

Elsevier required licence: © <2019>.

This manuscript version is made available under the CC-BY-NC-ND 4.0 license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

The definitive publisher version is available online at

<https://www.sciencedirect.com/science/article/pii/S2211926419305764?via%3Dihub>

**Novel endogenous promoters for genetic engineering of the marine microalga
Nannochloropsis gaditana CCMP526**

Margaret Ramarajan^a, Michele Fabris^{a, b *}, Raffaella M. Abbriano^a, Mathieu Pernice^a, Peter J. Ralph^a

^a University of Technology Sydney, Climate Change Cluster (C3), Faculty of Science, Broadway Campus, Ultimo NSW 2007, Australia

^b CSIRO Synthetic Biology Future Science Platform, GPO box 2583, Brisbane QLD 4001, Australia

* Corresponding author at: University of Technology Sydney, Climate Change Cluster (C3), Faculty of Science, Broadway Campus, Ultimo, NSW, 2007, Australia

Email: Michele.Fabris@uts.edu.au

Abstract

Nannochloropsis is a marine microalga from the *Eustigmatophyceae* stramenopile lineage that has been studied extensively due to a broad range of industrial applications, mostly related to their oil and pigment production. However, tools to genetically engineer members of this group, and therefore further understand and maximise their industrial potential are still limited. In order to expand the potential industrial uses of this organism, several molecular tools, including gene promoters of different strength, are needed. A comprehensive and diverse set of well-characterized promoters is key to a number of genetic engineering and synthetic biology applications, such as the assembly of complex biological functions or entire metabolic pathways.

In this study, we measured the promoter activity of three endogenous constitutive promoters from *N. gaditana* genes *EPPSII* (*Nga02101*); *HSP90* (*Nga00934*); *ATPase* (*Nga06354.1*) in driving the expression of a *Sh ble-mVenus* fluorescent reporter fusion protein. Through a combined approach that includes flow cytometry, RT-qPCR and immunoblotting, we profiled the activity of these promoters at both the transcript and protein level. Two promoters *HSP90* (*Nga00934*) and *EPPSII* (*Nga02101*) outperformed the widely used *β -tubulin* promoter, exhibiting 4.5 and 3.1-fold higher mVenus fluorescence, respectively. A third promoter *ATPase* (*Nga06354.1*) was also able to drive the expression of transgenes, albeit at lower levels. We show that the new promoters identified in this study are valuable tools, which can be used for genetic engineering and functional genetics studies in *N. gaditana*.

Keywords

Promoter, nuclear transformation, genetic engineering, *Nannochloropsis*, algae biotechnology

1. Introduction

Nannochloropsis gaditana is an *Eustigmatophyceae* (Stramenopiles) marine microalga [1] that has been extensively studied for its potential use in biofuel production [2–4] because it naturally accumulates high levels of triacylglycerol [5]. *N. gaditana* also produces valuable pigments such as violaxanthin, canthaxanthin, astaxanthin and zeaxanthin [6]. Moreover, it is a rich source of polyunsaturated fatty acids such as eicosapentaenoic acid (EPA) [7], hence it is often employed in nutraceutical production [8–10] and as a feedstock in aquaculture [11].

Genetic engineering is a powerful approach for industrial microbial applications, as it enables the editing of existing traits as well as the installation of novel new ones in production organisms. *Nannochloropsis* species are good platform for genetic engineering for several reasons: (1) they have an haploid genome [3,12,13], which simplifies genome editing of non-essential genes as only one chromosome must be modified; (2) their genome is fully sequenced [12–15]; (3) omics datasets are available [16,17]; (4) nuclear transformation is feasible [12,13,18–22]; and (5) homologous recombination (HR) is efficient [21] and can be exploited for targeted gene editing [3]. These advances in phycological research have laid an excellent foundation to engineer this species as an industrially relevant strain to produce high-value compounds.

However, unlike other model algal species that are rapidly entering the field of molecular biotechnology and synthetic biology, such as *Chlamydomonas reinhardtii* and *Phaeodactylum tricorutum* [23,24], there is a less well developed and diversified set of resources to genetically engineer *Nannochloropsis*. A wider array of molecular tools, including gene promoters are needed to enable complex metabolic engineering or synthetic biology strategies in *Nannochloropsis* [25]. A diverse set of promoters of different strength is

essential for fine-tuning the expression of transgenes and entire metabolic pathways, and can circumvent issues such as gene silencing due to promoter repetition, as previously reported in higher plants [26–28].

Several constitutive endogenous promoters have been identified in *Nannochloropsis* species [25] that drive the expression of antibiotic resistance genes (e.g. *Sh ble*) [12] and reporter genes (e.g. firefly *luciferase*, *NanoLuc* from a deep sea shrimp [29], and *Venus* [30]). To date, seven endogenous promoters have been successfully used for transgene expression in *N. gaditana* [12,31]. Of these, three endogenous promoters have been identified in *N. gaditana* strain CCMP526, which regulate the expression of genes putatively annotated as β -tubulin (*TUB*, *Nga00092*), heat shock protein 70 (*HSP*, *Nga07210*), and the ubiquitin extension protein (*UEP*, *Nga02115.1*) [12]. Although these were shown to effectively drive the expression of heterologous genes in *N. gaditana*, they have not been thoroughly profiled and characterized [12,31].

In this work, we set out to identify and study the activity of novel *N. gaditana* endogenous promoters that could efficiently drive the expression of transgenic constructs. We mined the publicly available transcriptomic data sets of *Nannochloropsis oceanica* CCMP1779 [16] to identify constitutively expressed genes, searched for their orthologs in *N. gaditana*, cloned their upstream putative promoter regions, and profiled their activity *in vivo* in *N. gaditana* at both the transcript and protein level. Here, we report the isolation and characterization of three novel promoters putatively annotated as *heat shock protein 90/ Molecular chaperon HtpG* (*HSP90*, *Nga00934*), *extrinsic protein in photosystem II* (*EPPSII*, *Nga02101*), and *P-type H⁺ ATPase* (*ATPase*, *Nga06354.1*). We parameterized the activity of these promoters *in vivo* with reporter genes, and determined that they drive transgene

expression at different levels in *N. gaditana*. We anticipate that these novel promoters represent useful resources to expand the genetic toolkit available for genetic engineering of this microalga, both for functional genetics studies and biotechnological applications.

2. Materials and methods

2.1. Cell culture

N. gaditana CCMP526 strain was purchased from the Provasoli-Guillard National Centre for Culture of Marine Phytoplankton, USA. The cells were grown in liquid modified F/2 medium [32] where the seawater component was replaced with Enriched Seawater, Artificial Water (ESAW) [33] and the vitamins concentration was doubled (primary vitamin stock solution: Thiamine · HCl- 400 mg, Biotin- 2 g/L) . Axenic cultures were maintained under cool white LED light panel set to a light intensity of $100 \mu\text{mol photons m}^{-2} \text{s}^{-1}$, with a photoperiod of light: dark 12hr: 12hr, at 21°C in a shaker incubator (Kühner, Switzerland). Batch cultures were maintained in 250 mL Erlenmeyer flasks closed with foam plugs to allow gas exchange. Cultures were checked for contamination every two months by streaking on marine broth agar plates. All transformation and promoter analysis experiments were carried out in the above mentioned conditions, but under continuous light regime.

2.2. Identification of promoters of high expressing genes in *N. gaditana*

To identify genes with constitutive high expression, a published transcriptomic data set of *Nannochloropsis oceanica* CCMP1779 [16] was used, as this closely represents commonly used laboratory growth conditions without nutrient stresses, and such dataset is currently not available for *N. gaditana*. Nine genes with Fragments Per Kilo base of exon per Million fragments mapped (FPKM) values greater than 60 across the different time points

were selected (Table 1) and their corresponding protein sequences were extracted from the Michigan State University, *Nannochloropsis oceanica* CCMP1779 resources [34]. The orthologs of these genes were obtained using the web-based *N. gaditana* BLASTp search tool using default parameters, in the CRIBI Genomics *Nannochloropsis* genome portal [35]. The β -tubulin TUB (*Nga00092*) promoter was selected as a control promoter for this study because of its utility in driving the constitutive expression of transgenes in *N. gaditana* has been demonstrated [12].

2.3. Gene expression analysis by Real-Time quantitative PCR (RT-qPCR)

Cell pellets from 50 mL of early log phase cultures of *N. gaditana* CCMP526 (1×10^7 cells mL⁻¹) were washed with 5 mL 1% PBS, re-pelleted at 3500 g for 10 mins, and flash-frozen in liquid nitrogen. Frozen cells were lysed with a tissue lyser (TissueLyser LT, Qiagen, Germany) with 2 RNase-free 1.6 mm stainless steel beads at 50 Hz for 2 mins at 0°C. Total RNA was extracted using Tri-reagent (Trizol, Thermo Fisher Scientific, USA) and an RNAeasy mini purification kit (Qiagen, Germany) including a gDNA elimination step following manufacturer's instructions. The RNA quantity and quality was assessed using a spectrophotometer (NanoDrop 2000, Thermo Fisher Scientific, USA) and RNA samples were stored at -80 °C for further RT-qPCR experiments. A total of 280 ng of RNA was used for each sample for cDNA synthesis using iScript™ gDNA clear cDNA Synthesis Kit (Biorad, USA). An Eppendorf epMotion® 5075 (Eppendorf, Germany) robot was programmed to prepare the RT-qPCR reactions in a Hard-Shell® 384-well, thin wall, skirted, white PCR plate (Biorad, USA). The RT-qPCR reaction was carried in a CFX384 Touch™ Real-Time PCR Detection System (Biorad, USA) with the following conditions: incubation at 95°C for 10 min, then 44 cycles of 95°C, 60°C, 68°C for 30 s each, followed by a melt curve to verify specific amplification of the target

amplicon. RT-qPCR reactions consisted of 10 μ L total volume with 4 μ L of cDNA and 6 μ L iTaq™ Universal SYBR® Green Supermix (Biorad, USA) together with gene specific primers of final concentration 300 nM (Table S1). RT-qPCR analysis was carried out in technical triplicate using the CFX Manager™ 3.1 software. RT-qPCR efficiency for each primer pair was determined from a calibration curve with a cDNA dilution gradient of 243, 81, 27, 9 and 3 ng and a linear regression model [36]. The corresponding real-time PCR efficiencies were calculated according to the equation described by Radonic et al [37]. All qRT-PCRs displayed efficiencies between 92% and 115% (R^2 of calibration curve >0.92), well within a range considered as valid (85%-115%) as previously described [38–41]. RT-qPCR primers used to amplify the selected orthologous genes from *N. gaditana* are given in Table S1 and their primer efficiency, calibration curve's R^2 , amplicon length and T_m are given in Table S2. A no template control as well as a no reverse transcription control was generated for each primer pair and RT-qPCR run to ensure that the cDNA samples were free of DNA contamination.

