



# Physiological and Genetic Regulation for High Lipid Accumulation by *Chlorella sorokiniana* Strains from Different Environments of an Arctic Glacier, Desert, and Temperate Lake under Nitrogen Deprivation Conditions

 Shanmei Zou,<sup>a</sup> Zheng Huang,<sup>a</sup> Xuemin Wu,<sup>a</sup> Xinke Yu<sup>a</sup>

<sup>a</sup>Jiangsu Provincial Key Laboratory of Marine Biology, College of Resources and Environmental Science, Nanjing Agricultural University, Nanjing, P. R. China

**ABSTRACT** Microalgae can adapt to extreme environments with specialized metabolic mechanisms. Here, we report comparative physiological and genetic regulation analyses of *Chlorella sorokiniana* from different environmental regions of an arctic glacier, desert, and temperate native lake in response to N deprivation, for screening the optimal strain with high lipid accumulation. Strains from the three regions showed different growth and biochemical compositions under N deprivation. The arctic glacier and desert strains produced higher soluble sugar content than strains from the native lake. The arctic glacier strains produced the highest levels of lipid content and neutral lipids under N deprivation compared with strains from desert and native lake. At a molecular level, the arctic strain produced more differentially expressed genes related to fatty acid biosynthesis, glycolysis gluconeogenesis, and glycerolipid metabolism. The important functional genes acetyl coenzyme A (acetyl-CoA) carboxylase (ACCase), fatty acid synthase complex, pyruvate dehydrogenase component, and fatty acyl-acyl carrier protein (acyl-ACP) thioesterase were highly expressed in arctic strains. More acetyl-CoA was produced from glycolysis gluconeogenesis and glycerolipid metabolism, which then produced more fatty acid with the catalytic function of ACCase and acyl-ACP thioesterase in fatty acid biosynthesis. Our results indicated that the *C. sorokiniana* strains from the arctic region had the fullest potential for biodiesel production due to special genetic regulation related to fatty acid synthesis, glycolysis gluconeogenesis, and glycerolipid metabolism.

**IMPORTANCE** It is important to reveal the physiological and genetic regulation mechanisms of microalgae for screening potential strains with high lipid production. Our results showed that *Chlorella sorokiniana* strains from arctic glacier, desert, and temperate native lake had different growth, biochemical composition, and genetic expression under N deprivation. The strains from an arctic glacier produced the highest lipid content (including neutral lipid), which was related to the genetic regulation of fatty acid biosynthesis, glycolysis gluconeogenesis, and glycerolipid metabolism. The functional genes for acetyl-CoA carboxylase, fatty acid synthase complex, pyruvate dehydrogenase component, and fatty acyl-ACP thioesterase in the three pathways were highly expressed in arctic strains. The revelation of physiological and genetic regulation of strains from different environmental regions will contribute to the microalgae selection for high lipid accumulation.

**KEYWORDS** microalgae, extreme environments, lipid accumulation, physiological process, genetic regulation, nitrogen deprivation

Overconsumption of fossil fuel by industry and transportation leads to decreasing world fossil fuel reserves and serious environment problems, e.g., atmospheric pollution and global warming (1). It is important to search for alternative and

**Editor** Eva C. Sonnenschein, Technical University of Denmark

**Copyright** © 2022 Zou et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Shanmei Zou, zousm912@njau.edu.cn.

The authors declare no conflict of interest.

**Received** 1 February 2022

**Accepted** 7 September 2022

**Published** 6 October 2022

ecologically friendly energy resources. Recently, biodiesel, which can be produced from a variety of sources (including plants, animals, and microbes) has been proposed as a renewable energy source (2). The search for sustainable sources of biofuels has led to renewed interest in microalgae as a potential feedstock, due to the potential to synthesize and accumulate large quantities of lipids in some species (3, 4). The advantage of microalgae is that they can accumulate large quantities of oils in the form of triacylglycerols (TAGs), which are preferred renewable oils because they possess a high molar ratio of hydrogen to carbon (4–9). Microalgae grown under nutrient limitation exhibit considerable variation in their biochemical composition, depending on the limiting nutrient and the degree of limitation. The nitrogen deprivation response is perhaps the best-characterized inducer of lipid accumulation in microalgae (10–13). In addition to increased total lipid content, N deprivation can also induce changes in fatty acid chain length and saturation (14, 15). Moreover, some species of microalgae can be efficiently transformed, which makes it possible to enhance the productivity of natural compounds through genetic strain engineering strategies (16, 17).

Among the oleaginous microalgae, the *Chlorella* genus of green algae (*Chlorophyceae*) is commonly considered a promising candidate, due to its high photosynthetic efficiency, lipid productivity, and fast growth ability (18). *Chlorella* microalgae are able to adapt to a variety of environmental conditions, including extreme habitats ranging from desert soil to the polar region (19). Although *Chlorella* has proved to be a potential industrial microalgae for biological engineering, it is unclear whether strains from some environments are more suitable for producing high levels of lipids (15, 20–24).

At present, screening oleaginous microalgae strains which have great potential to produce TAGs is of great importance for biodiesel production. The potentially oleaginous strains could also be used in transgenic microalgae, which could be used as industrial microalgae based on biological engineering (25). This biological engineering technology shows great promise to simplify the production process and significantly decrease biodiesel production costs. Many microalgae can adapt to extreme habitats ranging from desert soil to arctic regions (19, 26). Their long-term adaptation shows that the microalgae can adjust their physiological responses to adapt to such environments (27). In particular, microalgae from extreme environments possess special physiological features and molecular mechanisms for their adaptation (28). In this context, microalgae strains in various environments may have different physiological and genetic regulation metabolisms, including those for growth and lipid accumulation. For example, it was indicated that the increased total lipid content and fatty acid composition of Antarctic microalgae could provide its adaptability to low temperatures (29). Some desert microalgae have also been shown to have lipid productivity (30). Therefore, comparative analyses of microalgae strains from different environmental regions in response to nitrogen stress may be used to screen the potential industrial microalgae for biofuel production and also be used in transgenic microalgae for biofuel production. Presently, few studies have compared microalgae from extreme and native environmental regions, e.g., polar, desert, and temperate regions.

Here, we report the growth, lipid accumulation, and transcriptome analysis of three *Chlorella sorokiniana* strains from different environmental regions, an arctic glacier, desert soil, and temperate native lake, in response to nitrogen stress. The responses in terms of their cell growth, biochemical compositions, lipid productivity, fatty acid profiles, and functional genetic expression to N deprivation were compared. We aimed to reveal the physiological and genetic regulation of microalgae from different environmental regions with N deprivation for selecting potential industrial microalgae and providing the basis for transgenic microalgae in biofuels.

## RESULTS AND DISCUSSION

**Isolation and identification of *Chlorella* strains from different environmental regions of arctic glacier, desert soil, and temperate native lake.** Microalgal biomass, a promising source of biofuel, could contribute to the decrease in our dependence on fossil fuel while offering multiple environmental advantages compared with traditional

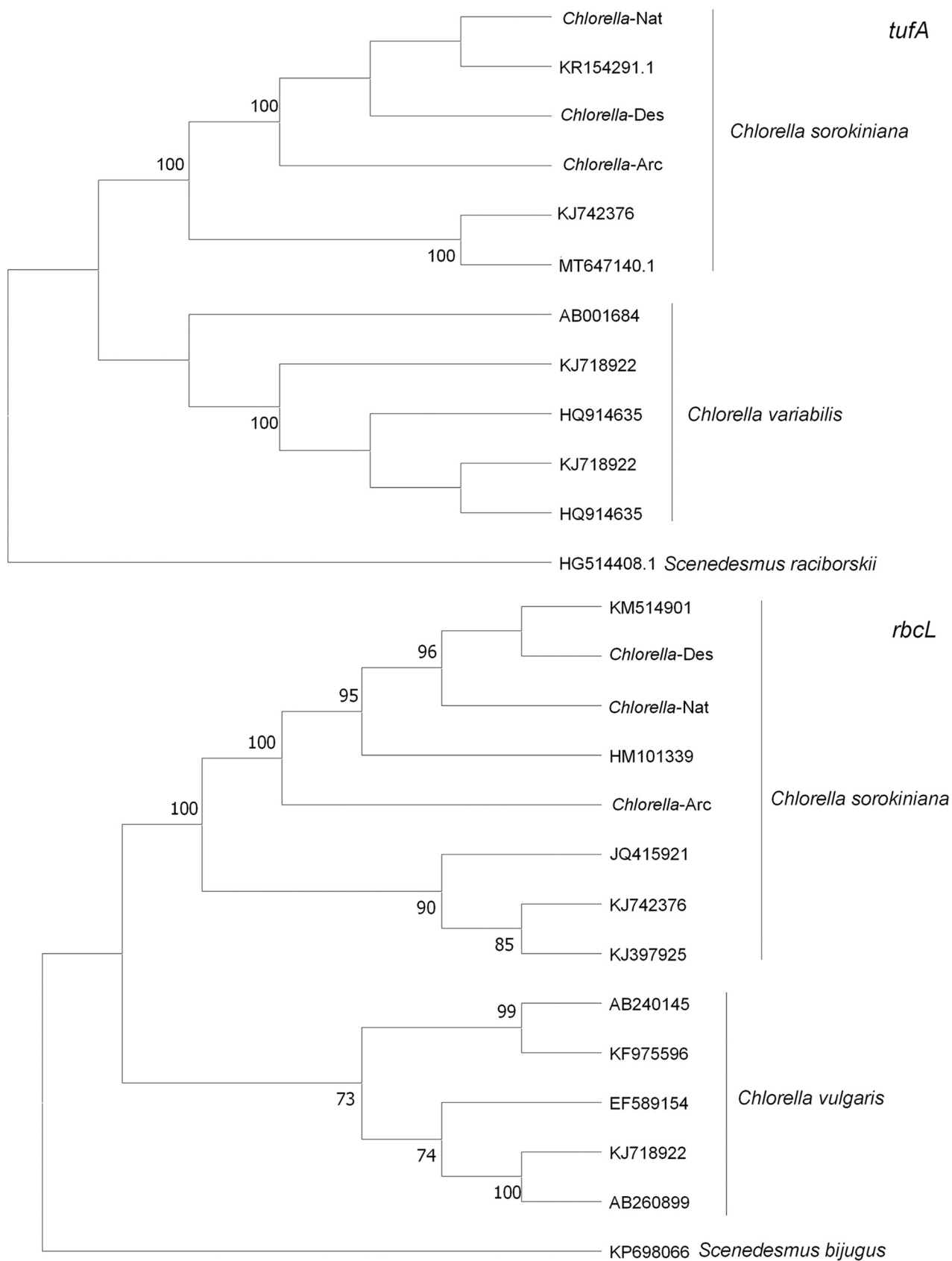
biofuel land crops. Long-term evolution has made possible the development in algae of physiological adaptations to cope with different natural environments (31). It was found that nitrogen deprivation plays an important role in lipid accumulation of microalgae (11–13). It is possible to screen microalgae strains with high lipid production under N deprivation from different environmental regions by comparing their physiological and genetic regulation mechanisms. *Chlorella*, a kind of oleaginous microalgae, is one of the genera of microalgae that has achieved commercial success on a large scale and can occupy extreme habitats. In the present research, three different *Chlorella* sp. strains were isolated from arctic glacier (referred to here as *Chlorella*-Arc), desert soil (*Chlorella*-Des), and temperate native lake (*Chlorella*-Nat) regions. They were first identified based on morphological characteristics. Then, both *rbcl* and *tufA* sequences were obtained from the three strains and were analyzed using BLAST against sequences in GenBank.

The identification of these *Chlorella* sp. strains from the three environmental regions is illustrated in Fig. 1. The corresponding GenBank accession numbers are as follows: for *tufA*, [KR154271.1](#) (*Chlorella*-Nat), [KR154255.1](#) (*Chlorella*-Des), and [ON602074](#) (*Chlorella*-Arc); for *rbcl*, [KM514865.1](#) (*Chlorella*-Nat), [KM514884.1](#) (*Chlorella*-Des), and [ON602073](#) (*Chlorella*-Arc). It was shown that the three strains clustered together with other *Chlorella sorokiniana* strains (from published sequences) in the *tufA* phylogenetic tree, which was independent from the *Chlorella variabilis* clade. The three strains also clustered together with other published *C. sorokiniana* sequences in the *rbcl* phylogenetic tree. Although *Chlorella*-Arc was not much closer to *Chlorella*-Des and *Chlorella*-Nat, it was in the same clade with the other two *C. sorokiniana* sequences. Thus, the results indicated that the three strains from desert, temperate native lake, and arctic glacier were identified as *C. sorokiniana* by molecular tools.

