Transformation of *Diplonema papillatum*, the type species of the highly diverse and abundant marine microeukaryotes Diplonemida (Euglenozoa)

Binnypreet Kaur,¹,² Matus Valach,³ Priscila Pena-Diaz,¹ Sandrine Moreira,³ Patrick J. Keeling,⁴ Gertraud Burger,³ Julius Lukes¹,² and Drahomı́ra Faktorová¹,²*

¹Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice (Budweis), Czech Republic.
²Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), Czech Republic.
³Department of Biochemistry and Robert-Cedergren Centre for Bioinformatics and Genomics, Université de Montréal, Montreal, Canada.
⁴Botany Department, University of British Columbia, Vancouver, Canada.

Summary

*Diplonema papillatum* is the type species of diplonemids, which are among the most abundant and diverse heterotrophic microeukaryotes in the world’s oceans. Diplonemids are also known for a unique form of post-transcriptional processing in mitochondria. However, the lack of reverse genetics methodologies in these protists has hampered elucidation of their cellular and molecular biology. Here we report a protocol for *D. papillatum* transformation. We have identified several antibiotics to which *D. papillatum* is sensitive and thus are suitable selectable markers, and focus in particular on puromycin. Constructs were designed encoding antibiotic resistance markers, fluorescent tags, and additional genomic sequences from *D. papillatum* to facilitate vector integration into chromosomes. We established conditions for effective electroporation, and demonstrated that electroporated constructs can be stably integrated in the *D. papillatum* nuclear genome. In *D. papillatum* transformants, the heterologous puromycin resistance gene is transcribed into mRNA and translated into protein, as determined by Southern hybridization, reverse transcription, and Western blot analyses. This is the first documented case of transformation in a euglenozoan protist outside the well-studied kinetoplastids, making *D. papillatum* a genetically tractable organism and potentially a model system for marine microeukaryotes.

Introduction

Diplonemids are biflagellate protists inhabiting marine ecosystems. Diplonemids, together with symbiontids (a small poorly studied group of species living in anoxic or low-oxygen marine environment, for which no molecular data are available [Breglia et al., 2010]), euglenids (important component of the freshwater ecosystems), and kinetoplastids (including free-living *Bodo*, as well as the highly pathogenic *Trypanosoma* and *Leishmania* spp.), form Euglenozoa. Both the kinetoplastids and euglenids have long been recognized as ecologically omnipresent and species-rich, while diplonemids were considered an insignificant group both in terms of diversity and ecology, and they remained poorly-studied (Lukeš et al., 2015).

The first diplonemid, *Diplonema breviciliata*, was described by Griessman in 1914 (Griessman, 1914), and only less than half a dozen species in two genera have been described since then (Skuja, 1948; Schuster et al., 1968; Porter, 1973; Larsen and Patterson, 1990; Schnepf, 1994; Simpson, 1997; Roy et al., 2007). More recent ecological surveys tracked down two clades of previously unrecognized diplonemids in deep-sea pelagic waters, notably the deep sea pelagic diplonemids clade I and II (DSPD I, II) (Łara et al., 2009). In a comprehensive marine survey, diplonemids emerged even as the 3rd most diverse and 6th most abundant group of oceanic eukaryotes (de Vargas et al., 2015; Lukeš et al., 2015) populating virtually all stations and all depths examined in the comprehensive Tara Oceans expedition (de Vargas et al., 2015; Flegontova et al., 2016). Analysis of the V9 18S rRNA sequence suggests that diplonemids may even be the most species-rich among all marine eukaryotes (Flegontova et al., 2016). At present, only diplonemids from the Diplonema/
Rhynchopus (D/R) clade can be grown axenically and are available from the American Type Culture Collection (ATCC). Representatives of the hyper-diverse DSPD clades are not yet in culture. An initial examination of 10 DSPD isolates has been performed by single-cell genomics (Gawryluk et al., 2016). Among the common structural features of diplonemids is their sac-like shape with cell sizes ranging from 5 to 50 μm, their highly metabolic movement, and the presence of two sub-apical flagella. They possess a single mitochondrial network lining the inside of the plasma membrane and containing a large amount of mitochondrial DNA (mtDNA) (Marande et al., 2005). Whether diplonemids have a parasitic, commensal, or free-living life-style is still an open question. Hence the ecological role of diplonemids in the oceanic ecosystem remains elusive (Lukeš et al., 2015). This important gap in knowledge reflects our ignorance about the ecological functions of marine protists in general (Worden and Wilken, 2016), and the lack of a suitable model system for diplonemids in particular.

Establishing a model organism requires several crucial steps to be fulfilled, the most prominent being their availability in culture, and the ability to genetically modify the species. Ideally, a genetic system includes a vector, a transformation system, and a selectable marker, all of which facilitate expression of introduced genes, including transcription, post-transcriptional processing, and translation. All these steps require extensive experimentation and optimization for each target organism, but once a protocol established, it is generally straightforward to create any range of constructs to address many functional questions.

