophyceae green algae, *Zygnema circumcarinatum*, Nuclei isolation, DNA extraction

*Correspondence:
Yanbin Yin
yyin@unl.edu
[1] Department of Food Science and Technology, Nebraska Food for Health Center, University of Nebraska-Lincoln, Lincoln, NE, USA
[2] Department of Botany, University of Innsbruck, Sternwartestrasse 15, Innsbruck 6020, Austria

Background

Zygnematophyceae green algae (ZGA) or formerly Conjugatophyceae green algae [1], having conquered the land about 450 million years ago [2], are pivotal in our understanding of terrestrial colonization. ZGA have sexual reproduction through conjugation, a unique characteristic compared to other Charophyceae green

algae (CGA). ZGA are the most diverse and species richest CGA, with more than 4,000 described species [3]. ZGA were known to have two orders, *Desmidiales* and *Zygnematales*, containing 3,500 and 800 species, respectively [3]. Recently, *Spirogloeophycidae* has been defined as a new subclass of ZGA [4]. Similar to land plants, ZGA exhibit complex cell wall structures and compositions. Their cell walls and extracellular matrix (ECM) usually consist of three layers: (i) an inner cell wall layer composed of celluloses, (ii) a middle layer with mostly pectins such as homogalacturonan (HG) and rhamnogalacturonan I (RGI), and (iii) an outer ECM layer of gel-like mucilage with complex anionic polysaccharides [3, 5–7]. The water-retraining ECM is crucial for ZGA's survival in challenging (semi)terrestrial environments, where they must endure low water availability, high UV radiation and cold stresses [8–11].

Sequencing algal genomes, particularly those of ZGA, is the key to unlocking insights into the origin and early evolution of land plants. Recently, we have successfully sequenced four genomes of the *Zygnema* genus [12, 13] employing both short read and long read sequencing technologies. While the short read sequencing offers advantages of being cheaper, faster and possessing low error rate, it leads to more fragmented and ambiguous genome assemblies [14, 15]. Such limitation can be overcome by the long read sequencing technology, represented by the Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT), which has become increasingly popular in the last 10 years [16]. In particular, the PacBio's high-fidelity (HiFi) sequencing platform has achieved the accuracy comparable to that of short read sequencing, which positions it as a preferred choice for extensive plant and algal genome sequencing [17, 18].

However, the third-generation long read sequencing necessitates the use of high purity and long genomic DNA molecules [16]. The presence of a substantial amount of gel-like mucilage in ZGA, enriched with complex anionic polysaccharides [9, 19], as well as polyphenols and flavonoids [20, 21], poses a significant challenge to the extraction of long and pure genomic DNA [22]. Several methods have been developed for the extraction of High Molecular Weight (HMW) genomic DNA from plants and algae [15, 22–26]. Among these, the cetyltrimethylammonium bromide (CTAB) method [27] and its modified version [28] have been widely employed for DNA extractions across various organisms, including plants, algae, bacteria, and fungi [29–31]. However, the efficiency of CTAB methods is compromised when used in mucilaginous algae [32], resulting in lower DNA purity. Moreover, these methods require further modifications or additional cleaning steps, consequently reducing the DNA yield [22].

In our *Z.- circumcarinatum* genome sequencing project [13], we developed a method for HMW DNA extraction from mucilaginous freshwater green algae. The extracted DNA showed high purity and HMW, sufficient for ONT or PacBio long read sequencing. In this paper, we provide the detailed description of the HMW DNA extraction method using *Z. circumcarinatum* SAG 698-1b as sample material.

Materials and methods

Algae and culturing conditions

Zygnema circumcarinatum SAG 698-1b was obtained from the Culture Collection of Algae at Göttingen University (SAG) (https://sagdb.uni-goettingen.de/detailedList.php?str_number=698-1b). It was cultured in modified Bold's Basal Medium (MBBM), supplemented with 0.02% L-arginine, 0.1% peptone and 0.5% sucrose [9]. The algae used for transmission electron microscopy (TEM) were cultivated in standard BBM. The filaments were grown for two weeks on a rotary shaker platform at 110 rpm in Precision Plant Growth Chamber (Thermo Fisher Scientific, USA) with conditions: 16/8 of light/dark cycle, 20 °C, ~ 50 µmol photons m^{-2} s^{-1} [9, 33]. Some cultures were also maintained on 1% agar or liquid MBBM.

Algae collection and treatment

The SAG 698-1b cultures were harvested using a vacuum filtration with Whatman #2 papers (GE Healthcare 47 mm) and washed with water for two or three times. The collected tissues were immediately frozen with liquid nitrogen. The frozen algae can be used for DNA extraction or stored in -80 °C until use. The algae were lyophilized before being tested for multiple DNA extraction methods or ground into fine powder in liquid nitrogen with pre-chilled mortar and pestle.

In total, 9 DNA extraction methods were tested in this study. These include four commercial kits, four non-kit methods commonly used for plant and algal DNA extraction, and one method that was developed by us.

DNA extraction with four commercial kits

Four commercial kits were used in SAG 698-1b genomic DNA extraction. These kits are GeneJET Plant Genomic DNA Purification Kit (Thermo Fisher Scientific, USA), NucleoSpin Plant II Mini kit (Macherey–Nagel, Germany), DNeasy Plant Mini Kit (Qiagen, Germany), and DNeasy PowerPlant Pro Kit (Qiagen, Germany). The DNA extraction steps followed manufacturer's protocols. DNA purification was performed with PowerClean Pro Cleanup Kit if used.

CTAB method

The CTAB method was described by Doyle and Doyle [27] and Porebski et al. [28]. In brief, ~ 50 mg of SAG 698-1b algal sample was ground to a fine powder under liquid nitrogen, and the powder was mixed with 60 °C pre-warmed 600 µl of CTAB buffer: 2% CTAB, 100 mM Tris–HCl pH 8.0, 25 mM EDTA, 1.4 M NaCl, 1% polyvinylpyrrolidone, 2% β-mercaptoethanol. This mixture was incubated at 60 °C for 15 min and homogenized well with plastic pestle every three minutes. After cooling down to room temperature, the mixture was extracted with CIA (chloroform: isoamyl alcohol, 24:1) and the supernatant was treated with RNase (Qiagen, Germany) for 30 min at room temperature. Then the solution was extracted with CIA two more times and precipitated with 0.7 volumes of isopropanol. The DNA pellet was washed with 70% ethanol and dissolved in 50 µl Tris–HCl, EDTA (TE) buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0).

