Characterization and Classification of Highly Productive Microalgae Strains Discovered for Biofuel and Bioproduct Generation

Peter Neofotis
Department of Biology, Graduate Center of the City University of New York, New York, New York 10016, USA
Department of Biology, Brooklyn College, Brooklyn, New York 11210-2889, USA

Andy Huang, Kiran Sury, William Chang, Florenal Joseph[1]*, Arwa Gabr*
Department of Biology, Brooklyn College, Brooklyn, New York 11210-2889, USA

Scott Twary
Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Weigang Qiu
Department of Biology, Hunter College, New York, New York 10065, USA

Omar Holguin
Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, New Mexico 88003, USA

and *Jürgen E.W. Polle*[2]
Department of Biology, Graduate Center of the City University of New York, New York, New York, 10016
Department of Biology, Brooklyn College, Brooklyn, New York 11210-2889

---

[1] Present Address: Molecular and Cellular Biology, School of Graduate Studies at State University of New York Downstate Medical Center, Brooklyn, NY 11203
[2] Author for correspondence: email jpolle@brooklyn.cuny.edu, Telephone 917-951-2025, Fax 718-951-4659

© 2016. This manuscript version is made available under the Elsevier user license
http://www.elsevier.com/open-access/userlicense/1.0/

ABSTRACT

This paper describes the characteristics of microalgal strains that originated out of an isolation and screening project included within the National Alliance for Advanced Biofuels and Bioproducts (NAABB). The project's goal was to identify new potential platform strains with high growth rates and/or lipid productivities. To classify the best performing strains, we conducted a combined microscopic and phylogenetic analysis. Among the best performing strains were many coccoid green algae. Several strains belong to the species *Acutodesmus (Scenedesmus) obliquus* and to the species *Chlorella sorokiniana*, thus expanding on existing germplasm. Identified at the genus level were some *Desmodesmus* strains and one *Ankistrodesmus* strain. Several strains were classified as belonging to the genus *Coelastrella*, a taxon reported for the first time for North America. Multiple additional strains had ambiguous identities, with some strains possibly representing novel species. Reporting on the above strains, some of which have been tested successfully in outdoor ponds and most of which are deposited at the University of Texas Culture Collection of Algae, is a step forward in expanding the biological resources available for algae biofuel production.

52 *Key index words*: biofuel, bioprospecting, green algae, *Scenedesmus*, *Coelastrella*,
53 *Chlorella*

54

55 *Abbreviations*: UHPLC, Ultra high performance liquid chromatography; PDA, Photo
56 diode array; MS, Mass Spectrometry; APCI, Atmospheric Pressure Chemical Ionization;
57 ITS, internal transcribed spacer; SSU, small-subunit; LSU, large-subunit; TAG,
58 triacylglycerols

59 **1. Introduction**

60 Algae are seen as a more appealing biofuel feedstock than land plants because of their
61 faster biomass doubling cycles, more accessible forms of stored carbon, and their ability
62 to thrive on waste or salt water sources [1, 2, 3]. But the best strains for such production
63 are unlikely to be known or utilized at large scales [4]. Therefore, a need exists for
64 "phyco-prospecting" [4] – the identification of novel algal "platform strains" – which
65 sustain high growth rates and accumulate lipids [5, 6, 7, 8, 9]. In addition, strains that
66 produce significant amounts of higher-value products, such as carotenoids, are also
67 desired [1, 2, 10, 11, 12, 13]. Such novel "platform strains" could, in turn, be developed
68 further via artificial selection, genetic engineering, or other "crop improvement" methods
69 [3, 14, 15] and used to generate biofuel feedstocks at lower costs.

70 Commercial scale algae ponds have been operated for more than a decade [3, 16, 17,
71 18], mainly to harvest pigments and metabolites that are desired as nutritional
72 supplements [10]. The algae cultivated in these ponds include Spirulina (*Arthrospira*), for
73 high protein powder [19, 20], *Haematococcus,* for the antioxidant astaxanthin [21, 22],
74 and *Dunaliella salina*, for pro-vitamin A production [23, 24]. Due to the success
75 demonstrated with these used species [17], the as of yet untapped diversity of microalgae
76 appears to offer possibilities in the field of biofuels [25, 26, 27], as well as the production
77 of high-value products [1], such as pharmaceuticals [10, 13, 28].

78 It is not clear how many microalgae species there are, with estimates running from
79 70,000 to one million [29, 30, 31]. Only about 44,000 have been described [29, 32]. New
80 species and genera are consistently being discovered; this consistent rate of discovery
81 indicates that a large proportion of undocumented species exists [29].

82 Based on the known utility of already characterized species and the known absence of
83 knowledge regarding the large amount of undocumented strains, it is hypothesized that
84 previously uncharacterized strains of microalgae exist which exhibit high growth rates
85 and lipid productivities. These novel "strains" may be new species or varieties of species
86 with more favorable growth characteristics for biofuel production. Some of these species
87 or strains may also produce high value products. To test this hypothesis, a large-scale
88 effort was conducted to isolate, screen, identify, and characterize microalgae strains that
89 could be used as platform strains for biofuel and/or high-value product generation. For a
90 number of the most promising strains to come out of the screen, a classification was
91 conducted. In deciding on an appropriate barcode, the nuclear rDNA Internal Transcribed
92 Spacer 2 Region (ITS2) was chosen because, unlike the 18S – which often does not vary
93 enough to distinguish algal species [33] – the ITS2 has proved to be a helpful tool for
94 discrimination at the genus [33, 34] and species level [35, 36, 37, 38, 39, 40, 41, 42]. At
95 the same time, its two dimensional structure is highly conserved throughout the
96 eukaryotes [43, 44, 45]. Combining the fast evolving sequences with its slowly evolving
97 structure has allowed for it to be used in high level classifications while still
98 descriminating at the low species level, within the same phylogenic tree [37].
99 The overall goal of our research was to discover new microalgal strains that could
100 enhance the genomic/biological resources available for algae biomass production. This
101 biomass would then be used as feedstock for biofuels and/or for bioproducts generation.
102 The focus of this report is on freshwater algae originating mainly from the NAABB
103 phyco-prospecting project [46]. The objective of this report is to document some of these
104 strains and begin to have a cladistic understanding of where they fall in the algal tree of

105 life. It is hoped that this work may provide some clues for targeted phyco-prospecting
106 approaches in the future.
107
108 **2. Materials and methods**
109
110 *2.1. Sampling, isolation, and screening.*
111
112 Cells were sampled and screened as described previously in detail [47]. Briefly,
113 environmental samples were collected over the three years of the NAABB project (2010-
114 2012) from a variety of habitats (e.g. freshwater lakes and soils) across the continental
115 U.S. and during all seasons. As described in detail in [47], for strain isolation an approach
116 was taken that combined the traditional plating method [48] with high-throughput flow
117 cytometry using fluorescence aided cell sorting [49, 50].
118 For the first-level screen, in summary, batch cultures of strains were grown in
119 Erlenmeyer flasks in defined minimal media suitable for this selection: C Medium, Bold
120 Basal Medium, and BG11 Medium [51], with the only CO_2 enrichment being the sodium
121 bicarbonate which was added to BG11. Flasks were grown under 50 μE m^{-2} s^{-1}
122 continuous light provided by daylight fluorescent lamps (See the SN-1 for more details
123 on screening methods) at room temperature.
124 Strains that passed the first-level screen were deemed to be potential producers and
125 were selected for further characterization (Figure 1). In a second-level screen, the strains
126 were cultivated under identical conditions as in the first screen in multiple different