For quantitative real-time PCR analysis of *N. gaditana* transformants, a total of 205 ng of RNA was used for each sample for cDNA synthesis using iScript™ gDNA clear cDNA Synthesis Kit (Biorad, USA). Data analysis was done in single threshold mode. RT-qPCR primers used to amplify *mVenus* and β *actin* genes are given in Table S1 and their primer efficiency was 103.8% with a R^2 of 0.994. The starting quantity (SQ) of *mVenus* was normalized using β *actin* SQ (which has been used as a housekeeping gene in previous studies [42]) to give the relative abundance of *Sh ble-mVenus* mRNA in all the samples analyzed.

2.4. Construction of plasmids pNaga4.mVenus and pNaga4.empty

Two transformation vectors, pNaga4.empty and pNaga4.mVenus (Fig. S1 and S2, respectively), were designed *in silico* using the NEBuilder Assembly Tool v1.12 18 software.

Phusion High Fidelity DNA polymerase (New England Biolabs, USA) was used to PCR amplify DNA fragments and the pUC19 backbone as per manufacturer's protocol using primers listed in Table S1. A PureLink Quick Gel Extraction kit (Invitrogen, Life Technologies, USA) was used to purify the vector and PCR insert fragments from 1% agarose gels. The following fragments were assembled into the pUC19 backbone using the Gibson Assembly[®] Master Mix (New England Biolabs, USA): *TUB* (Nga00092) promoter region, *Sh ble* gene, His-tag, multi-cloning site, Foot and Mouth Disease Virus (FMDV) self-cleaving 2A fragment, *mVenus* gene, and *TUB* (Nga00092) terminator region.

DH5-alpha *Escherichia coli* electrocompetent cells (New England Biolabs, USA) were used for DNA cloning, and were grown in Luria broth (Millers LB base) medium (Invitrogen, Life Technologies, USA) with 100 µg mL⁻¹ ampicillin. Plasmid DNA was isolated using the Zippy Plasmid Miniprep kit (Zymo Research, USA). The pNaga4.mVenus vector sequence was confirmed using restriction digestion using *EcoRI*, *BamHI*, and *HindIII* followed by DNA sequencing (Macrogen, South Korea).

2.4.1. Cloning of the putative promoter regions into pNaga4.mVenus

Five genes (*Nga01608*, *Nga21005*, *Nga02101*, *Nga00934*, *Nga06354.1*) with the highest relative mRNA levels measured using RT-qPCR were selected. The upstream putative promoter regions (600 bp–900 bp) of these genes were analyzed using PlantCARE [43] and Softberry TSSP [44] software and were amplified from *N. gaditana* genome using specifically designed primers with *PvuII* and *SacI* restriction sites (Table S1). The amplified putative promoter regions were purified using a PureLink Quick Gel Extraction kit (Invitrogen, Life Technologies, USA). The purified fragments were cloned into restriction digested pNaga4.mVenus replacing the *TUB* (Nga00092) promoter region at the *PvuII* and *SacI* sites

using T4 DNA ligase (New England Biolabs, USA). The *E. coli* clones were screened by colony PCR using GoTaq Flexi polymerase (Promega, USA) and promoter-specific primers (Table S1). The vector was sequenced to confirm the correct sequence (Macrogen, South Korea).

2.5. *N. gaditana* transformation

As the transformation of *N. gaditana* CCMP526 strain was not successful using previously published electroporation protocols [12,21,45], multiple standardization experiments (Table S3) were carried out, and some modifications to the culture conditions and pellet preparation were made as described below. *N. gaditana* CCMP526 strain cultures were grown to early log phase (approx. 2×10^7 cells mL⁻¹) at 21°C in continuous light (100 μ mol photons m⁻² s⁻¹) using 50% initial culture inoculum in modified F/2 medium, containing half the concentration of NaCl, and sub-cultured three times after 48 hrs of growth. 300 mL of culture was taken and centrifuged at 4000 g for 15 mins at 4°C. The cell pellet was washed three times in ice-cold 374 mM sorbitol and re-suspended to a final concentration of 2.5×10^9 cells mL⁻¹ in 375 mM sorbitol. Concentrated culture (200 μ L) was transferred to a 0.2 mm electroporation cuvette and 2 μ g of linearized (using *PvuII* or *NdeI*) transformation vector was added. The cuvette was incubated on ice for 10 mins before electroporation. A Gene Pulser Xcell™ Electroporation System (Biorad, USA) was used to electroporate the samples at 2.2 KV, 600 Ω , and 50 μ F. The pulse duration observed was 18—25 ms. The cells were transferred to 5 mL modified F/2-Si medium containing half the concentration of NaCl in a 50 mL tube and incubated for 48 hrs in a shaker incubator at 21°C under continuous low light (70 μ mol photons m⁻² s⁻¹) for recovery. The algal cell cultures were centrifuged at 3500 g for 10 mins at 4°C, plated onto half-salt modified F/2-Si plates with zeocin (5 μ g mL⁻¹) and incubated for 3 weeks until appearance of colonies.

To evaluate the transformation efficiency among constructs, the growth stage selected for cell harvest (log phase, OD₇₅₀ 0.3, approx. 1x10⁷ cells mL⁻¹), the final cell concentration (2–2.5x10⁹ cells mL⁻¹), the sorbitol wash process, and electroporation conditions were kept constant. The variables being the linearized plasmid amount (2–4.5 µg/electroporation) and the promoter constructs used. The transformation efficiencies were determined using the formula:

$$\text{Transformation efficiency} = \frac{(\text{Colonies}) / (\text{Cells plated out})}{(\text{Amount of plasmid in } \mu\text{g})}$$

Individual transformant colonies were placed into a Falcon[®] 96 Well Clear Assay plate, round bottom (Corning, USA) containing half-salt modified F/2-Si containing zeocin (5 µg mL⁻¹) and incubated at 21°C under 100 µmol photons m⁻² s⁻¹ light intensity for 4–5 days and subcultured thrice. Analysis of cells was performed by using a flow cytometer (CytoFLEX S, Beckman Coulter Life sciences, USA) to check for mVenus fluorescence using a 488 nM laser and 525/40 optical filter. PCR was performed with mVenus specific primers (Table S1) on three transgenic cell lines per promoter construct that possessed the strongest mVenus signal, to show the presence of the reporter gene (*mVenus*) in the genome of transformed *N. gaditana* using the PCR conditions as per manufacturer's protocol. *β-actin* primers (Table S1) were used as an internal positive control (Fig. S3).

2.6. Flow cytometry analysis of *N. gaditana* transformants

The selected *N. gaditana* clones expressing *Sh ble-mVenus* under the control of *EPPSII* (*Nga02101*); *HSP90* (*Nga00934*) and *ATPase* (*Nga06354.1*) promoters were placed into a Falcon[®] 96 Well Clear Assay plate (200 µL), round bottom (Corning, USA) in triplicate. Flow cytometric analysis of cells was performed once every day to measure mVenus fluorescence using a 488 nM laser and 525/40 optical filter and chlorophyll measurement using the 690/50

band pass filter. An empty vector control was used to compare between the chlorophyll and mVenus fluorescence. The gating used to select the mVenus positive cells is shown in Figure S4. The mean mVenus fluorescence intensity and the percentage of cells expressing mVenus fusion protein was calculated using *N. gaditana* wild-type strain and *N. gaditana* transformant without mVenus protein as negative controls. The net mVenus fluorescence was calculated as a product of mVenus intensity/cell and number of cells expressing mVenus fusion protein.

2.7. Immunoblotting

Nannochloropsis cell cultures (30 mL) were harvested by centrifugation at 4000 g for 15 mins, supernatant was removed, and pellet washed in 1 mL phosphate buffer saline (PBS). The cells were centrifuged again at 4000 g for 10 mins and the supernatant discarded. The pellet was re-suspended in lysis buffer (50 mM Tris HCl, 400 mM NaCl, 0.5% tween20 at pH 8) with 1X protease inhibitor cocktail (Sigma-Aldrich, USA) and frozen in liquid nitrogen for one hour. The samples were thawed and 200 µg of 2 mm glass beads were added. A TissueLyser LT (Qiagen, Germany) was used at 50 Hz for 5 mins to lyse the cells. The samples were centrifuged at 10,000 g for 10 mins at 4°C, and the clear supernatant was collected. The concentration of soluble crude protein was estimated using BCA reagent (Thermo Fisher Scientific, USA) following the manufacturer's guidelines. As the samples were dilute, the crude protein extracts were concentrated by lyophilization and re-suspension in 25 µL of PBS.

The fusion protein Sh ble-mVenus was detected by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting. The protein concentration in each sample was normalized to 20 µg prior to loading. The proteins were separated by 4–15% SDS-PAGE using a Tris/glycine/SDS buffer (Biorad, USA) and blotted onto a PVDF membrane (Biorad, USA) using the Trans-Blot® Turbo™ Transfer System (Bio-rad, USA).

The protein blot was incubated with mouse anti-His monoclonal antibody (Bio-rad, USA) diluted 1:1,000 in 5% (w/v) skim milk followed by rabbit anti-mouse IgG conjugated to horseradish peroxidase (Bio-rad, USA) diluted 1:3,000 in 5% (w/v) skim milk. The signals were detected with the chemiluminescence detection system, (Clarity Western ECL, Biorad, USA) according to the manufacturer's instructions and documented using the ChemiDoc™ XRS+ System (Biorad, USA).

2.8. Statistical analysis

The fluorescence intensity of the fusion protein *Sh ble-mVenus*, which expression was driven by the experimental promoter constructs *EPPSII (Nga02101)*; *HSP90 (Nga00934)* and *ATPase (Nga06354.1)* was compared to the fluorescence intensity of *Sh ble-mVenus* driven by the *TUB (Nga00092)* promoter construct. Paired *t*-tests were performed on flow cytometric data and quantitative RT-PCR data to estimate the significant difference in relative promoter activity between the (TUB) positive control and each of the three experimental constructs.