SDS method

The SDS method was described by Mayjonade et al. [24]. In short, ~ 50 mg grounded SAG 698-1b algal powder was mixed with 600 µl of lysis buffer (1.25% SDS, 1% polyvinylpolypyrrolidone (PVP) 40, 100 mM Tris–HCl pH 8.0, 50 mM EDTA, 1% sodium metabisulfite, 0.5 NaCl) and RNase A (Qiagen, Germany). The tube was incubated at 50 °C for 30 min and mixed well by shaking the tube with hand every 2 min. Then, add 1/3 volume of 5 M potassium acetate and incubate on ice for 5 min. The tube was centrifuged at room temperature with 15,000 rpm speed for 10 min. The resulting supernatant containing DNA was extracted with 2% Sera-Mag SpeedBead magnetic carboxylate beads solution (2% beads, 18% PEG 8000, 1 M NaCl, 10 mM Tris–HCl pH 8.0, 1 mM EDTA, 0.05% Tween 20), and gently agitated on a rotator for 10 min. The beads were washed with 70% ethanol and dissolved in 50 µl TE buffer to elute the DNA.

Chlamydomonas method

This method (CTAB/SDS/Phenol) was developed for *Chlamydomonas* DNA extraction by JGI (https://www.pacb.com/wp-content/uploads/2015/09/DNA-extraction-chlamy-CTAB-JGI.pdf). In brief, ~ 50 mg SAG 698-1b algal powder was mixed well with fresh prepared SDS-EB lysis buffer (2% SDS, 50 mM Tris–HCl pH 8.0, 20 mM EDTA, 0.2 M NaCl, 1 mg/ml Proteinase K). Then 1 × volume of pre-warmed CTAB buffer (2% CTAB, 50 mM Tris–HCl pH 8.0, 20 mM EDTA, 1.4 M NaCl, 1% polyvinylpyrrolidone, 2% β-mercaptoethanol) was added and incubated at 65 °C for 30 min. DNA was extracted with 1 × volume of phenol: chloroform: isoamyl alcohol (25:24:1). The resulting supernatant was treated with RNase A (Qiagen, Germany) and extracted two times with CIA. DNA precipitation and resuspension was the same as steps of aforementioned CTAB method.

Lithium chloride method

Hong et al. [34] developed a Lithium chloride based DNA extraction protocol for seaweeds. Briefly, ~ 50 mg SAG 698-1b algal powder was mixed with 500 µl of extraction buffer (1 M LiCl, 1% sarcosyl, 10 mM EDTA pH 9.0, 0.2% polyvinylpyrrolidone, 2.5% β-mercaptoethanol) and incubated at 55 °C for 5 min, then at 4 °C for 1 h. The lysate was centrifuged at 200 × g at 4 °C for 5 min. The resulting supernatant was treated with RNase A and precipitated with sodium acetate and ethanol. DNA was washed with 70% ethanol and resuspended with 50 µl TE buffer.

In addition to the eight previously reported methods, we have developed our own method, which integrates the algae nuclei isolation and modified CTAB extraction steps.

Our method: nuclei isolation followed by a modified CTAB method

We found that DNA extracted with abovementioned methods contained a very high content (> 60%) of chloroplast and mitochondria DNA. To remove these organellar DNA and enrich nuclear DNA, we employed a modified nucleus isolation method [35, 36]. Specifically, ~ 50 mg young fresh SAG 698-1b algal tissues were frozen in liquid nitrogen and then were grinded into fine powder with precooled mortar and pestle from -20 °C. The fine powder was transferred into a beaker containing 5 ml of freshly prepared 1 × nucleus isolation buffer (NIB) (0.5 M sucrose, 10 mM Tris–HCl, 10 mM EDTA pH 9.0, 80 mM KCl, 1 mM spermidine, 1 mM spermine, 0.5% Triton X-100, 0.2% β-mercaptoethanol). This mixture was gently homogenized well on ice with stir bar for 10 min, and then were filtered with two layers of Miracloth (MilliporeSigma, USA). The remaining nuclei were pelleted by centrifugation with speed of 800 × g at 4 °C for 10 min. Then, the crude nuclei pellet was washed with 500 µl of 1 × NIB and centrifuged with speed of 800 × g at 4 °C for 10 min.

HMW DNA was extracted with a modified method (CTAB buffer: 2% CTAB, 100 mM Tris–HCl pH 8.0, 25 mM EDTA, 2.5 M NaCl, 1% polyvinylpyrrolidone, 1% β-mercaptoethanol which was added just before use). Specifically, 500 µl of CTAB buffer was added into the nuclei pellet and mixed well. The mixture was homogenized with plastic pestle and incubated in heating block at 50 °C for 15 min. During the incubation, shake the tube vigorously with hand every 3 min. After the mixture was cooling down, 500 µl of CIA was added and extracted. The lysate was treated with 5 µl of RNase (Qiagen,

Germany) at 37 °C for 30 min, and then was extracted with CIA for two times. Then, 0.7 volume of isopropanol was added and incubated at -20 °C for 30 min, followed by centrifugation at room temperature with 15,000 rpm for 15 min. The pellet was washed with 0.5 ml of 70% ethanol and resuspend with 50 µl TE buffer.

Assessment of DNA quality and quantity

Quality and quantity of extracted and purified DNA were evaluated by using 1% agarose gel electrophoresis, NanoDrop 2000/2000c Spectrophotometers, and Qubit 3.0 Fluorometer (Thermo Fisher Scientific, USA).

DNA integrity analysis

The integrity of SAG 698-1b DNA was evaluated by DNA Integrity Number (DIN). DNA size was determined on 4200 TapeStation System (Agilent Technologies, USA) using the Genomic DNA ScreenTape Assay (reagents for 200 bp to 60,000 bp).

DNA sequencing

DNA samples were sequenced at Roy J. Carver Biotechnology Center at University of Illinois at Urbana-Champaign [13]. Oxford Nanopore DNA libraries were prepared with 1D library kit SQK-LSK109 and sequenced with SpotON R9.4.1 FLO-MIN106 flowcells for 48 h on a GridIONx5 sequencer. Base calling was performed with Guppy v1.5 (https://community.nanoporetech.com). DNA sequencing reads can be accessed through NCBI BioProject PRJNA917633.