127 media (C Medium, Bold Basal Medium, BG11) to determine if another growth medium
128 was better suited for use in continued strain characterization (Figure 1).
129 In a third-level screen for analysis of actual biomass and lipid production potential,
130 strains were grown at 28 Degree Celsius in glass columns (3 cm width with 500 mL
131 volume) into which CO_2 enriched air was bubbled through the cultures from the bottom.
132 Twenty-eight degrees Celsius was used at this point in the screen, because it was deemed
133 a good intermediate temperature to find strains that grow well both in the spring and fall
134 as well as in the summer. Cultures were illuminated by daylight fluorescent lamps from
135 one side at an intensity of 200 μmol photons m^{-2} s^{-1}.
136 Biomass was determined daily by measuring the Ash-Free-Dry-Weight (AFDW,
137 Supplemental Notes 1(SN-1)). The daily AFDW was graphed using the ggplot package,
138 within the R software environment for statistical computing and graphics (http://cran.r-
139 project.org/), to create weighted polynomial curves, which represented algal growth
140 curves. The doubling time was also calculated in R for the exponential growth phase with
141 the formula log(2)/k, with k being the slope of the growth curve.
142 When the algal cultures reached their stationary phase, biomass also was taken and
143 analyzed via a gravimetric method to determine the total lipid content [52].
144
145 2.2. *Metabolite analysis*.
146
147 Pigments and lipids were extracted using a modified version of an established
148 protocol [53] (See SN-2). Pigments and lipids were analyzed using a Ultra High-
149 Performance Liquid Chromatography (UHPLC) C18 reverse-phase column, connected to

150 a Photo Diode Array detector (Flexar FX-15 UHPLC system from Perkin Elmer), and
151 then passed through an Atmospheric-Pressure Chemical Ionization (APCI) probe
152 connected to a Time-of-Flight Mass Spectrometer (Perkin Elmer, Massachusetts)
153 (UHPLC-PDA-APCI-TOF-MS). The samples were run through the LC with solvent A
154 consisting of 39% water, 59% Acetonitrile, and 2% ammonium acetate and solvent B
155 consisting of 88.2% 2 propanol, 9.8% acetonitrile, and 1.9% ammonium acetate. The
156 column was first equilibrated for 15 minutes with 60% A and 40% B and a flow of .4
157 ml/min. A 10 minute gradient transition then brought solvent B to 100% — and it was
158 held for 6 minutes. For one minute, then, solvent B was brought down to 40% and
159 solvent A was brought to 60%. Following, the column was washed for 5 minutes. For the
160 PDA, the differential absorbance was set at 450nm/550nm to detect the carotenoids. For
161 the MS, the following settings were used – Corona 5uA, Endplate -2000 Volts, Cap
162 Entrance -3000 Volts, Drying Gas Flow 5.0, Drying Gas Heater 300°C, Auxilary Gas 25,
163 APCI Heater 200°C, Capillary Ext 100 Volts, Skimmer 25.
164
165 2.3. *Benchmark for comparison*.
166
167 Our 'gold-standard', or benchmark, for biomass and lipid productivity comparison of
168 newly isolated strains was the species *Nannochloropsis salina* strain CCMP1776. This
169 strain was used by the National Alliance for Advanced Biofuels and Bioproducts [46, 54]
170 in outdoor cultivation trials in raceway ponds [46]. It had also previously been evaluated
171 in outdoor photobioreactors [55].
172

173 *2.4. DNA Extraction, polymerase chain reaction (PCR), and sequencing.*

174

175 A 50 ml culture in the light-limited growth phase was concentrated by centrifugation,
176 and then the DNA of 1 ml of the concentrated liquid sample was extracted using a MO-
177 BIO DNA Isolation Kit (USA). The ITS2 rDNA region was amplified using the universal
178 primers ITS3 (GCATCGATGAAGAACGCAGC) and ITS4
179 (TCCTCCGCTTATTGATATGC) [56]. GoTaq® (Promega) master mix was used as the
180 PCR reaction mixture. We performed 30 cycles (1 minute 95°C, 1 minute 48°C, and 1
181 minute at 72°C – with an initial denaturing step of 5 minutes at 95°C). In cases where,
182 even with the ITS2 barcode, the identity of the species was still not clear, the entire18S,
183 ITS1, 5.8S, and ITS2 regions were also amplified using universal primers AL1500af
184 (GCGCGCTACACTGATGC) and LR 1850 (CCTCACGGTACTTGTTC). In these
185 instances, 30 cycles (30 seconds 95°C, 30 seconds 48°C, and 1.5 minutes at 72°C – with
186 an initial denaturing step of 2 minutes at 95°C) were performed. PCR products were then
187 purified and sequenced by Genewiz (www.genewiz.com). For one of these organisms,
188 DOE0101, the DNA was also extracted using a DNeasy Plant Maxi Kit (Qiagen), and the
189 genome was then sequenced using Illumina technology at Los Alamos National
190 Laboratory, NM.

191

192 *2.5. Alignment and phylogenetic analysis.*

193

194 All 28 ITS2 sequences obtained from our lab samples were delimited and cropped
195 with the hidden Markov model (HMM)-based annotation tool present at the ITS2

database [57] (E-Value <0.001, Viridiplantae HMMs). In addition to the 28 ITS2 sequences from our top performing strains, we also downloaded ~50 algae ITS2 sequences from the ITS2 database (http://its2.bioapps.biozentrum.uni-wuerzburg.de/). As the sequences were to be used to help determine the taxonomic position of our strains, sequences from algae branches were chosen that NCBI blast analyses suggested were close to our specimens – within the Chlorophyceae and Trebouxiophyceae clades. From the Trebouxiophycea, the Chlorellales order seemed particularly well represented. Within the Chlorophyceae, the Sphaeropleales and Chlamydomonadales orders were well represented. Since one of our strains appeared to be an *Ankistrodesmus*, yet no sequences of that genus were available in the ITS2 database, an *Ankistrodesmus* sequence was also downloaded from NCBI and cropped using the aforementioned model. As outgroups, two species from the prasinophytes (*Mantoniella squamata* and *Micromonas pusilla*) were chosen. *Mantoniella squamata* has been used previously to root phylogenetic trees showing the evolution relationship of chlorophytes based on SSU rRNA sequence data [58]. *Micromonas pusilla* is a well characterized prasinophyte whose complete genome had been sequenced [59].