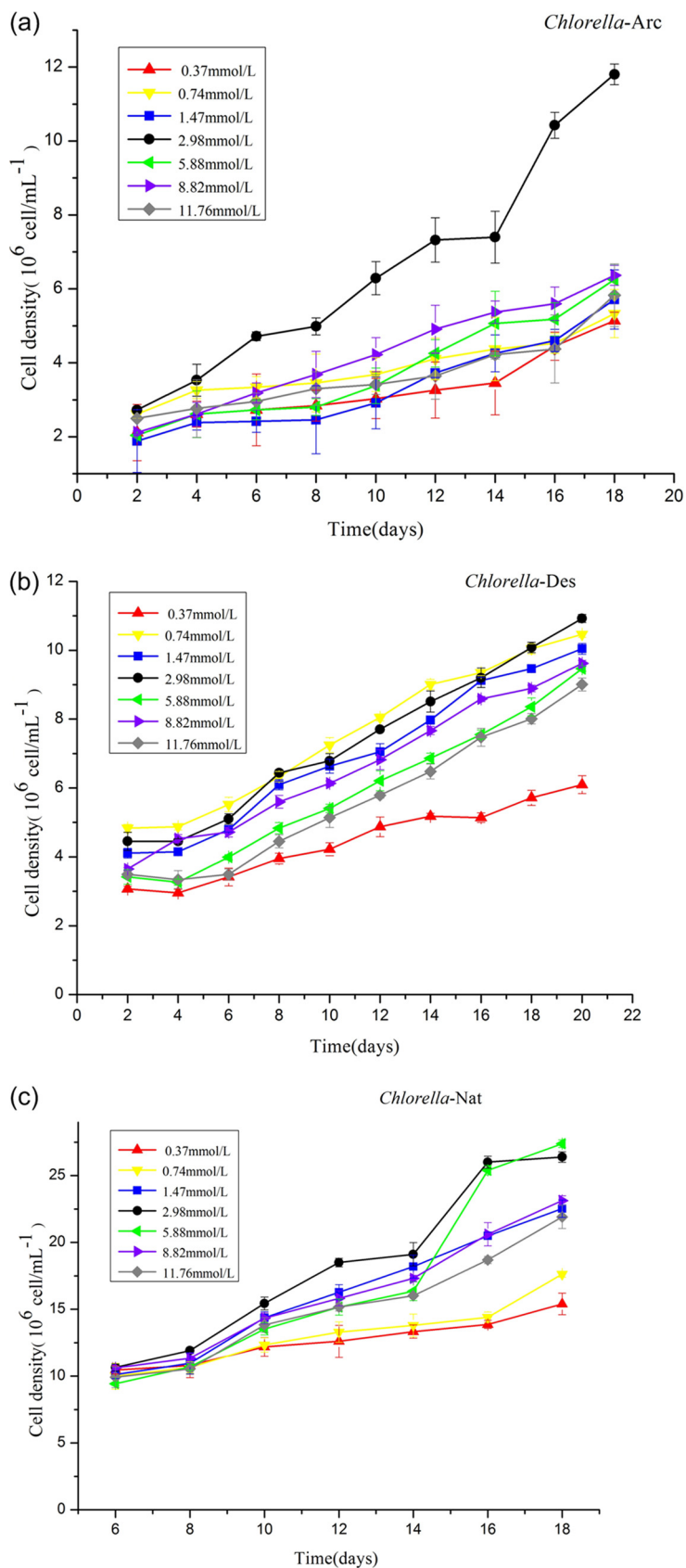
#### **Effect of nitrogen deprivation on growth of the three *C. sorokiniana* strains.**

The growth rates of the three *C. sorokiniana* strains from arctic glacier, desert, and temperate native lake in the medium deprived of different nitrogen are shown in Fig. 2, with numbers of cells provided. Generally, a lower or higher nitrogen concentration than 2.98 mmol/liter did not promote growth for the three strains. *Chlorella*-Des grew well in 0.74 mmol/liter, which may have been due to the extremely scarce nutrition in a desert. In addition, the growth of the three strains seemed to be region dependent under both N deprivation and N-replete cultural conditions. By comparison, the strain from the arctic glacier was more nitrogen sensitive than was the temperate native lake algae. The reason for the difference growth rates may be that the strains in different environmental regions have distinct mechanisms of growth adaptations and oil accumulation. So far, the mechanism of oil accumulation in microalgae, including N gene expression of nitrogen starvation-induced accumulation, is still not clear (13, 24, 32). The distinct responses to N deprivation and N-replete conditions among strains from different environmental regions indicated that microalgae could perform special adaptation to the nutritional limits in a regional environment, which is significant when selecting high-quality strains for producing biodiesel, especially based on genetic engineering.

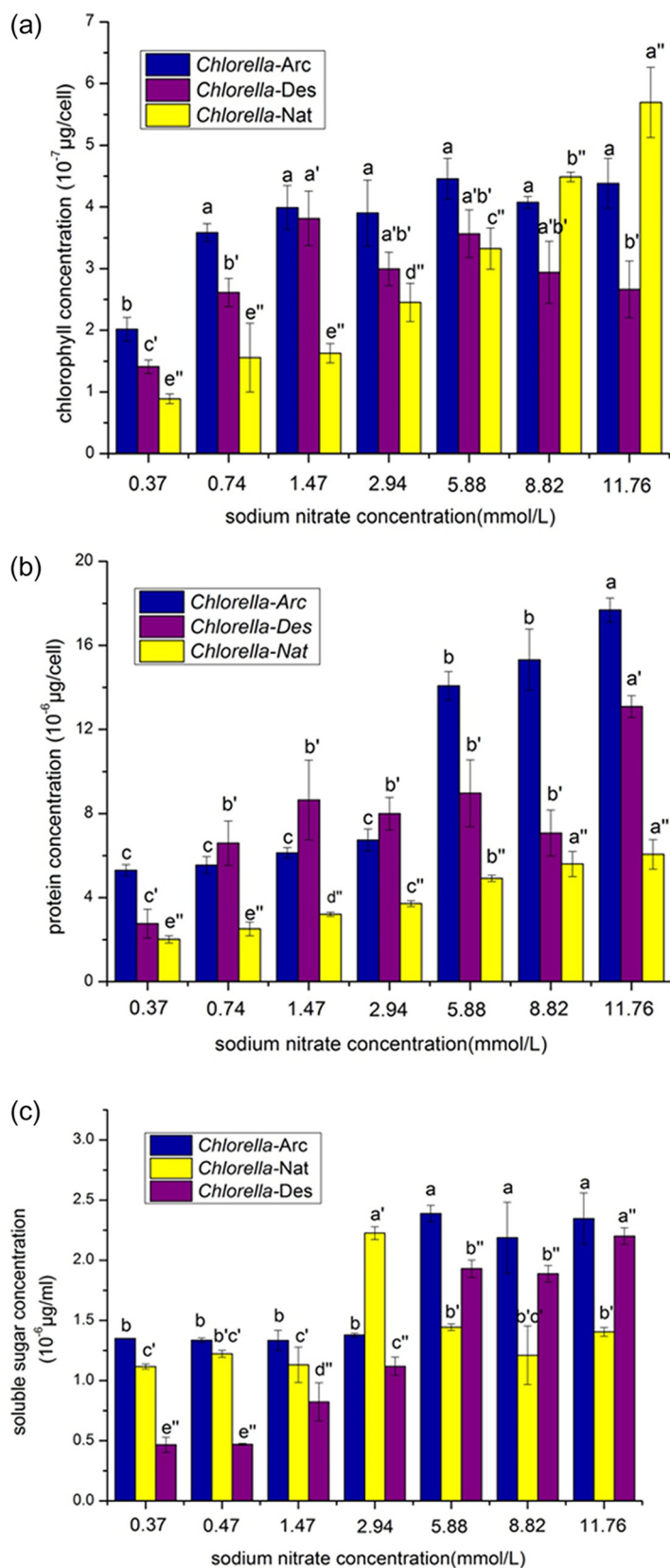
**Effect of nitrogen deprivation on biochemical composition.** The biochemical compositions of *Chlorella*-Arc, *Chlorella*-Des, and *Chlorella*-Nat were all measured at the end of exponential growth. The detailed time for *Chlorella*-Arc was on day 12, and the detailed time for both *Chlorella*-Des and *Chlorella*-Nat was day 10. There was obvious accumulation of chlorophyll in response to the increased nitrogen content for all three strains, but the chlorophyll contents of *Chlorella*-Arc and *Chlorella*-Des were generally constant after the nitrogen content reached 2.94 mmol/liter (Fig. 3a). This result was consistent with a previous report that degradation of chlorophyll could be induced by nitrogen deprivation in *Chlorella protothecoides* (33). Actually, chlorophyll breakdown is the most conspicuous symptom of leaf senescence and fruit ripening (34), and the character of chlorophyll breakdown in microalgae is similar to the visual degreasing observed during leaf senescence and fruit ripening. Compared with *Chlorella*-Nat, the *Chlorella*-Arc and *Chlorella*-Des strains showed higher chlorophyll contents in lower N concentrations between 0.37 and 8.82 mmol/liter. High chlorophyll content could be an indication of high photosynthetic capacity. The *Chlorella*-Arc and *Chlorella*-Des



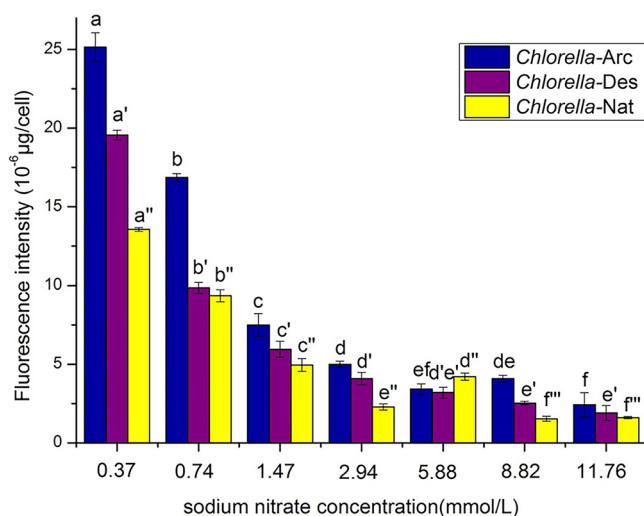
**FIG 1** Maximum-likelihood trees of *Chlorella* strains inferred from *tufA* and *rbcL* gene sequences (the three strains used in this study are underlined). Data are based on 1,000 replications; bootstrap values of >50% are shown above the internodes.



**FIG 2** Growth curves of *Chlorella-Arc* (a), *Chlorella-Des* (b), and *Chlorella-Nat* (c) cultured under different sodium nitrate concentrations. The points represent means  $\pm$  SD of triplicate samples.



**FIG 3** The changes of chlorophyll (a), protein (b), and soluble sugar (c) concentrations of *Chlorella-Arc*, *Chlorella-Des*, and *Chlorella-Nat* cultured under different sodium nitrate concentrations. The points represent means  $\pm$  SD of triplicate samples. Different letters above the bars indicate significant differences ( $P < 0.05$ ) between growth temperatures.