Fluorescence staining and microscopy

Methods for fluorescence staining were described previously [9]. In brief, 1 ml of young cultures (two weeks) of SAG 698-1b were filtered and washed with 1×PBS for 3 times. Then the algae were fixed with 4% formaldehyde (in 1×PBS) for 30 min and washed with 1×PBS for 3 times. The treated algae were then stained with 500 nM of Propidium Iodide (PI) (Thermo Fisher Scientific, USA) containing 0.3% of Triton-X 100. After staining for 10 min in the dark, the algae were washed with 1×PBS for 3 times. Fluorescent images were taken with a Confocal Laser Scanning Microscope (CLSM) with NIS-Element (Nikon Imaging Software, Nikon, Japan). Differential Interference Contrast (DIC) images and fluorescent images of isolated nuclei were investigated under an Axio Imager 2 microscope (Carl Zeiss Microscopy, LLC). Commercial Indian ink (Dr. Ph. Martin's Bombay blue) was used to visualize the ECM in a Zeiss Axiovert 200 M microscope (see Permann et al. [37]).

Transmission electron microscopy

Transmission electron microscopy (TEM) was performed as described before [38]. Briefly, four weeks old cultures of SAG 698-1b were fixed in 2.5% glutaraldehyde (in 20 mM cacodylate buffer, pH 7.0) for 1.5 h, rinsed extensively in 20 mM cacodylate buffer. Fixed cells were embedded in 3% agarose and post fixed in 1% OsO_4 in 20 mM cacodylate buffer at 4 °C overnight. Dehydration was performed using increasing ethanol concentrations, transferred to propylene oxide and embedded in modified Spurr's resin and polymerized at 70 °C for 8 h. Ultrathin sections (~60–90 nm) were prepared with a Reichert Ultracut (Leica Microsystems, Wien, Austria) and stained with 2% uranyl acetate and Reynold's lead citrate. Images were taken on a Zeiss Libra 120 transmission electron microscope (Carl Zeiss AG, Oberkochen, Germany) at 80 kV, which was equipped with a TRS 2 k SSCCD camera and operated by ImageSP software (Albert Tröndle Restlichtverstärker Systeme, Moorenweis, Germany).

Results

Nine DNA extraction methods

Zygnema circumcarinatum SAG 698-1b grows as a filamentous green algal organism (Fig. 1a, b). The cell width is 22 ~ 24 µm under the liquid MBBM [9]. The cell features two star-shaped chloroplasts (Fig. 1b, c and d), and the nucleus is located in the cell-center between the two chloroplasts (Fig. 1e, f and g). Staining with Indian ink illustrates the extracellular matrix (ECM) composed of a mucilage layer which remains white as the blue ink particles cannot penetrate (Fig. 1c, d). Transmission electron microscopy (TEM) further confirmed that the cell walls of the filaments are surrounded by a thick ECM, i.e., the mucilage layer (Fig. 2a-c). The mucilage layer consists of fibrils arranged perpendicularly to the cell wall (Fig. 2c), which is much thicker than cell walls (Fig. 2a, b). This layer exhibits a thickness of 5–10 µm and is rich in pectic polysaccharides, creating a significant challenge in obtaining high purity and high molecular weight (HMW) DNA.

In an effort to optimize HMW genomic DNA extraction from the highly mucilaginous SAG 698-1b, we evaluated nine different methods as summarized in Table 1 and Fig. 3. Our novel approach featured a nuclei isolation and enrichment step and followed by a modified CTAB step (Fig. 4). Among the nine tested methods, only DNeasy PowerPlant Pro kit and our method successfully obtained DNA (lanes 4 and 9 in Fig. 3). From the electrophoresis gel images, we observed that DNA extracted with the CTAB method (lane 5), the CTAB/SDS/Phenol method (lane 6) and the Lithium

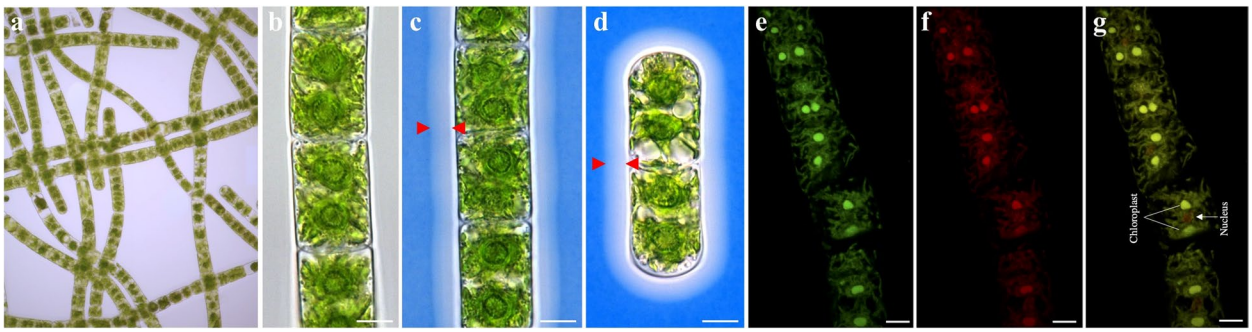

Fig. 1 Light and fluorescence microscopic images of *Zygnema circumcarinatum* SAG 698-1b. The algae contain two stellate chloroplasts and one nucleus (center position) in each unbranched filamentous cell (**a**-**g**). **a** overview of filaments, **b** individual filament showing two massive pyrenoids in each cell (three cells are shown), **c**-**d** indian ink stained cell illustrating the mucilage layer (indicated by red arrowheads), **e** autofluorescence, **f** PI staining, **g** merged image of (**e**) and (**f**). DIC (Differential Interference Contrast) images were taken with an Axio Imager 2 microscope (**a**) or a Zeiss Axiovert 200 M microscope (**b**-**d**) and fluorescent images were taken with a Nikon Confocal Microscope (**e**, **f**, **g**). Scale bars: **a**: 50 µm, **b**-**d**: 10 µm, **e**–**f**: 10 µm