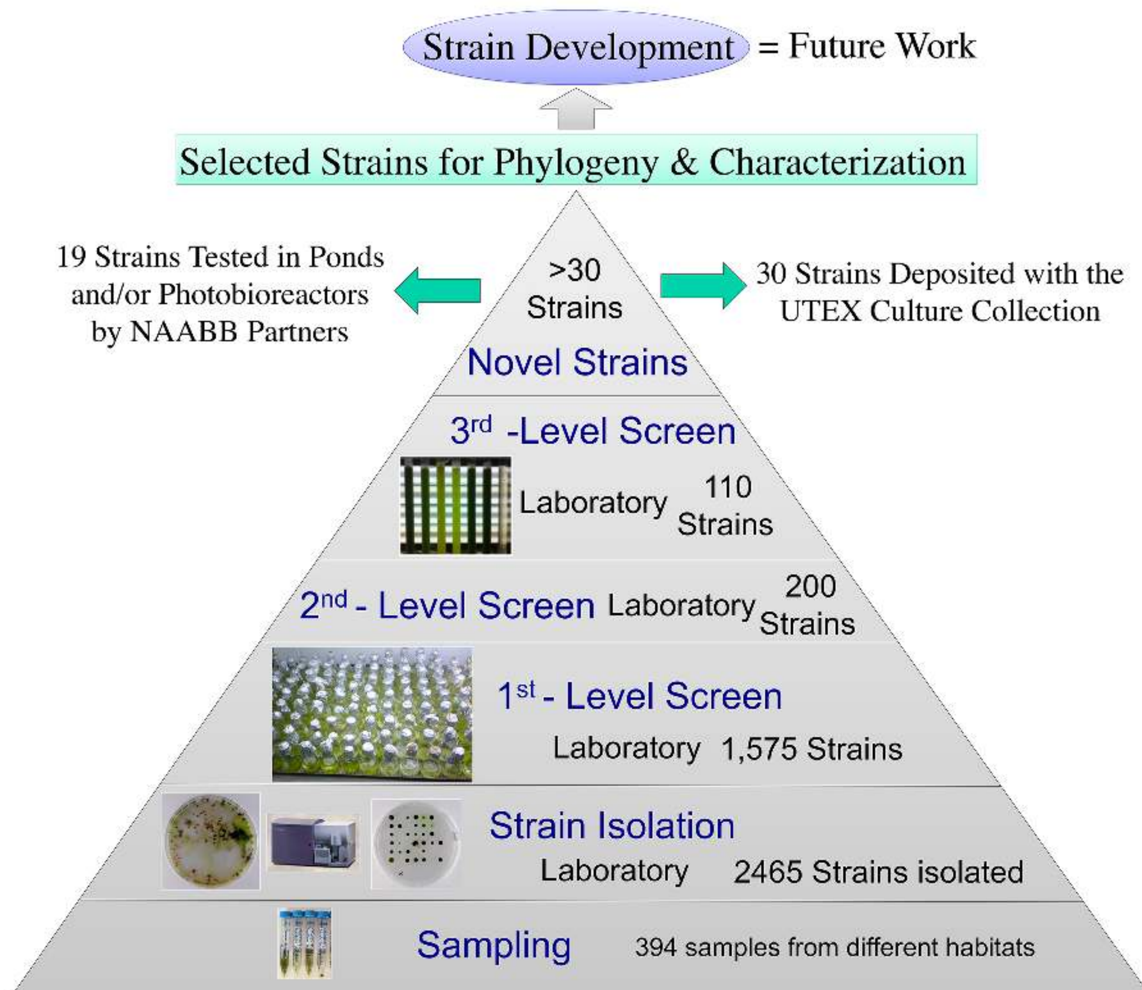
The cropped sequences and structures were then automatically aligned with 4SALE 1.7 using an ITS2-specific scoring matrix for sequences and structures [60, 61]. To find simultaneously the evolutionary distance between organisms according to sequence and secondary structures, profile neighbor joining (PNJ) was applied using ProfDistS 0.9.9 (Qt-Version) [62, 63]. An ITS2-specific general time-reversible substitution model was applied [60]. The resulting tree was displayed with FigTree (http://tree.bio.ed.ac.uk/) and further processed in R (http://cran.r-project.org/) using Ape [64].

219 The 18S, ITS1, 5.8S, ITS2, and 28S sequences were aligned with ClustalW and
220 reconstructed using Maximum Likelihood with a Tamura-Nei Model and Nearest-
221 Neighbor-Interchange (NNJ) heuristic method, with 500 bootstrap replications
222 performed. Significant bootstrap values are based on 500 replicates and mapped to the
223 appropriate internode. Branch lengths were drawn proportional to inferred changes.
224 Most of the strains described in this study were deposited at the UTEX Culture
225 Collection of Algae (http://web.biosci.utexas.edu/), and all sequences generated were also
226 deposited at the National Center for Biotechnology Information (NCBI) [65].
227
228 2.6. *Secondary Structure Prediction*.
229
230 The structures of most of the sequences downloaded from the ITS2 database were
231 determined either by direct-fold or homology modeling within the ITS2 database [41, 66,
232 67]. The secondary structures of most of the sequences from our strains were also
233 determined by the same homology model. For four of our sequences, that of DOE0101,
234 EN1423, DOE0259, and the *Ankistrodesmus* sequences from NCBI – because no related
235 organisms were found in the ITS2 database – RNAstructure 5.6 was used to predict the
236 ITS2 structure [68].
237
238 **3. Results and discussion**
239 The focus of our manuscript is on the phylogeny and characterization of some of the
240 most promising strains that were isolated in the course of our phyco-prospecting work for

241 the NAABB consortium [46]. Nevertheless, the first part of this section provides an
242 overview of the phyco-prospecting project.
243
244 *3.1. Summary of the NAABB phyco-prospecting project.*
245
246 The multi-year strain isolation and screening process, which yielded a variety of
247 novel strains, consisted of multiple stages [47]. For a better understanding of the results,
248 the overall approach that resulted in potential future platform strains is discussed below:
249 1. A dual spatial and temporal sampling strategy was employed. Sampling sites
250 represented very diverse, mostly aqueous but also several terrestrial environments,
251 with an emphasis on the southwestern United States, because the southwestern states
252 were viewed as the most promising potential sites for algal mass cultivation [9]. Many
253 shallow and temporal water bodies were sampled, because it was anticipated that
254 future use of strains would include raceway-type ponds.
255 2. Aliquots of environmental samples were plated onto different agar-solidified media in
256 Petri dishes. The rationale for this first step was that different types of algae could be
257 easily distinguished by colony color and size. Often microscopic examination of
258 colonies and its cells allowed a broad taxonomic categorization of potential isolates.
259 Furthermore, introduction of this first step allowed for the determination of isolates
260 that could grow on agar-solidified media. This was viewed as important for future
261 strain development, which might be done in high-throughput applications [69].

262 3. Transfer of colonies onto index plates was done with the rationale that more biomass
263 could be used for inoculation of liquid medium, which reduced the lag-phase of the
264 liquid cultures, thus speeding up the isolation process.
265 4. Transfer into liquid medium was done because the cell sorter could only process algal
266 cells provided in liquid. This had the additional benefit of demonstrating that the cells
267 did indeed grow in liquid culture, which was important as later applications were
268 anticipated to involve cultivation in open ponds or photobioreactors.
269 5. FACS cell sorting onto Petri dishes was applied with the rationale to provide a high-
270 throughput method to obtain a large number of unialgal strains rapidly. Sometimes
271 new strains were also already axenic following cell sorting onto plates. Sorting into
272 liquid wells did not prove as efficient, because determination of potentially unialgal
273 strains was more difficult in liquid medium.
274 6. Unialgal strains were then transferred into storage tubes consisting of agar-solidified
275 medium overlaid with liquid medium. This storage method was chosen because all
276 strains could be maintained in this manner. Also, we found that strains in agar/liquid-
277 overlay cultures under a 12 h light/dark regime could survive without transfer for up to
278 three years.
279 Figure 1 summarizes our phyco-prospecting efforts for the NAABB program. Within
280 the isolation and screening project, several bottlenecks were encountered at different
281 levels due to the labor-intensive steps involved. Out of our total of 2,465 isolated
282 NAABB strains, only 1,575 strains could be subjected to a first-level screen, of which
283 334 passed as candidates for potential use as future biofuel/nutraceutical platform
284 species. Subsequently, from the 334 candidate strains only 200 strains could be tested in a