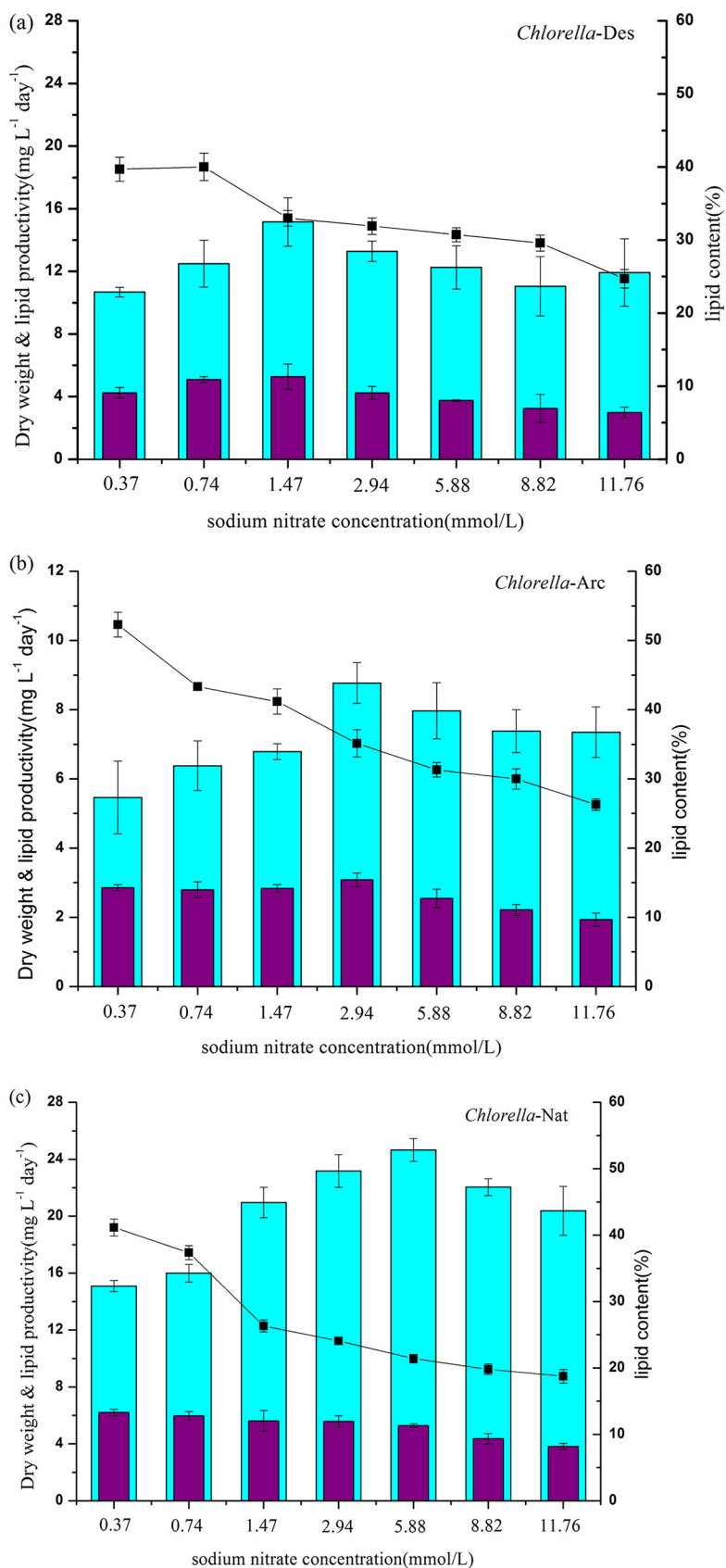
**FIG 4** Changes in neutral lipid concentrations based on fluorescence intensity of *Chlorella-Arc*, *Chlorella-Des*, and *Chlorella-Nat* cultured under different sodium nitrate concentrations. The bar charts represent means  $\pm$  SD of triplicate samples. Different letters above the bars indicate significant differences ( $P < 0.05$ ) between growth temperatures.

strains possibly evolved special regulation mechanisms for photosynthesis to producing chlorophyll under nutrition limitation in adapting to the extreme environments. The *Chlorella-Nat* strain has a normal regulation mechanism for producing chlorophyll. Thus, with increasing nutrition concentrations, the photosynthetic capacity of *Chlorella-Nat* is enhanced and thus the strain produces more chlorophyll. There were no significant changes in protein contents for any strains under the N-deprivation regimen, and the protein contents were just slightly higher for *Chlorella-Des* and *Chlorella-Nat* under the N-replete regimen (Fig. 3b). It was indicated that *Chlorella-Arc* had a significantly higher protein content at all treatments among the tested strains and had an obvious increased trend of protein content in response to the increased nitrogen content; this finding was consistent with a report that protein is the main product of photosynthesis in polar algae due to the lower temperatures (27).

The soluble sugar contents of the three strains did not show consistent trends with respect to N deprivation and N-replete conditions (Fig. 3c). The strains from the desert and the arctic glacier produced higher soluble sugar contents than the strains in the native lake region, indicating that this organism in desert and arctic regions tends to resist the extreme weather by increasing soluble sugar in cells, especially when sufficient nutrition is available. The seasonal and diurnal temperature fluctuations have a strong influence on the metabolic function and photosynthesis of microalgae in outdoor environments (35). For example, *Chlorella* strains in desert have a good capacity for sugar accumulation because of the diurnal temperature fluctuations (35). In addition, it has been demonstrated that microalgae in polar regions can resist the intense radiation by increasing the sugar content in cells (36), and this is consistent with the result here indicating that the strain from the arctic glacier had higher sugar content than the template native lake strain.

**Effect of nitrogen deprivation on neutral lipid and lipid content.** The changes of neutral lipids for *Chlorella-Arc*, *Chlorella-Des*, and *Chlorella-Nat* in response to the N deprivation and N-replete conditions are shown in Fig. 4. We found that the neutral lipid accumulated obviously in response to the N deprivation for all three strains from the different environments (Fig. 4), but it was apparent that the *Chlorella-Arc* strain accumulated the highest content of neutral lipids among the three strains.

The lipid content and lipid productivity of the three *Chlorella* strains are shown in Fig. 5. It can be clearly seen that the *Chlorella* strains from the three environmental regions of polar, desert, and temperate native lake had different changes in lipid contents in



**FIG 5** Biomass productivity, lipid productivity, and lipid content of *Chlorella-Arc*, *Chlorella-Des*, and *Chlorella-Nat* cultured under different sodium nitrate concentrations. The bar charts and points represent means  $\pm$  SD of triplicate samples.



**TABLE 1** Fatty acid compositions of *Chlorella*-Des cultured under different sodium nitrate concentrations

Fatty acid	% of total fatty acids (mean $\pm$ SD) when cultured in NaNO <sub>3</sub> concn of:						
	0.37 mmol/liter	0.74 mmol/liter	1.47 mmol/liter	2.94 mmol/liter	5.88 mmol/liter	8.82 mmol/liter	11.76 mmol/liter
<b>SFAs</b>							
C <sub>16:0</sub>	21.89 $\pm$ 2.40	16.89 $\pm$ 2.12	9.64 $\pm$ 5.81	1.30 $\pm$ 0.21	11.48 $\pm$ 7.38	2.27 $\pm$ 0.09	14.16 $\pm$ 0.88
C <sub>17:0</sub>	9.70 $\pm$ 2.02	12.95 $\pm$ 1.39	15.70 $\pm$ 0.25	14.92 $\pm$ 1.77	16.06 $\pm$ 3.78	15.66 $\pm$ 1.27	12.72 $\pm$ 0.20
C <sub>18:0</sub>	1.33 $\pm$ 0.09	0.20 $\pm$ 0.34	0.36 $\pm$ 0.63	0.31 $\pm$ 0.54	0.64 $\pm$ 0.71	ND	0.34 $\pm$ 0.48
C <sub>20:0</sub>	ND <sup>a</sup>	0.13 $\pm$ 0.22	0.17 $\pm$ 0.30	0.44 $\pm$ 0.38	0.28 $\pm$ 0.25	ND	0.28 $\pm$ 0.39
Total	32.92 $\pm$ 4.53	30.15 $\pm$ 2.65	29.30 $\pm$ 1.38	16.97 $\pm$ 2.29	28.46 $\pm$ 2.92	17.93 $\pm$ 1.37	27.49 $\pm$ 2.31
<b>MUFAs</b>							
C <sub>16:1</sub>	3.03 $\pm$ 1.45	2.26 $\pm$ 3.28	5.55 $\pm$ 0.91	7.58 $\pm$ 0.84	3.86 $\pm$ 3.49	4.52 $\pm$ 0.08	5.52 $\pm$ 1.78
C <sub>17:1</sub>	2.65 $\pm$ 0.31	2.30 $\pm$ 2.01	3.26 $\pm$ 0.30	3.52 $\pm$ 0.21	2.03 $\pm$ 1.76	3.44 $\pm$ 0.02	3.45 $\pm$ 0.05
C <sub>20:1</sub>	27.65 $\pm$ 1.54	36.14 $\pm$ 1.70	39.71 $\pm$ 7.16	40.50 $\pm$ 0.52	37.60 $\pm$ 1.32	40.30 $\pm$ 0.11	34.76 $\pm$ 2.02
C <sub>18:1n9t</sub>	2.57 $\pm$ 0.53	ND	ND	ND	ND	ND	ND
C <sub>18:1n9c</sub>	15.66 $\pm$ 0.15	17.66 $\pm$ 2.18	17.67 $\pm$ 4.20	17.92 $\pm$ 1.07	15.92 $\pm$ 1.47	18.55 $\pm$ 0.43	16.07 $\pm$ 0.90
Total	51.57 $\pm$ 2.92	58.36 $\pm$ 2.46	59.25 $\pm$ 1.31	69.52 $\pm$ 1.73	59.40 $\pm$ 3.47	66.81 $\pm$ 0.47	59.80 $\pm$ 1.13
<b>PUFAs</b>							
C <sub>18:2n6t</sub>	2.30 $\pm$ 0.41	2.93 $\pm$ 2.62	5.20 $\pm$ 1.96	5.20 $\pm$ 0.61	4.40 $\pm$ 0.17	5.81 $\pm$ 1.32	5.29 $\pm$ 0.15
C <sub>18:2n6c</sub>	11.52 $\pm$ 1.18	8.56 $\pm$ 1.00	7.96 $\pm$ 1.01	8.31 $\pm$ 0.26	7.73 $\pm$ 0.61	9.45 $\pm$ 2.22	7.41 $\pm$ 0.66
Total	13.82 $\pm$ 0.78	11.49 $\pm$ 1.62	11.44 $\pm$ 0.06	13.51 $\pm$ 0.63	12.13 $\pm$ 0.77	15.26 $\pm$ 0.89	12.70 $\pm$ 0.81
UFA/SFA ratio	2.01	2.12	2.42	4.89	2.51	4.59	2.64

<sup>a</sup>ND, not detected.

response to N deprivation. For all three strains, the total lipid content under the N deprivation condition was higher than that under the N-replete condition, and this was consistent with the changes of neutral lipids. It was apparent that all three strains had higher lipid contents at the lowest nitrogen concentrations of 0.37 mmol/liter and 0.74 mmol/liter, which indicated that the N deprivation had a positive effect on the lipid accumulation. Although the *Chlorella*-Arc strain had low lipid productivity, it had the highest lipid content among the three strains under N deprivation and N-replete culture conditions, as the lipid contents reached 43% and 54% under the nitrogen concentration of 0.74 mmol/liter and 0.37 mmol/liter, respectively. The *Chlorella*-Des and *Chlorella*-Nat strains had highest lipid contents 40% and 41%, respectively. The reason for the high lipid content in *Chlorella*-Arc may be that such a strain has to achieve high utilization of nitrogen and develop special lipid accumulation processes to produce lipid for storing energy for long-term survival in extremely cold regions (37).

It was clear that among the three strains, *Chlorella*-Arc from the polar region could be a promising candidate for biodiesel feedstock, since it could accumulate higher lipid content under the N deprivation condition, and in particular the highest neutral lipid contents, which are the main materials for biofuels. Although *Chlorella*-Arc had a low growth rate, it could be used in transgenic microalgae for biofuel production based on the genetic strain engineering strategies that offer a means to accelerate the commercialization of algal biofuels by improving the rate and total accumulation of microalga lipids (38, 39). The reason for the difference of growth rates and lipid contents of the three strains may be related to their adaptation to local environments in the long-term evolution on the genetic level. The different environmental factors could lead to their changes in physiology, which are controlled by gene regulation (40). At present, more studies are focusing on the molecular regulation of nitrogen starvation-induced lipid accumulation (11, 13, 24). For example, Goncalves and colleagues in 2016 (24) provided evidence to support the hypothesis that transcription factor ROC40 has a role in N-induced lipid accumulation, and they uncovered multiple previously unknown proteins modulated by short-term N in green algae.

