BMC Biology

Open Access

Genetic mechanisms underlying increased microalgal thermotolerance, maximal growth rate, and yield on light following adaptive laboratory evolution

Robin Barten[1*] , Dirk-Jan M. van Workum[2], Emma de Bakker[1], Judith Risse[2], Michelle Kleisman[2], Sofia Navalho[1], Sandra Smit[2], Rene H. Wijffels[1,3], Harm Nijveen[2] and Maria J. Barbosa[1]

Abstract

Background: Adaptive laboratory evolution (ALE) is a powerful method for strain optimization towards abiotic stress factors and for identifying adaptation mechanisms. In this study, the green microalga *Picochlorum* sp. *BPE23* was cultured under supra-optimal temperature to force genetic adaptation. The robustness and adaptive capacity of *Picochlorum* strains turned them into an emerging model for evolutionary studies on abiotic stressors such as temperature, salinity, and light.

Results: Mutant strains showed an expanded maximal growth temperature of 44.6 °C, whereas the maximal growth temperature of the wild-type strain was 42 °C. Moreover, at the optimal growth temperature of 38 °C, the biomass yield on light was 22.3% higher, and the maximal growth rate was 70.5% higher than the wild type. Genome sequencing and transcriptome analysis were performed to elucidate the mechanisms behind the improved phenotype. A de novo assembled phased reference genome allowed the identification of 21 genic mutations involved in various processes. Moreover, approximately half of the genome contigs were found to be duplicated or even triplicated in all mutants, suggesting a causal role in adaptation.

Conclusions: The developed tools and mutant strains provide a strong framework from whereupon *Picochlorum* sp. *BPE23* can be further developed. Moreover, the extensive strain characterization provides evidence of how microalgae evolve to supra-optimal temperature and to photobioreactor growth conditions. With this study, microalgal evolutionary mechanisms were identified by combining ALE with genome sequencing.

Keywords: Adaptive laboratory evolution, Microalgae, Temperature, *Picochlorum*, Photobioreactor, Genome assembly, Variant calling, Transcriptomics, Copy number variation

*Correspondence: robin.barten@wur.nl

[1] Bioprocess Engineering & AlgaePARC, Wageningen University and Research, PO Box 16, Wageningen 6700 AA, The Netherlands
Full list of author information is available at the end of the article

Background

Microalgal biomass can be processed into renewable lipids, proteins, carbohydrates, and specialty chemicals [1]. Such novel and renewable products become increasingly relevant in the battle against global warming. Despite the opportunities, microalgal production for the bulk product market is not yet competitive due to high operational costs [2, 3]. One major cost factor for

Barten *et al. BMC Biology* (2022) 20:242

Page 2 of 14

production is the need to cool the photobioreactor, due to high levels of solar irradiance. Strains with the capacity to cope with temperatures as found in photobioreactors placed in high-light regions can lead to a significant reduction in operational costs as less cooling is required [3, 4].

Adaptive laboratory evolution (ALE) has been proposed as a high potential tool for improving robustness against a wide variety of abiotic stress factors, such as high temperature [5]. ALE was adopted just recently in microalgal biotechnology but has been proposed as a powerful way to produce genetic mutations that lead to increased fitness in stressful environments [6, 7]. ALE studies commonly yield a rapid shift in the optimal and upper growth temperature of 2–3 °C, whereafter the rate of adaptation reduces significantly [8–11]. The initial rapid adaptation was hypothesized to be caused by pleiotropy [11]. Temperature as a stress factor impacts nearly every cellular process as it affects the protein configuration and membrane fluidity [12]. In addition to the general cellular metabolism, microalgae grow autotrophically through photosynthesis which is considered a thermosensitive process [4]. In prior studies, *Picochlorum* sp. *BPE23* was found to rapidly regulate its photosystems and the formation of storage compounds in order to align the photosynthetic energy production with the energy output to prevent the formation of harmful reactive oxygen species [13].

Significant steps in temperature tolerance are commonly generated by an accumulation of mutations in multiple genes [14]. Next to single nucleotide polymorphisms (SNPs) and genetic variants, copy number variations (CNVs) that are caused by partial genome duplications are hypothesized to drive evolution [15]. The extent to which CNVs contribute to the evolution and phenotypic variation is a topic of scientific interest [16]. In *Arabidopsis thaliana*, CNVs were found to cause significant genome diversity and physiologic differences, which consecutively drove adaptation towards an applied stress factor [15]. CNVs were also proposed to drive adaptation to wastewater in *Chlorella* [17].

In this study, we applied ALE to expand the upper-temperature boundary of the already thermotolerant microalga *Picochlorum* sp. *BPE23* [18]. *Picochlorum* has become one of the most relevant and promising microalgal species over the past years due to its robustness to environmental factors such as high temperature, salinity, and irradiance levels, in combination with its high growth rate. The relevance is also illustrated by the fact that the genome of *Picochlorum* sp. *SENEW3*, a closely related species, became the genome of the month in Trends in Genetics [19]. The species of *Picochlorum* have been used before in evolutionary studies on thermotolerance,
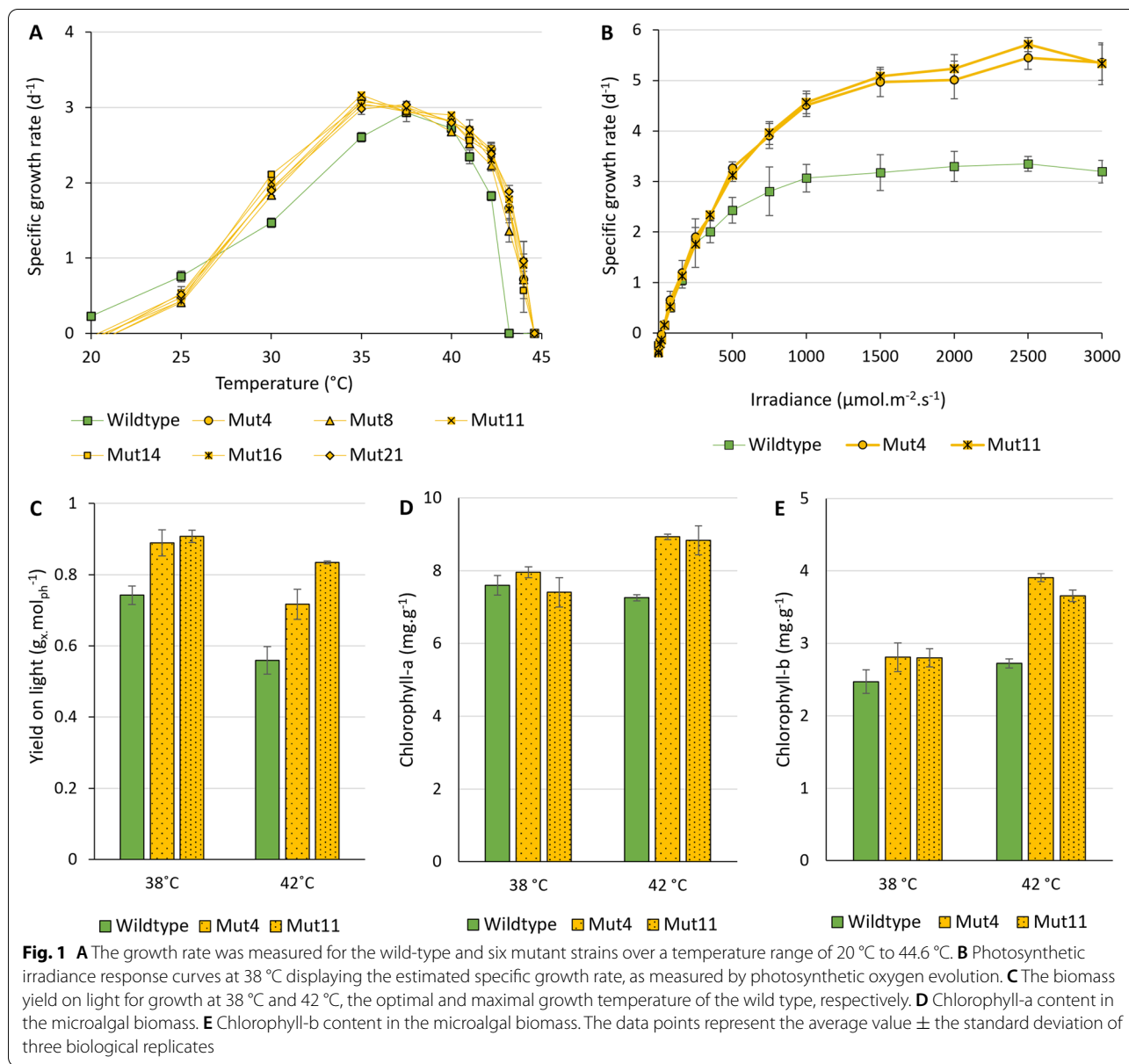
halotolerance, and irradiance tolerance due to their robustness and strong adaptation capacity in combination with a small genome size of 13–14 Mb [20–23]. In addition to the phenotypic characterization of growth and biomass composition, we present valuable genomic resources which can be used in future studies. Genome sequencing of both the wild-type and temperature-adapted mutant strains through PacBio HiFi sequencing was performed to generate a phased genome assembly. This assembly was then used to identify genomic mutations and for the subsequent transcriptome analysis. After analysis of all data, hypotheses were formed on the underlying mechanisms behind the improved phenotype of the mutants. This study provides an example of ALE applied to microalgae combined with extensive genetic characterization of mutations to identify potential evolutionary adaptation mechanisms for thermotolerance and domestication.