3. Results and discussion

3.1 Identification of constitutive promoter candidates in *N. gaditana*

Endogenous promoters for constitutive and efficient expression of transgenes are generally obtained from non-coding regions upstream of a transcription initiation site of genes [46], that are associated with high and steady expression throughout most experimental conditions. Thus, genome-wide transcriptomic analysis represent a particularly useful resource to identify such candidate genes based on their expression patterns. Since transcriptomic data for *N. gaditana* are limited, we mined a publicly available *N. oceanica*

transcriptomic dataset [16] to identify a set of putative constitutive promoters. This dataset was selected as it closely represents a batch cultivation growth condition free of other environmental stresses, such as nutrient depletion. We identified nine genes (Table 1) with Fragments per Kilo base of exon per Million fragments mapped (FPKM) values greater than 60 across all time points (Supp. data file 1). The putative annotation of most of these genes suggested that they could be involved in essential cellular functions such as photosynthesis, protein maintenance, ion transport and nutrient acquisition [Table 1, 45–62]. Promoters of homolog genes encoding heat-shock proteins (HSPs) and adenosinetriphosphatases (ATPases) have previously been used in mammalian cell lines [65] and in higher plants [66] for constitutive or cell-specific gene expression [67]. Similarly, glyceraldehyde-3-phosphate dehydrogenase (*G3P*) promoters have been used for protein expression in yeast and plants [68–70]. As these functions are generally highly conserved among different organisms, we performed a database search on sequence similarity to identify potential orthologs in the *N. gaditana* genome (Table 1).

To ensure that the selected genes were actively expressed in *N. gaditana* cultures, we performed real time quantitative PCR (RT-qPCR) on mid-log phase samples as a representative condition commonly used for transgene expression [29,71]. All tested genes were actively transcribed in the cell, and five genes with the lowest Cq value, which is inversely proportional to the transcript copy number in the sample, were selected for further analysis (Fig. S5).

<i>N. gaditana</i> CCMP526 gene ID	<i>N. oceanica</i> CCMP1779 gene ID	Abbreviation	Functional annotation	Percent identity	Protein size (aa)	E-value	Bit score
Nga02101	NannoCCMP1779 4578	EPPSII	extrinsic protein in photosystem ii	90.62%	64	4.00E-28	119
Nga00934	NannoCCMP1779 5366	HSP90	heat shock protein 90/ HTPG molecular chaperone	94.12%	714	0	1172
Nga06354.1	NannoCCMP1779 8848	ATPase	H ⁺ -transporting ATPase	85.00%	260	3.00E-133	471
Nga01608	NannoCCMP1779 8392	HYP	hypothetical protein	85.53%	159	3.00E-81	296
Nga21005	NannoCCMP1779 3236	G3P	glyceraldehyde-3- phosphate dehydrogenase	93.63%	377	0	716
Nga00164	NannoCCMP1779 8259	HSP20	heat shock protein hsp20	76.32%	152	5.00E-66	246
Nga21085	NannoCCMP1779 9984	FATPase	F-type H ⁺ -transporting ATPase subunit beta	93.09%	666	0	1224
Nga06853	NannoCCMP1779 9414	FCCP	light-harvesting protein	89.13%	184	2.00E-88	321
Nga20972	NannoCCMP1779 5503	AMT	ammonium transporter	87.12%	489	0	760

Table 1. *N. oceanica* genes selected from *N. oceanica* CCMP1779 transcriptomic data [16] and their corresponding orthologs identified in the *N. gaditana* genome using BLASTp and their annotation. The identity between the two amino acid sequences are given in percentage along with length of the amino acid sequence, E-value, and raw alignment score. The top five genes (highlighted in blue and green) were selected based on expression levels observed in Figure S5 and their promoters used for *N. gaditana* transformation. The promoters of the top three genes (highlighted in green) were successfully transformed into *N. gaditana*.

The upstream intergenic regions of *EPPSII* (*Nga02101*); *HSP90*, (*Nga00934*); *ATPase* (*Nga06354.1*); *HYP* (*Nga01608*); and *G3P* (*Nga21005*) were extracted using the *N. gaditana* CCMP526 genome browser [14]. Since computational tools to analyse the structure of algal promoter are not available, we used plant-specific tools PlantCARE [43] and Softberry TSSP [44,72] to predict transcription factor binding sites and cis-regulatory elements such as CAAT and TATA boxes (Supp. data file 1), and their position with respect to the transcription start site (TSS). All promoter sequences that were analyzed had at least one predicted promoter

region (Supp. data file 1). The computational results showed the presence of CAAT-boxes, common cis-acting elements in promoter and enhancer regions; TATA-boxes, which are core promoter elements generally 30 bp upstream of the start codon; and G-boxes, cis-acting regulatory elements involved in light responsiveness in plants [73–76]. Other putative transcription factor (TF) binding site like MYC (TF of the basic helix-loop-helix-leucine zipper family) [77], MYB (TF family involved in biotic and abiotic stresses, development, differentiation, metabolism, defense in plants) [78,79] and the stress response element (STRE) [80] were also found in all the tested promoters, whereas the promoter-specific Spl TF recognition site (Spl is a protein that binds to GC box promoter elements to selectively activate mRNA synthesis in mammalian cells) [81,82] was only found in *EPPSII* (*Nga02101*) and *HSP90* (*Nga00934*) gene promoters. Based on these promoter motif predictions, a 600 bp–900 bp of intergenic sequence was selected as a putative promoter for each gene (Table S4).

3.2 Transformation efficiencies vary with selected promoter region

To evaluate whether the selected putative promoter regions in *N. gaditana* are able to drive the expression of transgenes, *EPPSII* (*Nga02101*), *HSP90* (*Nga00934*), *ATPase* (*Nga06354.1*), *HYP* (*Nga01608*) and *G3P* (*Nga21005*) putative promoter regions were cloned upstream of a *Sh ble-mVenus* gene fusion (Fig. S2) in pNaga4.mVenus expression vector and introduced in *N. gaditana* by electroporation. Three out of five vectors containing *EPPSII*, (*Nga02101*), *HSP90* (*Nga00934*) and *ATPase* (*Nga06354.1*) yielded transformant colonies, whereas three separate transformation attempts were unsuccessful for the remaining constructs (*HYP*, *Nga01608* and *G3P*, *Nga21005*).

Failure to generate antibiotic resistant colonies may be due to low or no expression of

the fusion protein *Sh ble* – mVenus, possibly as a result of the misidentification of the putative promoter region or the absence of essential regulatory elements in the cloned sequence [83].

A previous study in *N. gaditana* showed that endogenous promoters improved the transformation efficiency of this species [12]. In this study, the six promoter constructs used yielded different transformation efficiencies (Table S5). The highest transformation efficiency was achieved using the *EPPSII (Nga02101)* promoter, which resulted in an efficiency of 8×10^{-8} colonies cells⁻¹ $\mu\text{g DNA}^{-1}$. The lowest was observed using the *HSP90 (Nga00934)* promoter, which resulted in an efficiency of 3×10^{-9} colonies cells⁻¹ $\mu\text{g DNA}^{-1}$. The control *TUB (Nga00092)* promoter achieved a transformation efficiency of 7×10^{-9} colonies cells⁻¹ $\mu\text{g DNA}^{-1}$. Radakovits et al. observed a transformation efficiency of 12.5×10^{-6} colonies cells⁻¹ $\mu\text{g DNA}^{-1}$ for *TUB (Nga00092)*[12]. Since each promoter directly drives the expression of the selectable marker *Sh ble*, it is plausible that the wide range of transformation efficiencies observed is due to intrinsic features of the cloned promoter regions that affect strength and expression profiles.

3.3 mVenus expression varies among and within transformant lines

Clones that maintained antibiotic resistance after two months of subculturing were considered stable and were screened by flow cytometry for mVenus fluorescence. Forty *N. gaditana* clones carrying *EPPSII (Nga02101)* and *ATPase (Nga06354.1)* promoters and ten clones transformed using *HSP90 (Nga00934)* promoters, due to low transformation efficiency, were evaluated (Fig. S6). As expected in case of random chromosomal integration of transgenes, the level of mVenus fluorescence varied largely among transformants, presumably due to multiple factors, including: i) position effects due to random chromosomal transgene integration and influence of neighboring regulatory elements on transgene expression [84–86]; ii) number of intact transgene integrations [73]; iii) integrative

fragmentation or DNA re-arrangement, including deletion or translocation, which can also influence the growth and fitness of the host [87–89]; and iv) various forms of epigenetic and homology-dependent gene silencing [90,91].

All the transformant colonies screened were positive for mVenus fluorescence, confirming that the selected upstream regions for the *EPPSII* (*Nga02101*); *ATPase* (*Nga06354.1*); and *HSP90* (*Nga00934*) genes are able to serve as functional promoter sequences for transgene expression in *N. gaditana*. However, the clonal population did not show uniform mVenus expression (Fig. S4 Heterogeneous transgene expression has been observed in different microalgae species and seems to be a common consequence of random chromosomal integration of the expression cassette [92,93]. Similar phenomena occur in mammalian model systems, where the transgene expression was found to be heterogeneous due to silencing position effects [84]. These can result in multiple cell populations with various levels of expression within the same clonal line [94]. For example, in mice the proportion of cells expressing a transgene can vary widely between lines of mice carrying the same transgene [95,96]. Variations within clonal lines can also be non-genetic in nature and dependent on other factors, such as unequal cell division, cell age and cell cycle stage in non-synchronized cultures [97,98].

Monitoring transgene expression in individual cells within a clonal population gives an accurate account of cellular response, rather than a bulk measurement [99]. To address this, the percentage of cells in a population expressing mVenus was used to calculate a net fluorescence intensity. Three independent transgenic cell lines per construct with the highest net mVenus fluorescence were selected for further experiments.

3.4 Multi-level analysis of promoter activity

To profile the activity of the three promoters throughout the growth cycle and to select time points for further analysis, we performed a small-scale growth experiment in multiwell plates to measure mVenus fluorescence over a period of 10 days. The net fluorescent intensity, as a proxy for fusion protein expression, was calculated in triplicate for each independent transgenic line, using the product of mVenus intensity/cell and the number of cells expressing mVenus (Fig. 1).