Only a single diplonemid species, Diplonema papillatum, has been examined at the molecular level in much detail. D. papillatum is a free-living protist with two short heterodynamic flagella that was first isolated from seawater at Friday Harbor, Washington (Porter, 1973). This species is available from the American Type Culture Collection (ATCC), can be easily cultivated axenically in the laboratory at high cell densities, and can be cryopreserved. The mitochondrial genome and transcriptome have been analyzed in detail, revealing novel modes of post-transcriptional gene expression (Marande and Burger, 2007; Kiethega et al., 2013; Moreira et al., 2016; Valach et al., 2016; Faktorová et al., 2018). Further, D. papillatum has been investigated regarding its compartmentalization of gluconeogenesis (Makiuchi et al., 2011; Morales et al., 2016). Finally, the nuclear genome has recently been sequenced and assembled, and functional annotation is under way (unpubl. data).

For all these reasons, establishing a tractable genetic system for D. papillatum and establishing it as a model system would be highly desirable. Here, we describe all key steps necessary for genetic manipulation of this marine flagellate.

Results

Optimization of cultivation and electroporation conditions

In the standard seawater medium (see Experimental Procedures), D. papillatum cultures reach the exponential phase at a cell density of 2–4 \times 10^6 cells/ml. The highest cell density obtained is 6 \times 10^6 cells/ml with a doubling time of 12 h (Fig. 1).

To introduce foreign DNA into D. papillatum, we first tested electroporation as the technique of choice, as it is being successfully and extensively used for transforming Trypanosoma brucei and several other trypanosomatids (Beverley and Clayton, 1993). Since diplonemids and trypanosomatids are closely related and share a number of fine-structural features including a corset of subpellicular microtubules (Roy et al., 2007; Tashyrev et al., 2018), we reasoned that electroporation may also be suitable for D. papillatum.

As a first step, we confirmed that D. papillatum cells survive for at least 10 min in electroporation buffers (Cytomix - used for BTX electroporator or Amaxa - used for Amaxa Nucleofector II) and thus can be electroporated in standard electroporation buffers. Survival rate was >90%. Next, two different electroporation apparatuses were tested: BTX electroporator and Amaxa Nucleofector II (see Experimental Procedures), for which we optimized several parameters, such as the composition of electroporation buffer, the electroporation program, and the amount of DNA used for transformation (Table 1). Specifically, cells were pelletted by a short centrifugation, re-suspended in electroporation buffer, subjected to the electric pulse, transferred back into seawater-containing medium, and then observed under the microscope. While electroporation with BTX killed the majority of cells, the Amaxa procedures yielded more than 20% survival (see Table 1), with cells retaining their original shape, and demonstrating full recovery after a few hours. The Amaxa Nucleofector II
programs proved to be best suited and therefore were chosen for further experiments (Table 1).

Identification of selectable markers

To identify selection markers, we tested the sensitivity of D. papillatum to seven antibiotics—hygromycin, geneticin, phleomycin, puromycin, blasticidin, nourseothricin and tetracycline. With the exception of tetracycline, these drugs have been extensively used for genetic manipulation of T. brucei and related trypanosomatids (http://tryps.rockefeller.edu). The antibiotic concentration for selection of D. papillatum transformants was assessed by determining cell viability by the Alamar Blue assay (Raz et al., 1997).

Diplonema papillatum was found to be sensitive to all the selection markers tested, except tetracycline. The effective concentrations for selection of electroporated D. papillatum are shown in Fig. 2 and Table 2.

The highest antibiotic sensitivity of Diplonema is to puromycin.

Rationale for construct design for Diplonema transformation

Our overall strategy for transformation is to incorporate foreign DNA via genomic integration, since nothing is known about DNA replication in this group and no native plasmids are known. Due to the high number of repetitive regions in the nuclear genome sequence of D. papillatum (unpubl. data), we restricted our selection of genomic regions to target for integration to genes that met several criteria. The genes must (i) be contained inside validated genomic contigs and (ii) be intron-less. Further the corresponding

| Electroporation machine | Used program | Survival rate |
|-------------------------|--------------|---------------|
| BTX                     | 1600V, 25Ω, 50 μF | 10%           |
| Amaxa Nucleofector II   | preset program X-001 | 50%-60%      |
| Amaxa Nucleofector II   | preset program X-014 | 20%-30%      |

Table 1. Used electroporation programs and survival rate of D. papillatum cells.