**Effect of nitrogen deprivation on fatty acid composition.** Fatty acid composition, which affects the quality of biodiesel production, is strongly strain specific (41). In our research, the fatty acid compositions of the three *Chlorella* strains from arctic glacier,

**TABLE 2** Fatty acid compositions of *Chlorella*-Arc cultured under different sodium nitrate concentrations

Fatty acid	% of total fatty acids (mean $\pm$ SD) when cultured in NaNO <sub>3</sub> concn of:						
	0.37 mmol·L <sup>-1</sup>	0.74 mmol·L <sup>-1</sup>	1.47 mmol·L <sup>-1</sup>	2.94 mmol·L <sup>-1</sup>	5.88 mmol·L <sup>-1</sup>	8.82 mmol·L <sup>-1</sup>	11.76 mmol·L <sup>-1</sup>
<b>SFAs</b>							
C <sub>16:0</sub>	8.12 $\pm$ 0.69	3.36 $\pm$ 1.11	5.39 $\pm$ 1.42	8.86 $\pm$ 0.49	3.95 $\pm$ 0.88	25.95 $\pm$ 1.57	24.41 $\pm$ 2.68
C <sub>20:0</sub>	12.51 $\pm$ 1.02	17.81 $\pm$ 0.58	11.53 $\pm$ 0.66	14.54 $\pm$ 0.53	16.39 $\pm$ 0.56	2.21 $\pm$ 1.30	12.08 $\pm$ 1.27
Total	20.63 $\pm$ 2.32	21.17 $\pm$ 0.46	16.92 $\pm$ 0.28	23.40 $\pm$ 0.46	20.34 $\pm$ 0.56	28.16 $\pm$ 6.85	36.49 $\pm$ 5.69
<b>MUFAs</b>							
C <sub>16:1</sub>	12.51 $\pm$ 0.34	6.64 $\pm$ 0.46	5.27 $\pm$ 0.03	3.3 $\pm$ 0.29	5.98 $\pm$ 0.67	7.04 $\pm$ 0.79	9.81 $\pm$ 3.27
C <sub>18:1n9t</sub>	6.23 $\pm$ 0.28	22.47 $\pm$ 0.23	16.35 $\pm$ 1.35	15.11 $\pm$ 1.23	15.62 $\pm$ 0.86	3.50 $\pm$ 4.5	0.46 $\pm$ 0.23
C <sub>18:1n9c</sub>	20.46 $\pm$ 0.08	18.34 $\pm$ 0.65	20.63 $\pm$ 2.56	23.89 $\pm$ 3.45	18.47 $\pm$ 1.32	11.57 $\pm$ 6.75	13.67 $\pm$ 4.20
C <sub>20:1</sub>	16.33 $\pm$ 0.45	15.93 $\pm$ 0.22	16.43 $\pm$ 0.73	21.41 $\pm$ 0.70	12.91 $\pm$ 2.13	39.92 $\pm$ 0.06	33.05 $\pm$ 3.83
Total	55.53 $\pm$ 0.60	63.38 $\pm$ 3.46	58.68 $\pm$ 0.70	53.71 $\pm$ 4.56	52.89 $\pm$ 0.45	62.03 $\pm$ 11.25	56.99 $\pm$ 12.41
<b>PUFAs</b>							
C <sub>18:2n6t</sub>	17.13 $\pm$ 0.8	8.65 $\pm$ 0.15	7.75 $\pm$ 0.56	2.67 $\pm$ 0.056	7.52 $\pm$ 0.05	9.82 $\pm$ 8.45	7.52 $\pm$ 5.79
C <sub>18:2n6c</sub>	2.98 $\pm$ 0.23	3.25 $\pm$ 2.25	12.77 $\pm$ 1.38	15.67 $\pm$ 1.34	13.53 $\pm$ 0.57	2.21 $\pm$ 3.13	11.07 $\pm$ 9.33
C <sub>18:3n3</sub>	3.73 $\pm$ 0.19	3.55 $\pm$ 0.41	3.88 $\pm$ 0.10	4.55 $\pm$ 0.55	5.72 $\pm$ 0.61	nd <sup>a</sup>	nd
Total	23.84 $\pm$ 0.22	15.45 $\pm$ 0.87	24.40 $\pm$ 3.56	22.89 $\pm$ 2.45	26.77 $\pm$ 4.23	12.02 $\pm$ 3.20	18.60 $\pm$ 8.27
UFA/SFA ratio	3.84	5.47	3.10	3.37	2.74	2.63	2.07

<sup>a</sup>ND, not detected.

desert, and temperate native lake regions in response to N deprivation were well compared. Different patterns of fatty acid composition were found among the three strains from the different environmental regions (Tables 1, 2, and 3). For each strain, the fatty acid composition was also different under N deprivation and N-replete conditions. The fatty acid compositions of the three *Chlorella* strains were mainly composed of C<sub>16</sub> and C<sub>18</sub> fatty acids; this was similar to findings reported elsewhere (35, 42, 43). C<sub>18:3</sub> (linolenic acid) contents in *Chlorella*-Arc and *Chlorella*-Nat strains were lower than 6% under all treatments (Table 2 and Table 3). Thus, the fatty acid compositions of all three *Chlorella* strains under N deprivation met the requirements of European Standard EN14214 for biodiesel production (44).

**TABLE 3** Fatty acid compositions of *Chlorella*-Nat cultured under different sodium nitrate concentrations

Fatty acid	% of total fatty acids (mean $\pm$ SD) when cultured in NaNO <sub>3</sub> concn of:						
	0.37 mmol/liter	0.74 mmol/liter	1.47 mmol/liter	2.94 mmol/liter	5.88 mmol/liter	8.82 mmol/liter	11.76 mmol/liter
<b>SFAs</b>							
C <sub>16:0</sub>	20.65 $\pm$ 1.46	28.81 $\pm$ 0.51	20.74 $\pm$ 2.52	31.4 $\pm$ 0.43	29.34 $\pm$ 0.77	16.08 $\pm$ 0.07	17.91 $\pm$ 2.14
C <sub>18:0</sub>	4.51 $\pm$ 0.55	2.09 $\pm$ 0.75	3.92 $\pm$ 0.79	0.63 $\pm$ 0.05	1.91 $\pm$ 0.21	10.18 $\pm$ 0.67	14.27 $\pm$ 3.21
Total	25.16 $\pm$ 0.38	30.90 $\pm$ 0.42	24.66 $\pm$ 1.67	32.03 $\pm$ 0.09	31.25 $\pm$ 0.18	26.26 $\pm$ 0.75	25.06 $\pm$ 6.81
<b>MUFAs</b>							
C <sub>16:1</sub>	1.29 $\pm$ 0.16	0.85 $\pm$ 0.07	1.27 $\pm$ 0.07	2.50 $\pm$ 0.32	0.59 $\pm$ 0.15	6.28 $\pm$ 0.53	6.66 $\pm$ 0.87
C <sub>18:1n9t</sub>	10.08 $\pm$ 0.90	26.70 $\pm$ 1.05	30.31 $\pm$ 0.59	18.37 $\pm$ 1.06	22.72 $\pm$ 1.28	0.45 $\pm$ 0.07	0.67 $\pm$ 0.42
C <sub>18:1n9c</sub>	12.26 $\pm$ 0.52	2.66 $\pm$ 1.14	5.48 $\pm$ 1.08	1.07 $\pm$ 0.19	1.16 $\pm$ 0.10	19.15 $\pm$ 0.38	17.31 $\pm$ 1.56
C <sub>20:1</sub>	3.99 $\pm$ 0.90	1.63 $\pm$ 0.25	5.52 $\pm$ 1.48	1.76 $\pm$ 0.48	2.09 $\pm$ 0.90	28.14 $\pm$ 0.75	19.64 $\pm$ 0.74
C <sub>20:1n9</sub>	18.19 $\pm$ 1.01	12.12 $\pm$ 1.55	9.64 $\pm$ 1.03	17.52 $\pm$ 0.57	19.26 $\pm$ 1.01	3.80 $\pm$ 0.54	3.91 $\pm$ 0.63
Total	45.84 $\pm$ 1.05	43.96 $\pm$ 1.36	52.22 $\pm$ 0.70	41.22 $\pm$ 3.28	45.82 $\pm$ 1.56	57.82 $\pm$ 0.07	55.29 $\pm$ 7.56
<b>PUFAs</b>							
C <sub>18:2n6t</sub>	5.95 $\pm$ 1.69	1.81 $\pm$ 0.88	3.73 $\pm$ 1.08	0.77 $\pm$ 0.19	2.70 $\pm$ 0.96	15.92 $\pm$ 0.73	ND
C <sub>18:2n6c</sub>	18.61 $\pm$ 1.38	21.78 $\pm$ 1.88	15.93 $\pm$ 1.60	25.38 $\pm$ 1.08	18.93 $\pm$ 1.66	ND <sup>a</sup>	ND
C <sub>18:3n3</sub>	4.24 $\pm$ 0.99	1.55 $\pm$ 0.39	3.46 $\pm$ 0.38	0.60 $\pm$ 0.09	1.31 $\pm$ 0.07	ND	ND
Total	28.81 $\pm$ 0.65	25.58 $\pm$ 1.45	23.12 $\pm$ 1.45	26.75 $\pm$ 2.35	22.93 $\pm$ 0.78	15.92 $\pm$ 0.74	19.65 $\pm$ 0.74
UFA/SFA ratio	2.97	2.25	3.06	2.12	2.2	2.81	2.99

<sup>a</sup>ND, not detected.

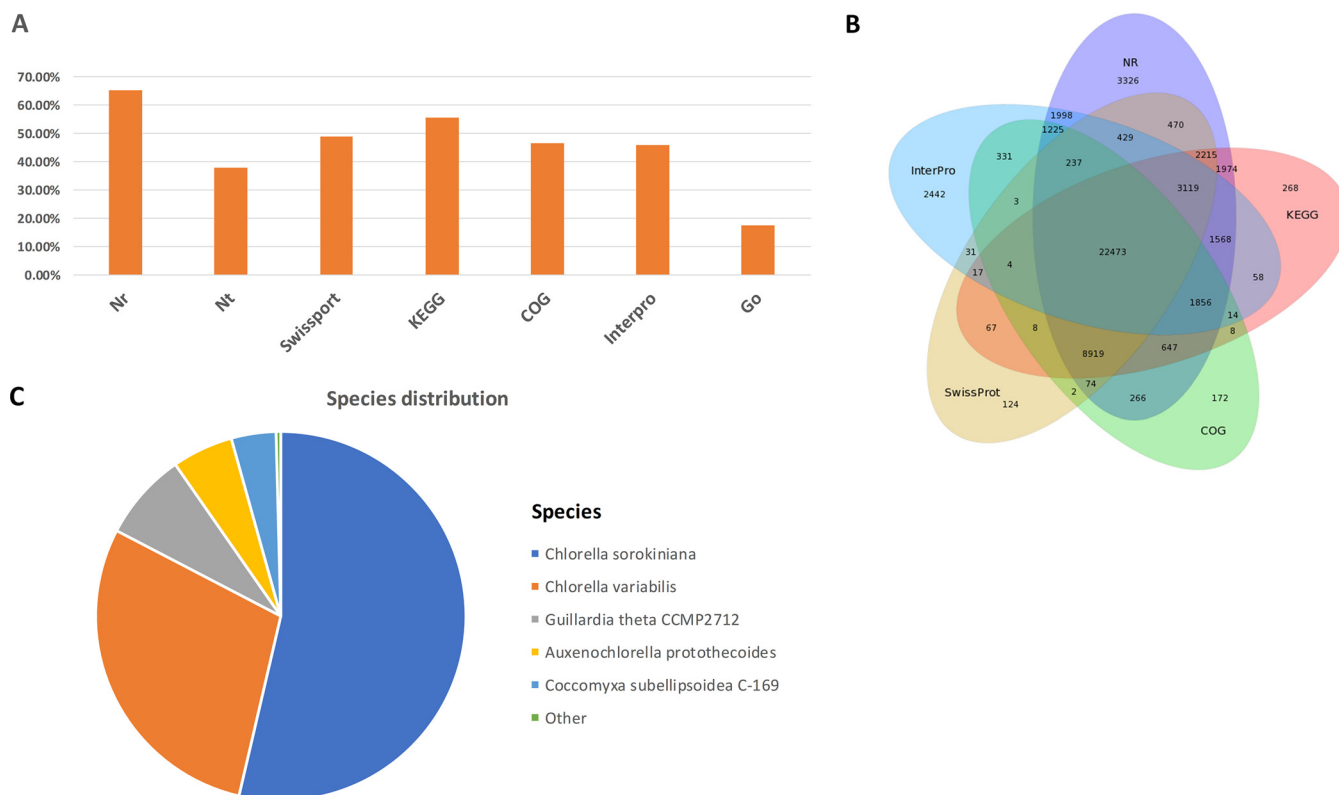
**TABLE 4** Quality statistics of transcripts and unigenes for each strain from three environmental regions cultured under N deprivation (1/8ND) compared to the control group (ND)

Sample	Total no.	Total length	Mean length	N50	N70	N90	% GC
Transcripts							
Nat-1/8ND	21,333	24,730,273	1,159	1,721	1,157	529	66.99
Nat-ND	25,771	28,764,323	1,116	1,721	1,131	480	66.96
Arc-1/8ND	44,358	40,980,088	923	1,433	873	378	61
Arc-ND	51,269	46,925,918	915	1,379	858	381	59.56
Des-1/8ND	26,359	28,928,276	1,097	1,730	1,140	446	66.29
Des-ND	26,770	29,424,456	1,099	1,733	1,133	447	66.15
All transcripts	195,860						
Unigenes							
Nat-1/8ND	18,441	23,193,984	1,257	1,857	1,269	605	67.04
Nat-ND	22,145	26,846,722	1,212	1,872	1,251	546	67.01
Arc-1/8ND	39,597	38,202,965	964	1,524	935	389	61.07
Arc-ND	45,602	43,438,848	952	1,457	913	392	59.65
Des-1/8ND	22,354	26,864,972	1,201	1,912	1,288	503	66.26
Des-ND	22,789	27,187,125	1,192	1,879	1,264	505	66.1
All unigenes	77,917	87,573,379	1,123	1,841	1,170	446	62.47

The lipids containing high amounts of saturated fatty acids (SFAs) and monounsaturated fatty acids (MUFAs) with low polyunsaturated fatty acid (PUFA) content, especially  $C_{18:3}$ , are more suitable for biodiesel production, since PUFA can lead to low oxidative stability during storage (45). In our study, the sum of SFA and MUFA content was generally high (67.08 to 83.08%) in strains from all three environmental regions, and the sum of PUFA content in *Chlorella*-Arc and *Chlorella*-Des strain was generally lower than that in *Chlorella*-Nat under different sodium nitrate concentrations (Tables 1 to 3). The proportion of  $C_{18:3}$  was low (<6%) in the *Chlorella* strain from the arctic and native lake regions and was absent in the strain of the desert region (Tables 1 to 3). In addition, high UFA and SFA are important for stress resistance, since unsaturated fatty acids are more easily oxidized by oxygen radicals than are saturated fatty acids (19). *Chlorella*-Arc possessed the highest UFA/SFA ratios under the treatment of N deprivation, which indicated that this *Chlorella* strain might be the most suitable strain that not only could accumulate high lipids but also could have suitable fatty acid composition for biodiesel production under the N deprivation condition.