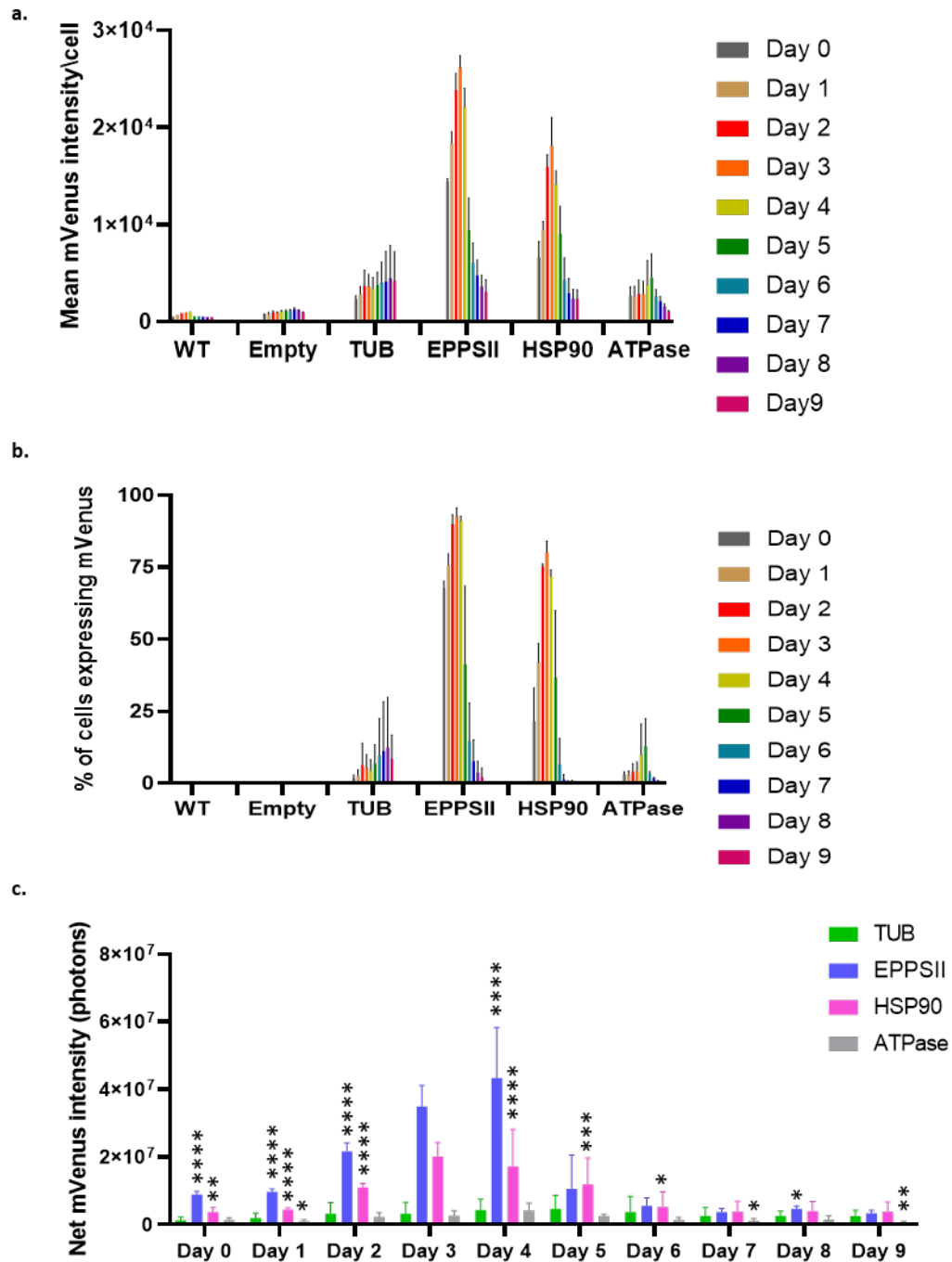


Figure 1. mVenus fluorescence measured over ten days in a multiwell plate. a. Mean mVenus fluorescence per cell; **b.** percentage of cells expressing mVenus; **c.** Net mVenus fluorescence calculated as a product of mVenus intensity/cell and number of cells expressing mVenus. Error bars represent standard deviation for independent transgenic cell lines (n=9) and wild-type *N. gaditana* (n=3). Significant differences between control promoter *TUB*,

(*Nga00092*) and experimental promoters *HSP90* (*Nga00934*), *EPPSII* (*Nga02101*), *ATPase*, (*Nga06354.1*) were calculated using paired t-test: * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$, **** $P < 0.0001$.

For all 10 days tested, net mVenus fluorescence intensities for *EPPSII* (*Nga02101*) and *HSP90* (*Nga00934*) transformants were significantly higher than those measured for *TUB* (*Nga00092*) transformants. Averaged *EPPSII* (*Nga02101*) putative promoter activity was 4.8 fold higher, and the *HSP90* (*Nga00934*) putative promoter activity was 2.7 fold higher than the *TUB* (*Nga00092*) control promoter. As maximum expression of mVenus fusion protein was observed between Day 4 and 6 (early-mid log phase) of culture (Fig. 1), the time point (Day 5) was used for further evaluation of reporter expression based on mRNA and protein analysis.

To assess transcript and transgenic protein abundance, three independent transgenic cell lines expressing the constructs described above were scaled-up to include RNA and protein extraction. Clonal lines were cultivated in batch mode over a period of 5 days while keeping the culture parameters consistent with the previous pilot experiment. The growth curves (Fig. 2b) of all the transgenic and wild-type *N. gaditana* lines are shown in Figure 2b. Growth rate in the exponential phase was calculated for both transgenic and wild-type cell lines, and no significant difference was observed among all transgenic lines except *EPPSII*, which was found to have a significantly higher exponential growth rate compared to wild type *N. gaditana* (Fig. S7, Table S6). However, wild type cells had a higher initial growth rate (from day 0 to 1) and maintained a higher maximum cell density than all transgenic lines throughout the experiment (Fig 2b). Differences in growth between wild-type and transgenic lines suggests that random genomic integration may have had detrimental effects to algal fitness due to random insertional mutagenesis, such as genetic rearrangements and disruption of

relevant genetic loci [87–89,100].

mVenus fluorescence was tracked daily to follow the expression level of the fusion protein *in vivo*, until Day 4 (Fig. 2a, Fig. S8). A temporal expression pattern similar to the initial pilot experiment was observed, with increasing net mVenus fluorescence as the cultures approached exponential phase (Fig. 1c, Fig. 2a). The exponential phase promoter activity of *EPPSII* (*Nga02101*) again surpassed the *TUB* (*Nga00092*) control promoter, (3.1 fold higher), although it was higher (4.8 fold) in the pilot experiment. The activity of *HSP90* (*Nga00934*) putative promoter was 4.5 fold higher than the *TUB* (*Nga00092*) control promoter, while previously we observed only a 2.7 fold efficiency in multiwell plates. The discrepancies observed in the two experiments may be due to the marked environmental differences associated with cultivation in shake flasks and microwell plates, even though the culture parameters were kept consistent.

Day 5 was chosen as suitable mid-exponential phase time-point to evaluate the activity of the tested promoters, both at the transcript and protein level. The relative transcriptional promoter activity for each promoter construct was tested by quantification of transgene transcripts on Day 5 using RT-qPCR. At the transcriptional level, the average relative promoter activities measured were 1.7 and 1.8-fold higher than that of the control promoter *TUB* (*Nga00092*) for *ATPase* (*Nga06354.1*) and *HSP90* (*Nga00934*), respectively (Fig. 2c). Conversely, no significant difference between the average relative promoter activity of *TUB*, (*Nga00092*) and the *EPPSII* (*Nga02101*) promoter was observed (Fig. 2c).

To evaluate the activity of the promoters at the protein level, we evaluated the abundance of the recombinant Sh ble-mVenus fusion protein in the same samples. The immunoblot analysis on protein samples harvested on Day 5 revealed a 47 kDa band

corresponding to the Sh ble-mVenus fusion protein, in transformants carrying the *HSP90* (*Nga00934*) promoter and *EPPSII* (*Nga02101*) promoter (Fig. 2d), the two promoters associated with highest net mVenus fluorescence (Fig. 2a) based on flow cytometry. No 47 kDa band was observed in transgenic cell lines containing the *ATPase* (*Nga06354.1*) promoter and the control *TUB*, (*Nga00092*) promoter. This may be due to low protein production, which may have sufficient protein to convey antibiotic resistance and for detection by flow cytometry, but not enough protein to be detected in an immunoblot. The flow cytometer used in this study has a sensitivity of <30 molecules of equivalent soluble fluorochrome in the FITC channel (CytoFLEX S, Beckman Coulter Life sciences, USA) whereas the Clarity Western ECL substrate used for chemiluminescence based immunoblot detection can detect as low as 0.6 ng protein in a sample (Biorad, USA). As expected, no bands were detected in lanes corresponding to the empty vector control and wild-type samples.

Transcript levels (Fig. 2c), mVenus expression, and recombinant fusion protein detected by immunoblot (Fig. 2d) were all detected in *EPPSII* (*Nga02101*) and *HSP90* (*Nga00934*) clonal lines. The same cannot be said for *ATPase* (*Nga06354.1*) and *TUB* (*Nga00092*) clonal lines as no 47 kDa protein band was observed to correlate with mRNA abundance observed. As stated earlier, reasons for this discrepancy can include low protein abundance or protein degradation. As described in various studies, mRNA abundance does not always correlate with protein level [101–103]. A number of factors, can interfere with transgene expression, such as position effects or post-transcriptional/post-translational silencing, both of which can be a consequence of random transgene integration [85,86,88,104]. These largely uncharacterised mechanisms could have plausibly affected Sh ble-mVenus protein concentrations in transgenic cell lines. Hence, high mRNA levels observed

can be sometimes misleading, which makes it key to study promoter activity with an integrated approach including both the transcriptional and translational level.

Finally, no apparent relationship was observed between transformation efficiency and full-scale transcript and protein abundance for each promoter. The *EPPSII* (*Nga02101*) promoter yielded the highest transformation efficiency and had the second highest expression of the fusion reporter protein by mVenus fluorescence. In contrast, the *HSP90* (*Nga00934*) promoter was associated with the strongest expression levels both at the transcript and protein level, despite being associated with poor transformation efficiency. The *ATPase* (*Nga06354.1*) promoter was the least efficient of the three at driving transgene expression, despite of having high transformation efficiency.