Fig. 2. Antibiotic sensitivity of D. papillatum. Effect of various concentrations of antibiotics (mg/ml; x axis) on survival of Diplonema cells, as determined by the Alamar blue assay, which measures viability by fluorescence (see Experimental Procedures). The fluorescence intensity corresponding to 1% cell survival is indicated by a dotted horizontal line.
mRNAs must (iii) have a high steady state level (rank among the top hundred by their expression level) and (iv) carry a spliced leader (SL) at their 5’ end (Sturm et al., 2001).

In order to confirm proper integration, the target gene was tagged with a cassette that is composed of a fluorescent protein tag and a resistance gene, both flanked by 5’ and 3’ UTRs from Diplonema. To direct the integration of this cassette into the proper genomic position via homologous recombination, we appended to the cassette appropriate sequences of D. papillatum genome (targeting regions). This strategy, successfully used in numerous model systems (Stretton et al., 1998; Janke et al., 2004; Lai et al., 2010; Wang et al. 2017), should lead to antibiotic-resistant transformants expressing a fluorescently labelled protein that can be detected with a fluorescent microscope or by commercially available antibodies.

**Endogenous N-terminal tagging of α-tubulin**

Based on the above described strategy, we have designed a construct for tagging the N-terminus of the Diplonema α-tubulin gene. The cassette contains the puromycin resistance gene (pac), here called puromycinR (encoding the puromycin N-acetyltransferase of Streptomyces alboniger), which from all tested drugs was selectable at the lowest concentrations (Table 2). Moreover, it also contains the sequence encoding the fluorescent protein mCherry which lacks the stop codon, as we intended to create a fused mCherry-α-tubulin protein. We chose mCherry, which emits red light, to avoid overlap with the quite strong green autofluorescence of D. papillatum (our unpubl. data). The puromycinR-mCherry cassette (~2 kbp) was flanked by D. papillatum sequences including 5’ and 3’ UTRs and homology regions about 500 bp long, which ought to enhance the integration of the construct into the targeted locus. A schematic representation of this construct is shown in Fig. 3A. A similar tagging approach was recently used for cytoskeleton studies in T. brucei (Sheriff et al., 2014).

**Table 2.** Tested resistance markers and their concentration used for selection.

| Antibiotic      | Concentration (µg/ml) |
|-----------------|-----------------------|
| Puromycin       | 20                    |
| Blasticidin     | 50                    |
| Geneticin       | 75                    |
| Hygromycin      | 125                   |
| Nourseothricin  | 400                   |
| Phleomycin      | 500                   |
| Tetracycline    | not sensitive         |

**Introduced heterologous gene can be transcribed, post-transcriptionally processed and translated in Diplonema**

Both NotI-linearized and circular constructs were then electroporated into Diplonema cells in parallel and the transfectants were subjected to selection with increasing concentrations of puromycin to ensure stringent selection.

A total of 10 puromycin-resistant Diplonema clones were recovered after 8–10 days, at which time point a negative-control culture represented by wild type cells without the construct and in the presence of the drug did not display any viable cells.

Clones A3 and A4 (labelled according to their position in the 24-well plate) were investigated in detail by PCR and amplicon sequencing showing the presence of the puromycinR-mCherry cassette in Diplonema genomic DNA (Fig. 3B; Supporting Information Figs. S1 and S4). To verify the cassette’s presence in the genome, we performed Southern blot analysis using radioactively labelled probes against mCherry (Fig. 3C). The expected size of the SpeI restriction fragment containing the mCherry CDS is 1969 bp, which corresponds to the detected band size (Fig. 3C, right). Digestion of genomic DNA with PacI should produce a fragment of 9699 bp in case the cassette has integrated into the expected site by homologous recombination; however, we detected bands of ~3.5 kbp (Fig. 3C, left). We, therefore, conclude that the complete cassette has been incorporated in the two D. papillatum transformants, but at heterologous sites in the genome. In line with this observation, amplification of the cassette with primers outside the cassette failed (data not shown).

**Transfection of Diplonema leads to stable chromosomal integration of foreign DNA**

To examine whether mCherry and puromycinR are transcribed in the D. papillatum transformants A3 and A4, we conducted reverse transcription followed by PCR amplification of the first-strand cDNA (RT-PCR). This experiment produced a single band of expected size confirming transcription of the two heterologous genes (Fig. 4B; Supporting Information Fig. S5A). We also confirmed, by nested RT-PCR, that the mCherry and puromycinR mRNAs are properly processed post-transcriptionally by addition of the SL RNA to their 5’ end (Fig. 4C; Supporting Information Figs. S2 and S5B).

To verify translation of the heterologous genes in A3 and A4 clones, we checked mCherry fluorescence and performed Western blots with an anti-mCherry antibody. However, in both cases no signal was detected, as expected due to the integration of the cassette into a different site. In contrast, translation of the puromycinR gene could be demonstrated by two different immunoassays,
one using the anti-Puromycin antibody, and the other using the anti-Puromycin N-acetyltransferase antibody.