285 second-level screen to determine the optimal growth medium. In the following third-level
286 screen only 110 strains could be examined for actual biomass and lipid productivities in
287 bubbling columns. Of the 110 strains tested in bubbling columns, more than 30
288 freshwater strains outperformed our 'gold' standard, *Nannochloropsis salina*, in terms of
289 growth rate and final biomass yield. In addition to having high growth rates, several of
290 our best performing strains also accumulated carotenoids in cytosolic oil bodies.
291



292
293 Figure 1: Schematic summarizing the overall phyco-prospecting process performed
294 in our laboratory as part of the NAABB Consortium to identify the best performing
295 microalgae strains for biofuel production.
296

297 Overall, 30 strains were deposited with the UTEX culture collection. Currently, these
298 strains are not in the main UTEX collection, but they are in a special collection. Out of
299 these 30 deposited strains, general information on 25 strains is contained in Table 1. Most
300 strains were isolated from samples collected in summer. Three strains came out of fall
301 samples (DOE0686, DOE0717, DOE1357) and strain DOE1135 originated from a
302 February sample. Additionally, Table 1 includes information on three strains from a
303 previous phyco-prospecting effort.

Strain Number	**Taxon**	**State**	**Latitude**	**Longitude**	**Habitat**	**NCBI Accession**
DOE0013	*Scenedesmus*	CA	37.97	-122.56	Fountain	KJ434964
DOE0043	*Desmodesmus*	CA	37.9	-122.25	Lake	KJ434974
DOE0088	*Coelastrella*	CA	37.96	-122.34	Birdbath	KJ43972
DOE0101*	Chlamydomonadales	TX	29.92	-95.82	Roadside Mud	KJ434990
DOE0111	*Scenedesmus obliquus*	CA	37.96	-122.34	Water from a Garden Pot	KJ439665
DOE0152*	*Scenedesmus obliquus*	NY	40.65	-74.01	Water from a Turtle Tank	KJ434966
DOE0155	*Coelastrella*	TX	29.84	-95.87	Bayou	KJ434970
DOE0202*	*Coelastrella*	CA	37.96	-122.34	Water from a Garden Pot	KJ434971
DOE0259	*Ankistrodesmus*	TX	29.87	-95.81	Roadside Ditch	KJ434991
DOE0314	*Dichloster-related*	TX	29.87	-95.81	Roadside Ditch	KJ434977
DOE0369	*Coelastrella*	CA	38.89	-121.97	Fountain	KJ43967
DOE0623	Chlorellaceae	TX	29.77	-95.62	Bayou	KJ434983
DOE0686	*Chlorella sorokiniana*	CT	41.6	-72.7	Greenhouse Hydroponics	KJ434981
DOE0700	Chlorellaceae	NM	32.31	-106.78	Wastewater Pond	KJ434979
DOE0717	Chlorellaceae	CT	41.6	-72.7	Greenhouse Hydroponics	KJ434986
DOE1041	Chlorellaceae	TX	30.29	-97.84	Roadside Ditch	KJ434978
DOE1044	Chlorellaceae	TX	30.78	-95.95	Unknown	KJ434987
DOE1051	*Desmodesmus*	TX	30.04	-101.2	Ditch	KJ434975
DOE1070	Chlorellaceae	TX	27.71	-97.18	Bay	KJ434985
DOE1095	Chlorellaceae	TX	29.01	-95.22	Estuary	KJ434984
DOE1116	*Chlorella sorokiniana*	TX	29.28	-94.83	Estuary	KJ434982

DOE1135	Chlorellaceae	TX	27.71	-97.18	Bay	KJ434988
DOE1357	*Desmodesmus*	TX	30.36	40.63	Lake	KJ43976
DOE1412*	*Chlorella sorokiniana*	NY	40.63	-73.95	Culture Contaminant	KJ434980
DOE1418	*Desmodesmus*	IL	39.5	-88.18	Lagoon	KJ34973
EN1234	*Coelastrella*	CA	32.77	-117.19	Bird Bath	KJ434969
EN1235	*Coelastrella*	NY	Unknown		Bird Bath	KJ43968
EN1423	*Borodinellopsis texensis*	NM	34.34	-105.57	Roadside Saline Soil	KJ434989

304
305
306 Table 1: Listed are the top performing strains that were classified in the work reported
307 here. Included in the table are the strain numbers, taxon, and location along with the
308 habitat of collection. BG11 was the best growth media for all strains in the list. Strains
309 with the ID DOE#### are from the NAABB phycoprospecting project. All these strains
310 have been deposited at the University of Texas Culture Collection of Algae and are
311 available to researchers. Strains with the ID EN#### resulted from an Airforce Office of
312 Science funded phyco-prospecting effort. The asterisk next to Strain ID numbers
313 indicates that the strain was tested successfully in outdoor ponds at NAABB testbed sites
314 [46].
315
316

317 3.2. *Growth evaluation of strains.*

318

319 Following sampling and strain isolation a multi-level screen was applied to identify
320 potential lipid producers. The first-level screen evaluated growth by optical density
321 changes and lipid productivity using Nile Red fluorescence [47]. The second-level screen
322 evaluated performance of strains in different growth media [47]. In the third-level screen
323 strains were cultivated in bubbling columns [47] and biomass productivities and lipid
324 levels were determined gravimetrically.

325 For most strains the third-level screen only consisted of one culture having been
326 tested. Only for a few selected groups of strains was more than one culture grown. Figure
327 2 shows growth curves of selected groups of the best performing microalgal strains
328 identified: what our phylogenetic analysis would classify as *Chlamydomonadales* strain

329 DOE0101, *Chlorella sorokiniana* strain DOE1412, three novel *Coleastrella* sp., and two

330 strains of the species *Scenedesmus obliquus* - also known as *Acutodesmus obliquus* [34].

331 Because all bubbling column cultures were inoculated from pre-cultures that were in the

332 light-limited growth phase, the lag-phase for all the bubbling column cultures were

333 usually very short. For all strains shown in Figure 2 the exponential growth phase was

334 followed, between about 50 to 100 hours, by the light-limited growth phase. During the

335 light-limited growth phase the biomass productivity of the best strains identified was

336 between 0.6 to 0.9 g AFDW l^{-1} d^{-1} (Figure 2). In comparison, *N. salina*, our 'standard',

337 only produced about 0.6 g AFDW l^{-1} d^{-1}. A recent literature review discussed the

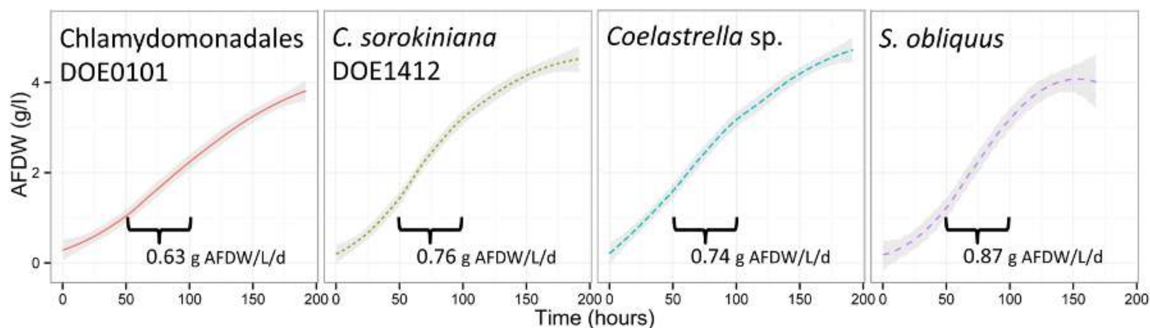
338 difficulties of making rigorous comparisons between different algae growth experiments

339 because of variations in culture conditions [70]. Nevertheless, the biomass productivities

340 of our strains were also within the range of the reviewed most promising species of algae

341 for lipid productivity [70].