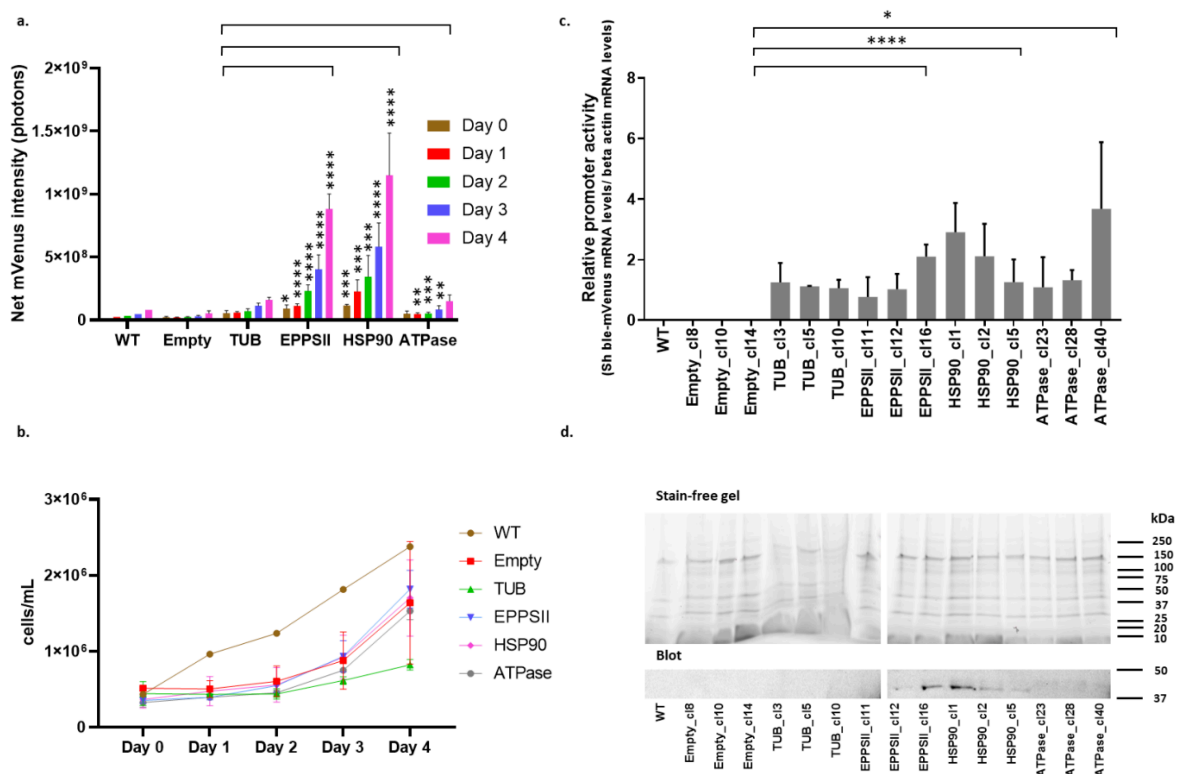


Figure 2. Relative mRNA and protein expression of *Sh ble-mVenus* in *N. gaditana* transformants **a.** Net mVenus fluorescence calculated as a product of mVenus intensity per cell, and number of cells expressing mVenus. Significant differences between control

promoter *TUB* (*Nga00092*) and experimental promoters: *HSP90* (*Nga00934*) promoter, *EPPSII* (*Nga02101*) promoter, *ATPase* (*Nga06354.1*) promoter calculated using paired t-test: * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$, **** $P < 0.0001$. **b.** Growth of wild type and transgenic *N. gaditana* cultures (Days 0–4). **c.** Relative quantification of Sh *ble-mVenus* transcript abundance normalized to β -*actin* mRNA in all *N. gaditana* transformant lines, driven by different promoters and in the empty vector control. Error bars represent standard deviation of 3 independent cell lines ($n=3$). Significant differences between control *TUB* (*Nga00092*) promoter and experimental promoters *HSP90* (*Nga00934*), *EPPSII* (*Nga02101*), *ATPase* (*Nga06354.1*) calculated using paired t-test: * $P < 0.05$, **** $P < 0.0001$. **d.** Immunoblot with anti-His monoclonal antibody confirmed the presence of a 47 kDa band corresponding to Sh *ble-mVenus* fusion protein. Analysis of Sh *ble-mVenus* protein in the wild-type (WT), empty vector controls (E8, E10 and E14), promoter control *TUB* (*Nga00092*) (cl03, cl05, cl10), and experimental promoters *EPPSII* (*Nga02101*) (cl11, cl12, cl16), *HSP90* (*Nga00934*) (cl01, cl02, cl05), *ATPase* (*Nga06354.1*) (cl23, cl28, cl40). Total cellular protein separated by SDS-PAGE (20 μ g of total soluble proteins were loaded per lane) on a 4–15% polyacrylamide gel.

This study identified three novel endogenous promoters that efficiently and constitutively drive the expression of reporter transgenes. Due to random chromosomal integration, the observed expression levels were variable among independent cell lines carrying the same transgenic construct. As mentioned, copy number and integration loci are key factors influencing transgene expression [85,86,104,105]. These factors cannot be controlled using the current transgenic method but other strategies, such as targeted knock-in using homologous recombination, are possible in this microalga and could overcome this issue [21]. Moreover, in the course of this work, the episome-based, non-integrative

expression of transgenes has been demonstrated to be feasible with *Nannochloropsis spp.*, [106,107] and could provide a more stable platform for future studies and synthetic biology applications.

In this work, promoters were profiled in the experiments carried out in batch mode under continuous light and nutrient replete conditions and not under variable culture conditions such as heat and light stress. In other organisms, the responsiveness of putatively similar promoters varies in stress conditions. For example, the heat shock proteins (HSPs) are molecular chaperones functionally involved in fundamental cellular processes and are induced by heat shock and other stimulants such as heavy metals, nitric oxide, hormones, microbial infections etc [108]. The orthologue of *HSP90* (*Nga00934*) in *E. coli* (*htpG*) is regulated by temperature and an increase in gene expression was observed in response to temperature under different growth conditions [48]. In humans, there are two major cytoplasmic isoforms of *Hsp90*: *Hsp90 α* , and *Hsp90 β* which are generally inducible and constitutive, respectively [57]. In higher plants, the expression of some genes in the *HSP90* gene family are affected by temperature and heavy metal stress, while others are constitutive [109,110].

Similarly, expression of *EPPSII* (*Nga02101*) may also vary under specific conditions. In higher plants, *EPPSII* (*PsbQ*) is a part of the oxygen-evolving complex of photosystem II (*PSII*) and is required for the PSII reaction center accumulation under low light [60,61]. *PsbQ* is vital for photoautotrophic growth under low light conditions [60,61] and may be upregulated during this state. In *Arabidopsis*, the expression of this gene was found to be light-inducible and organ specific [111]. The behaviour of these promoters under light, temperature, and heavy metal stress in *N. gaditana* may improve transgene expression further but are beyond

the scope of this paper. These are avenues that could be explored for future research.

4. Conclusions

Complex genetic engineering and synthetic biology resources are rapidly advancing and being applied to microalgae, widening their industrial potential. However, limitations include 1) a limited number of promoters experimentally shown to drive transgene expression; 2) a narrow repertoire of functional promoters with limited diversity in terms of expression range and inducibility; and 3) a lack of knowledge of specific regulatory elements that control transcription in *Nannochloropsis* that would allow for predictable expression and rational design of promoter regions. To date, several studies involving gene overexpression have been conducted in *Nannochloropsis* using endogenous promoters [25] but in-depth promoter analysis done in other model microalgae like *Phaeodactylum* [112–116] and *Chlamydomonas* [117] is still missing with *Nannochloropsis sp.*

In this work we identified and profiled three endogenous promoters *HSP90* (*Nga00934*); *EPPSII* (*Nga02101*) and *ATPase* (*Nga06354.1*) and we demonstrated that they can efficiently drive the expression of transgenes at different levels in *N. gaditana*, expanding the current genetic toolbox for this industrially relevant algal species. The availability of multiple promoters with similar strength, such as *HSP90* (*Nga00934*) and *EPPSII* (*Nga02101*), is a key requisite in synthetic biology. For example, when designing and assembling multi-gene constructs, the repeated use of the same promoter might incur in genetic rearrangements during the assembly phase in *E. coli* and/or *S. cerevisiae*, as well after the transformation in *Nannochloropsis spp*, given that all these organisms have highly efficient homologous recombination machineries. In other scenarios, such as the use of gene-stacking and combinatorial assembly of metabolic pathways, it is often required to express transgenes

at different levels in order to reach the optimal balance necessary for maximizing the product yield. In these cases, the availability of weak, constitutive promoters such as *ATPase* (*Nga06354.1*) is relevant. In conclusion, the new promoters characterized here broaden the repertoire of molecular tools that can be used for genetic engineering of this species, both for biotechnological applications such as metabolic engineering or recombinant protein production, and to better understand the biology of *Nannochloropsis*.

Author contributions

M.R. performed the experiments; M.F. and P.J.R. conceived the research project; M.F, R.A. and M. P. contributed in data analysis; M.R. wrote the manuscript; all authors edited and reviewed the manuscript.

Acknowledgements

This research was supported by UTS: C3 funding. MF is supported by a CSIRO Synthetic Biology Future Science Fellowship, co-funded by CSIRO and the University of Technology Sydney. The authors would like to thank Kun Xiao for technical assistance with flow cytometry, Stephen Woodcock for statistical analysis and Christian Evenhuis for ImageJ custom script to automate microscopic cell count.

Statement of informed consent, human/animal rights

No conflicts, informed consent, human or animal rights applicable.

References

- [1] D.J. Hibberd, Notes on the taxonomy and nomenclature of the algal classes Eustigmatophyceae and Tribophyceae (synonym Xanthophyceae), *Bot. J. Linn. Soc.* 82 (1981) 93–119. doi:10.1111/j.1095-8339.1981.tb00954.x.
- [2] P. Bondioli, L. Della, G. Rivolta, G. Chini, N. Bassi, L. Rodolfi, D. Casini, M. Prussi, D. Chiaramonti, M.R. Tredici, Bioresource Technology Oil production by the marine microalgae *Nannochloropsis* sp . F & M-M24 and *Tetraselmis suecica* F & M-M33, *Bioresour. Technol.* 114 (2012) 567–572. doi:10.1016/j.biortech.2012.02.123.
- [3] R.E. Jinkerson, R. Radakovits, M.C. Posewitz, Genomic insights from the oleaginous model alga *Nannochloropsis gaditana*, *Bioengineered.* 4 (2013). doi:10.4161/bioe.21880.
- [4] N. Moazami, A. Ashori, R. Ranjbar, M. Tangestani, R. Eghtesadi, A.S. Nejad, Large-scale biodiesel production using microalgae biomass of *Nannochloropsis*, *Biomass and Bioenergy.* 39 (2012) 449–453. doi:10.1016/j.biombioe.2012.01.046.
- [5] L. Rodolfi, G.C. Zittelli, N. Bassi, G. Padovani, N. Biondi, G. Bonini, M.R. Tredici, Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor, *Biotechnol. Bioeng.* 102 (2009) 100–112. doi:10.1002/bit.22033.
- [6] L.M. Lubi, O. Montero, I. Moreno-garrido, I.E. Huertas, C. Sobrino, M. Gonz, G. Par, *Nannochloropsis* (Eustigmatophyceae) as source of commercially valuable pigments, (2000) 249–255.