Results

Adaptive laboratory evolution yields an improved mutant strain

ALE was performed for 322 days, during which 293 generations were cultivated (Additional file 1: Fig. S1). The temperature was increased stepwise in small increments after the cell culture fitness recovered from the previous temperature step. In this way, the maximal temperature tolerated by the cell culture shifted from 42 to 44.6 °C at the end of the ALE trajectory.

After mutant strain isolation, six mutant strains and the wild-type strain were subjected to growth characterization over a temperature range of 20 to 44.6 °C (Fig. 1A). During these screening experiments, a continuous irradiance of 200 μmol m^{-2} s^{-1} was applied. The mutant strains showed an expanded upper-temperature boundary which was expected based on the growth behavior of the cell culture during the ALE trajectory. ALE-induced shifts in the upper-temperature boundary are commonly accompanied by a trade-off at the lower-temperature boundary [24]. Such a trade-off was observed for the mutant strains as they were not able to grow at 20 °C and, in addition, showed reduced growth at 25 °C compared to the wild-type strain. Interestingly, the growth rate at 30 °C and 35 °C was higher for the mutant strains than for the wild-type strain. This increase in growth rate was not expected for these temperatures as selective pressure during ALE was at a supra-optimal temperature. Nonetheless, a broader temperature optimum is advantageous during application for commercial production in photobioreactors due to the fluctuating temperatures in such systems when not controlled. The applied light intensity of 200 μmol m^{-2} s^{-1} is considered a low to average intensity for microalgal cultivation since solar light intensities

Fig. 1 **A** The growth rate was measured for the wild-type and six mutant strains over a temperature range of 20 °C to 44.6 °C. **B** Photosynthetic irradiance response curves at 38 °C displaying the estimated specific growth rate, as measured by photosynthetic oxygen evolution. **C** The biomass yield on light for growth at 38 °C and 42 °C, the optimal and maximal growth temperature of the wild type, respectively. **D** Chlorophyll-a content in the microalgal biomass. **E** Chlorophyll-b content in the microalgal biomass. The data points represent the average value ± the standard deviation of three biological replicates

can fluctuate up to values larger than 2000 µmol m^{-2} s^{-1}. Photosynthetic irradiance curves were made under the optimal growth temperature (38 °C) to compare the maximal specific growth rates of the mutant and wild-type strains (Fig. 1B). Mut4 and mut11 displayed a maximal specific growth rate of 5.45 ± 0.23 days^{-1} and 5.71 ± 0.14 days^{-1}, respectively, which is significantly higher than the maximal specific growth rate of the wild type, 3.35 ± 0.15 day^{-1}. The growth rate of the wild-type strain, grown on nitrate, is comparable to values reported in previous research [25].

Mut4, mut11, and the wild type were further characterized for their growth rate and biomass yield on light

in controlled photobioreactors. Biomass yield on light measured at steady state is displayed in Fig. 1C, whereas growth and adaptation trends prior to steady state are presented in Additional file 1: Fig. S2. The growth rate of the wild-type and mutant strains decreased after the temperature change from 38 to 42 °C, as expected based on the growth data displayed in Fig. 1A. Interestingly, the mutant strains reached a new steady state within days without a strong growth rate decrease or a decreased quantum yield, whereas the wild-type strain required 19 days to reach a new steady state with a significant temporary drop in growth and a decreased quantum yield. In fact, two steady-state situations were recognized for the

Barten *et al. BMC Biology*   (2022) 20:242

Page 4 of 14

wild-type strain, being at 9 and 19 days after the increase in temperature, with biomass yield on light values of 0.31 $\pm$ 0.11 g_x mol^{-1} and 0.56 $\pm$ 0.04 g_x mol^{-1}, respectively (Additional file 1: Fig. S2). In prior research, transcriptomic and compositional analysis after a temperature increase to 42 °C revealed an immediate system-wide stress response, followed by gradual acclimation [12]. A similar response is hypothesized to have taken place in this study, whereby the first steady-state situation at day 9 is considered to be part of the acclimation phase towards the actual steady-state phase at day 19.

The concentration (in mg g^{-1}) of both chlorophyll-a and chlorophyll-b, but also all xanthophyll cycle pigments and β-carotene, did not significantly differ between wild-type and mutant strains at 38 °C, despite the higher biomass yield on light (Fig. 1D, E, Additional file 1: Fig. S3). Surprisingly, both mutant strains increased their chlorophyll-a and chlorophyll-b content at 42 °C, indicating that the mutant cells actively increased their photosynthetic light absorbance in response to the temperature increase. Supra-optimal temperature commonly causes photoinhibition to prevent cellular damage through reactive oxygen species (ROS) [26]. This would lead to a severe stress response with decreased growth and reduced photosystems as a result [13]. However, the opposite was observed. Hypothesizing the cellular machinery which copes with excessive heat may have mutated to be more active under this condition, thereby allowing more ROS formation without direct cellular damage. Such waste of light comes at the cost of a reduced photosynthetic biomass yield on light. Microalgal strains naturally exhibit different biomass yields on light [25]. Potential causes for the increased photosynthetic efficiency are increased photochemical efficiency, altered antenna size and filtering pigment composition, and altered biomass composition [4, 27].

Based on growth data and literature examples, potential mechanisms behind the expanded temperature tolerance, the increased maximal specific growth rate, and the increased biomass yield on light are difficult to designate. To elucidate the evolutionary mechanisms underlying the observed physiologic changes, biomass samples were taken during the aforementioned steady states to analyze the mRNA expression levels, gDNA variations, pigment composition, and fatty acid composition).
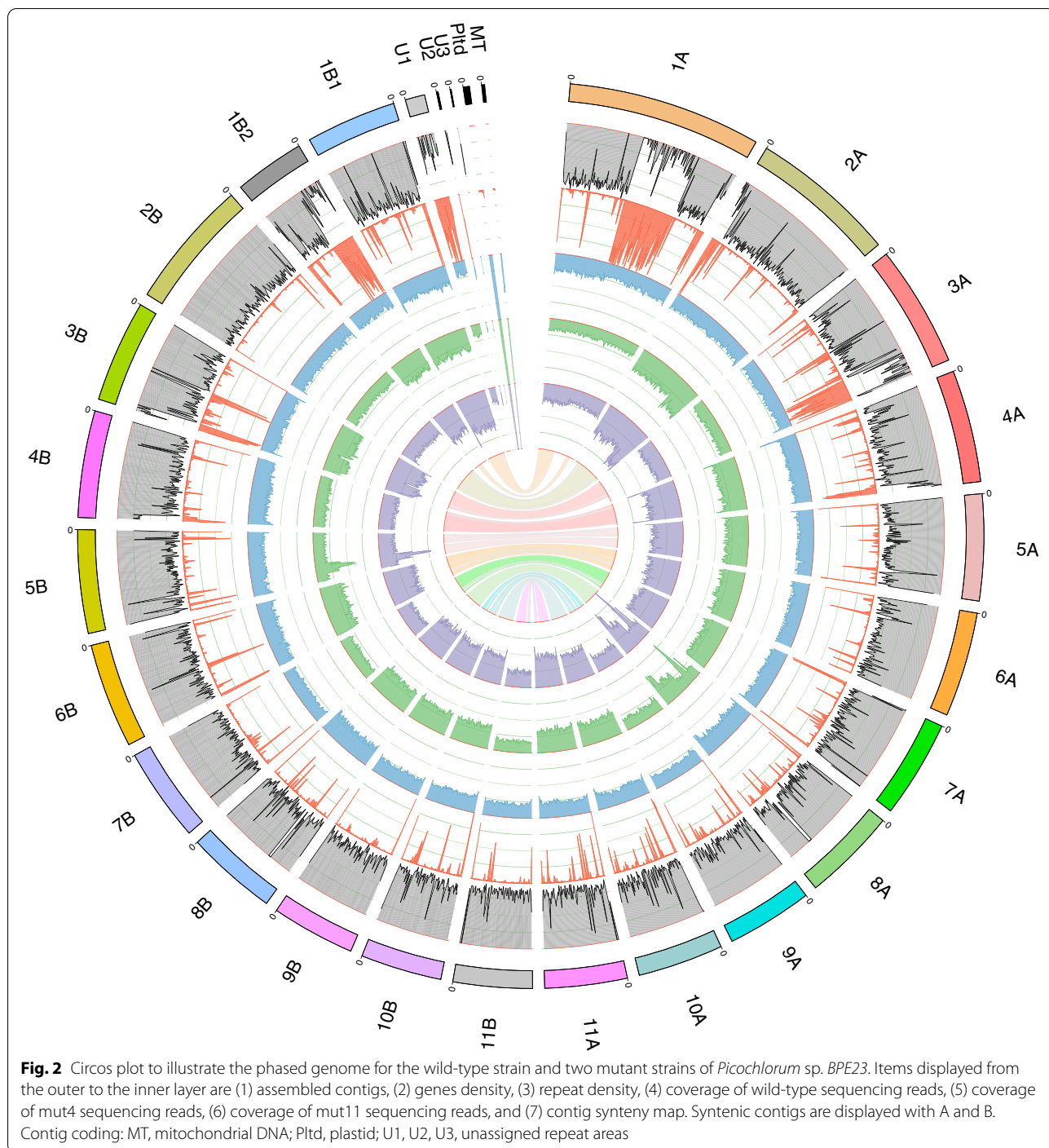
Partial genome multiplication and gene variants were identified