342



343

344 Figure 2: Shown are growth curves based on the AFDW (grams/liter) of some of the most
345 productive strains. Biomass productivities are given for the light-limited growth phase
346 between 50-100 hours. At least two biological replicates were included for each strain.
347 Shading represents the 95% confidence region for the regression fit of the biological
348 replicates.
349

350 For the data shown in Figure 2, for *S. obliquus* the growth curve shown represents

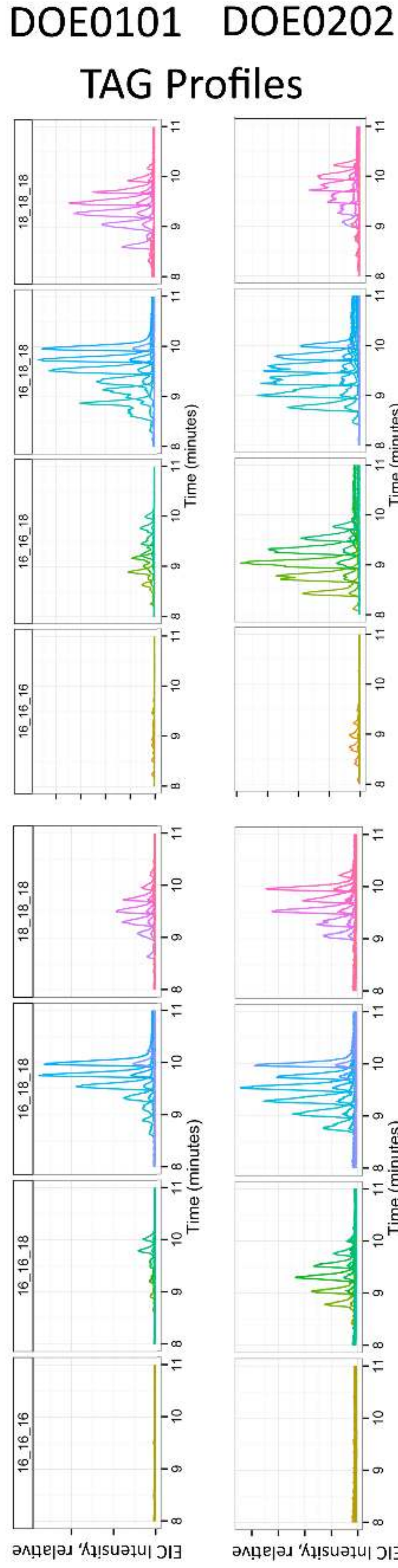
351 two strains DOE0111 and DOE0152. The growth curve of *Coelastrella* sp. represents the
352 summary results of three independently isolated strains DOE0088, DOE0202, and
353 DOE0369. Most of our tested strains reached the stationary growth phase with total
354 biomass yields of sometimes more than 4.0 g AFDW L^{-1}. Noteworthy is that at about 200
355 hours, when the batch cultures were terminated, the strains of the *Coelastrella* had still
356 not reached the stationary growth phase. At that time, the biomass yield was about 4.5 g
357 AFDW l^{-1}. When grown for a longer time period, the biomass yields were sometimes
358 upwards of 6.0 g AFDW l^{-1} and the cultures had still not reached the stationary phase.

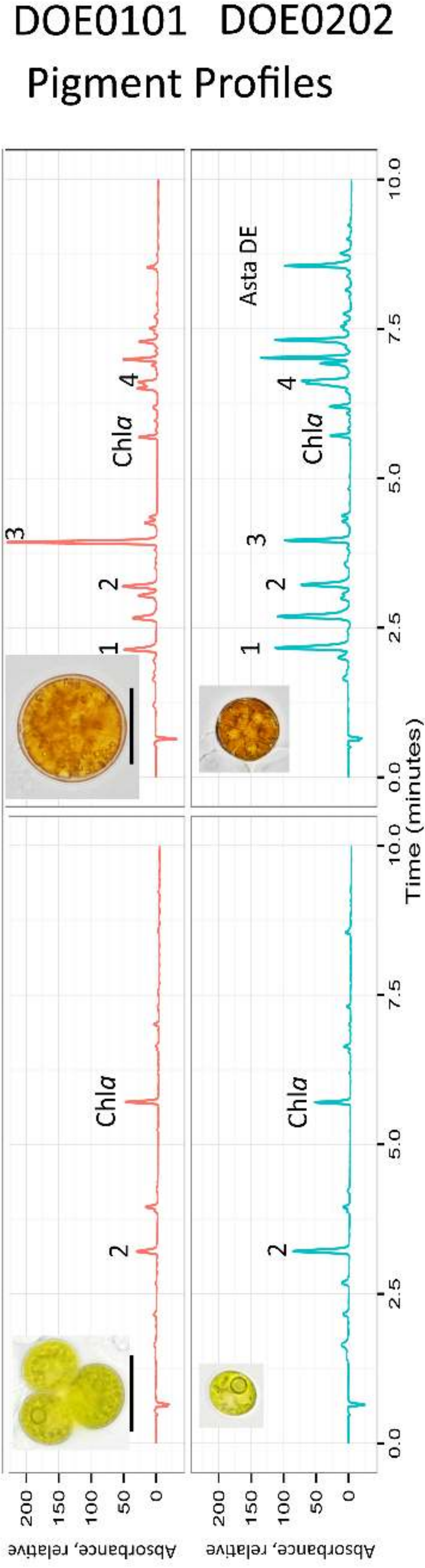
359 When in the stationary growth phase at 180 hours the strains had accumulated the
360 following amounts of total lipids as part of their biomass: *Chlamydomonadales sp*. strain
361 DOE0101 (25%); *Chlorella sorokiniana* DOE1412 (25%); *Coelastrella* sp. strain
362 DOE0088 (17%) and strain DOE0202 (23%); *S. obliquus* strain DOE0111 (29%) and
363 strain DOE0152 (28%). It should be noted that due to their thick cell walls it did not
364 appear that all lipids could be extracted from cells of the *Coelastrella* strains, as the
365 remaining biomass in the extraction was often still green. Thus, total lipid contents was
366 likely underestimated. Overall, the total lipid productivities of all strains shown in Figure
367 2 were similar at about 0.2 g lipids l^{-1} d^{-1}.