**Quality of RNA sequencing data and gene functional annotation.** A total of 40.37 Gb bases were generated for all strains. The clean reads quality metrics after filtering sequences that contained low-quality, adaptor-polluted, and or a high content of unknown base (N) reads, are shown in Table S1 in the supplemental material. The abundance of raw reads of the 6 strains were from 50.63 to 52.26 Mb. The clean reads ratios of all strains ranged from 84.38% to 89.74%. The number of transcripts for the six samples were 21,333, 25,771, 44,358, 51,269, 26,359, and 26,770, with a mean length ranging from 915 bp to 1,159 bp and N50 length from 1,379 bp to 1,733 bp (Table 4; see also Fig. S1 to S6). The numbers of unigenes for the six samples were 18,441, 22,145, 39,597, 45,602, 22,354, and 22,789, respectively, with a mean length of 952 bp to 1,257 bp and an N50 length of 1,457 to 1,912 (Table 4; see also Fig. S1 to S6). These reads and assembly qualities indicated that the transcriptome sequencing performed well for functional analysis.

All the unigenes assembled were annotated to the NR, NT, GO, COG, KEGG, Swissprot, and Interpro databases, respectively. For all annotated unigenes, the NR, Swissprot, COG, and KEGG databases showed higher annotation proportions (Fig. 6A). Among the 7 databases, the NR database had the most unique annotation (Fig. 6B). There were 22,473 shared annotated genes among the 7 databases. The top annotated species included *C. sorokiniana*, *Chlorella variabilis*, *Guillardia theta*, *Auxenochlorella protothecoides*, and *Coccomyxa subellipsoidea*, whereas *Chlorella variabilis* and *Chlorella*



**FIG 6** (A) Annotation showing proportions of unigenes in seven databases. (B) Annotation showing differences in seven databases for *C. sorokiniana*. (C) Species distribution for annotation of *C. sorokiniana*.

*sorokiniana* showed higher annotation proportions (Fig. 6C). This was not strange, since the genomes of *C. sorokiniana* and *C. variabilis* strains are available in public databases (46, 47). The COG functional enrichment was performed for all unigenes where the lipid and amino acid transport and metabolism were included (Fig. 7). The translation, ribosomal structure and biogenesis, cell cycle control, cell division, chromosome partitioning, and general function predictions accounted for the top classification categories (Fig. 7). The KEGG database annotated the unigenes into five categories, including cellular processes, environmental information processing, genetic information processing, metabolism, and organism systems, whereas the translation and global and the overview maps accounted for the top annotation.

#### Genetic expression patterns and pathways of differentially expressed genes.

First, the gene expression distributions of the six samples from the three strains were compared, including their shared and unique unigenes. From the principal-component analysis (PCA) plot of gene expression (Fig. 8), we noted that the strains DesND and Des1/8ND from desert clustered together closely and the NatND and Nat1/8ND strains from the native lake clustered together loosely. However, the ArcND and Arc1/8ND strains were separated. The comparison of the shared and unique unigenes among the strains are shown in Fig. 9, where the distributions of the unigenes were generally consistent with the PCA analysis. These indicated that the strains under different nitrogen stress in the desert region have similar gene expression patterns. The strains under different nitrogen stress in temperate regions also had similar gene expression patterns. However, the arctic strains under different nitrogen concentrations showed significant differentially expressed genes (DEGs).

The DEGs among the six samples of the three strains revealed upregulated and downregulated genes among them (Fig. 10). It was apparent that the Nat1/8ND versus Des1/8ND, NatND versus DesND, and NatND versus ArcND comparisons showed more DEGs, but the DesND versus Des1/8ND comparison showed fewer DEGs. The ArcND

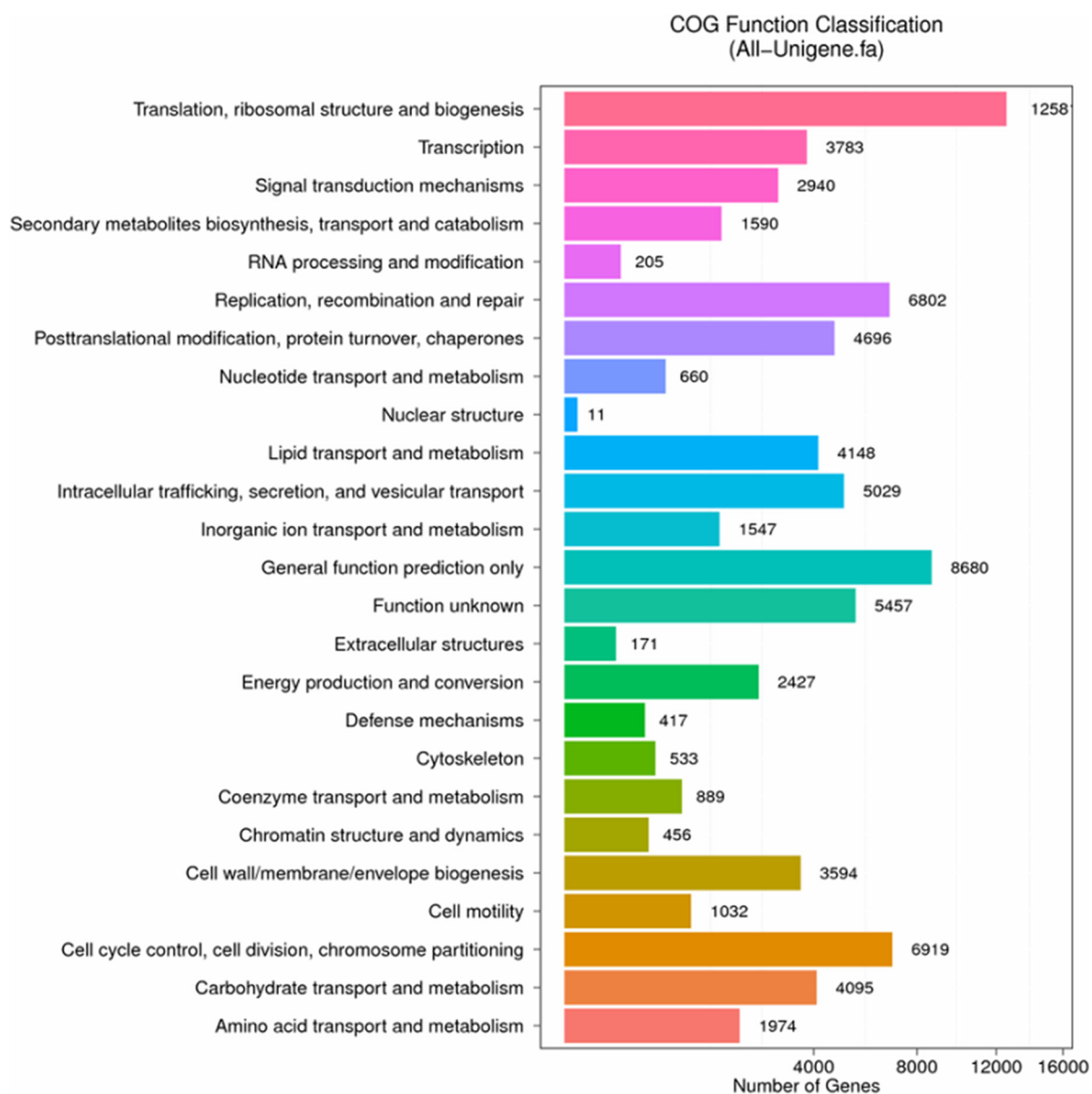
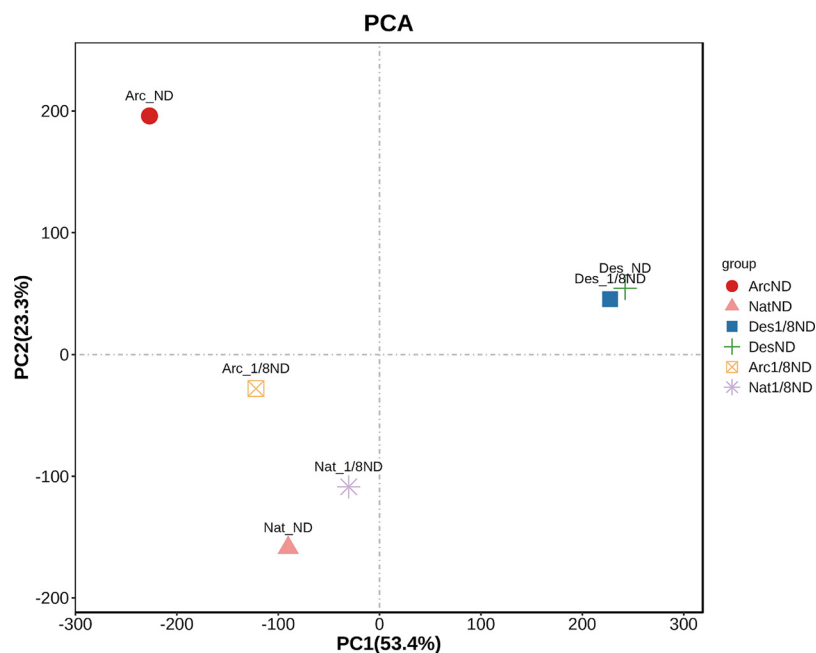


FIG 7 Distribution of COG function classifications for *C. sorokiniana*.

versus Arc1/8ND comparison also showed more DEGs than the DesND versus Des1/8ND comparison. This was consistent with the result from the PCA plot that showed strains from the same environmental regions had different genetic expression patterns, except for strains from the arctic region, which showed significantly different genes. These results suggest that the genetic regulated mechanisms of *C. sorokiniana* from different environmental regions differ from each other. In particular, the strains from arctic regions under the 1/8 N deprivation condition showed significantly different expression genes from the control group, which further suggested that the *C. sorokiniana* strain from the arctic region possessed a special mechanism for high lipid production (discussed below). Extreme environments like the polar region possibly make the microalgae possess special regulation mechanisms for nutrient production due to adaptive evolution (28). Multiple DEGs among the six samples were selected randomly for quantitative PCR (qPCR). The 18S gene was selected as the reference gene. In general, the statistical results indicated that the relative expression levels of the unigenes



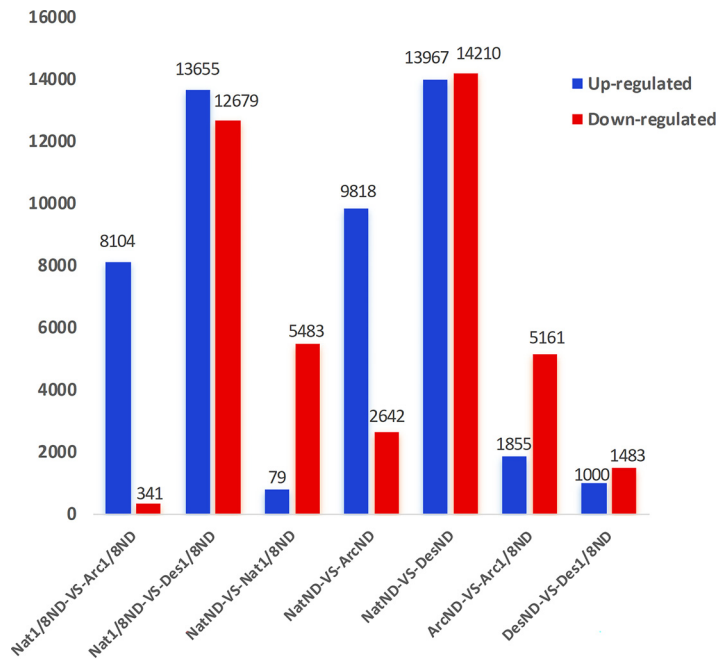
**FIG 8** PCA plot for the differences in expression levels among the strains of the three environmental regions.

by qPCR were generally consistent with the actual expression levels (fragments per kilobase of transcript per million mapped reads [FPKM]) (Fig. 11).