Puromycin is an aminonucleoside antibiotic that functions as an inhibitor of protein synthesis by disrupting peptide transfer on ribosomes, causing premature chain termination during translation. Since its structure resembles the 3’ end of aminoacylated tRNA, puromycin enters the acceptor site of the ribosome and is added to the growing polypeptide chain, which leads to premature termination of the newly synthesized proteins (Pestka, 1971). Anti-Puromycin antibody binds puromycin-containing newly synthesized proteins, visible in Western blot as a smear due to their different molecular weights. Diplonema wild type cells show indeed a strong signal of a broad molecular weight range, while the A3 and A4 clones do not (Supporting Information Fig. S3). This indicates that puromycin is not incorporated in the proteins of the clones A3 and A4, providing indirect evidence for the translation

Fig. 3. Confirmation that the electroporated construct is integrated in the *D. papillatum* genome.
A. Scheme of the puromycin<sup>1</sup> (puro<sup>1</sup>) + mCherry cassette (2995 bp NotI fragment) including restriction sites, positions of the primers and expected sizes of the amplicons.
B. PCR of total DNA of *D. papillatum* wild type (WT) and selected transformants (A3 and A4) using specific primers for amplification of puromycin<sup>1</sup> and mCherry CDSs. Negative control PCR (NC) was performed without template DNA. The explicit sequences are shown in Supporting Information Fig. S1 and the whole gels are shown in Supporting Information Fig. S4.
C. Southern hybridization of total DNA from *D. papillatum* wild type (WT) and transformants A3 and A4 using a DNA fragment of mCherry as a radiolabelled probe. Right panel, total DNA digested with SpeI+NdeI, which cut inside the cassette (+), and undigested (-). The SpeI+NdeI band is of expected size—1969 bp. Left panel, total DNA digested PciI, which cuts outside the cassette (+), and undigested (-). In case of homologous integration, the expected size of the PciI band is 9699 bp, however, we detected bands of ~3.5 kbp.
of the puromycin\textsuperscript{R} gene in the transformants (Supporting Information Fig. S3).

Similarly, Western immunoassays with the anti-Puromycin N-acetyltransferase antibody show clear expression of the heterologous puromycin\textsuperscript{R} gene in A3 and A4, revealing distinct bands of the expected size (Fig. 5; Supporting Information Fig. S6). In both experiments, puromycin-resistant and sensitive cell lines from \textit{T. brucei} served as negative and positive controls.

Discussion

Very few marine protists have been successfully transformed at present, most notably members of the green and red microalgae, diatoms, chlorarachniophytes (for review see Gong \textit{et al}., 2011), and further the alveolate \textit{Perkinsus marinus} (Fernández-Robledo \textit{et al}., 2008), the prasinophyte \textit{Ostreococcus tauri} (van Ooijen \textit{et al}., 2012), the haptophyte \textit{Pleurochrysis carterae} (Endo \textit{et al}., 2016) and recently the kinetoplastid \textit{Parabodo caudatus} (Gomaa \textit{et al}., 2017). Different transfection methods were used, notably polyethylene glycol-mediated approach for the haptophyte, otherwise microprojectile bombardment or electroporation.

\textbf{The challenge of electroporating marine organisms}

Electroporation induces a transient destabilization of the cell membrane, which then becomes highly permeable to
foreign DNA, proteins or small molecules. Electroporation is used particularly for suspension cultures as all cells are essentially transfected simultaneously (Heiser, 2000). For transforming *D. papillatum*, we have opted for electroporation, which is an established technique in kinetoplastids, the sister group of diplonemids (e.g., Adl et al., 2012), specifically in the medically important *Trypanosoma* and *Leishmania* (Beverley and Clayton, 1993).

The technical challenge of electroporating marine organisms is that salts in the medium cause electrical discharge (arching), which reduces the viability of the organism (Potter and Heller, 2003). In the case of *Parabodo caudatus* (Gomaa et al., 2017), electroporation succeeded because this microeukaryote lives in both freshwater and marine environments and is, therefore, only moderate sensitive to low salt concentrations. Fortunately, *D. papillatum* is also tolerant to brackish water, and thus endures short-term exposure to the low-salt electroporation buffer.

**Integration and expression of heterologous genes in the Diplonema nucleus**

For two transformed *D. papillatum* clones, we demonstrated by PCR, DNA sequencing, and Southern blot analysis that the transfection constructs are readily integrated into the nuclear genome. However, the constructs integrated at apparently random genomic positions, a common issue in many genetic systems including mammalian and plant systems (Sargent et al., 1997; Gorbunova and Levy, 1999). Ectopic integration may be due to the fact that the genome of *Diplonema* is highly repetitive, or, alternatively, because homologous recombination (HR) may be less efficient in this organism than microhomology-mediated end-joining (MMEJ) or the classical non-homologous end-joining (NHEJ) DNA repair/recombination pathways (for a review see Michael, 2010). Note that the genes known to be involved in the corresponding machineries are present in the *D. papillatum* genome (our unpubl. data).