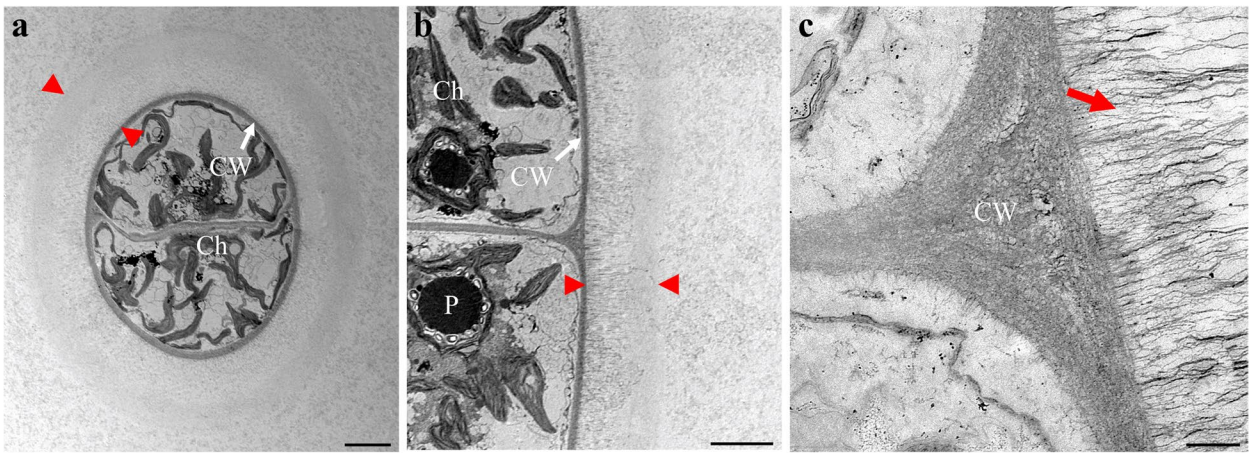

Fig. 2 Transmission electron microscopy of *Zygnema circumcarinatum* SAG 698-1b. **a** cross section through a filament illustrating the massive extracellular matrix (ECM, arrowheads) outside of the cell wall (CW), **b** longitudinal section with pyrenoid (P) and chloroplasts (Ch), and the cell wall (CW) surrounded by ECM (arrowheads), **c** detailed view of the cell wall (CW) in the attachment area of two cells, covered by a fibrillar ECM layer (arrow). Scale bars: **a**, **b**: 5 µm, **c**: 500 nm

Table 1 Nine methods for HMW genomic DNA extraction from SAG 698-1b

No	Methods	Absorbance A260:280	Absorbance A260:230	Concentration (ng/µl) (Nanodrop)	Concentration (ng/µl) (Qubit)
1	GeneJET Plant Genomic DNA purification kit	1.10 ± 0.002	0.57 ± 0.006	43.38 ± 2.32	0.56 ± 0.02
2	NucleoSpin Plant II Mini kit	1.39 ± 0.003	1.34 ± 0.01	134.58 ± 1.18	0.89 ± 0.01
3	DNeasy Plant Mini kit	1.10 ± 0.002	0.79 ± 0.01	51.01 ± 1.02	0.84 ± 0.01
4	DNeasy PowerPlant Pro kit	1.03 ± 0.02	0.51 ± 0.03	40.26 ± 3.22	9.12 ± 0.04
5	CTAB method	1.12 ± 0.001	0.54 ± 0.002	112.57 ± 1.82	2.55 ± 0.02
6	CTAB/SDS/Phenol method	1.04 ± 0.005	0.54 ± 0.006	226.15 ± 1.89	5.33 ± 0.03
7	SDS method	1.02 ± 0.002	0.75 ± 0.001	27.63 ± 0.33	1.72 ± 0.01
8	Lithium chloride method	1.17 ± 0.006	0.56 ± 0.006	171.72 ± 0.64	4.83 ± 0.35
9	Nuclei isolation/ modified CTAB (our method)	1.79 ± 0.005	1.90 ± 0.03	49.14 ± 1.23	38.07 ± 0.35

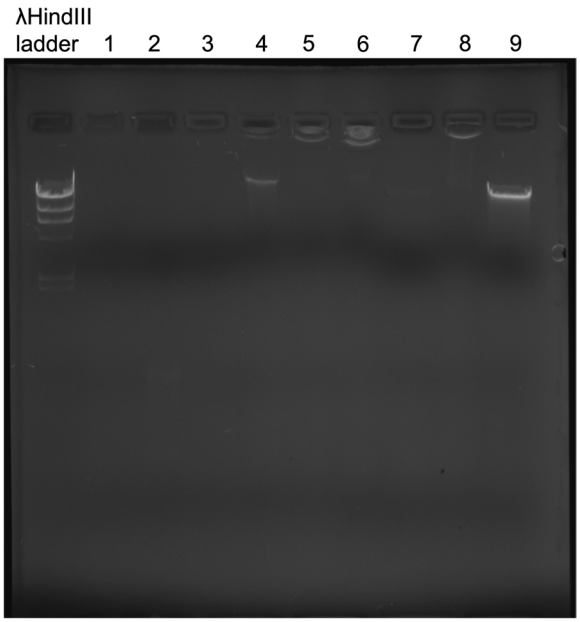

Fig. 3 DNA electrophoresis gel image of genomic DNA extracted with nine methods. The labeled nine DNA extraction methods: (1) GeneJET Plant Genomic DNA Purification Kit; (2) NucleoSpin Plant II Mini kit; (3) DNeasy Plant Mini kit; (4) DNeasy PowerPlant Pro Kit; (5) CTAB method; (6) CTAB/SDS/Phenol method; (7) SDS method; (8) Lithium chloride method; (9) Nuclei isolation/modified CTAB method

chloride method (lane 8) were all stuck in the wells and failed to migrate. These are likely due to the sticky contaminants from the cell wall and ECM polysaccharides holding the DNA in the gel. For other kits/methods (lanes 1,2,3,7), the DNA bands were almost invisible likely because the DNA extraction yield was too low for detection.

In addition, the 8 existing DNA extraction methods had low ratios of A260:280 and A260:230, which are commonly used for assessing the purity of DNA. Although the methods 5, 6 and 8 demonstrated high Nanodrop concentrations, 112.57 ng/µl, 226.15 ng/µl and 171.12 ng/µl, respectively (Table 1), the Qubit concentrations were much lower, 2.55 ng/µl, 5.33 ng/µl and 4.83 ng/µl, respectively (Table 1). These discrepancies indicated the presence of contaminants in the samples (electrophoresis gel wells of 5, 6 and 8, respectively, Fig. 3).

DNA integrity

We further assessed the integrity of SAG 698-1b DNA extracted by the DNeasy PowerPlant Pro Kit and our method (lanes 4 and 9 in Fig. 3). The DNA from the DNeasy PowerPlant Pro Kit showed a DIN (DNA Integrity Number) of 8.0 and a peak of ~ 24.3 kb (Fig. 5a and b). In contrast, the DNA from our method yielded a higher DIN of 9.6 and a higher peak of ~ 55.7 kb (Fig. 5a and d). This indicates our method enhanced DNA integrity and molecular weight (fragment length). To improve the DNA quality from the DNeasy PowerPlant Pro Kit, the DNA was further purified with the DNeasy PowerClean Pro Cleanup Kit. However, the purified DNA became more fragmented, with a lower DIN of 7.4 and a decreased peak of ~ 17.6 kb (Fig. 5a and c).