368 Although growing slower, the lipid content of *N. salina* could reach over 30% lipids
369 within the five days in batch culture, thus ultimately also reaching about 0.2 g total lipid l^{-}
370 ^{1} day^{-1}. However, compared to *N. salina*, our top strains had shorter lag-phases and higher
371 daily biomass productivities. This is important, as the latest biofuels applications aim at
372 use of hydrothermal liquefaction technologies (HLT), which convert biomass into crude
373 oils rather than extracting lipids only [71, 72, 73, 74]. Consequently, rather than

374 technologies that are dependent on the lipid content of algal cells, Hydrothermal
375 liquefaction technologies can rely on fast growing, biomass-accumulating strains to
376 increase higher overall crude oil productivities.
377
378 3.3. *Characterization of carotenogenic capabilities of selected strains.*
379
380 When cultivated in shaker flasks during the first–level screening process, it was
381 noticed that a variety of strains over-accumulated secondary carotenoids in cytosolic lipid
382 bodies [47] similar to *Haematococcus pluvialis* [75, 76]. All *Coelastrella* strains listed in
383 Table 1, the strain DOE0101 belonging to the Chlamydomonadales, and *Borodinellopsis*
384 *texensis* strain EN1423 displayed such carotenogenic capability. This over-accumulation
385 of secondary carotenoids, including Astaxanthin, may be of value for future
386 biotechnological application in carotenoid production. This is especially the case for
387 DOE0101 and DOE0202, as both strains were successfully cultivated in outdoor ponds at
388 the NAABB testbed site in Pecos, TX [46]. Therefore, these two strains were further
389 characterized. As the secondary carotenoids accumulate in cytosolic oil bodies, the
390 carotenoid and triacylglycerol (TAG) profiles were compared for both strains for actively
391 growing (green) cultures and carotenogenic (orange) cultures (Figure 3). In the center,
392 Figure 3 displays exemplary UHPLC chromatogram traces of extracted pigments. Peaks
393 shown in the UHPLC chromatogram traces are either carotenoids or chlorophylls.
394 Identification of carotenoids was done based on their absorbance spectra and based on
395 their molecular mass (http://metlin.scripps.edu/index.php). Nevertheless, for some of the
396 peaks, identifying the molecules as one carotenoid rather than another could not be done,

397 as they were either stereoisomers or different ions with the same mass. However,
398 carotenoids of pharmaceutical relevance are present in these algae. As indicated by the
399 orange color of cells and cultures, the three strains investigated contained different
400 carotenoids such as Peak 1 with Adonirubin or Astaxanthin, Peak 2 with Lutein or
401 Zeaxanthin, Peak 3 with Canthaxanthin or Adonixanthin, Peak 4 with Echinenone or
402 Cryptoxanthin. and minor amounts of Astaxanthin diesters (Peaks at later retention
403 times). In brief, strain DOE0101 seemed to have a distinctly different profile with a
404 particularly high accumulation of either canthaxanthin or adinoxanthin when compared to
405 the *Coelastrella* species strain DOE0202 (Figure 3).
406





407

408
409 Figure 3: Compared are cells and metabolites of the *Chlamydomonadales* sp. strain
410 DOE0101 and the *Coelastrella* sp. strain DOE0202. The upper panel shows EIC traces
411 obtained by Mass Spectrometry representing Triacylglycerol (TAG) profiles. The lower
412 panel displays PDA profiles (absorbance traces at 450 nm) of extracts from cells of
413 actively growing cultures and of carotenogenic (orange) cultures. Photos for
414 representative cells from these cultures are shown for both strains. The black bars
415 underneath the cells of strain DOE0101 indicate 10μm. The same scale was used for
416 photos of cells from strain DOE0202. Extractions were conducted from the same volume
417 of cells for cultures of the same age (SN-3).
418

419 Figure 3 also displays profiles of overlaid Extracted Ion Counts (EIC) traces for a
420 variety of triacylglycerol molecules identified for both strains. The individual EIC traces
421 were obtained by UHPLC separation followed by APCI-TOF-MS detection of TAG
422 molecules from lipids extracted from algal strains. Each EIC peak represents a different
423 specific TAG molecule [77]. All strains investigated contained dozens of different
424 molecular classes of TAGs (SN-4). Superficially, under the light microscope the cells
425 look somewhat similar (Figure 3). Nevertheless, stressed cells of DOE0101 have
426 comparatively very little C_{16}-C_{16}-C_{18} TAGs, and have mostly saturated C_{16}-C_{18}-C_{18}
427 molecules. In contrast, cells of the *Coelastrella* species have more C_{16}-C_{16}-C_{18} TAGs.
428 Also, its C_{16}-C_{18}-C_{18} s are weighted, if anything, toward the unsaturated molecules
429 (Figure 3). These differences among the cells of the two species may be due to the
430 different amounts of the various secondary carotenoids, in the two strains, that are
431 sequestered in the oil bodies. Investigating this further is of pharmaceutical/cosmetic
432 relevance, but is beyond the scope of this work.

433

434 3.4. *Strain classification.*

435

436 For the phylogenetic analysis, rDNA ITS2 was used as the molecular marker because

437 both it's sequence and secondary structure can be used for classification, and the

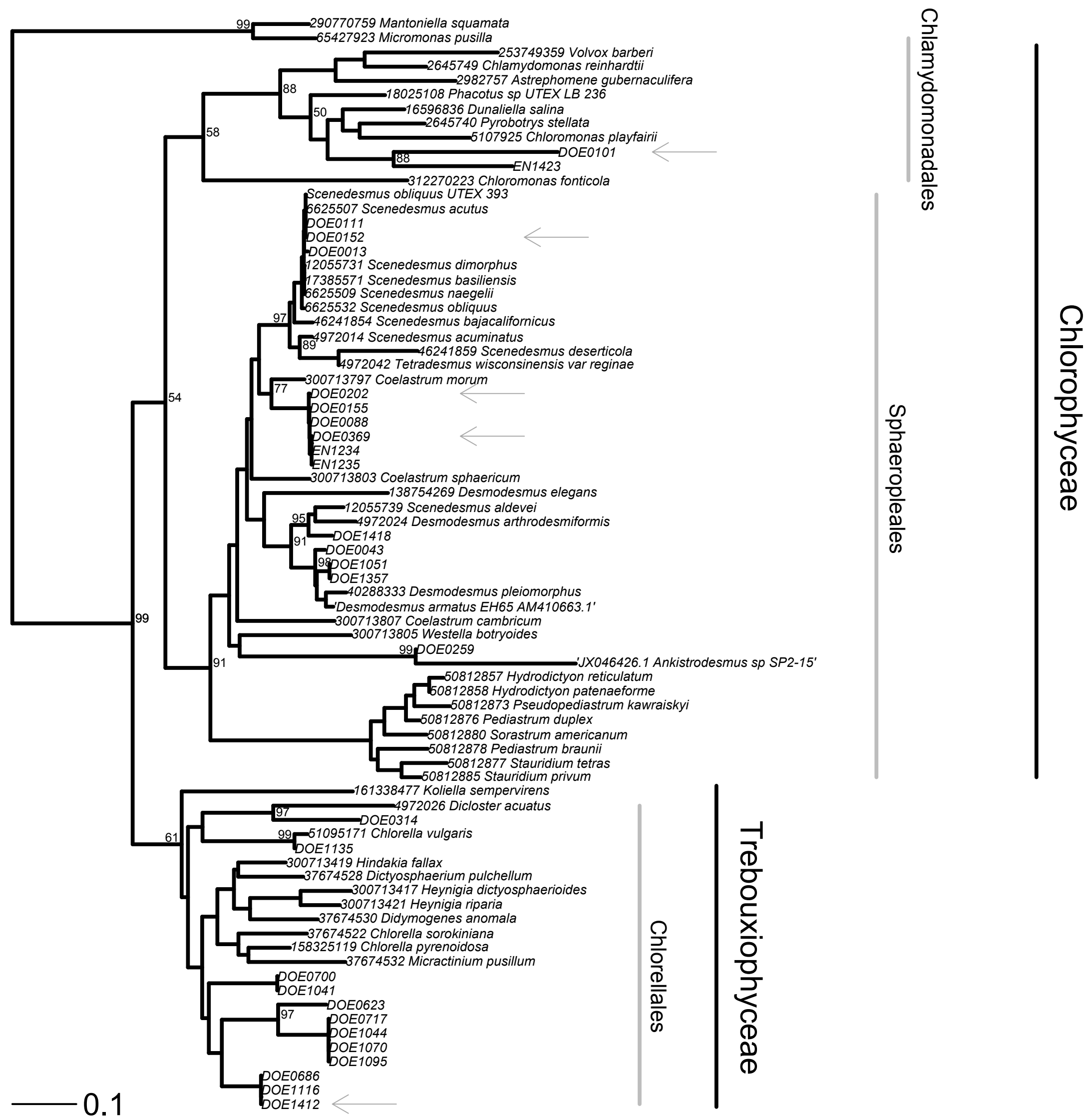
438 presences of many other ITS2 sequences in the ITS2 database