Using the 199× coverage PacBio HiFi sequencing data with an average length of 18 kb, we produced a fully phased diploid genome assembly for wild-type *Picochlorum* sp. *BPE23*. The assembled contigs, with a total length of 28,372,001 bp, are displayed in Fig. 2. For contig 1A, the long sequencing reads bridged a repeat area, whereas contig 1B1 and 1B2 were not connected. Contig U1, U2, and U3 contain unassigned repeat areas. The plastid and mitochondrial genomes were both assembled in one contig. In the assembly, 14,247 genes were identified that code for 14,443 predicted transcripts, of which 12,283 were annotated with at least one functional protein domain. This is the second phased genome for a species of *Picochlorum* [28].

In addition, mut4 and mut11 were sequenced (PacBio HiFi) with a coverage of 231× and 315× and an average read length of 17 kb, to characterize the genomic differences with the wild type. Interestingly, both mutant strains show a double or even higher coverage of the baseline signal for specific genomic regions, indicating a multiplication of at least 2-fold for nearly half the genome (Fig. 2, tracks 5 and 6). The sequencing coverage of the wild-type strain was constant over nearly the complete genome, indicating accurate assembly and no polyploidy. Both mutant strains exhibited identical genomic multiplication patterns. The genomes of mut8 and mut16 were sequenced using Illumina PE150 and showed the same genome multiplication patterns (Additional file 1: Fig. S5). Microscopic karyotyping is suggested as a follow-up method to study how the partial genome multiplication materialized. Hypothetically, whole chromosomes or chromosome arms were duplicated.

While multiplication of genes with particular functions might be expected in the mutant strains, GO term enrichment analysis did not reveal any significant gene functions. Even the seemingly narrow peaks on contig 5B and 8A with 3 to 4 times genetic multiplication contain genes with a large variety in gene functionality. CNVs are hypothesized to drive adaptation and evolution, as shown in *Arabidopsis thaliana* [15]. Multiple gene copies give mutational freedom as a defective mutation does not immediately result in decreased fitness or cell death, allowing for the accumulation of mutations. In a thermotolerance ALE study with *Saccharomyces cerevisiae*, large genetic duplications were observed in 5 out of 7 mutant strains [8]. However, these genetic duplications were said to have little effect on thermotolerance, and the thermotolerant phenotype was dedicated largely to gene mutations. There is a consensus that after significant genome multiplication, a large number of the non-beneficial duplicated gene copies are lost over time [29]. In a different study with *S. cerevisiae*, such chromosomal duplications disappeared over time as they were unstable and imposed an energetic burden on the cell [30]. The authors speculate that large genetic duplications are acquired as a crude solution to stress, after which more subtle and

Barten *et al. BMC Biology*   (2022) 20:242

Page 5 of 14

Fig. 2 Circos plot to illustrate the phased genome for the wild-type strain and two mutant strains of *Picochlorum* sp. *BPE23*. Items displayed from the outer to the inner layer are (1) assembled contigs, (2) genes density, (3) repeat density, (4) coverage of wild-type sequencing reads, (5) coverage of mut4 sequencing reads, (6) coverage of mut11 sequencing reads, and (7) contig synteny map. Syntenic contigs are displayed with A and B. Contig coding: MT, mitochondrial DNA; Pltd, plastid; U1, U2, U3, unassigned repeat areas

efficient mutations are accumulated over time that make the chromosomal duplication obsolete.

As all mutant strains show similar growth and thermotolerance patterns, we hypothesize that the improved phenotype originated from mutations that both strains have in common. An overview of in-gene SNPs and nucleotide insertions and deletions is presented in Table 1. In total, 220 mutations were identified of which 154 were intergenic mutations shared by both mutant strains and 45 were found in only one mutant strain. In addition, 15 gene mutations were shared by both mutant strains, and 6 were found in only 1 mutant. Of all mutations, 76.8% are shared between mutant strains, whereas 71.4% of the genic mutations are shared.

Barten *et al. BMC Biology*   (2022) 20:242

Page 6 of 14

By deducing from the fact that most shared mutations are present on only one of the multiplicated gene copies, it can be concluded that the partial genomic replication took place close to the beginning of the ALE trajectory. Moreover, it is likely that the introduction of the shared mutations between the two mutant strains dates to the moment before the cell lines separated. The large number of shared mutations compared to the non-shared mutations indicates that the parental strain of mut4 and mut11 was dominant in the ALE cell culture until just before the end of the ALE trajectory.

In ALE strain improvement trajectories, the most significant gain in terms of temperature tolerance is gained at the beginning, after which additional mutations have a diminished effect on fitness improvement due to network complexity [7]. Therefore, it is likely that the shared mutations gained at the start of the ALE experiment caused the thermotolerant phenotype.

Most of the identified mutations were found in intergenic non-coding regions at the ends of a contig while few mutations were found in regions between coding genes. Most intergenic DNA has no currently known function. However, some intergenic DNA stretches are known to enhance the translation of nearby genes [31]. Non-coding RNA was demonstrated to regulate acclimation to heat [32]. There were a few intergenic mutations present close to coding genes. We found 249 non-coding RNAs in the assembled genome of *Picochlorum* sp. *BPE23*, but none of the identified mutations was present in a non-coding RNA region. It is therefore unlikely that mutations in intergenic regions were responsible for the improved phenotype. However, the current state of knowledge on intergenic regions does not allow for finite conclusions about the impact of intergenic mutations.

Of the 21 mutated genes, only few showed significant up- or downregulation with log2 fold change values of 1–2 times (Table 1). Most mutations involve a SNP that introduces an amino acid change in the protein whereas a few mutations involve nucleotide insertions/deletions. Three genic mutations and one intronic mutation are unique to mut4, and two intronic mutations are unique to mut11 (PicoBPE23_07590 and PicoBPE23_11307). Moreover, two intronic mutations occurred at the same location in both mutant strains with different nucleotide substitutions. The differences between mutant strains are of special interest as mut11 showed a higher biomass yield on light at 42 °C and performed slightly better than mut4, in general. Mut4 had extra mutations in introns of genes associated with formiminotransferace catalytic domain superfamily, pumilio RNA-binding repeat, ribosomal protein S12 methylthiotransferase RimO, and RNA recognition motif domain. Lastly, the ploidy of polyketide synthase is 2 for mut11. None of these genes exhibits differential expression.

mRNA expression levels

Principal component analysis confirmed that the mRNA expression patterns of the two mutant strains are very similar to each other, both at 38 °C and 42 °C, but clearly distinct from the mRNA expression patterns of the wild-type strain (Additional file 1: Fig. S6). GO term enrichment was performed to investigate the mRNA expression activity at a process scale (Fig. 3). The mutant strains showed comparable transcription patterns, with minor differences. It must be noted that most GO terms enriched in only one mutant strain have a *p*-value that is relatively close to the cutoff value.

Both wild-type and mutant strains suppressed the processes associated to photosynthesis and the central metabolism in response to a supra-optimal temperature of 42 °C. In addition, many RNA processing and protein processing processes were activated. Regardless of the expanded maximal growth temperature of the mutant strains, 42 °C is still a supra-optimal temperature for both the wild-type and the mutant strains by which reduced metabolic activity was expected. While photosynthesis and metabolism were downregulated in both the wild-type and the mutant strains in response to increased temperature, the wild-type strain showed stronger downregulation.