Nucleus DNA enrichment

In plant and algal DNA extraction, DNA are often from both nuclear and organelles, and usually organellar DNA dominate due to their high copy numbers. Given the complexity and size of nuclear genome relative to plastome and mitogenome, it is crucial to reduce DNA of organelles prior to sequencing. This enrichment for

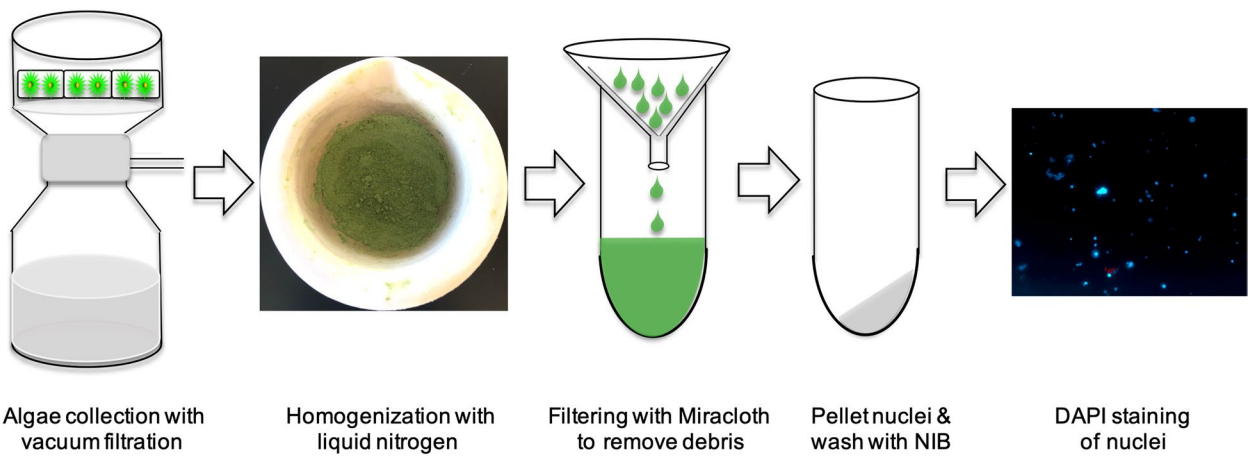

Fig. 4 Illustration of our modified method for nuclei isolation from *Zygnema circumcarinatum* SAG 698-1b. Integrity of isolated crude nuclei were examined by staining with DAPI and viewed under Axio Imager 2 microscope. The nuclei were uniform, approximately 2 µm in diameter

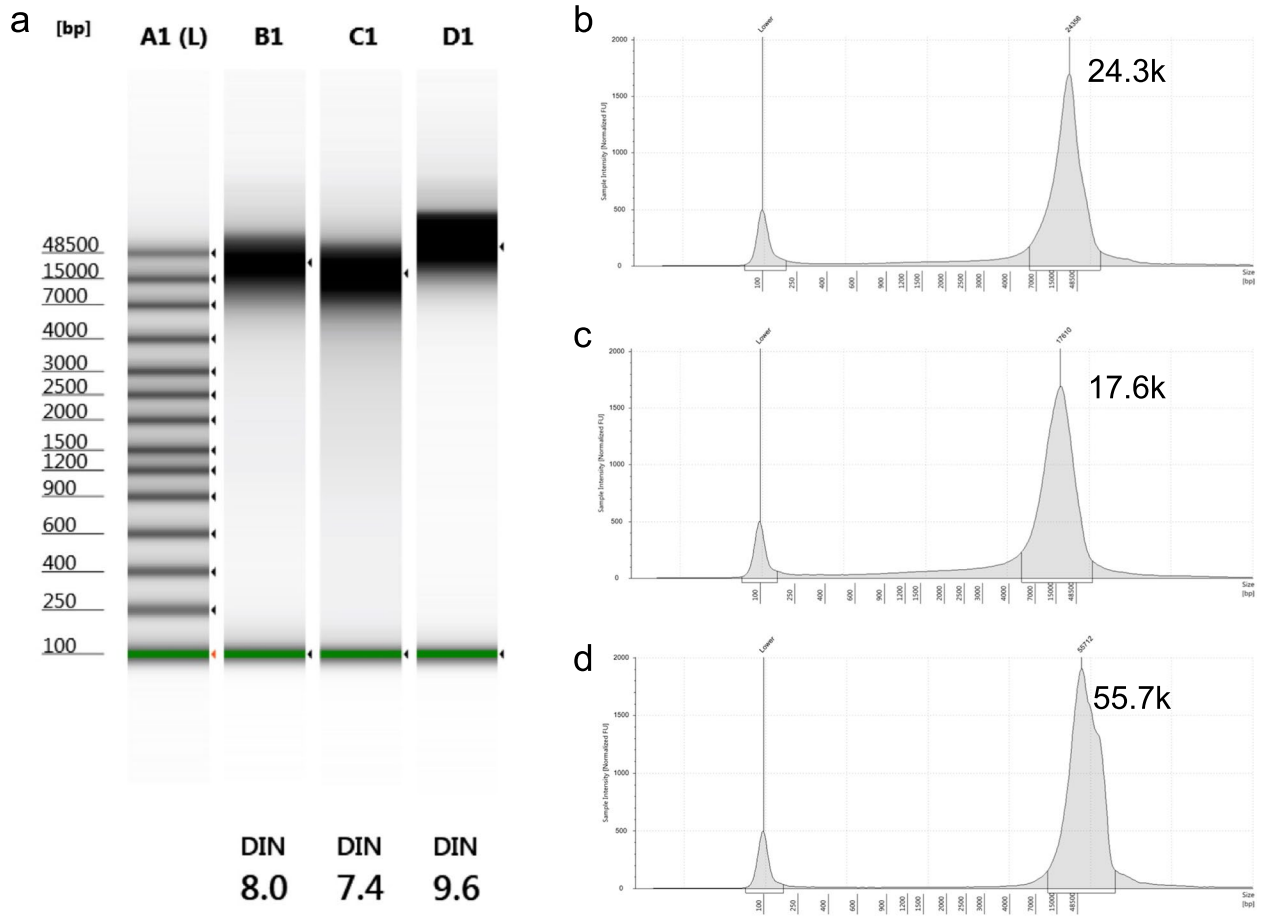

Fig. 5 DNA integrity of *Zygnema circumcarinatum* SAG 698-1b. **a** Automated electrophoresis analysis (Tapestation 4200) of extracted DNA samples. DIN was indicated. B1: DNeasy PowerPlant Pro Kit; C1: DNeasy PowerClean Pro Cleanup Kit; D1: our method. **b**, **c**, **d** Electrophoresis densitometry of B1, C1 and D1

nuclear DNA can significantly decrease the sequencing cost by removing organellar DNA.