439 (http://its2.bioapps.biozentrum.uni-wuerzburg.de/) [33, 78, 79, 80]. Sequences of 49

440 previously identified strains of Chlorophytes and three strains from a previous

441 phycoprospecting project were compared with sequences of 25 of our most promising

442 strains (Figure 4).



443

444 Figure 4: ITS2 structure-aligned sequence tree of novel strains for biofuel production and
445 also known strains of the Chlorophyta – rooted with two prasinophyceae. Significant
446 bootstrap values (>50%) are based on 500 replicates and mapped to the appropriate
447 internode. Unlabeled internal branches are not significant (<50% bootstrap support).
448 Branch lengths are drawn proportional to the number of nucleotide substitutions. Arrows
449 point to strains discussed in this paper in more detail.
450

451 Within the ITS2 phylogenetic tree - including 25 newly isolated strains from the
452 NAABB project, 3 previously isolated novel strains, and the 49 known strains – some
453 clades were not resolved, because of low support in some internal branches (Figure 4).
454 Nevertheless, the major branches were well supported in their monophyly, and for the
455 known sequences from the ITS2 database, our phylogeny is consistent with the
456 established relationship among the Chlorophyta [81]. Similarly, in the systematics of the
457 major clade within our study, the Scenedesmaceae, there is low bootstrap support for the
458 bases of the trees and only good support for the groups of species within genera [33, 34,
459 82]. Therefore, despite the presence of some low bootstrap values in our tree, that the
460 known taxa fall in accordance to other published ITS1, ITS2, and 18S trees yields some
461 confidence for the phylogenetic information revealed in the branches. Our investigation
462 corroborates conclusions of other studies [83], which showed that sequence-structure
463 analysis of ITS2 provides a taxon-rich means of testing phylogenetic hypothesis at high
464 taxonomic levels.

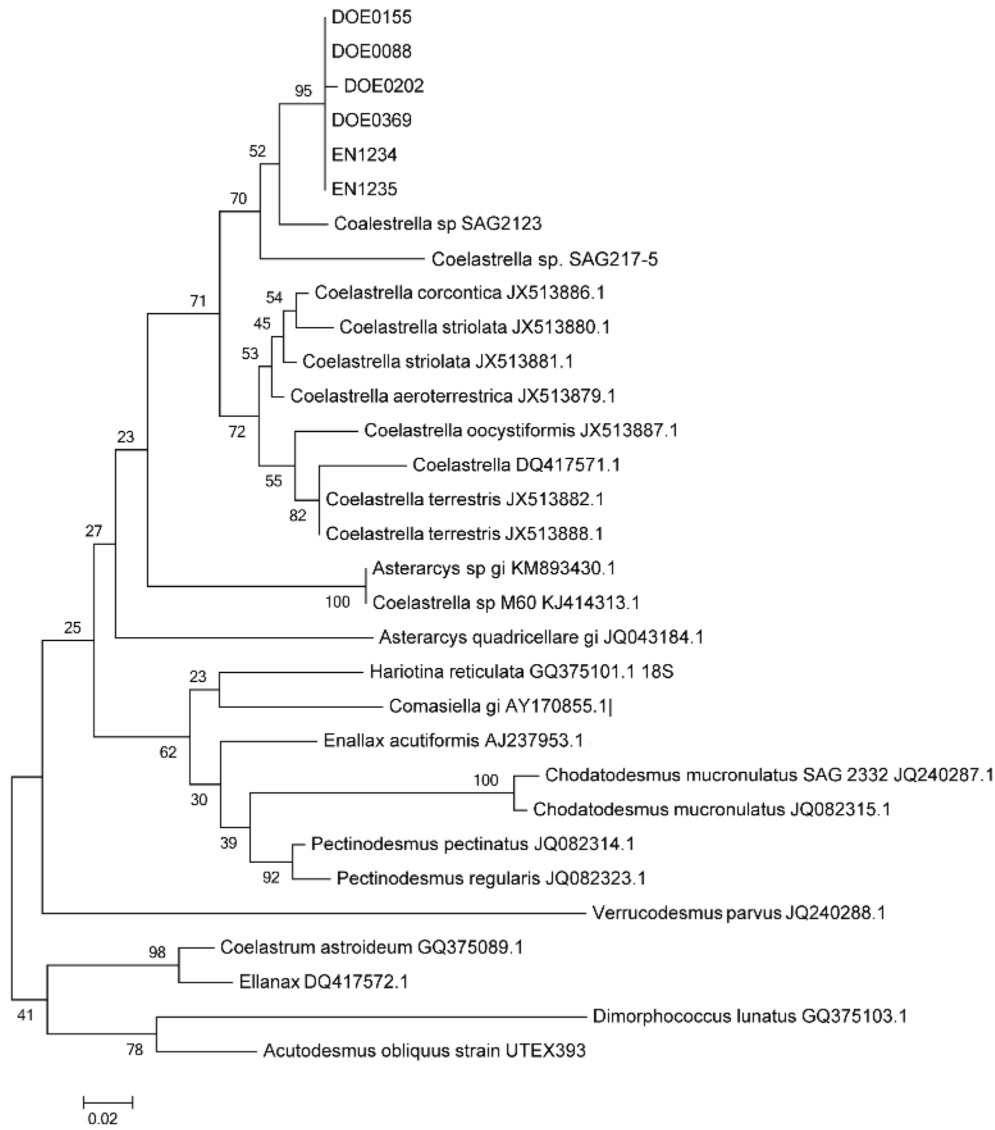
465 In general, our best performing strains were coccoid green algae that either fell within
466 the green algal order of the Sphaeropleales (Chlorophyceae), the Chlamydomonadales
467 (Chlorophyceae), or the Chlorellales (Trebouxiophyceae) (Figure 4). Among the
468 thousands of strains that were isolated preliminary microscopic investigation of many
469 strains during the isolation and screening process demonstrated presence of flagellate,