Among their DEGs, the most common were those obtained among the six *C. sorokiniana* samples (Fig. 12). The genes that were related to carbon and lipid metabolism were included in the DEGs, including acetyl coenzyme A (acetyl-CoA) carboxylase (ACCase), fatty acid synthase complex (FAS), diacylglycerol acyltransferase (DGAT), pyruvate dehydrogenase component (PDCom), phosphoenolpyruvate carboxylase (PEPC), long-chain acyl-CoA synthetase (LC-FACS), fatty acid desaturase (35, 40, 48, 49). These DEGs were highly expressed in the *C. sorokiniana* strain from the arctic regions (especially under 1/8 N deprivation), e.g., ACCase, LC-FACS, DGAT, PEPC, PDCom, and FAS. These findings suggested that the *C. sorokiniana* strain from the arctic region expressed more genes related to carbon and lipid metabolism to produce a high amount of lipids, consistent with our physiological results. First of all, the accumulation of large quantities of lipids in microalgae requires a continuous supply of acetyl-CoA for fatty acid biosynthesis ([https://www.kegg.jp/kegg-bin/show\\_pathway?cvr00061](https://www.kegg.jp/kegg-bin/show_pathway?cvr00061)) (40) (Fig. 13). In the glycolysis gluconeogenesis process of *Chlorella variabilis*, pyruvate, which is a key precursor in central carbon metabolism and lipid synthesis, can be transferred to acetyl-CoA under the catalytic action of PDCom (Fig. 13) ([https://www.genome.jp/kegg-bin/show\\_pathway?cvr00010](https://www.genome.jp/kegg-bin/show_pathway?cvr00010)) (49). Meanwhile, PEPC is important for the production of pyruvate. Here, we showed that the arctic strain expressed more PDCom than the desert or temperate strains, and Arc1\_8ND produced more PEPC than all other strains of the three regions (Fig. 12). The physiological results also indicated that the arctic strains produced more sugar, which can then produce more acetyl-CoA through the function of PDCom and PEPC in glycolysis gluconeogenesis. DGAT plays an important role in acetyl-CoA and triacylglycerol (TAG) biosynthesis, also referred to as catabolism of fatty acids (35, 48). In glycerolipid metabolism, TAG and acetyl-CoA are finally produced with the catalysis of DGAT (Fig. 13). In this study, DGAT was highly expressed in Arc1\_8ND (Fig. 12). ACCase initiates fatty acid biosynthesis in microalgae and plays an important role for catalyzing the reaction converting acetyl-CoA to malonyl-CoA in the initial step (40). It was indicated that ACCase was highly expressed in Arc1\_8ND (Fig. 12). The fatty acyl-ACP thioesterase transfers long-chain acyl-ACP to



**FIG 9** The shared and unique unigene distributions of strains among the three environmental regions.



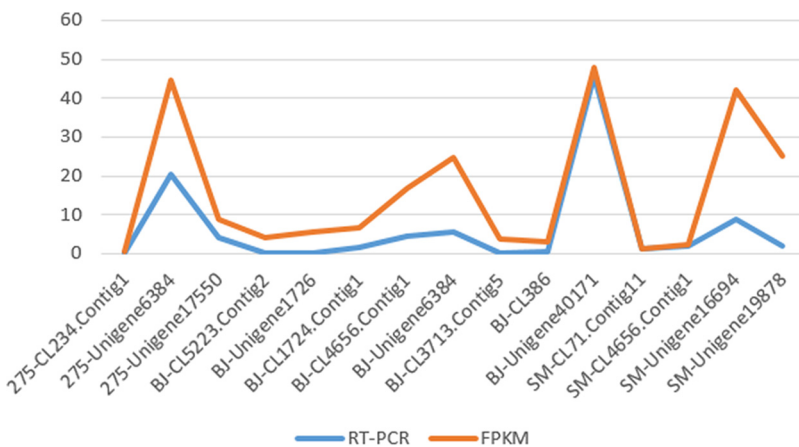
**FIG 10** Statistics of upregulated and downregulated genes among six samples of *Chlorella-Arc*, *Chlorella-Des*, and *Chlorella-Nat*, where ND and 1/8ND indicate the concentration of NaNO<sub>3</sub> (with 1 set for the control, 2.94 mmol/liter) and 1/8 times lower than the control, respectively. For example, Nat1/8ND versus Arc1/8ND comparison shows that Arc1/8ND generated more upregulated genes (blue) than downregulated genes (red) in comparison with Nat1/8ND.

long-chain fatty acids in the last step of fatty acid biosynthesis (Fig. 13). We found that Arc1\_8ND expressed the highest fatty acyl-ACP thioesterase compared to other strains.

In conclusion, the high lipid production of arctic strains is related to pathways of glycolysis gluconeogenesis, glycerolipid metabolism, and fatty acid biosynthesis. PDCOm and DGAT were expressed highly in arctic strains, and thus more acetyl-CoA was produced in the glycolysis gluconeogenesis and glycerolipid metabolism. Finally, more acetyl-CoA initiated fatty acid biosynthesis through the catalytic function of ACCase and acyl-ACP thioesterase, which were highly expressed in the arctic strain.

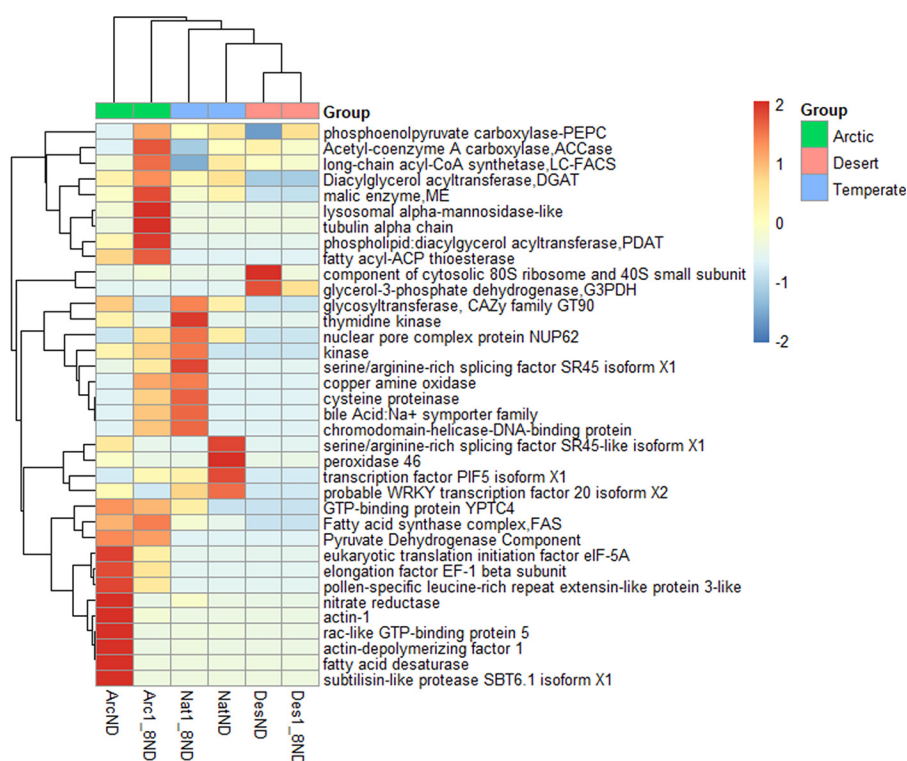
**MATERIALS AND METHODS**

**Strain isolation and identification.** The three *Chlorella* strains used in this research were isolated from glacial melted water around the Spitsbergen Bergen Islands arctic region (78°3'N, 16°6'E; *Chlorella-Arc*), desert soil from the Gurbantünggüt Desert region (45°2'N, 87°6'E; *Chlorella-Des*), and temperate



**FIG 11** Consistency verification of relative expression levels from RT-PCR and actual expression levels from RNA-Seq.





**FIG 12** Expression comparisons of the highest DEGs among the strains of the three environmental regions.

native Lake Xuanwuhu, Nanjing, China (32°1'N, 118°8'E; *Chlorella*-Nat). They were first observed based on morphological characteristics by microscope after isolation. Then, molecular tools were used to identify them to the species level. Both *rbcl* and *tufA* genes were amplified and sequenced for each strain of the different regions. Phylogenetic trees were constructed by the maximum-likelihood method to identify the three strains.

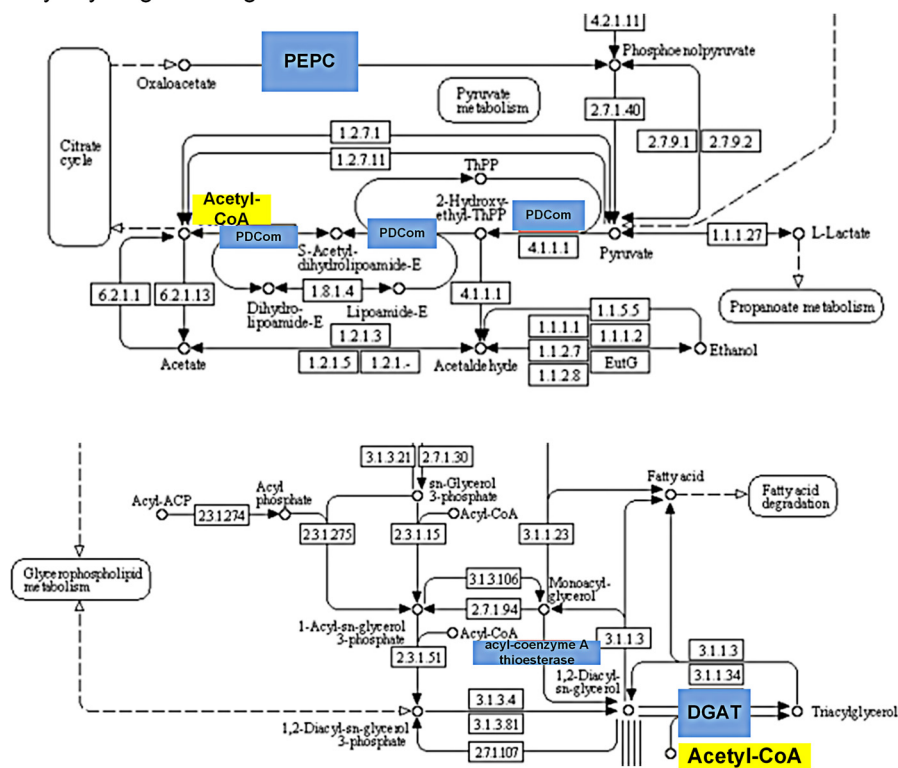
**Cultivation of microalgae and experimental design.** The optimal culture temperatures for the strains from arctic glacier, desert soil, and temperate native lake were 15°C, 25°C, and 28°C, respectively, based on our previous experiments (35, 50). All the strains were grown under illumination of ~120 photons  $\mu\text{mol}/\text{m}^2/\text{s}$  with a 14-h/12-h light/dark photoperiod. All *Chlorella* strains were cultivated in 250-mL flasks containing 200 mL BB medium. The initial cell density was  $\sim 5 \times 10^6$  cell/mL for all three *Chlorella* strains from arctic glacier, desert soil, and temperate native lake. For the control group, the BB medium consisted of 250 mg  $\text{NaNO}_3$ , 75 mg  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 25 mg  $\text{NaCl}$ , 75 mg  $\text{K}_2\text{HPO}_4$ , 175 mg  $\text{KH}_2\text{PO}_4$ , 25 mg  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 1 mL of trace elements solution which contained 11.42 mg  $\text{H}_3\text{BO}_3$ , 50 mg EDTA, 31 mg KOH, and 4.98 mg  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  in 1 liter of fresh water, in which the concentration of  $\text{NaNO}_3$  was 2.94 mmol/liter. For the culture condition of N deprivation, the concentration of  $\text{NaNO}_3$  was set 1/8 $\times$ , 1/4 $\times$ , and 1/2 $\times$  that for the control group (2.94 mmol/liter). For a better comparison, the concentration of  $\text{NaNO}_3$  was set 2 $\times$ , 4 $\times$ , and 8 $\times$  higher than the control condition (2.94 mmol/liter) for the N-replete condition. Each condition of every strain was evaluated in triplicate. Each flask was shaken more than five times every day.

We also took steps to avoid contamination of the algae strains. Firstly, we made sure of the identification accuracy by both morphological and molecular tools. Second, during culturing, the microalgae composition was observed from time to time. We also conducted Sanger sequencing once for all the strains during culturing, to make sure that the culturing was not contaminated by other microalgae. Finally, all culturing environments were clean and isolated. All the strains from different conditions were cultured separately in their own containers.