The SAG 698-1b DNA samples extracted using both the DNeasy PowerPlant Pro kit and our method were individually sequenced with Oxford Nanopore long read sequencing technology. Subsequent reads mapping to the SAG 698-1b plastome (MT040697, [13]) and mitogenome (OQ319605, [13]) revealed that 53.6% and 18.9% of the DNA reads from DNeasy PowerPlant Pro kit were derived from plastome and mitogenome, respectively (Fig. 6a). In contrast, using our method with the nuclei isolation step (Fig. 4), only 4.2% and 5.4% of the DNA reads were mapped to the SAG 698-1b plastome and mitogenome (Fig. 6b). Therefore, our method significantly increased the fraction of the nuclear DNA reads in the final long read sequencing data from 27.5% to 90.4%.

Discussion

Prior to 2023, four ZGA genomes (*Spirogloea muscicola*, *Mesotaenium endlicherianum, Penium margaritaceu, Closterium*) have been published [4, 20, 39]. All these genomes were sequenced using the short read sequencing technologies and thus the genome assemblies are largely fragmented. Recently four additional ZGAs from the *Zygnema* genus were successfully sequenced using long read PacBio HiFi and NanoPore sequencing technologies [13]. These genomes thus have substantially better genome continuity and completeness. To obtain high molecular weight and high purity DNA for sequencing the four *Zygnema* genomes, we developed the method that combines a nuclei isolation process and an improved CTAB step. This approach fills a gap that no workable DNA extraction method

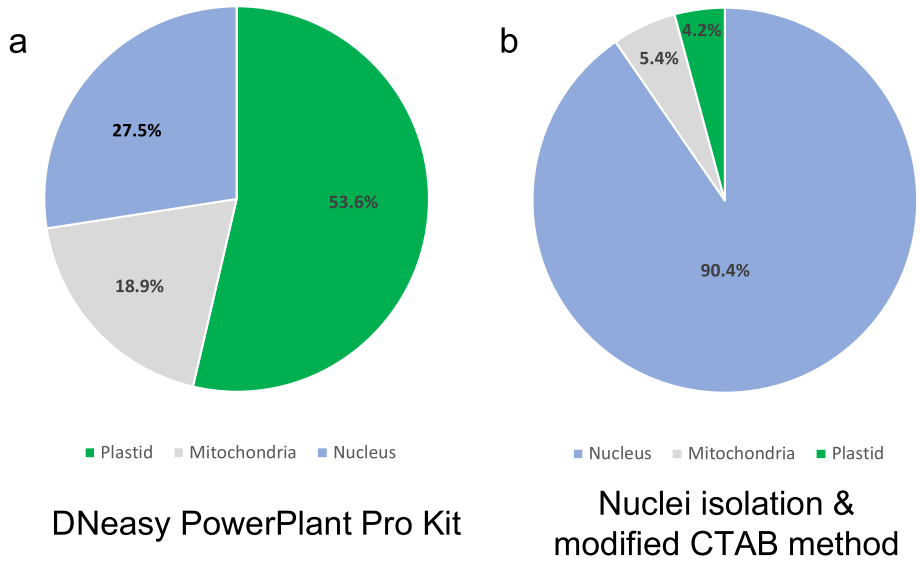

Fig. 6 Nanopore DNA sequencing read mapping percentages. **a** DNeasy PowerPlant Pro Kit sample. **b** our method sample

is available for high quality and HMW genomic DNA extraction from the recalcitrant mucilaginous algae.

SAG 698-1b exhibits a thick ECM layer composed of mucilages as visualized by Indian ink staining (Fig. 1) and TEM (Fig. 2), which presents a great challenges for HMW genomic DNA extraction [9, 40]. Palacio-López et al. [41] demonstrated that the ECM in *Zygnema* contains arabinogalactan proteins (AGPs). Domozych et al. [42] reported that the mucilage contains HMW acidic polysaccharides, rich in glucuronic acids and fucoses. Our observation indicated that SAG 698-1b produced a larger amount of sticky mucilage in older cultures or in agar medium with limited nutrients. Consequently, to reduce the mucilage levels, we preferred young cultures of algae (less than two weeks) grown in full strength liquid medium (MBBM, supplemented with 0.02% L-arginine, 0.1% peptone and 0.5% sucrose). We failed to generate protoplasts by using cell wall enzyme digestion likely due to the excessive mucilage present in the algae [9]. This suggests that enzyme treatment is not a viable option for HMW DNA extraction in ZGA. Furthermore, a previous report showed that dark treatments can help reduce polysaccharides and polyphenolics [15]. However, in our study such approach had no success in improving the SAG 698-1b genomic DNA extraction.

Commercial DNA extraction kits are widely used in plant genomics for their convenience and standardized procedures. However, out of the four tested kits, only the DNeasy PowerPlant Pro kit succeeded in extracting DNA from SAG 698-1b. This kit has major limitations though: (i) the extracted DNA were sheared into short fragments (Figs. 3 and 5a), (ii) the DNA purity was low (Fig. 5), and (iii) the DNA were dominated by organellar DNA (Fig. 6a). All these made it not suitable for long read sequencing. In contrast, we incorporated a nuclei purification step that effectively removed most DNA coprecipitating compounds, such as polysaccharides, polyphenols and other secondary metabolites. Additionally our protocol avoided embedding nuclei in agarose [36], which may introduce additional contaminations. This nuclei isolation step eliminated excessive organellar DNA and thus significantly enriched the nuclear DNA.