When comparing the mutant strains to the wild-type strain, many biological processes associated to photosynthesis and energy metabolism were upregulated in the mutant strains, both at 38 °C and at 42 °C (Fig. 3). The increased transcriptional activity of photosynthesis and the energy metabolism corresponds to the observation that mutant cells grew faster with a higher yield on light. While not significantly upregulated as a pathway, several genes associated with nitrogen assimilation and amino-acid biosynthesis were upregulated in the mutant strains. On the other hand, the downregulation of response to oxygen-containing compounds in both mutant strains indicates a lower necessity for ROS deactivation mechanisms. Both mutant strains have downregulated mRNA processing activity. Moreover, mut11 displays downregulation of a wide variety of RNA processing-associated biological processes. Among the downregulated RNA-associated biological processes was non-coding RNA, which is assumed to be involved in regulating gene activity in response to abiotic stress such as heat [33]. Moreover, mut11 shows downregulation of protein processing processes, associated to the proteasome complex, peptidase complex, and endopeptidase complex. The general perception is that protein degradation is necessary in the

Barten *et al. BMC Biology*   (2022) 20:242

Page 7 of 14

Table 1 An overview of all mutations that are identified within a gene. Mutations shared by both mutants are listed first, whereas unique mutations follow. The REF and ALT columns display the nucleotide substitutions from reference to alternative. The columns mut4 and mut11 display which REF/ALT variant is present on the haplotypes of the genome where 0 represents the wild-type variant and 1/2 represents the alternative variants. An mRNA differential expression heatmap is displayed showing log2 fold changes of the mutated genes. Lastly, a putative functional gene annotation is presented. Priority was given to BlastP matches, and when no good matches were found, InterPro homology domain annotation was used to infer a gene function

gene ID	Genomic position	REF	ALT	effect	mut4	mut11	Ploidy mut4/mut11	Differential expression 38 °C Mut4/mut11	42 °C Mut4/mut11	Putative functional annotation (BlastP / Interpro domain)
PicoBPE23_05040	contig_4A (850888)	A	G	Synonymous	0/1	0/1	2			IPR011004 (Trimeric LpxA-like superfamily)
PicoBPE23_07467	contig_6A (275034)	C	T	Intronic	0/1	0/1	2			-
PicoBPE23_08649	contig_6B (1138611)	C	A	Gly273Cys	0/1	0/1	2			IPR041667 (Cupin-like domain 8)
PicoBPE23_07264	contig_5B (1135004)	C	A	Pro17Thr	1/1	1/1	1			BlastP: Superoxide dismutase / PANTHER: IPR044624 (repeat protein Mbb1-like)
PicoBPE23_04067	contig_3B (126306)	C	T	Asp353Asn	0/1	0/1	2			IPR036279 (5'-3' exonuclease, C-terminal domain superfamily)
PicoBPE23_08830	contig_7A (268737)	G	A	Pro34Ser	0/1	0/1	2			IPR006634 (TRAM/LAG1/CLN8 homology domain)
PicoBPE23_00798	contig_1B1 (69387)	C	T	Gln193	0/1	0/1	2			IPR040122 (Importin beta family)
PicoBPE23_11836	contig_9B (455009)	C	T	Synonymous	0/1	0/1	2			IPR003439 (ABC transporter-like, ATP-binding domain)
PicoBPE23_13246	contig_11A (185704)	T	C	Met277Val	0/1	0/1	2			BlastP: RNA polymerase II C-terminal domain phosphatase-like
PicoBPE23_00222	contig_1A (498474)	GGT	G	Frame shift	1/1	1/1	1			IPR036955 (AP2/ERF domain superfamily)
PicoBPE23_02273	contig_2A (1365932)	T	TG,TGG	Intronic	1/2	0/1	2			IPR022384 (Formiminotransferase catalytic domain superfamily)
PicoBPE23_03888	contig_3A (994228)	CATAT	C,CAT	Intronic	1/1	2/2	1			IPR001313 (Pumilio RNA-binding repeat)
PicoBPE23_09102	contig_7A (770005)	T	TCAA	Intronic	0/1	0/1	2			IPR038071 (UROD/MetE-like superfamily)
PicoBPE23_10084	contig_8A (331679)	C	CAT,CATAT	Intronic	1/2	1/2	1			IPR026854 (Vacuolar protein sorting-associated protein 13-like, N-terminal)
PicoBPE23_01916	contig_2A (766369)	C	T	Arg401Trp	0/1	0/0	2			IPR001012: UBX domain / IPR018997: PUB domain
PicoBPE23_07590	contig_6A (491187)	G	GTA	Intronic	0/0	0/1	2			IPR005840: Ribosomal protein S12 methylthiotransferase RimO
PicoBPE23_04568	contig_3B (1111698)	G	C	Met393Ile	0/1	0/0	2			IPR027806: Harbinger transposase-derived nuclease domain
PicoBPE23_10884	contig_8B (697809)	C	CAT	Intronic	0/1	0/0	2			IPR001680: WD40 repeat
PicoBPE23_00982	contig_1B1 (457245)	A	G	His257Arg	0/1	0/0	1/2			BlastP: Polyketide synthase
PicoBPE23_11307	contig_9A (463100)	GATA	G	Intronic	0/0	0/1	1			IPR000504: RNA recognition motif domain

wild-type strain to break down damaged proteins as a result of heat stress [4, 32].

Interestingly, we did not observe significant over-expression in genes associated with a heat-shock response. In previous research, we exposed the wild-type strain of *Picochlorum* sp. *BPE23* to 5 days of 42 °C and monitored the cell physiology whereby we observed a severe system-wide heat-shock response on a transcription, growth, and cell composition level [12]. The absence of a heat-stress response in this study confirms that our approach to wait for a steady state (19 days) before sampling was required to be able to compare the growth of the mutant strain to the wild-type strains at an optimal and supra-optimal temperature without disturbance of a heat-shock response.
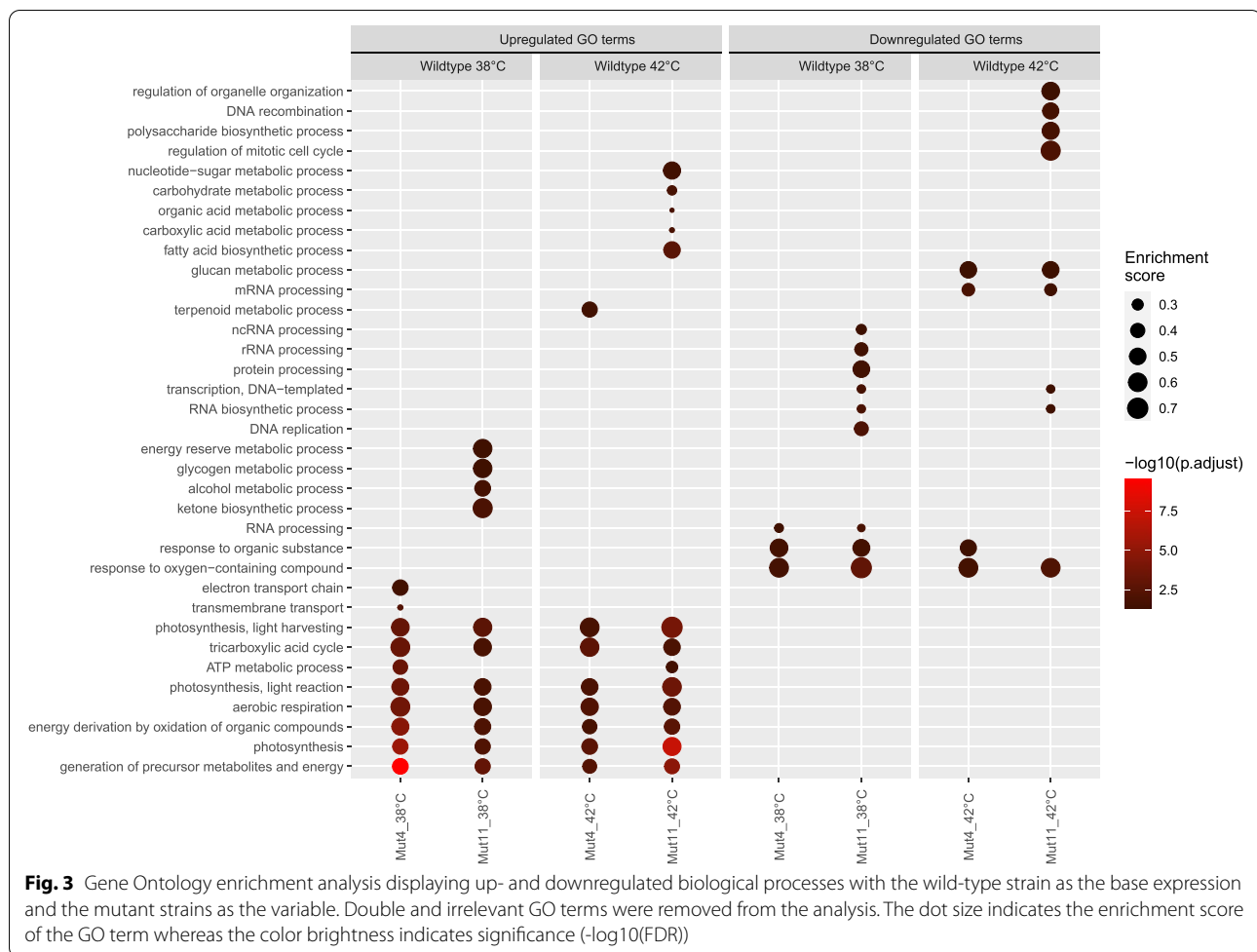
Discussion
Adapted mechanisms underlying the physiology improvement
Both the CNVs and the gene mutations can theoretically have caused the thermotolerant phenotype. As several hard selection events (i.e., temperature increases) were applied throughout the ALE trajectory, it is hypothesized that not one but multiple consecutive adaptation events are responsible for the final improved phenotype. Increased selective pressure causes a wipe-out event after which part of the cell lines in the culture are eliminated [34]. Based on cell function, a selection of mutated genes is hypothesized to impact microalgal thermotolerance and growth in general. Due to the diploid nature of *Picochlorum* sp. *BPE23* deleterious mutations can be compensated for by the second copy of the gene on the second allele. This effect induces uncertainty in the impact of the observed mutations. Moreover, microalgal species do not have the same annotation quality as industrialized model organisms such as *Escherichia coli* and *S. cerevisiae*. Nonetheless, strong predictions on gene function can be made based on protein homology and protein domains.