- [7] G.C. Zittelli, F. Lavista, A. Bastianini, L. Rodolfi, M. Vincenzini, M.R. Tredici, Production of eicosapentaenoic acid by *Nannochloropsis* sp . cultures in outdoor tubular photobioreactors, 70 (1999) 299–312.
- [8] M. Hoffmann, K. Marxen, R. Schulz, K.H. Vanselow, TFA and EPA productivities of *Nannochloropsis salina* influenced by temperature and nitrate stimuli in turbidostatic controlled experiments, Mar. Drugs. 8 (2010) 2526–2545. doi:10.3390/md8092526.
- [9] X.N. Ma, T.P. Chen, B. Yang, J. Liu, F. Chen, Lipid production from *Nannochloropsis*, Mar. Drugs. 14 (2016). doi:10.3390/md14040061.
- [10] T. Tonon, D. Harvey, T.R. Larson, I.A. Graham, Long chain polyunsaturated fatty acid production and partitioning to triacylglycerols in four microalgae, Phytochemistry. 61 (2002) 15–24. doi:10.1016/S0031-9422(02)00201-7.
- [11] M. Ferreira, P. Coutinho, P. Seixas, J. Fábregas, A. Otero, Enriching Rotifers with “ Premium ” Microalgae . *Nannochloropsis gaditana*, (2009) 585–595. doi:10.1007/s10126-008-9174-x.
- [12] R. Radakovits, R.E. Jinkerson, S.I. Fuerstenberg, H. Tae, R.E. Settlege, J.L. Boore, M.C. Posewitz, Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*, Nat. Commun. 3 (2012) 686. doi:10.1038/ncomms1688.
- [13] A. Vieler, G. Wu, C.H. Tsai, B. Bullard, A.J. Cornish, C. Harvey, I.B. Reca, C. Thornburg, R. Achawanantakun, C.J. Buehl, M.S. Campbell, D. Cavalier, K.L. Childs, T.J. Clark, R. Deshpande, E. Erickson, A. Armenia Ferguson, W. Handee, Q. Kong, X. Li, B. Liu, S. Lundback, C. Peng, R.L. Roston, Sanjaya, J.P. Simpson, A. TerBush, J. Warakanont, S. Zauner, E.M. Farre, E.L. Hegg, N. Jiang, M.H. Kuo, Y. Lu, K.K. Niyogi, J. Ohlrogge, K.W.

- Osteryoung, Y. Shachar-Hill, B.B. Sears, Y. Sun, H. Takahashi, M. Yandell, S.H. Shiu, C. Benning, Genome, Functional Gene Annotation, and Nuclear Transformation of the Heterokont Oleaginous Alga *Nannochloropsis oceanica* CCMP1779, *PLoS Genet.* 8 (2012). doi:10.1371/journal.pgen.1003064.
- [14] E. Corteggiani Carpinelli, A. Telatin, N. Vitulo, C. Forcato, M. D'Angelo, R. Schiavon, A. Vezzi, G.M. Giacometti, T. Morosinotto, G. Valle, Chromosome scale genome assembly and transcriptome profiling of *nannochloropsis gaditana* in nitrogen depletion, *Mol. Plant.* 7 (2014) 323–335. doi:10.1093/mp/sst120.
- [15] D. Wang, K. Ning, J. Li, J. Hu, D. Han, H. Wang, X. Zeng, X. Jing, Q. Zhou, X. Su, X. Chang, A. Wang, W. Wang, J. Jia, L. Wei, Y. Xin, Y. Qiao, R. Huang, J. Chen, B. Han, K. Yoon, R.T. Hill, Y. Zohar, F. Chen, Q. Hu, J. Xu, *Nannochloropsis* Genomes Reveal Evolution of Microalgal Oleaginous Traits, 10 (2014). doi:10.1371/journal.pgen.1004094.
- [16] E. Poliner, N. Panchy, L. Newton, G. Wu, A. Lapinsky, B. Bullard, A. Zienkiewicz, Transcriptional coordination of physiological responses in *Nannochloropsis oceanica* CCMP1779 under light / dark cycles, (2015) 1097–1113. doi:10.1111/tpj.12944.
- [17] A. Mühlroth, P. Winge, A. El Assimi, J. Jouhet, E. Marechal, M.F. Hohmann-Marriott, O. Vadstein, A.M. Bones, Mechanisms of phosphorus acquisition and lipid class remodelling under P limitation in a marine microalga, *Plant Physiol.* (2017) pp.00621.2017. doi:10.1104/pp.17.00621.
- [18] F. Li, D. Gao, H. Hu, High-efficiency nuclear transformation of the oleaginous marine *Nannochloropsis* species using PCR product., *Biosci. Biotechnol. Biochem.* 78 (2014)

- 812–817. doi:10.1080/09168451.2014.905184.
- [19] D. Moog, S. Stork, S. Reislöhner, C. Grosche, U.G. Maier, In vivo localization studies in the stramenopile alga *Nannochloropsis oceanica*, *Protist.* 166 (2015) 161–171. doi:10.1016/j.protis.2015.01.003.
- [20] N.K. Kang, G.G. Choi, E.K. Kim, S.E. Shin, S. Jeon, M.S. Park, K.J. Jeong, B.R. Jeong, Y.K. Chang, J.W. Yang, B. Lee, Heterologous overexpression of sfCherry fluorescent protein in *Nannochloropsis salina*, *Biotechnol. Reports.* 8 (2015) 10–15. doi:10.1016/j.btre.2015.08.004.
- [21] O. Kilian, C.S.E. Benemann, K.K. Niyogi, B. Vick, High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis* sp., *Proc. Natl. Acad. Sci.* 108 (2011) 21265–21269. doi:10.1073/pnas.1105861108.
- [22] E. Bertrand Vick, A. Oliver Kilian, Transformation of Algal cells, US 8,759,615 B2, 2014.
- [23] M. Hlavova, Z. Turoczy, K. Bisova, Improving microalgae for biotechnology - From genetics to synthetic biology, *Biotechnol. Adv.* 33 (2015) 1194–1203. doi:10.1016/j.biotechadv.2015.01.009.
- [24] M.A. Scaife, A.G. Smith, Towards developing algal synthetic biology, *Biochem. Soc. Trans.* 44 (2016) 716–722. doi:10.1042/bst20160061.
- [25] E. Poliner, E.M. Farré, C. Benning, Advanced genetic tools enable synthetic biology in the oleaginous microalgae *Nannochloropsis* sp., *Plant Cell Rep.* 0 (2018) 0. doi:10.1007/s00299-018-2270-0.
- [26] C. De Wilde, H. Van Houdt, S. De Buck, G. Angenon, G. De Jaeger, A. Depicker, *Plants*

- as bioreactors for protein production : avoiding the problem of transgene silencing, (2000) 347–359.
- [27] Y.-D. Park, I. Papp, E.A. Moscone, V.A., Iglesias, H. Vaucheret, A.J.M. Matzke, M.A. Matzke, Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity, *Plant J.* 9 (1996) 183–194. doi:<https://doi.org/10.1046/j.1365-313X.1996.09020183.x>.
- [28] A.J.M. Matzke, F. Neuhuber, Y.D. Park, P.F. Ambros, M.A. Matzke, Homology-dependent gene silencing in transgenic plants: epistatic silencing loci contain multiple copies of methylated transgenes, *MGG Mol. Gen. Genet.* 244 (1994) 219–229. doi:[10.1007/BF00285449](https://doi.org/10.1007/BF00285449).
- [29] E. Poliner, J.A. Pulman, K. Zienkiewicz, K. Childs, C. Benning, E.M. Farré, A toolkit for *Nannochloropsis oceanica* CCMP1779 enables gene stacking and genetic engineering of the eicosapentaenoic acid pathway for enhanced long-chain polyunsaturated fatty acid production, *Plant Biotechnol. J.* (2017) 1–12. doi:[10.1111/pbi.12772](https://doi.org/10.1111/pbi.12772).
- [30] T. Nagai, K. Ibata, E.S. Park, M. Kubota, K. Mikoshiba, A. Miyawaki, A variant of YFP with fast and efficient maturation for cell-biological applications, *Nature.* 20 (2002) 1585–1588.
- [31] I. Ajjawi, J. Verruto, M. Aqui, L.B. Soriaga, J. Coppersmith, K. Kwok, L. Peach, E. Orchard, R. Kalb, W. Xu, T.J. Carlson, K. Francis, K. Konigsfeld, J. Bartalis, A. Schultz, W. Lambert, A.S. Schwartz, R. Brown, E.R. Moellering, Lipid production in *Nannochloropsis gaditana* is doubled by decreasing expression of a single

- transcriptional regulator, *Nat. Biotechnol.* 35 (2017) 647–652. doi:10.1038/nbt.3865.
- [32] R.R.L. Guillard, Culture of Phytoplankton for Feeding Marine Invertebrates, in: W.L. Smith, M.H. Chanley (Eds.), *Cult. Mar. Invertebr. Anim. Proc. --- 1st Conf. Cult. Mar. Invertebr. Anim.* Greenport, Springer US, Boston, MA, 1975: pp. 29–60.
doi:10.1007/978-1-4615-8714-9_3.
- [33] J.A. Berges, D.J. Franklin, P.J. Harrison, Evolution of an artificial seawater medium: Improvements in enriched seawater, artificial water over the last two decades, *J. Phycol.* 37 (2001) 1138–1145. doi:10.1046/j.1529-8817.2001.01052.x.
- [34] C. Benning, [Bmb.natsci.msu.edu](http://bmb.natsci.msu.edu), (n.d.).
https://bmb.natsci.msu.edu/sites/_bmb/assets/File/benning/protein_sequence.txt.
- [35] Giorgio Valle group, *Nannochloropsis gaditana* Genomes portal - CRIBI, University of Padua, (2013). <http://www.nannochloropsis.org/index.php>.
- [36] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT – PCR, 29 (2001) 16–21.
- [37] A. Radonić, S. Thulke, I.M. Mackay, O. Landt, W. Siegert, A. Nitsche, Guideline to reference gene selection for quantitative real-time PCR, *Biochem. Biophys. Res. Commun.* 313 (2004) 856–862. doi:10.1016/j.bbrc.2003.11.177.
- [38] P. Fernandez, J.A.D. Rienzo, S. Moschen, G.A.A. Dosio, L.A.N. Aguirrezábal, H.E. Hopp, N. Paniego, R.A. Heinz, Comparison of predictive methods and biological validation for qPCR reference genes in sunflower leaf senescence transcript analysis, *Plant Cell Rep.* 30 (2011) 63–74. doi:10.1007/s00299-010-0944-3.