Several steps are essential to successfully obtain high purity and HMW DNA from mucilage-rich algae and plants. Firstly, younger cultures (less than two weeks) in liquid MBBM are recommended, as older cultures grown on the agar MBBM tend to produce more mucilage [9]. These viscous materials will co-precipitate with DNA in subsequent steps by ethanol or isopropanol. Secondly, 0.5% Triton X-100 (non-ionic detergent) is recommended in the NIB (nucleus isolation buffer) for crude nuclei isolation. This will help lysis the membranes of cells, chloroplasts, and mitochondria, and release most organellar DNA into the solution, thus preventing their co-pelleting with the nuclei [43]. It is worth noted that concentration of Triton X-100 can be adjusted depending on the sample types. For example, 1% Triton X-100 was suggested for lysing the tomato organelle and cell membrane; however, 0.3% of Triton X-100 was recommended for the Tobacco, Spinach and *Arabidopsis* (CelLytic PN Plant Nuclei Isolation/Extraction Kit, Sigma-Aldrich, https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/product/documents/259/383/cellytpn1pis-ms.pdf). Thirdly, in the nuclei isolation step, we pelleted the nuclei at relatively

low speed of 800 × g. Li et al. [15] and Zhang et al. [36] showed that in *Arabidopsis thaliana*, nuclei could be pelleted at a higher speed (3,840 × g). However, in our algae, we found that with a higher speed more sticky materials co-precipitated with the extracted nuclei. Finally, we found that increasing the NaCl concentration from 1.4 M to 2.5 M can prevent the co-precipitation of polysaccharides and DNA during the ethanol or isopropanol precipitation steps [44].

In addition, it is worth mentioning that the Genomic DNA ScreenTape Assay can analyze the DNA size range from 200 bp to > 60,000 bp, but it has sizing accuracy less than 15,000 bp. Our results using Tapestation 4200 showed that the DNA from our method has peak around 56 kb, which is still suitable for the PacBio HiFi and Oxford Nanopore DNA library preparation and sequencing. To obtain more accurate sizing and higher separation resolution of large DNA samples, the pulsed-field electrophoresis, such as Agilent Femto Pulse system, is recommended. The pulsed-field electrophoresis can reduce the compression of large DNA (usually greater than 20 kb) and separate the DNA to true size.

Conclusion

To summarize, our modified method was specifically designed to extract high purity and HMW DNA from mucilage rich ZGA. It was the only DNA extraction method that succeeded and was employed in our recent long read DNA sequencing of four *Zygnema* genomes [13]. The method could be applicable to other mucilage rich plants and algae. In addition, our method can preferentially remove most of the organelles and effectively enrich nuclear DNA. This will reduce the sequencing cost by avoiding the excessive amount of organellar DNA reads, and the final DNA sequencing reads will be primarily from the nuclear genome.

Acknowledgements
We thank You Zhou of UNL for taking the fluorescent images at the Microscopy Research Core Facility of the Center for Biotechnology, which is partially funded by the Nebraska Center for Integrated Biomolecular Communication COBRE grant (P20 GM113126 and NIGMS) and the Nebraska Research Initiative. We thank Chi Zhang of UNL for supplying the reagents for DNA concentration and integrity analysis. We would like to thank Sabrina Obwegeser of University Innsbruck for the help in TEM sectioning.

Authors' contributions
This study was designed and performed by XF and YY. XF performed the development of the DNA extraction protocols. Light microscopy and Indian ink staining was performed by CP. Transmission electron microscopy was performed by CP and AH. The draft manuscript was written by XF and YY and all other authors contributed their sections. All authors read and approved the final manuscript.

Funding
This work was supported by the National Science Foundation (NSF) CAREER award (DBI-1933521) and the United States Department of Agriculture (USDA) award (58–8042-9–089) to YY. This research was funded in part by the Austrian Science Fund (FWF)https://doi.org/10.55776/P34181to AH.

Data availability
Oxford Nanopore DNA sequencing reads are available through NCBI BioProject PRJNA917633, BioSample SAMN32543366, accession SRX20173906.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 21 August 2024 Accepted: 8 November 2024
Published online: 29 November 2024

References

1. Guiry MD. Taxonomy and nomenclature of the Conjugatophyceae (= Zygnematophyceae). Algae. 2013;28(1):1–29.
2. Delwiche CF, Cooper ED. The evolutionary origin of a terrestrial flora. Curr Biol. 2015;25(19):R899–910.
3. Hall JD, McCourt RM: Conjugating green algae including desmids. In: Freshwater Algae of North America. Elsevier; 2015;429–457.
4. Cheng S, Xian W, Fu Y, Marin B, Keller J, Wu T, Sun W, Li X, Xu Y, Zhang Y, et al. Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. Cell. 2019;179(5):1057–1067. e1014.
5. Youssef NH, Farag IF, Rinke C, Hallam SJ, Woyke T, Elshahed MS. *In Silico* analysis of the metabolic potential and niche specialization of candidate phylum" *Latescibacteria*"(WS3). PLoS ONE. 2015;10(6):e0127499.
6. Domozych DS, Bagdan K. The cell biology of charophytes: Exploring the past and models for the future. Plant Physiol. 2022;190(3):1588–608.
7. Domozych DS, LoRicco JG. The extracellular matrix of green algae. Plant Physiol. 2024;194(1):15–32.
8. Herburger K, Xin A, Holzinger A. Homogalacturonan accumulation in cell walls of the green alga *Zygnema* sp. (Charophyta) increases desiccation resistance. Front Plant Sci. 2019;10:540.
9. Feng X, Holzinger A, Permann C, Anderson D, Yin Y. Characterization of two *Zygnema* strains (*Zygnema circumcarinatum* SAG 698–1a and SAG 698–1b) and a rapid method to estimate nuclear genome size of zygnematophycean green algae. Front Plant Sci. 2021;12:610381.
10. Holzinger A, Albert A, Aigner S, Uhl J, Schmitt-Kopplin P, Trumhová K, Pichrtová M. Arctic, Antarctic, and temperate green algae *Zygnema* spp. under UV-B stress: vegetative cells perform better than pre-akinetes. Protoplasma. 2018;255:1239–52.
11. Trumhová K, Holzinger A, Obwegeser S, Neuner G, Pichrtová M. The conjugating green alga *Zygnema* sp. (Zygnematophyceae) from the Arctic shows high frost tolerance in mature cells (pre-akinetes). Protoplasma. 2019;256:1681–94.
12. Czurda V. Experimentelle Untersuchungen über die Sexualitätsverhältnisse der Zygnemalen. Bot Zbl. 1930;47:15–68.
13. Feng X, Zheng J, Irisarri I, Yu H, Zheng B, Ali Z, de Vries S, Keller J, Fürst-Jansen JM, Dadras A, et al. Genomes of multicellular algal sisters to land plants illuminate signaling network evolution. Nat Genet. 2024;56:1018–31.
14. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The third revolution in sequencing technology. Trends Genet. 2018;34(9):666–81.
15. Li Z, Parris S, Saski CA. A simple plant high-molecular-weight DNA extraction method suitable for single-molecule technologies. Plant Methods. 2020;16(1):1–6.
16. Jung H, Winefield C, Bombarely A, Prentis P, Waterhouse P. Tools and strategies for long-read sequencing and de novo assembly of plant genomes. Trends Plant Sci. 2019;24(8):700–24.
17. Wenger AM, Peluso P, Rowell WJ, Chang P-C, Hall RJ, Concepcion GT, Ebler J, Fungtammasan A, Kolesnikov A, Olson ND. Accurate circular consensus