Adaptation to temperature stress
The effect of heat on cells is system-wide but impacts some cellular components and processes more severely, whereby cells have evolved adaptation mechanisms [12]. The cell membranes are among these components as temperature affects membrane fluidity and therefore unbalances many biological processes. In response, cells remodel the cell membrane composition. Three gene mutations were associated with temperature tolerance specifically. We identified a SNP in PicoBPE23_07264 (ceramide synthase), inducing a Pro34Ser substitution. This gene encodes for the waxy sphingolipid ceramide. In *S. cerevisiae*, de novo synthesis of ceramide synthase was shown to be elevated after heat stress [35].

Fig. 3 Gene Ontology enrichment analysis displaying up- and downregulated biological processes with the wild-type strain as the base expression and the mutant strains as the variable. Double and irrelevant GO terms were removed from the analysis. The dot size indicates the enrichment score of the GO term whereas the color brightness indicates significance (-log10(FDR))

Due to heat-induced unbalancing in the photosynthetic machinery and the electron transport chain, ROS are formed causing oxidation of cellular components [4]. Photosynthetic organisms downregulate their photosystems to prevent this negative impact which naturally leads to reduced growth. Superoxide dismutase performs an important stress-relieving function by degrading ROS that form due to heat stress [36]. The Pro17Thr substitution to PicoBPE23_07264 could affect its function or rate. This gene's expression levels increased slightly in the mutant strains. Also, PicoBPE23_13246, encoding for a putative RNA polymerase II C-terminal domain phosphatase-like 4, was found mutated. While intuitively one does not associate this enzyme to heat tolerance, several studies in *A. thaliana* mutator strains showcased its role in tolerance to ROS-inducing abiotic stress factors such as heat through interaction with specific transcription factors and regulatory proteins [37].

Regulatory genes

Several mutations were identified in genes associated with mRNA transcription and regulation. The mRNA expression of PicoBPE23_03888, associated with mRNA stability regulation, was halved in the mutant strains indicating a lower requirement of the specific mRNA. The transcriptional activator PicoBPE_00222 was found to have undergone a frameshift mutation in both haplotype versions. A study in *A. thaliana* showed that mutagenesis of this gene caused temperature sensitivity [38]. Arguably, mutations in this gene could contribute to thermotolerance. Transcriptional activators and regulators are often found mutated in thermotolerant ALE-derived mutant strains [8].

Photosynthesis, carbon fixation, and the energy metabolism

Transcriptome analysis revealed many upregulated biological processes associated with photosynthesis, carbon fixation, and energy metabolism in the mutant strains. These were upregulated both at 38 °C and 42

Barten *et al. BMC Biology*   (2022) 20:242

Page 9 of 14

°C compared to the wild type. While most differentially expressed genes had a log2 fold change below 2, a selection of genes was upregulated more severely. Both haplotype versions of RuBisCO, PicoBPE23_01132, and PicoBPE23_00347 were among the most upregulated transcripts with approximately 8 times higher expression. Transcript expression levels were equal for both genes while only the gene on contig 1B1 had undergone a copy number duplication, indicating that transcription was regulated based on the cellular demand for the enzyme. Comparable observations were made for other carbon fixation-associated genes and photosynthesis, glycolysis, and citric acid cycle-associated genes.

Two genes putatively involved in photosynthesis were mutated: PicoBPE23_09102 and PicoBPE23_07264. The first gene contains a UROD domain which catalyzes a step in chlorophyll synthesis whereas the second gene is putatively involved in the regulation of multiple photosystem II subunits, as shown by homology to *Chlamydomonas* [39]. However, both genes show even stronger sequence similarity to other genes, showcasing the difficulty of interpreting genetic data for non-model organisms. None of the gene mutations points exclusively at photosynthesis. While large-scale CNVs do not directly lead to increased enzyme production due to product-regulated mRNA expression, such an event inevitably has a significant effect on cellular physiology. Ultimately, microalgae are capable of regulating energy uptake through antenna size alteration to suit the growth conditions and prevent metabolic imbalance [13, 22]. The observed upregulation of transcription photosynthesis, carbon fixation, and energy metabolism and the increased chlorophyll concentration at 42 °C could also result from pathways downstream or stress-relieving mechanisms that became more efficient under specific growth conditions, thereby allowing all energy uptake mechanisms to be upregulated by the cell.

Nitrogen metabolism and amino acid biosynthesis

We hypothesize that in addition to evolution to supra-optimal temperature, evolution took place towards growth in the photobioreactor systems and to the microalgal growth medium. The wild-type *Picochlorum* sp. *BPE23* was isolated from Bonairean waters and was not cultured extensively in the laboratory prior to ALE, by which we assume that the wild-type strain did not have mutations that benefit growth on microalgal growth medium. It is known that ALE reproducibly yields side effect mutations that are beneficial for growth in bioreactors on a minimal medium [5]. In the case of *E. coli*, increased substrate utilization efficiency in adaptation to glucose is a recurring observation, among other mechanisms [14]. Light was the energy-providing substrate for our photoautotrophic cultivations, whereas nitrate was added to provide nitrogen for de novo amino acid assimilation. Several mutations can be associated with nitrogen metabolism and amino acid biosynthesis, and several genes associated with these pathways were severely overexpressed. Two mutated genes were involved in amino acid synthesis (PicoBPE23_09102 and PicoBPE23_02273). Moreover, three mutated genes were associated with the transport of cellular compounds and proteins (PicoBPE23_00798, PicoBPE23_11836, and PicoBPE23_10084).

Confirmation of mutations

Recurring ALE mutations in microalgal strains are not yet known, but their identification becomes increasingly feasible due to the increasing number of ALE studies. Many recent ALE studies have applied genetic sequencing methods such as transcriptomics or genome resequencing to identify mutations [5]. However, these studies were performed with varying microalgal species and ALE targets. With this elaborate case study, we provide strong arguments for evolutionary mechanisms towards both temperature and growth in a minimal medium. Looking forward, more ALE studies with closely related species are required to elucidate recurring adaptation mechanisms to supra-optimal temperature and growth in photobioreactors.

To validate the hypothesized impact of the mutated genes on thermotolerance, they should be genetically engineered in the wild-type strain. Recently, genetic modification tools have been developed for a closely related strain, *Picochlorum renovo* [23]. A strong example of an efficient engineering approach to study the impact of ALE mutations on thermotolerance is demonstrated in *E. coli* by the use of the MAGE technology [14]. After genetic engineering of *E. coli* with the mutations as identified after ALE, an enrichment under selective pressure was followed by resequencing of the cell culture to establish the presence of the mutated genes in the cell culture for their impact on fitness.

Conclusions

ALE with supra-optimal temperature as a selective pressure was used to create thermotolerant strains of *Picochlorum* sp. *BPE23*. The mutant strains had an expanded maximal growth temperature of 44.6 °C, whereas the maximal growth temperature of the wild type was 42 °C. The mutant strains have a superior maximal growth rate. Moreover, the mutant strains have an increased biomass yield on light at the optimal and maximal growth temperature. The mutant strains showed partial genome multiplication and 21 genic

mutations. Based on protein homology, several mutations were associated with mechanisms that could be responsible for increased growth or thermotolerance. Transcriptome analysis showed that photosynthesis and metabolic processes were upregulated in the mutant strains compared to the wild type, while the heat stress-associated processes were downregulated. We hypothesize that evolution towards supra-optimal temperature occurred and was accompanied by evolution towards microalgal growth medium.