Importantly, however, transfected heterologous genes are nevertheless expressed in *Diplonema*. The protein tag (mCherry) and resistance gene (puromycinR) used here are transcribed, and the SL is correctly trans-spliced onto the 5' end of the transcripts. For the puromycinR gene, we were able to demonstrate translation into protein and observe the expected resistant phenotype. In contrast, translation of the mCherry transcript was neither observed nor expected, because a stop codon and a 3' UTR were absent, as it was designed to be fused to the α-tubulin gene as an N-terminal tag. Due to the lack of a stop codon and/or poly-(A) tail, the mRNA may have been degraded by one of the quality control mechanisms of the eukaryotic cell (Klauser and van Hoof, 2012).

Since the purpose of this construct was to create a fused protein composed of N-terminally tagged α-tubulin with mCherry, the stop codon at the end of the mCherry gene was intentionally removed. However, as this construct was not integrated into a proper position, no expression of mCherry could be observed.

An alternative approach to obtain resistant cell line emitting red fluorescence would be to design a cassette containing intact genes and corresponding UTRs to replace an endogenous gene such as tubulin, or insert the cassette into a particular position in the *Diplonema* genome, f.e. in a transcriptionally silent region, a routinely strategy in other organisms. While this approach would not produce a labelled endogenous protein, it would allow the expression of the mCherry protein. In our future work, we will apply different strategies with the aim to achieve correct integration of the cassette.

**Future prospects for *D. papillatum* as a genetic system**

Taken together, these data show that *D. papillatum* has all prerequisites for becoming a genetically tractable organism. Successful transformation provides first example of heterologous gene expression in an euglenozoan protist outside the kinetoplastids, and raises questions about how widely these methods can be applied within the group, and what other tools can now be used in model diplonemids. Currently, only half a dozen diplonemids are available in the ATCC collection and can be cultivated in the laboratory. For the time being, with a representative of the species-rich DSPD I clade yet to be brought into culture, the next candidate for transformation is *Hemistasia phaeocystica* (Yabuki and Tame, 2015; Yabuki et al., 2016), which is more closely related to the most abundant diplonemids. However, this species is also much more challenging to work with, as it prefers live diatoms as a food source, reaches only low cell densities and so far could not be cryopreserved (our unpubl. data).

The availability of a methodology to transfect *Diplonema* will also facilitate investigation of the machineries that drive the unique post-transcriptional processes in their mitochondria, such as U-appendage RNA editing and trans-splicing of fragmented genes (for reviews see Valach et al., 2016; Faktorová et al., 2018). For example, fluorescence tagging of terminal uridytransferases or RNA ligases would identify the respective enzyme that acts in mitochondria, and open the venue for uncovering other components in the hypothetical editosome or trans-spliceosome.

Future work may achieve targeted integration by further extension of the 5' and 3' homologous regions of the constructs to more than 1500 bp, as shown in trypanosomes (Barnes and McCulloch, 2007). Furthermore, it would be worthwhile to explore the CRISPR/Cas9 strategy, which was recently successfully implemented in kinetoplastids (Lander et al., 2016; Beneke et al., 2017). In this strategy, a circular plasmid that is in the transformed organism
Transformation of a diplonemid flagellate

Experimental procedures

Strains, cultivation and growth curves

D. papillatum (ATCC 50162) was cultivated axenically at 27°C in an artificial sea salt mixture 40 g/l (Sigma, S9883), 0.1% (w/v) tryptone, 1% (v/v) fetal bovine serum and 100 µg/ml of chloramphenicol. Cell density was measured manually by the Neubauer cell chamber. Before measurement, cells had to be fixed in 3.7% (v/v) formaldehyde in SSC to retain their shape and to prevent them from moving prior to counting.

Determination of resistance to antibiotics using Alamar Blue assay

The inhibition concentration where 99% of the cell population is dead (IC99) was determined using the fluorescence viability indicator Alamar blue (Resazurin sodium salt, Sigma, R7017) as described in Gould et al. (2013). It is based on resazurin, a non-toxic, permeable and weakly fluorescent dye used as a redox indicator. Its reduced form or resorufin is pink and highly fluorescent, with its fluorescence intensity being proportional to the number of respiring (e.g., metabolically active) cells. Cells were inoculated into the 96-well flat-bottomed microtiter plate (Costar) at a concentration 1 × 10^5 cells/ml. Seven drugs (hygromycin, puromycin, phleomycin, geneticin, blasticidin, nourseothricin and tetracycline) were tested in triplicates with inhibitors of the non-homologous end joining pathway using inhibitors of the non-homologous end joining pathway, as recently described in Cryptococcus neoformans (Arras et al., 2016). As a high quality genome and transcriptome of D. papillatum shall soon be available, it would be beneficial to have a transformation protocol in place. We envisage its application in functional analysis of proteins with potential ecological significance, such as those involved in pathways that respond to environmental stress.