- [39] F. Ponton, M.P. Chapuis, M. Pernice, G.A. Sword, S.J. Simpson, Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*, *J. Insect Physiol.* 57 (2011) 840–850. doi:10.1016/j.jinsphys.2011.03.014.
- [40] M. Schliep, M. Pernice, S. Sinutok, C. V. Bryant, P.H. York, M.A. Rasheed, P.J. Ralph, Evaluation of reference genes for RT-qPCR studies in the seagrass *Zostera muelleri* exposed to light limitation, *Sci. Rep.* 5 (2015) 1–11. doi:10.1038/srep17051.
- [41] S. Goyen, M. Pernice, M. Szabó, M.E. Warner, P.J. Ralph, D.J. Suggett, A molecular physiology basis for functional diversity of hydrogen peroxide production amongst *Symbiodinium* spp. (Dinophyceae), *Mar. Biol.* 164 (2017) 1–12. doi:10.1007/s00227-017-3073-5.
- [42] S. Cao, X. Zhang, N. Ye, X. Fan, S. Mou, D. Xu, C. Liang, Y. Wang, W. Wang, Evaluation of putative internal reference genes for gene expression normalization in *Nannochloropsis* sp. by quantitative real-time RT-PCR, *Biochem. Biophys. Res. Commun.* 424 (2012) 118–123. doi:10.1016/j.bbrc.2012.06.086.
- [43] M. Lescot, P. Dehais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouze, S. Rombauts, PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, *Nucleic Acids Res.* 30 (2002) 325–327. doi:10.1093/nar/30.1.325.
- [44] V. V Solovyev, I.A. Shahmuradov, A.A. Salamov, Identification of Promoter Regions and Regulatory Sites, in: I. Ladunga (Ed.), *Comput. Biol. Transcr. Factor Bind.*, Humana Press, Totowa, NJ, 2010: pp. 57–83. doi:10.1007/978-1-60761-854-6_5.

- [45] L. Wei, Y. Xin, Q. Wang, J. Yang, H. Hu, J. Xu, RNAi-based targeted gene-knockdown in the model oleaginous microalgae *Nannochloropsis oceanica*, *Plant J.* (2016) 1–15. doi:10.1111/tpj.13411.
- [46] R.C. Perier, V. Praz, T. Junier, C. Bonnard, P. Bucher, The Eukaryotic Promoter Database (EDP), *Nucleic Acids Res.* 28 (2000) 302–303. doi:10.1093/nar/28.1.302.
- [47] J. Buchner, Hsp90 and Co. - A holding for folding, *Trends Biochem. Sci.* 24 (1999) 136–141. doi:10.1016/S0968-0004(99)01373-0.
- [48] A. Heitzer, C.A. Mason, M. Snozzi, G. Hamer, Some effects of growth conditions on steady state and heat shock induced htpG gene expression in continuous cultures of *Escherichia coli*, *Arch. Microbiol.* 155 (1990) 7–12. doi:10.1007/BF00291266.
- [49] M. Arango, F. Gévaudant, M. Oufattole, M. Boutry, The plasma membrane proton pump ATPase: the significance of gene subfamilies., *Planta.* 216 (2003) 355–365. doi:10.1007/s00425-002-0856-8.
- [50] F. Révillion, V. Pawlowski, L. Hornez, J.P. Peyrat, Glyceraldehyde-3-phosphate dehydrogenase gene expression in human breast cancer, *Eur. J. Cancer.* 36 (2000) 1038–1042. doi:10.1016/S0959-8049(00)00051-4.
- [51] L. Mcalisters, M.J. Holland, Differential Expression of the Three Yeast Glyceraldehyde-3-phosphate Dehydrogenase Genes, *J. Biol. Chem.* 260 (1985) 15019–15027. <http://www.jbc.org/content/260/28/15019.abstract>.
- [52] J.D. Beckmann, P.O. Ljungdahl, J.L. Lopez, B.L. Trumpower, Isolation and characterization of the nuclear gene encoding the Rieske iron-sulfur protein (RIP1)

- from *Saccharomyces cerevisiae*., *J. Biol. Chem.* 262 (1987) 8901–8909.
- [53] S. Boxhammer, S. Glaser, A. Kühn, A.K. Wagner, C.L. Schmidt, Characterization of the recombinant Rieske [2Fe-2S] proteins HcaC and YeaW from *E. coli*, *BioMetals*. 21 (2008) 459–467. doi:10.1007/s10534-008-9134-y.
- [54] S. de Goër, M. Quinet-Szely, L. Caron, C. Berkaloff, D. Douady, Gene Structure of a Chlorophyll a/c-Binding Protein from a Brown Alga: Presence of an Intron and Phylogenetic Implications, *J. Mol. Evol.* 43 (2006) 270–280. doi:10.1007/pl00006086.
- [55] A.M. Marini, S. Soussi-Boudekou, S. Vissers, B. Andre, A family of ammonium transporters in *Saccharomyces cerevisiae*., *Mol. Cell. Biol.* 17 (2015) 4282–4293. doi:10.1128/mcb.17.8.4282.
- [56] N. von Wirén, J. Yamaguchi, Y. Sonoda, S. Saiki, T. Yamaya, A. Ikeda, Distinct Expression and Function of Three Ammonium Transporter Genes (*OsAMT1;1 – 1;3*) in Rice, *Plant Cell Physiol.* 44 (2003) 726–734. doi:10.1093/pcp/pcg083.
- [57] A.S. Sreedhar, É. Kalmár, P. Csermely, Y.F. Shen, Hsp90 isoforms: Functions, expression and clinical importance, *FEBS Lett.* 562 (2004) 11–15. doi:10.1016/S0014-5793(04)00229-7.
- [58] O. Hoegh-Guldberg, M. Pernice, S. Dunn, S. Dove, N.N. Rosic, Gene expression profiles of cytosolic heat shock proteins Hsp70 and Hsp90 from symbiotic dinoflagellates in response to thermal stress: possible implications for coral bleaching, *Cell Stress Chaperones*. 16 (2010) 69–80. doi:10.1007/s12192-010-0222-x.
- [59] H. Ohta, J.-R. Shen, M. Ueno, I. Enami, T. Suzuki, S. Yoshihara, A. Okumura, Extrinsic

- proteins of photosystem II, *Eur. J. Biochem.* 270 (2003) 4156–4163.
doi:10.1046/j.1432-1033.2003.03810.x.
- [60] J.L. Roose, L.K. Frankel, M.P. Mummadisetti, T.M. Bricker, The extrinsic proteins of photosystem II: update, *Planta*. 243 (2016) 889–908. doi:10.1007/s00425-015-2462-6.
- [61] T.M. Bricker, L.K. Frankel, Auxiliary functions of the PsbO, PsbP and PsbQ proteins of higher plant Photosystem II: A critical analysis, *J. Photochem. Photobiol. B Biol.* 104 (2011) 165–178. doi:10.1016/j.jphotobiol.2011.01.025.
- [62] G. Duby, M. Boutry, The plant plasma membrane proton pump ATPase: A highly regulated P-type ATPase with multiple physiological roles, *Pflügers Arch. Eur. J. Physiol.* 457 (2009) 645–655. doi:10.1007/s00424-008-0457-x.
- [63] N.N. Ewing, A.B. Bennett, Assessment of the Number and Expression of P-Type H⁺-ATPase Genes in Tomato, *Plant Physiol.* 106 (2002) 547–557.
doi:10.1104/pp.106.2.547.
- [64] K. Harms, R. V. Wöhner, B. Schulz, W.B. Frommer, Isolation and characterization of P-type H⁺-ATPase genes from potato, *Plant Mol. Biol.* 26 (1994) 979–988.
doi:10.1007/BF00028864.
- [65] C. Prodromou, Mechanisms of Hsp90 regulation, *Biochem. J.* 473 (2016) 2439–2452.
doi:10.1042/bcj20160005.
- [66] H. Struve, T. Rausch, P. Bernasconi, L. Taiz, Structure and function of the promoter of the carrot V-type H⁺-ATPase catalytic subunit gene, *J. Biol. Chem.* 265 (1990) 7927–7932.

- [67] A. Muraoka, M. Kaise, Y.-J. Guo, J. Yamada, I. Song, J. Delvalle, A. Todisco, T. Yamada, Canine H⁺-K⁺ -ATPase alpha subunit gene promoter : studies with canine parietal cells in primary culture, *Am. J. Physiol.* 271 (1996) G1104–G1113.
doi:<https://doi.org/10.1152/ajpgi.1996.271.6.G1104>.
- [68] H.R. Waterham, M.E. Digan, P.J. Koutz, S. V. Lair, J.M. Cregg, Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter, *Gene*. 186 (1997) 37–44. doi:10.1016/S0378-1119(96)00675-0.
- [69] G.A. Bitter, K.M. Egan, Expression of heterologous genes in *Saccharomyces cerevisiae* from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter, *Gene*. 32 (1984) 263–274. doi:10.1016/0378-1119(84)90002-7.
- [70] S. Manjunath, M.M. Sachs, Molecular characterization and promoter analysis of the maize cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its expression during anoxia, *Plant Mol. Biol.* 33 (1997) 97–112.
doi:10.1023/A:1005729112038.
- [71] L. Wei, Q. Wang, Y. Xin, Y. Lu, J. Xu, Enhancing photosynthetic biomass productivity of industrial oleaginous microalgae by overexpression of RuBisCO activase, *Algal Res.* 27 (2017) 366–375. doi:10.1016/j.algal.2017.07.023.
- [72] I.A. Shahmuradov, R.K. Umarov, V. V. Solovyev, TSSPlant: A new tool for prediction of plant Pol II promoters, *Nucleic Acids Res.* 45 (2017). doi:10.1093/nar/gkw1353.
- [73] C. Benoist, K. O’Hare, R. Breathnach, P. Chambon, The ovalbumin gene - sequence of putative control regions, *Nucleic Acids Res.* 8 (1980) 127–142.
doi:10.1093/nar/8.1.127.