From an applied point of view, the improved mutant strains can replace the wild-type strain for biomass production. The biomass yield on light of mut11 increased from 0.74 to 0.91 g_x mol^{-1} at 38 °C, which directly improves biomass productivity. The broader temperature optimum of the mutant strains stimulates the productivity levels over a wider temperature range, whereby less strict photobioreactor process control is required. Moreover, the maximal specific growth rate of mut11 increased from 3.35 ± 0.15 day^{-1} to 5.71 ± 0.14 day^{-1}, allowing for faster upscaling of cell cultures in addition to faster strain improvement trajectories.

With this study, a strong platform was created to enhance the development of *Picochlorum* sp. *BPE23*. A phased genome was assembled to enable future projects using targeted genetic strain modification and fundamental genetic research.

Methods
Cell cultivation
Growth media and inoculum preparation
Picochlorum sp. *BPE23*, isolated from a coastal bay on Bonaire, was pre-cultivated in shake flasks in an orbital shaker incubator (Multitron, Infors HT) under continuous light at an intensity of 100 µmol m^{-2} s^{-1} photosynthetically active radiation (PAR) [18]. The temperature was set at 40 °C, and the relative humidity of the air in the headspace was set to 60% and enriched with 2% CO_2. Cells were cultured in artificial seawater enriched with nutrients and trace elements. Elements were provided at the following concentrations (in g L^{-1}): NaCl, 24.50; MgCl$_2$·6H$_2$O, 9.80; Na$_2$SO$_4$, 3.20; NaNO$_3$, 2.12; K$_2$SO$_4$, 0.85; CaCL$_2$.2H$_2$O, 0.80; KH$_2$PO$_4$, 0.23; Na$_2$EDTA·2H$_2$O, 0.11; NaFeEDTA, 3.96·10^{-2}; MnCl$_2$·2H$_2$O, 1.71·10^{-3}; ZnSO$_4$·7H$_2$O, 6.60·10^{-4}; Na$_2$MoO$_4$·2H$_2$O, 2.42·10^{-4}; Co(NO$_3$)$_2$·6H$_2$O, 7.00·10^{-5}; NiSO$_4$·6H$_2$O, 2.63·10^{-5}; CuSO$_4$·5H$_2$O, 2.40·10^{-5}; K$_2$CrO$_4$, 1.94·10^{-5}; Na$_3$VO$_4$, 1.84·10^{-5}; and H$_2$SeO$_3$, 1.29·10^{-5}. HEPES (4.77 g L^{-1}) was added to Erlenmeyer cultures as a pH buffer. The medium pH was adjusted to 7.0, after which it was filter sterilized before use. During photobioreactor cultivation, Antifoam B (J.T.Baker, Avantor, USA) was added at a concentration of 0.5 mL L^{-1} from a 1% w/w% stock. Also,

0.168 g L^{-1} Sodium bicarbonate (NaHCO$_3$) was added at inoculation to provide sufficient CO_2 at the start of the cultivation. The photobioreactor was inoculated at a starting cell density of OD$_{750}$ 0.2.

Adaptive laboratory evolution and mutant isolation
ALE was performed in heat sterilized flat-panel photobioreactors with a 1.8-L working volume, a 20.7-mm light path, and a 0.08-m^2 surface area for irradiation (Labfors 5 lx, Infors HT, Switzerland) [13]. The incident irradiance was set at a continuous rate of 813 µmol m^{-2} s^{-1} (PAR). This irradiance rate was chosen as the total quantity of light is then equal to the total light of an average day in the Caribbean but distributed homogeneously over 24 h instead of as a diel pattern. The cell density was kept within a certain range through automatic photobioreactor dilution based on the penetration of light through the cell culture (Turbidostat mode). The photobioreactor was diluted to maintain a light penetration level of 10 µmol m^{-2} s^{-1} (PAR). The photobioreactor was aerated through sparging of compressed air at a rate of 980 mL min^{-1}. CO_2 was provided on demand through indication of an increased pH. The reactor pH was set at 7. The temperature was increased stepwise throughout the evolution process for 322 consecutive days as follows: 30 °C, 42 °C, 42,5 °C 43 °C, 43.5 °C, 44 °C, 44.2 °C, 44.7 °C, 44.2 °C, 44.4 °C, 44.6 °C, and 44.8 °C (Additional file 1: Fig. S1). At 44.6 °C, a sample was taken from the photobioreactor, and cells were plated on agar plates followed by incubation at 42 °C for 1 week. Colonies were picked, transferred into liquid growth medium, and subjected to a growth screening using the HT24 incubating unit (Algenuity, UK) in which a screening temperature of 42 °C and a light intensity of 200 µmol m^{-2} s^{-1} (PAR) were applied. The 6 strains with the highest growth rate were selected for more elaborated growth rate and thermotolerance characterization.

Growth rate screening—Erlenmeyer cultivation
The wild-type and six mutant strains with numbers 4, 8, 11, 14, 16, and 21 were cultured at 20 °C, 25 °C, 30 °C, 35 °C, 37.5 °C, 40 °C, 41 °C, 42.2 °C, 43.2 °C, 44 °C, and 44.6 °C to determine the maximum growth rate over a temperature range. Cells were cultured in 250 mL Erlenmeyer flasks with a liquid volume of 100 mL. Repeated batch cultivation was applied in which cultures were diluted daily to a cell concentration of OD$_{750}$ 0.05. The OD$_{750}$ was measured at 6 and 24 h after inoculation. Broad-spectrum white fluorescent light (Philips master TL-D reflex, color temperature 840) was set at a continuous intensity of 200 µmol m^{-2} s^{-1} (PAR). A 2% CO_2

Barten *et al. BMC Biology*   (2022) 20:242

Page 11 of 14

headspace was created. Data acquisition was made at steady-state [18].

Characterization of mutant strains—photobioreactor operation

Sterilized flat panel photobioreactors (Algaemist; Technical Development Studio, WUR) with a 0.38-L working volume and a 14-mm optical depth were used for characterization and comparison of growth characteristics and RNA expression patterns of the wild-type and two evolved strains of *Picochlorum* sp. *BPE23* [25]. Cells were grown in turbidostat mode with similar growth conditions as in the "Adaptive laboratory evolution and mutant isolation" section, with the exception that air was supplied at a rate of 200 mL min^{-1}. Experiments were performed as biological triplicates, and samples were taken in a steady-state for both 38 °C and 42 °C. Temperature was initially set at 38 °C until samples were taken, after which it was increased to 42 °C until a new steady state was reached.

Offline measurements and analysis

The quantum yield, cell size, cell number, and biomass concentration were measured as described in [13]. Biomass harvest and biomass lyophilization, determination of pigments through HPLC, and determination of fatty acids through GC was performed as described in [12]. The photosynthesis irradiance curve and the maximal specific growth rate were measured as described in [25].

Genomic DNA analysis

Genome sequencing

gDNA extraction and sequencing were performed by the genomics facility of WUR Bioscience (Wageningen, The Netherlands). gDNA of the wild type, mut4, and mut11 was sequenced through PacBio sequencing. Fifty milligrams of washed and lyophilized biomass was used for gDNA extraction using MagMax Plant DNA kit (Applied Biosystems). Extracted gDNA was sheared using a megadisruptor device (Diagenode), aiming for a 25-kb fragment size. Barcoded SMRT bell template libraries were made for each microalgal strain using SMRT bell express template prep kit 2.0 (PacBio). gDNA was eluted in 20 µL elution buffer and pooled for size selection on a Blue Pippin instrument (Saga Sciences). The 8 to 13 kb and ≥ 13 kb fractions were collected for each lane after which purification and concentration were performed with Ampure XP beads (Beckman Coulter). Purified SMRT bells were finally analyzed for size and used for primer annealing, polymerase binding, and purified using AMPure PB beads (SMRTLink10.1). Sequencing was done on one SMRT cell of a PacBio Sequel IIe system

with 65 pM on plate loading by diffusion with standard settings and 30-h movie time. Data demultiplexing and HiFi read processing were done by SMRT Link 10.1.

Illumina paired-end sequencing of the wild type, mut4, mut8, mut11, and mut16 with 150 bp reads was performed by Novogene (Novogene, UK). For library preparation, DNA was randomly fragmented to a size of 350 nt, then end polished, A-tailed, ligated with Illumina sequencing adapters, followed by PCR enrichment of P5 and P7 oligos. The PCR product was purified (AMPure XP system) and sequenced.