Electroporation and obtaining of the transformants

A total of 5 × 10^7 cells (2 × 10^6 cells/ml) was harvested by centrifugation at 1300 g for 10 min at 25°C and re-suspended in 100 µl of cytoxim buffer (van den Hoff et al., 1992) and electroporated with the BTX machine (1600V, 250, 50 µF), or in 100 µl of AMAXA buffer (81.8 µl of Human T-cell nucleofector solution + 18.2 µl of Supplement) for the electroporation by Amaxa Nucleofector II.

Ten to 15 µg of DNA were mixed with electroporation buffer and Diplonema cells before transferring into the electroporation cuvette and electroporation. After the electroporation pulse was applied to the cuvette, the mixture was immediately transferred into 10 ml of Diplonema growth media. Subsequently, the cells were allowed to recover for about 8 h and the transfectants were subjected to selection with increasing concentrations (12–40 μg/ml) of puromycin. While this wide range makes the experiment more time consuming, it ensured stringent selection of transformants. Clones A3 and A4 examined in details were cultivated in 24 and 28 μg/ml of puromycin, respectively.

Following an expansion of each clone to a volume of 20 ml transformants had been cultured for up to 8 weeks prior to testing by PCR, which proved that all of them indeed contain integrated constructs. We were able to freeze-store the transformants in 10% glycerol-containing medium at −80°C or in liquid nitrogen for several months.

PCR using genomic DNA

Genomic DNA was isolated using Qiagen DNA isolation kit (Qiagen, 69504). Primer pairs used for verification of the integration are shown in Fig. 3A, for primer sequences see Supporting Information Table S1. PCR amplification was done using OneTaq polymerase (NEB Biolabs, M0486L) and the following program: initial denaturation 94°C for 3 min, denaturation at 94°C for 30 s, annealing at 58°C for 60 s, extension at 68°C for 3 min and a final extension at 68°C for 10 min, for 30 cycles. To amplify mCherry and puromycin, a lower extension of 68°C for 1 min 30 s and a final extension of 68°C for 5 min were used. Cassette integration in the genome of D. papillatum was confirmed by sequencing of the PCR products (Eurofins Genomics).

Design and preparation of transformation cassette

Design of the puromycin cassette

The following fragments ordered in the 5′–3′ direction, starting with a NotI restriction site: 499 nt of the sequence just before the beginning of α-tubulin gene DpTUB1 including its 5′ UTR, puromycin N-acetyl transferase gene (puromycin) with the stop codon, 3′ UTR of DpRPL11, 5′ UTR of DpRPL14, a TY tag and the mCherry CDS followed by a linker, 510 nt of the 5′ of the α-tubulin gene DpTUB1, followed by a second NotI site. The cassette was synthesized by Biomatik (http://www.biomatik.com), cloned into pBluescript II SK(+) vector and the sequence was submitted to GenBank (Acc. No. MG490656). The vector was isolated using QiAprep Spin Miniprep Kit (Qiagen, 27106) and either directly used for electroporation or cut with NotI restriction enzyme, isolated from the gel using QiAquick Gel Extraction Kit (Qiagen, 28706), and re-suspended in 10 µl of deionized distilled water before electroporation.

Design of the puromycinR cassette

Design and preparation of transformation cassette

The following fragments ordered in the 5′–3′ direction, starting with a NotI restriction site: 499 nt of the sequence just before the beginning of α-tubulin gene DpTUB1 including its 5′ UTR, puromycin N-acetyl transferase gene (puromycin) with the stop codon, 3′ UTR of DpRPL11, 5′ UTR of DpRPL14, a TY tag and the mCherry CDS followed by a linker, 510 nt of the 5′ of the α-tubulin gene DpTUB1, followed by a second NotI site. The cassette was synthesized by Biomatik (http://www.biomatik.com), cloned into pBluescript II SK(+) vector and the sequence was submitted to GenBank (Acc. No. MG490656). The vector was isolated using QiAprep Spin Miniprep Kit (Qiagen, 27106) and either directly used for electroporation or cut with NotI restriction enzyme, isolated from the gel using QiAquick Gel Extraction Kit (Qiagen, 28706), and re-suspended in 10 µl of deionized distilled water before electroporation.