**Measurements of growth, biochemical composition, lipid productivity, and fatty acids.** The biomass was detected based on cell number, which was counted using a hemacytometer (improved double-Neubauer). Samples were harvested by centrifugation at  $8,000 \times g$  for 10 min, and pellets were dried by vacuum freeze drier. Chlorophyll was analyzed spectrophotometrically after extraction with methanol overnight at 4°C in the dark (51). The soluble sugar was measured by the phenol-sulfuric acid spectrophotometric method (52). The protein was measured using the bicinchoninic acid assay (53).

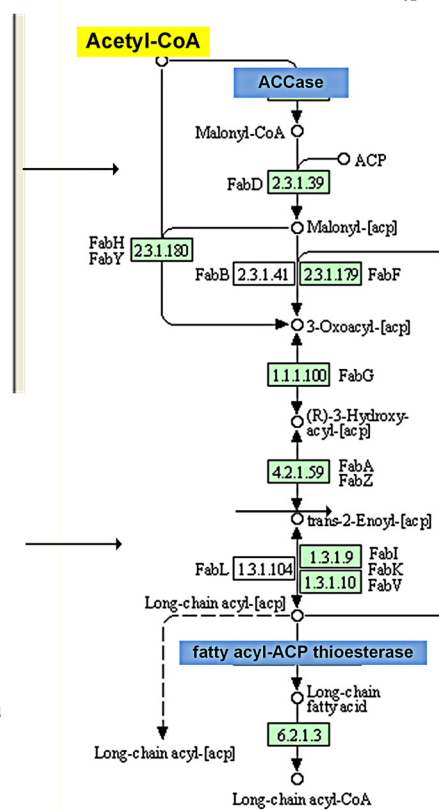
The total lipid content was measured after 120-mL algal cultures at late logarithmic growth phase were harvested using a modified version of the Bligh and Dyer method (54). A rapid neutral lipid determination was performed based on the Nile red method. First, the microalgae samples were pretreated with 20% (vol/vol) dimethyl sulfoxide for 20 min at 40°C. Then, the stock solution of Nile red was prepared in acetone (0.1 mg/mL), and 15  $\mu\text{L}$  Nile red dye was added into a 1-mL microalgae suspension, which was stained for 5 min. The fluorescence intensity was measured by a spectrophotometer

## Glycolysis gluconeogenesis



## Glycerolipid metabolism

## Fatty acid biosynthesis



**FIG 13** KEGG pathways of fatty acid biosynthesis, glycolysis gluconeogenesis, and glycerolipid metabolism. The genes highlighted in blue are functional catalyzing enzymes.

(Molecular Devices, USA) with the excitation and emission wavelengths of 480 nm and 575 nm, respectively. The gas chromatography with mass selective detector (GC-MS; Agilent 7890A), equipped with an HP-5msi column (30 m × 0.25 mm × 0.25 μm; Agilent Technologies), was used to analyze fatty acids. Two microliters of each sample was injected into the splitless injection mode. The injection temperature and MS detector temperature were set at 250°C and 280°C, respectively. The temperature program was as follows: maintaining at the initial temperature of 50°C for 1 min, followed by increasing at a rate of 40°C/min up to 170°C, maintained for 1 min, then raising to 210°C at 18°C/min, and maintained at 210°C for 28 min. Nitrogen gas was used as the carrier gas. Finally, fatty acids were discriminated by comparison of their retention times with those of standards (Sigma). Heptadecanoic acid (C<sub>17:0</sub>) was used as an internal standard.

**Statistical analysis.** All samples were performed in triplicate. The mean values with their standard deviations were also determined. The significant differences were assessed by one-way analysis of variance (ANOVA) with SPSS statistical software, followed by Duncan tests. Different letters in the figures indicate significant differences between treatments ( $P < 0.05$ ).

**RNA isolation, library construction, and RNA-Seq.** For each environmental region, arctic glacier, desert, and temperate native lake, the strains cultured in 1/8 N deprivation and the control group were selected for RNA sequencing (RNA-Seq). We did not conduct RNA-Seq of biological replicates separately for each strain, but three replicates of each strain were cultured under the same condition and were mixed first for RNA extraction and sequencing. Then, the 1/8 N deprivation condition and the control group from each environmental region were sequenced separately. Thus, a total of six samples of three strains from the three environmental regions were finally sequenced separately (see Table S1 in the supplemental material). An ethanol precipitation protocol and CTAB-PBIOZOL reagent were used for the extraction of total RNA (55). Subsequently, total RNA was qualified and quantified using a NanoDrop and Agilent 2100 bioanalyzer (Thermo Fisher Scientific, MA, USA).

Oligo(dT)-attached magnetic beads were used to purify mRNA. Purified mRNA was fragmented into small pieces with fragment buffer, which was then used as a template to generate the first-strand cDNA for a first-strand reaction system by PCR, followed by a second-strand cDNA synthesis. The reaction product was purified by magnetic beads. Next, A-Tailing mix and RNA index adapters were added, and mixtures were incubated to end-repair the cDNA fragments with adaptors amplified by PCR, and the products were purified by Ampure XP Beads. The product was validated on the Agilent Technologies 2100 bioanalyzer for quality control. The qualified library was amplified on cBot to generate the cluster on the flow cell. And the amplified flow cell was sequenced single-end on the HiSeq X-10 platform (BGI-Shenzhen, China).

**RNA-Seq data preprocessing.** For each of the six samples, their clean reads were obtained after filtering the raw reads via SOAPnuke (v1.4.0) by removing reads containing adaptors, poly(N), or low-quality reads (56). Then, the clean reads of each of the six samples were assembled separately (Table S1). Trinity (v2.0.6) (57) was used to assemble the clean reads, and Tgicl (v2.0.6) (58) was used to perform clustering and eliminate redundant data in the assembled transcripts to obtain unique genes. We compared the unigene distributions of seven sets: NatND versus ArcND, NatND versus DesND, Nat1/8ND versus NatND, Nat1/8ND versus Arc1/8ND, Nat1/8ND versus Des1/8ND, Arc1/8ND versus ArcND, and Des1/8ND versus DesND, including their shared and unique unigenes, where ND and 1/8ND indicate the concentration of  $\text{NaNO}_3$  in the control (2.94 mmol/liter) and 1/8 times lower than the control, respectively. Then, based on these sets, their assembled transcripts were processed for further annotation and expression analysis. Clean reads were mapped to the assembled unique genes by Bowtie2 (v2.2.5) (59), and the expression levels of genes were calculated by RSEM (v1.2.8) (60) and normalized to the FPKM. Functional annotation of genes was achieved by mapping genes to different databases (NT, NR, GO, COG, KEGG, Swiss-Prot, and InterPro) using the software BLAST (v2.2.23) (61). GO annotation was performed by Blast2GO (v 2.5.0) with NR annotations. PossionDis (62) was used to detect DEGs, and DEGs with a fold change of  $>2$  or  $<-2$  and a false-discovery rate (FDR) of  $<0.001$  were considered significantly differentially expressed genes. GO enrichment analysis and KEGG enrichment analysis were performed using Phyper, a function of R. The significant levels of terms and pathways were corrected by  $Q$  value with a rigorous threshold ( $Q < 0.05$ ).

**Validation of DEGs by qRT-PCR.** The confidence of the high-throughput transcriptome sequencing was further validated by quantitative real-time PCR (qRT-PCR). A total of 16 random differentially expressed genes were selected for qPCR using the SG Fastq PCR master mix (High Rox), and the primer Premier 5.0 was used for designing primers for these genes. PCRs were carried out in a total volume of 20  $\mu\text{L}$ , including 10  $\mu\text{L}$  of SybrGreenq PCR master mix. PCR conditions for all primer sets were 95°C for 3 min, 45 cycles of 95°C for 7 s, 57°C for 10 s, and 72°C for 15 s, and this was followed by disassociation curve analysis in ABI Steponeplus (Applied Biosystems, Foster City, CA, USA). The relative gene expression levels of the samples were calculated by the  $2^{-\Delta\Delta\text{CT}}$  threshold cycle method, and these levels were normalized using the 18S mRNA level. The SPSS was used to determine whether the expression data included any significant differences at the  $P < 0.05$  level. The fold changes of these genes were calculated via FPKM, and the gene's  $\log_2$  fold change values from qRT-PCR and RNA-Seq were used for graphical presentations.

**Data availability.** The next-generation sequencing (NGS) data of all samples obtained in this study have been deposited with NCBI. The SRA accession for the NGS sequences is [PRJNA797923](https://www.ncbi.nlm.nih.gov/sra/PRJNA797923). The submission information for each sample mainly includes the reads sequences in fastq form, the sampling location, and the sequence platform.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, XLSX file, 0.02 MB.

**SUPPLEMENTAL FILE 2**, PDF file, 0.9 MB.

## ACKNOWLEDGMENTS

Financial support from the China Postdoctoral Science Foundation (2014M561661, 2508BH001) and Fundamental Research Funds for the Central Universities of Ministry of Education of China (Y0201600141, KJQN201742) for this research is gratefully acknowledged. The bioinformatics analysis of the project was supported by the Bioinformatics Center of Nanjing Agricultural University. We thank the Jiangsu Provincial Key Laboratory of Marine Biology for supporting our obtaining the arctic strain.

## REFERENCES

1. Tredici MR, Biondi N, Ponis E, Rodolfi L. 2009. Advances in microalgal culture for aquaculture feed and other uses, p 610–676. In Burnell G, Allan G (ed), *New technologies in aquaculture: improving production efficiency, quality and environmental management*. CRC Press, Boca Raton, FL.
2. Williams PJB, Laurens LML. 2010. Microalgae as biodiesel & biomass feedstocks: review & analysis of the biochemistry, energetics & economics. *Energy Environ Sci* 3:554–590. <https://doi.org/10.1039/b924978h>.
3. Griffiths MJ, Harrison ST. 2009. Lipid productivity as a key characteristic for choosing algal species for biodiesel production. *J Appl Phycol* 21: 493–507. <https://doi.org/10.1007/s10811-008-9392-7>.
4. Wijffels RH, Barbosa MJ. 2010. An outlook on microalgal biofuels. *Science* 329:796–799. <https://doi.org/10.1126/science.1189003>.
5. Chisti Y. 2007. Biodiesel from microalgae. *Biotechnol Adv* 25:294–306. <https://doi.org/10.1016/j.biotechadv.2007.02.001>.
6. Li X, Xu H, Wu Q. 2007. Large-scale biodiesel production from microalga *Chlorella protothecoides* through heterotrophic cultivation in bioreactors. *Biotechnol Bioeng* 98:764–771. <https://doi.org/10.1002/bit.21489>.
7. Hu Q, Sommerfeld M, Jarvis E, Ghirardi M, Posewitz M, Seibert M, Darzins A. 2008. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *Plant J* 54:621–639. <https://doi.org/10.1111/j.1365-3113.2008.03492.x>.
8. Harwood JL, Guschina IA. 2009. The versatility of algae and their lipid metabolism. *Biochimie* 91:679–684. <https://doi.org/10.1016/j.biochi.2008.11.004>.
9. Wu YH, Yu Y, Hu HY. 2013. Potential biomass yield per phosphorus and lipid accumulation property of seven microalgal species. *Bioresour Technol* 130:599–602. <https://doi.org/10.1016/j.biortech.2012.12.116>.
10. Zhekisheva M, Boussiba S, Khozin GI, Zarka A, Cohen Z. 2002. Accumulation of oleic acid in *Haematococcus pluvialis* (Chlorophyceae) under nitrogen starvation or high light is correlated with that of astaxanthin esters. *J Appl Phycol* 38:325–331. <https://doi.org/10.1046/j.1529-8817.2002.01107.x>.