- [74] S.L. McKnight, R. Kingsbury, Transcriptional control signals of a eukaryotic protein-coding gene, *Science* (80-.). 217 (1982) 316–324. doi:10.1126/science.6283634.
- [75] A.E. Menkens, U. Schindler, A.R. Cashmore, The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins, *Trends Biochem. Sci.* 20 (1995) 506–510. doi:10.1016/S0968-0004(00)89118-5.
- [76] K. Yamaguchi-shinozaki, K. Shinozaki, Organization of cis -acting regulatory elements in osmotic- and cold-stress- responsive promoters, *Trends Plant Sci.* 10 (2005) 88–94. doi:10.1016/j.tplants.2004.12.012.
- [77] A. Sabò, B. Amati, Genome recognition by MYC, *Cold Spring Harb. Perspect. Med.* 4 (2014) 1–13. doi:10.1101/cshperspect.a014191.
- [78] H. Abe, K. Yamaguchi-Shinozaki, T. Urao, T. Iwasaki, D. Hosokawa, K. Shinozaki, Role of Arabidopsis MYC and MYB Homologs in Drought- and Abscisic Acid-Regulated Gene Expression, *Plant Cell.* 9 (2007) 1859. doi:10.2307/3870530.
- [79] S. Ambawat, P. Sharma, N.R. Yadav, R.C. Yadav, MYB transcription factor genes as regulators for plant responses: An overview, *Physiol. Mol. Biol. Plants.* 19 (2013) 307–321. doi:10.1007/s12298-013-0179-1.
- [80] E. Moskvina, E.M. Imre, H. Ruis, Stress factors acting at the level of the plasma membrane induce transcription via the stress response element (STRE) of the yeast *Saccharomyces cerevisiae*, *Mol. Microbiol.* 32 (1999) 1263–1272. doi:10.1046/j.1365-2958.1999.01438.x.
- [81] M.R. Briggs, J.T. Kadonaga, S.P. Bell, R. Tjian, Purification and biochemical

- characterization of the promoter-specific transcription factor, Sp1, *Science* (80-.). 234 (1986) 47–52. doi:10.1126/science.3529394.
- [82] J.T. Kadonaga, K.R. Carner, F.R. Masiarz, R. Tjian, Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain, *Cell*. 51 (1987) 1079–1090. doi:10.1016/0092-8674(87)90594-0.
- [83] C.M. Hernandez-Garcia, J.J. Finer, Identification and validation of promoters and cis-acting regulatory elements, *Plant Sci*. 217–218 (2014) 109–119. doi:10.1016/j.plantsci.2013.12.007.
- [84] Y.-Q. Feng, M.C. Lorincz, S. Fiering, J.M. Greally, E.E. Bouhassira, Position Effects Are Influenced by the Orientation of a Transgene with Respect to Flanking Chromatin, *Mol. Cell. Biol.* 21 (2001) 298–309. doi:10.1128/MCB.21.1.298-309.2001.
- [85] A.J. Clark, P. Bissinger, D.W. Bullock, S. Damak, R. Wallace, C.B.A. Whitelaw, F. Yull, Chromosomal position effects and the modulation of transgene expression, *Reprod. Fertil. Dev.* 6 (1994) 589–594. doi:10.1071/RD9940589.
- [86] C. Peach, J. Velten, Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters, *Plant Mol. Biol.* 17 (1991) 49–60. doi:10.1007/BF00036805.
- [87] S.B. Maqbool, P. Christou, Multiple traits of agronomic importance in transgenic indica rice plants: analysis of transgene integration pattern, expression levels and stability, *Mol. Breed.* 5 (1999) 471–480. doi:https://doi.org/10.1023/A:1009634226797.

- [88] R. al-Shawi, J. Kinnaird, J. Burke, J.O. Bishop, Expression of a foreign gene in a line of transgenic mice is modulated by a chromosomal position effect., *Mol. Cell. Biol.* 10 (2015) 1192–1198. doi:10.1128/mcb.10.3.1192.
- [89] M. Takano, H. Egawa, J.E. Ikeda, K. Wakasa, The structures of integration sites in transgenic rice, *Plant J.* 11 (1997) 353–361. doi:10.1046/j.1365-313X.1997.11030353.x.
- [90] F.F. Assaad, K.L. Tucker, E.R. Signer, Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*, *Plant Mol. Biol.* 22 (1993) 1067–1085. doi:10.1007/BF00028978.
- [91] Cerutti H, A. Johnson, N. Gillham, J. Boynton, Epigenetic Silencing of a Foreign Gene in Nuclear Transformants of *Chlamydomonas*, *Plant Cell.* 9 (2002) 925–945. doi:10.1105/tpc.9.6.925.
- [92] R.P. Shrestha, M. Hildebrand, Development of a silicon limitation inducible expression system for recombinant protein production in the centric diatoms *Thalassiosira pseudonana* and *Cyclotella cryptica*, *Microb. Cell Fact.* 16 (2017) 1–14. doi:10.1186/s12934-017-0760-3.
- [93] A. Davis, L.T. Crum, L.B. Corbeil, M. Hildebrand, Expression of *Histophilus somni* IbpA DR2 protective antigen in the diatom *Thalassiosira pseudonana*, *Appl. Microbiol. Biotechnol.* 101 (2017) 5313–5324. doi:10.1007/s00253-017-8267-8.
- [94] D.I.K. Martin, E. Whitelaw, The vagaries of variegating transgenes, *BioEssays.* 18 (1996) 919–923. doi:10.1002/bies.950181111.
- [95] G. Robertson, D. Garrick, W. Wu, M. Kearns, D. Martin, E. Whitelaw, Position-

- dependent variegation of globin transgene expression in mice., *Proc. Natl. Acad. Sci.* 92 (2006) 5371–5375. doi:10.1073/pnas.92.12.5371.
- [96] R. Alami, beta-globin YAC transgenes exhibit uniform expression levels but position effect variegation in mice, *Hum. Mol. Genet.* 9 (2000) 631–636. doi:10.1093/hmg/9.4.631.
- [97] S.R. Laney, R.J. Olson, H.M. Sosik, Diatoms favor their younger daughters, *Limnol. Oceanogr.* 57 (2012) 1572–1578. doi:10.4319/lo.2012.57.5.1572.
- [98] H. Aguilaniu, L. Gustafsson, M. Rigoulet, T. Nyström, Asymmetric inheritance of oxidatively damaged proteins during cytokinesis, *Science* (80-.). 299 (2003) 1751–1753. doi:10.1126/science.1080418.
- [99] J.C. Traller, M. Hildebrand, High throughput imaging to the diatom *Cyclotella cryptica* demonstrates substantial cell-to-cell variability in the rate and extent of triacylglycerol accumulation, *Algal Res.* 2 (2013) 244–252. doi:10.1016/j.algal.2013.03.003.
- [100] K. V. Dmytruk, A.Y. Voronovsky, A.A. Sibirny, Insertion mutagenesis of the yeast *Candida famata* (*Debaryomyces hansenii*) by random integration of linear DNA fragments, *Curr. Genet.* 50 (2006) 183–191. doi:10.1007/s00294-006-0083-0.
- [101] C. Vogel, E.M. Marcotte, Insights into the regulation of protein abundance from proteomic and transcriptomic analyses, *Nat Rev Genet.* 13 (2013) 227–232. doi:10.1038/nrg3185.Insights.
- [102] Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA

- Abundance, *Cell*. 165 (2016) 535–550. doi:10.1016/j.cell.2016.03.014.
- [103] N. Kendrick, A gene's mRNA level does not usually predict its protein level, *Www.Kendricklabs.Com*. 2 (2012) 1–6. doi:10.1146/annurev.ento.49.061802.123354.
- [104] a J. Matzke, M. a Matzke, Position effects and epigenetic silencing of plant transgenes., *Curr. Opin. Plant Biol.* 1 (1998) 142–148. doi:10.1016/S1369-5266(98)80016-2.
- [105] Q. Kong, M. Wu, Y. Huan, L. Zhang, H. Liu, G. Bou, Y. Luo, Y. Mu, Z. Liu, Transgene expression is associated with copy number and cytomegalovirus promoter methylation in transgenic pigs, *PLoS One*. 4 (2009). doi:10.1371/journal.pone.0006679.
- [106] E. Poliner, T. Takeuchi, Z.-Y. Du, C. Benning, E. Farre, Non-transgenic marker-free gene disruption by an episomal CRISPR system in the oleaginous microalga, *Nannochloropsis oceanica* CCMP1779, *ACS Synth. Biol.* (2018) acssynbio.7b00362. doi:10.1021/acssynbio.7b00362.
- [107] J. Verruto, K. Francis, Y. Wang, M.C. Low, J. Greiner, S. Tacke, F. Kuzminov, W. Lambert, J. McCarren, I. Ajjawi, N. Bauman, R. Kalb, G. Hannum, E.R. Moellering, Unrestrained markerless trait stacking in *Nannochloropsis gaditana* through combined genome editing and marker recycling technologies , *Proc. Natl. Acad. Sci.* 115 (2018) E7015–E7022. doi:10.1073/pnas.1718193115.
- [108] J.G. Kiang, G.C. Tsokos, Heat shock protein 70 kDa: Molecular biology, biochemistry, and physiology, *Pharmacol. Ther.* 80 (1998) 183–201. doi:10.1016/S0163-7258(98)00028-X.

- [109] D. Milioni, P. Hatzopoulos, Genomic organization of hsp90 gene family in Arabidopsis, *Plant Mol. Biol.* 35 (1997) 955–961. doi:10.1023/A:1005874521528.
- [110] P. Krishna, G. Gloor, The Hsp90 family of proteins in *Arabidopsis thaliana.*, *Cell Stress Chaperones.* 6 (2001) 238–46.
<http://www.ncbi.nlm.nih.gov/pubmed/11599565><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC434405>.
- [111] M. Grover, T. Gaur, A. Kochhar, S.C. Maheshwari, A.K. Tyagi, Nucleotide sequence of psbQ gene for 16-kDa protein of oxygen-evolving complex from *Arabidopsis thaliana* and regulation of its expression, *DNA Res.* 6 (1999) 173–177.
doi:10.1093/dnares/6.3.173.
- [112] Y. Watanabe, T. Kadono, N. Kira, K. Suzuki, O. Iwata, Marine Genomics Development of endogenous promoters that drive high-level expression of introduced genes in the model diatom *Phaeodactylum tricornutum*, *Mar. Genomics.* 42 (2018) 41–48.
doi:10.1016/j.margen.2018.06.003.
- [113] H.-Y. Lin, S.-C. Yen, P.-C. Kuo, C.-Y. Chung, K.-L. Yeh, C.-H. Huang, J. Chang, H.-J. Lin, Alkaline phosphatase promoter as an efficient driving element for exogenic recombinant in the marine diatom *Phaeodactylum tricornutum*, *Algal Res.* 23 (2017) 58–65. doi:10.1016/j.algal.2017.01.007.
- [114] E. Erdene-Ochir, B.K. Shin, M.N. Huda, D.H. Kim, E.H. Lee, D.G. Song, Y.M. Kim, S.M. Kim, C.H. Pan, Cloning of a novel endogenous promoter for foreign gene expression in *Phaeodactylum tricornutum*, *Appl. Biol. Chem.* 59 (2016) 1–7. doi:10.1007/s13765-016-0235-y.

- [115] Z. Adler-Agnon (Shemesh), S. Leu, A. Zarka, S. Boussiba, I. Khozin-Goldberg, Novel promoters for constitutive and inducible expression of transgenes in the diatom *Phaeodactylum tricornutum* under varied nitrate availability, *J. Appl. Phycol.* (2017) 1–10. doi:10.1007/s10811-017-1335-8.
- [116] S. Seo, H. Jeon, S. Hwang, E. Jin, K.S. Chang, Development of a new constitutive expression system for the transformation of the diatom *Phaeodactylum tricornutum*, *Algal Res.* 11 (2015) 50–54. doi:10.1016/j.algal.2015.05.012.
- [117] M. Schroda, D. Blöcker, C.F. Beck, The HSP70A promoter as a tool for the improved expression of transgenes in *Chlamydomonas*, *Plant J.* 21 (2000) 121–131. doi:10.1046/j.1365-313X.2000.00652.x.