Design and preparation of transformation cassette

The following fragments ordered in the 5′–3′ direction, starting with a NotI restriction site: 499 nt of the sequence just before the beginning of α-tubulin gene DpTUB1 including its 5′ UTR, puromycin N-acetyl transferase gene (puromycin) with the stop codon, 3′ UTR of DpRPL11, 5′ UTR of DpRPL14, a TY tag and the mCherry CDS followed by a linker, 510 nt of the 5′ of the α-tubulin gene DpTUB1, followed by a second NotI site. The cassette was synthesized by Biomatik (http://www.biomatik.com), cloned into pBluescript II SK(+) vector and the sequence was submitted to GenBank (Acc. No. MG490656). The vector was isolated using QiAprep Spin Miniprep Kit (Qiagen, 27106) and either directly used for electroporation or cut with NotI restriction enzyme, isolated from the gel using QiAquick Gel Extraction Kit (Qiagen, 28706), and re-suspended in 10 µl of deionized distilled water before electroporation.
Southern blot hybridization

Southern blot analysis was performed as previously described (Lai et al., 2008). Total DNA was isolated from wild type cells and transformed clones, digested with selected restriction enzymes and subjected to agarose gel electrophoresis. Samples were blotted overnight to a Zeta-Probe membrane (Bio-rad) by capillarity and subsequently cross-linked with UV light as described (Vondrůšková et al., 2005). The mCherry probe was PCR amplified using mCherry_Fw and mCherry_Rv primers and radiolabelled with [α-32P]dATP using Radioactive DNA Labelling kit (Thermo Scientific DecaLabel DNA Labelling Kit, #K0622) and hybridization with the probe was performed overnight at 60°C according to the Zeta-Probe membrane manual. The membrane was exposed to a Fuji Imaging phosphor screen and hybridization with the probe was performed overnight at 0.5% (v/v) Tween 20 and probed with the primary monoclonal antibody conjugated with horseradish peroxidase (1:1000). Monoclonal anti-α-Tubulin antibody produced in mouse (1:1000) (Sigma, T9026) was used as a loading control.

Acknowledgements

We thank Jorge Morales (Heinrich Heine University Düsseldorf) for discussion in the initial stages of the project; Anzhelika Butenko and Pavel Flegontov (Institute of Parasitology) for help with the selection of genes and design of the constructs; Galina Prokopchuk and Daria Tashyreva (Institute of Parasitology) for sharing unpublished data. Support from the Grant Agency of University of South Bohemia (050/2016/P to BK), the Czech Grant Agency (15–219745 and 16–18699S to JL), the ERC CZ (LL1601 to JL), the Canadian Institute of Health Research (CIHR, MOP 70309 to GB) and from the Gordon and Betty Moore Foundation (GBMF4983.01 to JL, GB and PK) are kindly acknowledged.

Conflict of Interest

Authors have no conflict of interest to declare.

References

Adl, S.M., Simpson, A.G.B., Lane, C.E., Lukeš, J., Bass, D., Bowser, S.S., et al. (2012) The revised classification of eukaryotes. J Eukaryot Microbiol 59: 429–493.
Arras, S.D.M., Fraser, J.A., and Lustig, A.J. (2016) Chemical inhibitors of non-homologous end joining increase targeted construct integration in Cryptococcus neoformans. PLoS One 11: e0163049.
Barnes, R.L., and McCulloch, R. (2007) Trypanosoma brucei homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases. Nucleic Acids Res 35: 3478–3493.
Beneke, T., Madden, R., Makin, L., Valli, J., Sunter, J., and Gluenz, E. (2017) A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. R Soc Open Sci 4: 170095.
Beverley, S.M., and Clayton, C.E. (1993) Transfection of Leishmania and Trypanosoma brucei by electroporation. Methods Mol Biol 21: 333–348.
Breglia, S.A., Yubuki, N., Hoppenrath, M., and Leander, B.S. (2010) Ultrastructure and molecular phylogenetic position of a novel euglenozoan with extrusive episymbiotic bacteria: Bihospites bacati n. gen. et sp. (Symbiontida). BMC Microbiol 10: 145.
de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., et al. (2015) Eukaryotic plankton diversity in the sunlit global ocean. Science 348: 1261605.
Endo, H., Yoshida, M., Uji, T., Saga, N., Inoue, K., and Nagasawa, H. (2016) Stable nuclear transformation system for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Faktorová, D., Valach, M., Kaur, B., Burger, G., and Lukeš, J. (2018) Mitochondrial RNA editing and processing in diplommid protists. “RNA Metabolism in Mitochondria,” Springer series “Nucleic Acids and Molecular Biology,” in press.

Fernández-Robledo, J.A., Lin, Z., and Vasta, G.R. (2008) Transfection of the protozoan parasite *Perkinsus marinus*. *Mol Biochem Parasitol* 157: 44–53.

Flegontova, O., Flegontov, P., Malviya, S., Audic, S., Wincker, P., de Vargas, C., et al. (2016) Extreme diversity of diplommid eukaryotes in the ocean. *Curr Biol* 26: 3060–3065.