11. Wang ZT, Ullrich N, Joo S, Waffenschmidt S, Goodenough U. 2009. Algal lipid bodies: stress induction, purification and biochemical characterization in wild-type and starchless *Chlamydomonas reinhardtii*. *Eukaryot Cell* 8:1856–1868. <https://doi.org/10.1128/EC.00272-09>.
12. Moellering ER, Benning C. 2010. RNA interference silencing of a major lipid droplet protein affects lipid droplet size in *Chlamydomonas reinhardtii*. *Eukaryot Cell* 9:97–106. <https://doi.org/10.1128/EC.00203-09>.
13. Miller R, Wu G, Deshpande RR, Vieler A, Gärtner K, Li X, Moellering ER, Zäuner S, Cornish AJ, Liu B, Bullard B, Sears BB, Kuo MH, Hegg EL, Shachar-Hill Y, Shiu SH, Benning C. 2010. Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. *Plant Physiol* 154:1737–1752. <https://doi.org/10.1104/pp.110.165159>.
14. Guarnieri MT, Nag A, Smolinski SL, Darzins A, Seibert M, Pienkos PT. 2011. Examination of triacylglycerol biosynthetic pathways via de novo transcriptomic and proteomic analyses in an unsequenced microalga. *PLoS One* 6:e25851. <https://doi.org/10.1371/journal.pone.0025851>.
15. Stephenson AL, Dennis JS, Howe CJ, Scott SA, Smith AG. 2010. Influence of nitrogen-limitation regime on the production by *Chlorella vulgaris* of lipids for biodiesel feedstocks. *Biofuels* 1:47–58. <https://doi.org/10.4155/bfs.09.1>.
16. Cordero BF, Couso I, León R, Rodríguez H, Vargas MÁ. 2011. Enhancement of carotenoids biosynthesis in *Chlamydomonas reinhardtii* by nuclear transformation using a phytoene synthase gene isolated from *Chlorella zofingiensis*. *Appl Microbiol Biotechnol* 91:341–351. <https://doi.org/10.1007/s00253-011-3262-y>.
17. Steinbrenner J, Sandmann G. 2006. Transformation of the green alga *Haematococcus pluvialis* with a phytoene desaturase for accelerated astaxanthin biosynthesis. *Appl Environ Microbiol* 72:7477–7484. <https://doi.org/10.1128/AEM.01461-06>.
18. Xu J, Hu H. 2013. Screening high oleaginous *Chlorella* strains from different climate zones. *Bioresour Technol* 144:637–643. <https://doi.org/10.1016/j.biortech.2013.07.029>.
19. Teoh ML, Phang SM, Chu WL. 2013. Response of Antarctic, temperate, and tropical microalgae to temperature stress. *J Appl Phycol* 25:285–297. <https://doi.org/10.1007/s10811-012-9863-8>.
20. Yeh KL, Chang JS. 2011. Nitrogen starvation strategies and photobioreactor design for enhancing lipid content and lipid production of a newly isolated microalga *Chlorella vulgaris* ESP-31: implications for biofuels. *Biotechnol J* 6:1358–1366. <https://doi.org/10.1002/biot.201000433>.
21. Wan M-X, Wang R-M, Xia J-L, Rosenberg JN, Nie Z-Y, Kobayashi N, Oyler GA, Betenbaugh MJ. 2012. Physiological evaluation of a new *Chlorella sorokiniana* isolate for its biomass production and lipid accumulation in photoautotrophic and heterotrophic cultures. *Biotechnol Bioeng* 109:1958–1964. <https://doi.org/10.1002/bit.24477>.
22. Guarnieri MT, Nag A, Yang S, Pienkos PT. 2013. Proteomic analysis of *Chlorella vulgaris*: potential targets for enhanced lipid accumulation. *J Proteomics* 93:245–253. <https://doi.org/10.1016/j.jprot.2013.05.025>.
23. Rosenberg JN, Kobayashi N, Barnes A, Noel EA, Betenbaugh MJ, Oyler GA. 2014. Comparative analyses of three *Chlorella* species in response to light and sugar reveal distinctive lipid accumulation patterns in the microalga *C. sorokiniana*. *PLoS One* 9:e92460. <https://doi.org/10.1371/journal.pone.0092460>.
24. Goncalves EC, Koh J, Zhu N, Yoo M-J, Chen S, Matsuo T, Johnson JV, Rathinasabapathi B. 2016. Nitrogen starvation-induced triacylglycerol accumulation in the green algae: evidence for a role for ROC40, a transcription factor involved in circadian rhythm. *Plant J* 85:743–757. <https://doi.org/10.1111/tpj.13144>.
25. Gong Y, Hu H, Gao Y, Xu X, Gao H. 2011. Microalgae as platforms for production of recombinant proteins and valuable compounds: progress and prospects. *J Ind Microbiol Biotechnol* 38:1879–1890. <https://doi.org/10.1007/s10295-011-1032-6>.
26. Mayfield SP, Manuell AL, Chen S, Wu J, Tran M, Siefker D, Muto M, Marin-Navarro J. 2007. *Chlamydomonas reinhardtii* chloroplasts as protein factories. *Curr Opin Biotechnol* 18:126–133. <https://doi.org/10.1016/j.copbio.2007.02.001>.
27. McMinn A, Martin A. 2013. Dark survival in a warming world. *Proc Biol Sci* 280:20122909. <https://doi.org/10.1098/rspb.2012.2909>.
28. Morales-Sánchez D, Schulze PSC, Kiron V, Wijffels RH. 2020. Temperature-dependent lipid accumulation in the polar marine microalga *Chlamydomonas malina* RCC2488. *Front Plant Sci* 11:619064. <https://doi.org/10.3389/fpls.2020.619064>.
29. Chen M, Tang H, Ma H, Holland TC, Ng KY, Salley SO. 2011. Effect of nutrients on growth and lipid accumulation in the green algae *Dunaliella tertiolecta*. *Bioresour Technol* 102:1649–1655. <https://doi.org/10.1016/j.biortech.2010.09.062>.
30. Men X, Xiaolin X, Chun L. 2012. Effects of nitrogen and phosphorus on the mixotrophic growth and total lipid of high-lipid desert microalgae. *J Shihezi Univ Nat Sci* 30:401–406.
31. Michel C, Legendre L, Theriault JC, Demers S. 1989. Photosynthetic responses of Arctic sea-ice microalgae to short-term temperature acclimation. *Polar Biol* 9:437–442. <https://doi.org/10.1007/BF00443230>.
32. Hernández-Torres A, Zapata-Morales AL, Alfaro AEO, Soria-Guerra RE. 2016. Identification of gene transcripts involved in lipid biosynthesis in *Chlamydomonas reinhardtii* under nitrogen, iron and sulfur deprivation. *World J Microb Biotechnol* 32:1–10.
33. Hörtensteiner S, Chinner J, Matile P, Thomas H, Donnison IS. 2000. Chlorophyll breakdown in *Chlorella protothecoides*: characterization of degreening and cloning of degreening-related genes. *Plant Mol Biol* 42:439–450. <https://doi.org/10.1023/A:1006380125438>.
34. Brown SB, Houghton JD, Hendry GAF. 1991. Chlorophyll breakdown, p 465–489. In Scheer H (ed), *Chlorophylls*. CRC Press, Boca Raton, FL.
35. Yang W, Zou S, He M, Fei C, Luo W, Zheng S, Chen B, Wang C. 2016. Growth and lipid accumulation in three *Chlorella* strains from different regions in response to diurnal temperature fluctuations. *Bioresour Technol* 202:15–24. <https://doi.org/10.1016/j.biortech.2015.11.070>.
36. Duval B, Shetty K, Thomas WH. 1999. Phenolic compounds and antioxidant properties in the snow alga *Chlamydomonas nivalis* after exposure to UV light. *J Appl Phycol* 11:559–566. <https://doi.org/10.1023/A:1008178208949>.
37. Praveenkumar R, Shameera K, Mahalakshmi G, Akbarsha MA, Thajuddin N. 2012. Influence of nutrient deprivations on lipid accumulation in a dominant indigenous microalga *Chlorella* sp., BUM11008: evaluation for biodiesel production. *Biomass Bioenergy* 37:60–66. <https://doi.org/10.1016/j.biombioe.2011.12.035>.
38. Rosenberg JN, Oyler GA, Wilkinson L, Betenbaugh MJ. 2008. A green light for engineered algae: redirecting metabolism to fuel a biotechnology revolution. *Curr Opin Biotechnol* 19:430–436. <https://doi.org/10.1016/j.copbio.2008.07.008>.
39. Niederholtmeyer H, Wolfstatter BT, Savage DF, Silver PA, Way JC. 2010. Engineering cyanobacteria to synthesize and export hydrophilic products. *Appl Environ Microbiol* 76:3462–3466. <https://doi.org/10.1128/AEM.00202-10>.
40. Rismani-Yazdi H, Haznedaroglu BZ, Bibby K, Peccia J. 2011. Transcriptome sequencing and annotation of the microalga *Dunaliella tertiolecta*: pathway description and gene discovery for production of next-generation biofuels. *BMC Genomics* 12:148. <https://doi.org/10.1186/1471-2164-12-148>.
41. Converti A, Casazza AA, Ortiz EY, Perego P, Del Borghi M. 2009. Effect of temperature and nitrogen concentration on the growth and lipid content of *Nannochloropsis oculata* and *Chlorella vulgaris* for biodiesel production. *Chem Eng Process* 48:1146–1151. <https://doi.org/10.1016/j.ccep.2009.03.006>.
42. Ahn JW, Hwangbo K, Lee SY, Choi HG, Park YI, Liu JR, Jeong WJ. 2012. A new Arctic *Chlorella* species for biodiesel production. *Bioresour Technol* 125:340–343. <https://doi.org/10.1016/j.biortech.2012.09.026>.
43. Li T, Zheng Y, Yu L, Chen S. 2013. High productivity cultivation of a heat-resistant microalga *Chlorella sorokiniana* for biofuel production. *Bioresour Technol* 131:60–67. <https://doi.org/10.1016/j.biortech.2012.11.121>.
44. Canha N, Felizardo P, Correia MJN. 2018. Controlling the oxidative stability of biodiesel using oils or biodiesel blending or antioxidants addition. *Environ Prog Sustain Energy* 37:1031–1040. <https://doi.org/10.1002/ep.12778>.
45. Belarbi EH, Molina E, Chisti Y. 2000. A process for high yield and scalable recovery of high purity eicosapentaenoic acid esters from microalgae and fish oil. *Enzyme Microb Technol* 26:516–529. [https://doi.org/10.1016/S0141-0229\(99\)00191-X](https://doi.org/10.1016/S0141-0229(99)00191-X).
46. Blanc G, Duncan G, Agarkova I, Borodovsky M, Gurnon J, Kuo A, Lindquist E, Lucas S, Pangilinan J, Polle J, Salamov A, Terry A, Yamada T, Dunigan DD, Grigoriev IV, Claverie J-M, Van Etten JL. 2010. The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* 22:2943–2955. <https://doi.org/10.1105/tpc.110.076406>.
47. Arriola MB, Velmurugan N, Zhang Y, Plunkett MH, Hondzo H, Barney BM. 2018. Genome sequences of *Chlorella sorokiniana* UTEX 1602 and *Micractinium conductrix* SAG 241.80: implications to maltose excretion by a green alga. *Plant J* 93:566–586. <https://doi.org/10.1111/tpj.13789>.
48. Dahlqvist A, Stahl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S. 2000. Stymne H:phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc Natl Acad Sci U S A* 97:6487–6492. <https://doi.org/10.1073/pnas.120067297>.

49. Zuñiga C, Li C-T, Huelsman T, Levering J, Zielinski DC, McConnell BO, Long CP, Knoshaug EP, Guarnieri MT, Antoniewicz MR, Betenbaugh MJ, Zengler K. 2016. Genome-scale metabolic model for the green alga *Chlorella vulgaris* UTEX 395 accurately predicts phenotypes under autotrophic, heterotrophic, and mixotrophic growth conditions. *Plant Physiol* 172:589–602. <https://doi.org/10.1104/pp.16.00593>.
50. Cao K, He M, Yang W, Chen B, Luo W, Zou S, Wang C. 2015. The eurythermal adaptivity and temperature tolerance of a newly isolated psychrotolerant Arctic *Chlorella* sp. *J Appl Phycol* 28:1–12. <https://doi.org/10.1007/s10811-015-0627-0>.
51. Dere S, Günes T, Sivaci R. 1998. Spectrophotometric determination of chlorophyll-a, b and carotenoid contents of some algae species using different solvents. *Turk J Bot* 22:13–17.
52. Dubois M, Gilles KA, Hamilton JK, Rebers P, Smith F. 1956. Colorimetric method for determination of sugars and related substances. *Anal Chem* 28:350–356. <https://doi.org/10.1021/ac60111a017>.
53. Walker JM. 2009. *The protein protocols handbook*, 3rd ed, p 11–15. Springer, New York, NY.
54. Bligh EG, Dyer WJ. 1959. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 37:911–917. <https://doi.org/10.1139/o59-099>.
55. Shi CY, Wan XC, Jiang CJ. 2007. Method for high-quality total RNA isolation from tea plants (*Camellia sinensis* (L.) O. Kuntze). *J Anhui Agric Univ* 34:360–363.
56. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, Zhang X, Wang J, Yang H, Fang L, Chen Q. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* 7:1–6. <https://doi.org/10.1093/gigascience/gix120>.
57. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Muceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat Biotechnol* 29:644–652. <https://doi.org/10.1038/nbt.1883>.
58. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J. 2003. TIGR Gene Indices Clustering Tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19:651–652. <https://doi.org/10.1093/bioinformatics/btg034>.
59. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
60. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. <https://doi.org/10.1186/1471-2105-12-323>.
61. Miller W, Myers EW, Lipman DJ. 2008. Blast (basic local alignment search tool). *Encycl Genet Genom Proteom Inform*: 215–221.
62. Audic S, Claverie JM. 1997. The significance of digital gene expression profiles. *Genome Res* 7:986–995. <https://doi.org/10.1101/gr.7.10.986>.