Genome assembly

The genome assembly of wild-type *Picochlorum* sp. *BPE23* was reconstructed by assembling the HiFi reads using Flye 2.9-b1768 with default settings [40]. Next, Illumina sequencing data of the wild-type strain were used to polish the Flye genome assembly with Pilon v1.24 with argument "--fix all" [41]. For polishing, the paired-end Illumina reads were trimmed with Trimmomatic v0.39 "ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36" and mapped to the assembly with BWA-MEM2 v2.2.1 [42, 43]. Sorting and indexing of mapping files were done with SAMtools v1.12 [44]. The chloroplast and mitochondrial gDNA were labeled by mapping to the assembly of *Picochlorum* sp. *"soloecismus"* strain DOE 101 with minimap2 "-xasm20" [45]. The identity of the chloroplast and mitochondrial gDNA fragments was then confirmed by a BLASTN 2.11.0+ hit to the NCBI nt database [46]. Homologous contigs were identified by leveraging the location of the duplicated BUSCO genes followed by self-alignment with minimap2 v2.10 ("-k19 -w19 -A1 -B9 -O16,41 -E2,1 -s200 -z200 -X") [47]. To assess the completeness of the genome assembly, BUSCO v5.2.2 was used with the chlorophyta_odb10 set as well as KAT v2.4.1 [48, 49]. To further assess the structure of the genome assembly, it was aligned to the closely related genomes of *Picochlorum SENEW3*, *Picochlorum oklahomensis*, and *Picochlorum celeri* [20, 28]. In addition, the original HiFi data was mapped back with minimap2 v2.22. The coverage of the HiFi data was calculated using BEDtools v2.30.0 [50]. HiFi and Illumina data were compared for wild type, mut4, and mut11. Further study towards genome ploidy was performed through KMC v3.1.1 and Smudgeplot v0.2.3 [51, 52].

Genome annotation

Before annotation of the wild-type genome assembly, repeats were identified in the polished assembly with RepeatModeler v2.0.2 with argument "-LTRStruct" and

subsequently masked with RepeatMasker 4.1.2-p1 [53]. The RNA-seq samples of the wild type were mapped to the genome with HISAT2 v2.2.1 to serve as RNA evidence, whereas the proteome of *Chlamydomonas reinhardtii* (GCF_000002595.2) served as protein evidence for gene prediction [54]. Gene prediction was performed with BRAKER v2.1.6 with argument "--etpmode" [55]. Mitochondrial and chloroplast sequences that were included in RNA-seq data mapping were excluded for gene prediction. Functional annotation was done with InterProScan 5.52-86.0 [56]. The included analyses were TIGRFAM, SUPERFAMILY, PANTHER, Gene3D, Coils, Pfam, and MobiDBLite. In addition, InterPro and Gene Ontology (GO) terms were listed. An organism-specific annotation R package was created using the MakeOrgPackage function from the AnnotationForge Package v1.32.0 [57]. Lastly, non-coding RNAs were identified by searching the genome with Cmscan from the Infernal toolkit against the Rfam database [58].

Genome analysis

Variant calling to identify mutations in the ALE strains was performed by aligning HiFi reads to the phased genome assembly with Minimap v2.22 with argument "-ax map-hifi --MD --secondary=no." Variants were called with GATK v4.3.2.0 HaplotypeCaller, combined with CombineGVCFs and genotyped with GenotypeGVCFs [59]. Filtering variants with GATK SelectVariants was done on QUAL < 1000, homozygous REF allele for wild type, and non-homozygous REF allele for either mutant strain. We used SnpEff 5.0 with the novel wild-type genome and annotation to determine the effect of the mutations on the amino acid sequence [60].

Enrichment analyses were performed with genes that had doubled or tripled coverage in mut4 or mut11 compared to the wild type using ClusterProfiler in R v4.0.4 [61]. MCScanX was used to identify syntenic blocks based on an all-vs-all BLASTP search (Wang et al. 2012) [62].

Circos v0.69-8 was used to visually combine all data on the framework of the wild-type genome assembly [63]. Gene and repeat density were calculated in bins of 10 kb.

mRNA extraction, sequencing, and analysis

Biologically triplicate biomass samples of wild type, mut4, and mut11 steady-state cell cultures at 38 °C and 42 °C (paragraph 2.1.4) were analyzed for mRNA levels. Biomass pre-treatment, mRNA extraction and sequencing, and data quality control were performed as described by Barten et al. [12]. Fragments of the mRNA library with an average size of 300 bp were sequenced using the Illumina Novoseq PE150 platform, yielding paired-end reads of 150 nt (Novogene, UK).

Paired-end reads were mapped to the phased wild-type genome using HISAT2 v 2.2.1. Transcripts were assembled using StringTie v1.3.2d. StringTie's prepDE Python script was used to extract raw read counts per gene. Raw counts were normalized using DESeq2 v1.30.1. Pairwise differential expression between mutant strains and wild type at different temperatures was calculated using DESeq2, setting alpha to 0.05. Enriched GO terms were detected using ClusterProfiler's gseGO function and visualized using the Ggplot2 R package.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-022-01431-y.

Additional file 1: Figure S1. Photobioreactor dilution rate and temperature during the adaptive laboratory evolution process. **Figure S2.** The photobioreactor dilution rate during characterization of the wildtype, mut4, and mut11. **Figure S3.** The concentration of pigments in the wildtype and mutant strains at 38 °C and 42 °C. **Figure S4.** The concentration of polar fatty acids in the wildtype and mutant strains at 38 °C and 42 °C. **Figure S5.** Coverage of genome sequencing reads for the wildtype, mut4, mut8, mut11, and mut16. **Figure S6.** PCA bi-plot of the sequenced mRNA samples collected from steady-state cultures at 38 °C and 42 °C.

Authors' contributions

RB: conceptualization, methodology, data acquisition, data analysis, and writing—original manuscript. DJvW: methodology, data acquisition, data analysis, and writing—review and editing. EdB: methodology, data acquisition, and data analysis. JR: methodology and data analysis. MK: methodology, data acquisition, and data analysis. SN: methodology and data acquisition. SS: methodology and writing—review and editing. RHW: writing—review and editing and funding acquisition. HN: methodology, data analysis, and writing—review and editing. MJB: conceptualization, methodology, writing—review and editing, and funding acquisition. The author(s) read and approved the final manuscript.

Availability of data and materials

Wild-type and mutant strains of *Picochlorum* sp. *BPE23* are stored in liquid nitrogen and are available upon reasonable request. A web portal (https://www.bioinformatics.nl/picochlorum_bpe23), including a genome browser, is available for scientists to use in their research [64]. Included items are the genomes of the wild type, mut4, and mut11. Moreover, all mRNA sequencing data and the identified mutations are included. Moreover, genomic sequence read archives and assemblies are available at NCBI under the umbrella project number: PRJNA872549 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA872549) [65].

Declarations

Consent for publication

Not applicable.

Barten *et al. BMC Biology*   (2022) 20:242

Page 13 of 14

Author details
[1]Bioprocess Engineering & AlgaePARC, Wageningen University and Research, PO Box 16, Wageningen 6700 AA, The Netherlands. [2]Bioinformatics Group, Wageningen University and Research, PO Box 633, Wageningen 6700 AP, The Netherlands. [3]Biosciences and Aquaculture, Nord University, N-8049 Bodø, Norway.