Gawryluk, R.M.R., del Campo, J., Okamoto, N., Strassert, J.F.H., Lukeš, J., Richards, T.A., et al. (2016) Morphological identification and single-cell genomics of recombinant diplommids. *Curr Biol* 26: 3053–3059.

Gong, Y., Hu, H., Gao, Y., Xu, X., and Gao, H. (2011) Microalgae as platforms for production of recombinant proteins and valuable compounds: progress and prospects. *J Ind Microbiol Biotechnol* 38: 1879–1890.

Gomaa, F., Garcia, P.A., Delaney, J., Girguis, P.R., Buie, C.R., and Edgcomb, V.P. (2017) Toward establishing model organisms for marine protists: successful transfection protocols for *Parabodo caudatus* (Kinetoplastida: Excavata). *Environ Microbiol* 19: 3487–3499.

Gorbunova, V.V., and Levy, A.A. (1999) How plants make less, phagotrophic Euglenozoon with concealed flagella. *Eukaryot Cell* 89: 1–78.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.

Heiser, W.C. (2000) Optimizing electroporation conditions for the coccolithophorid alga *Pleurochrysis carterae*. *Sci Rep* 6: 22252.
with the reference sequence. Beginning parts of the amplicon sequences are aligned to the reference sequence.

Fig. S1. List of primers used.

**Table S1.** PCR amplicons (from Fig. 3B) obtained with primers that amplify the puromycin<sup>R</sup> region (A) or mCherry region (B) of clones A3 and A4 were verified by sequencing. Beginning parts of the amplicon sequences are aligned with the reference sequence.

**Fig. S2.** SL-PCR amplicons (from Fig. 4C) were verified by sequencing. Part of the amplicon sequences obtained with primers that amplify the puromycin<sup>R</sup> region (A) or mCherry region (B) are aligned to the reference sequence.

**Fig. S3.** Western blot analysis of wild type (WT) *D. papillatum* and transformants A3 and A4 after 24 h incubation in a medium containing 20 μg/ml puromycin. *Trypanosoma brucei* SMOX P9 procyclic stage cells (SmOX) expressing the puromycin<sup>R</sup> gene were used as a positive control. Monoclonal mouse anti-Puromycin antibodies (1:1000) and secondary anti-mouse antibodies coupled to horseradish peroxidase (1:1000) were used for visualization employing the ECL kit.

**Fig. S4.** Confirmation that the electroporated construct is integrated in the *D. papillatum* genome. (A) Scheme of the puromycin<sup>R</sup> (puro<sup>R</sup>) + mCherry cassette (2995 bp NcoI fragment) including restriction sites, positions of the primers and expected sizes of the amplicons. The whole gels with PCR of total DNA of *D. papillatum* wild type (WT), selected transformants (A3 and A4; underlined) and some other obtained transformants (A2, A5, B4, B2, C2) using specific primers for amplification of puromycin<sup>R</sup> (B), mCherry (C) or puromycin<sup>R</sup> + mCherry (D). Negative control PCR (NC) was performed without template DNA.

**Fig. S5.** Validation of proper transcription and post-transcriptional 5' end-processing of the transcripts produced from heterologous genes in *D. papillatum*. (A) The whole gels of RT-PCR shown in Fig. 4B. RNA from *D. papillatum* wild type (WT) and selected transformants (A3 and A4; underlined) and some other obtained transformants (C3, B3, B4, A5) was used as a template for RT-PCR. Expected sizes of the RT-PCR products are indicated. Reactions with reverse transcriptase added and negative controls without reverse transcriptase are indicated by (+) and (-). (B) The amplification of the 5' region from the puromycin<sup>R</sup> (puro<sup>R</sup>) transcript and the mCherry transcript respectively. Nested SL RT-PCR using total RNA from the transformants A3 and A4 as templates, and two sets of primers that anneal to the conserved SL sequence of *D. papillatum* (forward primers) and to the 5' end of the transcripts (reverse primers), see Fig. 4A. The whole gels of both SL RT-PCR reaction are shown here together with the sizes of expected product.

**Fig. S6.** Whole gels of western blot analysis (shown in Fig. 5) of *D. papillatum* wild type (WT) and transformants A3 and A4 that express the puromycin<sup>R</sup> gene (puromycin N-acytlytransferase). Monoclonal rabbit anti-Puromycin N-acetytyltransferase antibodies (1:500) and secondary anti-rabbit antibodies coupled to horseradish peroxidase (1:1,000). 29-13, *T. brucei* procyclic-stage cell line 29-13, which does not express the puromycin<sup>R</sup> gene, is used as a control. SMOX, *T. brucei* cell line SMOX P9, which does express the puromycin<sup>R</sup> gene, is used as a positive control. Anti-α-tubulin antibodies were used as a loading control.