long-read sequencing improves variant detection and assembly of a human genome. Nat Biotechnol. 2019;37(10):1155–62.
18. Hon T, Mars K, Young G, Tsai Y-C, Karalius JW, Landolin JM, Maurer N, Kudrna D, Hardigan MA, Steiner CC. Highly accurate long-read HiFi sequencing data for five complex genomes. Sci Data. 2020;7:399.
19. Domozych D, Ciancia M, Fangel JU, Mikkelsen MD, Ulvskov P, Willats WG. The cell walls of green algae: a journey through evolution and diversity. Front Plant Sci. 2012;3:82.
20. Jiao C, Sørensen I, Sun X, Sun H, Behar H, Alseekh S, Philippe G, Lopez KP, Sun L, Reed R, et al. The *Penium margaritaceum* genome: hallmarks of the origins of land plants. Cell. 2020;181:1–15.
21. Pichrtová M, Remias D, Lewis LA, Holzinger A. 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. Microb Ecol. 2013;65:68–83.
22. Vaillancourt B, Buell CR. High molecular weight DNA isolation method from diverse plant species for use with Oxford Nanopore sequencing. Preprint at bioRxiv. 2019;783159. https://doi.org/10.1101/783159.
23. Jose J, Usha R. Extraction of geminiviral DNA from a highly mucilaginous plant (*Abelmoschus esculentus*). Plant Mol Biol Rep. 2000;18(4):349–55.
24. Mayjonade B, Gouzy J, Donnadieu C, Pouilly N, Marande W, Callot C, Langlade N, Muños S. Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. Biotechniques. 2016;61(4):203–5.
25. Arimoto A, Nishitsuji K, Higa Y, Arakaki N, Hisata K, Shinzato C, Satoh N, Shoguchi E. A siphonous macroalgal genome suggests convergent functions of homeobox genes in algae and land plants. DNA Res. 2019;26(2):183–92.
26. Stark JR, Cardon ZG, Peredo EL. Extraction of high-quality, high-molecular-weight DNA depends heavily on cell homogenization methods in green microalgae. Appl Plant Sci. 2020;8(3):e11333.
27. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytoch Bull. 1987;19(1):11–5.
28. Porebski S, Bailey LG, Baum BR. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Rep. 1997;15(1):8–15.
29. Tel-Zur N, Abbo S, Myslabodski D, Mizrahi Y. Modified CTAB procedure for DNA isolation from epiphytic cacti of the genera Hylocereus and Selenicereus (Cactaceae). Plant Mol Biol Rep. 1999;17(3):249–54.
30. Minas K, McEwan NR, Newbold CJ, Scott KP. Optimization of a high-throughput CTAB-based protocol for the extraction of qPCR-grade DNA from rumen fluid, plant and bacterial pure cultures. FEMS Microbiol Lett. 2011;325(2):162–9.
31. Almakarem ASA, Heilman KL, Conger HL, Shtarkman YM, Rogers SO. Extraction of DNA from plant and fungus tissues in situ. BMC Res Notes. 2012;5(1):1–11.
32. Sarrazola H, Alzate F. Obtaining DNA from *Urticaceae*: overcoming the challenges associated with chemical compounds and herbarium specimens. Int J Mol Biol. 2019;4(5):158–65.
33. Fitzek E, Orton L, Entwistle S, Grayburn WS, Ausland C, Duvall MR, Yin Y. Cell wall enzymes in *Zygnema circumcarinatum* UTEX 1559 respond to osmotic stress in a plant-like fashion. Front Plant Sci. 2019;10:732.
34. Hong YK, Coury DA, Polne-Fuller M, Gibor A. Lithium chloride extraction of DNA from the seaweed *Porphyra perforata* (Rhodophyta). J Phycol. 1992;28(5):717–20.
35. Varela-Álvarez E, Andreakis N, Lago-Lestón A, Pearson GA, Serrao EA, Procaccini G, Duarte CM, Marba N. Genomic DNA isolation from green and brown algae (caulerpales and fucales) for microsatellite library construction. J Phycol. 2006;42(3):741–5.
36. Zhang M, Zhang Y, Scheuring CF, Wu C-C, Dong JJ, Zhang H-B. Preparation of megabase-sized DNA from a variety of organisms using the nuclei method for advanced genomics research. Nat Protoc. 2012;7(3):467–78.
37. Permann C, Herburger K, Niedermeier M, Felhofer M, Gierlinger N, Holzinger A. Cell wall characteristics during sexual reproduction of *Mougeotia* sp. (Zygnematophyceae) revealed by electron microscopy, glycan microarrays and RAMAN spectroscopy. Protoplasma. 2021;258(6):1261–75.
38. Holzinger A, Roleda MY, Lütz C. The vegetative arctic freshwater green alga *Zygnema* is insensitive to experimental UV exposure. Micron. 2009;40(8):831–8.
39. Sekimoto H, Komiya A, Tsuyuki N, Kawai J, Kanda N, Ootsuki R, Suzuki Y, Toyoda A, Fujiyama A, Kasahara M. A divergent RWP-RK transcription factor determines mating type in heterothallic Closterium. New Phytol. 2023;237(5):1636–51.
40. Miller RD. A developmental and physiological comparison of two mating strains of *Zygnema circumcarinatum* (Czurda). Ph.D. Thesis. Tucson: The University of Arizona; 1973.
41. Palacio-López K, Tinaz B, Holzinger A, Domozych DS. Arabinogalactan proteins and the extracellular matrix of charophytes: a sticky business. Front Plant Sci. 2019;10:447.
42. Domozych CR, Plante K, Blais P, Paliulis L, Domozych DS. Mucilage processing and secretion in the green alga *Closterium*. I. Cytology and biochemistry. J Phycol. 1993;29(5):650–9.
43. Watson JC, Thompson WF. Purification and restriction endonuclease analysis of plant nuclear DNA. Methods enzymol. 1986;18:57–75.
44. Fang G, Hammar S, Grumet R. A quick and inexpensive method for removing polysaccharides from plant genomic DNA. BioTechniques. 1992;13(1):52–56.

Publisher's Note