References

1. Wijffels RH, Barbosa MJ. An outlook on microalgal biofuels. Science. 2010;329(5993):796–9.
2. Benvenuti G, Ruiz J, Lamers PP, Bosma R, Wijffels RH, Barbosa MJ. Towards microalgal triglycerides in the commodity markets. Biotechnol Biofuels. 2017;10(1):1–10.
3. Ruiz J, Olivieri G, de Vree J, Bosma R, Willems P, Reith JH, et al. Towards industrial products from microalgae. Energy Environ Sci. 2016;9(10):3036–43.
4. Ras M, Steyer J-P, Bernard O. Temperature effect on microalgae: a crucial factor for outdoor production. Rev Environ Sci Biotechnol. 2013;12(2):153–64.
5. LaPanse AJ, Krishnan A, Posewitz MC. Adaptive laboratory evolution for algal strain improvement: methodologies and applications. Algal Res. 2020;53:102122.
6. Cooper VS. Experimental evolution as a high-throughput screen for genetic adaptations. MSphere. 2018;3(3):e00121–18.
7. Sandberg TE, Salazar MJ, Weng LL, Palsson BO, Feist AM. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. Metab Eng. 2019;56:1–16.
8. Caspeta L, Chen Y, Ghiaci P, Feizi A, Buskov S, Hallström BM, et al. Altered sterol composition renders yeast thermotolerant. Science. 2014;346(6205):75–8.
9. Tillich UM, Wolter N, Franke P, Dühring U, Frohme M. Screening and genetic characterization of thermo-tolerant Synechocystis sp. PCC6803 strains created by adaptive evolution. BMC Biotechnol. 2014;14(1):1–15.
10. O'Donnell DR, Hamman CR, Johnson EC, Kremer CT, Klausmeier CA, Litchman E. Rapid thermal adaptation in a marine diatom reveals constraints and trade-offs. Glob Chang Biol. 2018;24(10):4554–65.
11. Cooper VS, Bennett AF, Lenski RE. Evolution of thermal dependence of growth rate of Escherichia coli populations during 20,000 generations in a constant environment. Evolution. 2001;55(5):889–96.
12. Barten R, Kleisman M, D'Ermo G, Nijveen H, Wijffels RH, Barbosa MJ. Short-term physiologic response of the green microalga Picochlorum sp. (BPE23) to supra-optimal temperature. Sci Rep. 2022;12(1):1–12.
13. Barten R, Djohan Y, Evers W, Wijffels R, Barbosa M. Towards industrial production of microalgae without temperature control: the effect of diel temperature fluctuations on microalgal physiology. J Biotechnol. 2021;336:56–63.
14. Sandberg TE, Pedersen M, LaCroix RA, Ebrahim A, Bonde M, Herrgard MJ, et al. Evolution of Escherichia coli to 42 C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. Mol Biol Evol. 2014;31(10):2647–62.
15. DeBolt S. Copy number variation shapes genome diversity in Arabidopsis over immediate family generational scales. Genome Biol Evol. 2010;2:441–53.
16. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. Science. 2004;305(5683):525–8.
17. Wu T, Li L, Jiang X, Yang Y, Song Y, Chen L, et al. Sequencing and comparative analysis of three Chlorella genomes provide insights into strain-specific adaptation to wastewater. Sci Rep. 2019;9(1):1–12.
18. Barten RJ, Wijffels RH, Barbosa MJ. Bioprospecting and characterization of temperature tolerant microalgae from Bonaire. Algal Res. 2020;50:102008.
19. da Roza PA, Goold HD, Paulsen IT. Picochlorum sp. SENEW3. Trends Genet. 2022;38(2):209–10.
20. Foflonker F, Mollegard D, Ong M, Yoon HS, Bhattacharya D. Genomic analysis of Picochlorum species reveals how microalgae may adapt to variable environments. Mol Biol Evol. 2018;35(11):2702–11.
21. Krasovec M, Vancaester E, Rombauts S, Bucchini F, Yau S, Hemon C, et al. Genome analyses of the microalga Picochlorum provide insights into the evolution of thermotolerance in the green lineage. Genome Biol Evol. 2018;10(9):2347–65.
22. Weissman JC, Likhogrud M, Thomas DC, Fang W, Karns DA, Chung JW, et al. High-light selection produces a fast-growing Picochlorum celeri. Algal Res. 2018;36:17–28.
23. Dahlin LR, Gerritsen AT, Henard CA, Van Wychen S, Linger JG, Kunde Y, et al. Development of a high-productivity, halophilic, thermotolerant microalga Picochlorum renovo. Commun Biol. 2019;2(1):1–9.
24. Caspeta L, Nielsen J. Thermotolerant yeast strains adapted by laboratory evolution show trade-off at ancestral temperatures and preadaptation to other stresses. MBio. 2015;6(4):e00431–15.
25. Barten R, Chin-On R, de Vree J, van Beersum E, Wijffels RH, Barbosa M, et al. Growth parameter estimation and model simulation for three industrially relevant microalgae: Picochlorum, Nannochloropsis, and Neochloris. Biotechnol Bioeng. 2022;119(6):1416–\.
26. Mathur S, Agrawal D, Jajoo A. Photosynthesis: response to high temperature stress. J Photochem Photobiol B Biol. 2014;137:116–26.
27. Masojídek J, Ranglová K, Lakatos GE, Silva Benavides AM, Torzillo G. Variables governing photosynthesis and growth in microalgae mass cultures. Processes. 2021;9(5):820.
28. Becker SA, Spreafico R, Kit JL, Brown R, Likhogrud M, Fang W, et al. Phased diploid genome sequence for the fast-growing microalga Picochlorum celeri. Microbiol Resour Announc. 2020;9(20):e00087–20.
29. Kaltenegger E, Leng S, Heyl A. The effects of repeated whole genome duplication events on the evolution of cytokinin signaling pathway. BMC Evol Biol. 2018;18(1):1–19.
30. Yona AH, Manor YS, Herbst RH, Romano GH, Mitchell A, Kupiec M, et al. Chromosomal duplication is a transient evolutionary solution to stress. Proc Natl Acad Sci. 2012;109(51):21010–5.
31. Jegousse C, Yang Y, Zhan J, Wang J, Zhou Y. Structural signatures of thermal adaptation of bacterial ribosomal RNA, transfer RNA, and messenger RNA. PLoS One. 2017;12(9):e0184722.
32. Zhao J, He Q, Chen G, Wang L, Jin B. Regulation of non-coding RNAs in heat stress responses of plants. Front Plant Sci. 2016;7:1213.
33. Jha UC, Nayyar H, Jha R, Khurshid M, Zhou M, Mantri N, et al. Long non-coding RNAs: emerging players regulating plant abiotic stress response and adaptation. BMC Plant Biol. 2020;20(1):1–20.
34. Barrick JE, Lenski RE. Genome dynamics during experimental evolution. Nat Rev Genet. 2013;14(12):827–39.
35. Wells GB, Dickson RC, Lester RL. Heat-induced elevation of ceramide in Saccharomyces cerevisiae via de novo synthesis. J Biol Chem. 1998;273(13):7235–43.
36. Arriola MB, Velmurugan N, Zhang Y, Plunkett MH, Hondzo H, Barney BM. Genome sequences of Chlorella sorokiniana UTEX 1602 and Micractinium conductrix SAG 241.80: implications to maltose excretion by a green alga. Plant J. 2018;93(3):566–86.
37. Thatcher LF, Foley R, Casarotto HJ, Gao L-L, Kamphuis LG, Melser S, et al. The Arabidopsis RNA polymerase II carboxyl terminal domain (CTD) phosphatase-like1 (CPL1) is a biotic stress susceptibility gene. Sci Rep. 2018;8(1):1–14.
38. Jofuku KD, Den Boer B, Van Montagu M, Okamuro JK. Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. Plant Cell. 1994;6(9):1211–25.
39. Vaistij FE, Boudreau E, Lemaire SD, Goldschmidt-Clermont M, Rochaix J-D. Characterization of Mbb1, a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast psbB/psbT/psbH gene cluster in Chlamydomonas reinhardtii. Proc Natl Acad Sci. 2000;97(26):14813–8.
40. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37(5):540–6.
41. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 2014;9(11):e112963.
42. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

Barten *et al. BMC Biology*   (2022) 20:242

Page 14 of 14

43. Vasimuddin M, Misra S, Li H, Aluru S, editors. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. IEEE International Parallel and Distributed Processing Symposium (IPDPS). 2019:314–24.

44. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.

45. Gonzalez-Esquer CR, Twary SN, Hovde BT, Starkenburg SR. Nuclear, chloroplast, and mitochondrial genome sequences of the prospective microalgal biofuel strain Picochlorum soloecismus. Genome Announc. 2018;6(4):e01498–17.

46. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):1–9.

47. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100.

48. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021;38(10):4647–54.

49. Mapleson D, Garcia Accinelli G, Kettleborough G, Wright J, Clavijo BJ. KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics. 2017;33(4):574–6.

50. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–2.

51. Kokot M, Długosz M, Deorowicz S. KMC 3: counting and manipulating k-mer statistics. Bioinformatics. 2017;33(17):2759–61.

52. Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020;11(1):1–10.

53. Chen N. Using Repeat Masker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics. 2004;5(1):4–10.

54. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.

55. Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genom Bioinform. 2021;3(1):lqaa108.

56. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–40.

57. Carlson M, Pages H. AnnotationForge: tools for building SQLite-based annotation data packages. R Packag. version 1.32. 0; 2020.

58. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics. 2013;29(22):2933–5.

59. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. BioRxiv. 2018:201178.

60. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012;6(2):80–92.

61. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation. 2021;2(3):100141.

62. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.

63. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4.

64. Barten R, Workum DJM, Bakker E, Risse J, Kleisman M, Navalho S, Smit S, Wijffels RH, Barbosa MJ. Picochlorum BPE23 WUR Web Portal 2022. Available from: https://www.bioinformatics.nl/picochlorum_bpe23/.

65. Barten R, Workum DJM, Bakker E, Risse J, Kleisman M, Navalho S, Smit S, Wijffels RH, Barbosa MJ. Picochlorum sp. BPE23 genome assembly. Available from: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA872549.

Publisher's Note