470 coccoid, and even some filamentous strains of the green algae. In addition, many
471 cyanobacteria, diatoms, and other algae (e.g. Eustigmatophytes) were identified based on
472 microscopic examination of cells. Therefore, our overall isolation protocol did not per se
473 discriminate against non-green algal species. This report focuses on the green algae,
474 because among the best performing strains emerging from our multi-level screen the
475 majority were coccoid green algal species. This may be because non-motile coccoid cells
476 do not 'waste' resources on active movement, a process that consumes large amounts of
477 energy. Moreover, coccoid cells cannot determine their location and may have a need to
478 store high-energy containing molecules as reserve substances. In consequence, many
479 coccoid species could have evolved to store energy in oil bodies, specifically under
480 environmentally unfavorable conditions.
481
482 *3.4.1. Sphaeropleales (Class of the Chlorophyceae)*
483 With regard to the newly isolated strains that are predicted to be within the class of
484 the Chlorophyceae, many strains belong to the order of the Sphaeropleales and the family
485 of the Scenedesmaceae. Several of our top performing strains (DOE0013, DOE0111,
486 DOE0152) are similar to strains identified in the ITS2 database as *S. obliquus*. Although
487 the species *S. obliquus* is synonymous to *A. obliquus* [34], in this paper, we refer to it as
488 *S. obliquus* to keep consistent with the papers we are comparing our results to. Cells of
489 our newly isolated strains have the characteristic oblong shape and coenobia. As an
490 example, we show cells of *S. obliquus* strain DOE0152 in Figure SM-1.
491 That several top performers independently isolated were strains of *S. obliquus* was
492 not surprising. Strains of *S. obliquus* have been known to grow well in mass cultures –

493 even demonstrating potential in outdoor raceway ponds in the developing world [84, 85].
494 Strains of the genus *Scenedesmus* have been found to be able to accumulate lipids [86]
495 even under colder temperatures such as 10°C [87], which is important as winter strains
496 are needed for biofuel production. Other studies have found that *S. obliquus* has a high
497 tolerance and growth rates under elevated levels of CO_2, suggesting that the species has
498 potential to reduce CO_2 in the flue gases emitted by thermoelectric power plants [88].

499 In bubbling columns with 3% CO_2 at 25°C, biomass productivity for *S. obliquus*
500 strains was about 1.0 g L^{-1} d^{-1} [89], and another report used bubbling column with 0.5%
501 CO_2 at 25°C resulting in about 0.6 g L^{-1} d^{-1} [90]. Both reported productivities for bubbling
502 columns were in our range of productivity results (Figure 2). Further, a comparison can
503 be found in a study looking at eight microalgae from a total of 33 isolated cultures of
504 water samples from freshwater rivers and livestock wastewater treatment plants at
505 Wunju, South Korea [91]. One of the highest performers was a strain of *S. obliquus*. In
506 that study, among multiple isolated *S. obliquus* strains, which were the same species
507 according to LSU rDNAD1-D2 regions, significant differences in lipid productivity
508 existed. Such different productivities within strains of a species from the same
509 geographical region demonstrate that surveying culture collections may not be sufficient,
510 as different strains of the same species may have different biomass or lipid productivities.
511 Similarly, showing that different strains of the same species can have widely varied
512 physiological parameters, some *S. obliquus* isolates have been found to have 3-4 times
513 higher carbon fixation efficiencies then other *S. obliquus* isolates [92]. In agreement with
514 these studies, our work also resulted in several independent strains of *S. obliquus*, that not

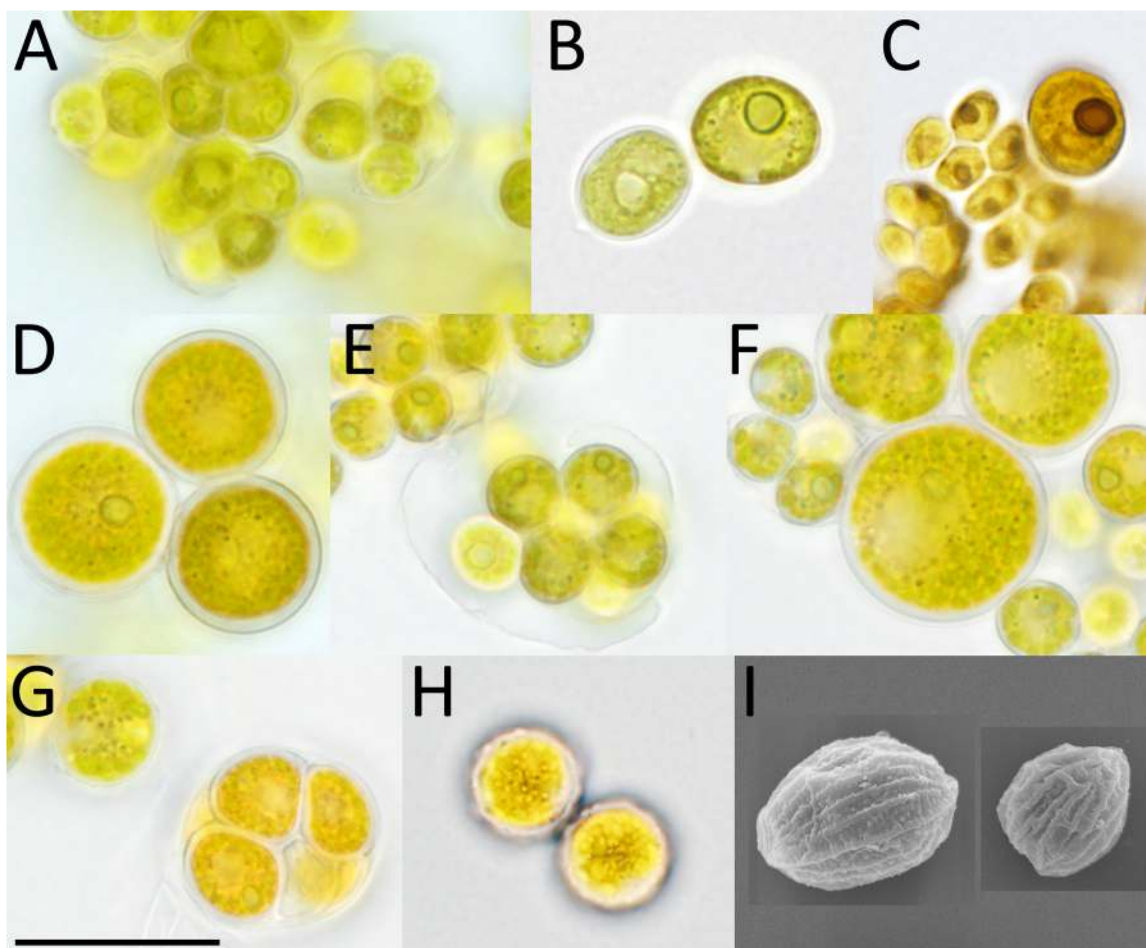
515 only grew well under laboratory conditions, but also outdoors even when compared to *N.*
516 *salina* [46].
517 Another group of strains (DOE0088, DOE0155, DOE0202, DOE0369, EN1234,
518 EN1235) nested closely together in a clade with *Coelastrella* sp. strain SAG217-5 [33,
519 93] being the sister group (Figure 5). To gain a better taxonomic overview, a second and
520 specific phylogenetic analysis was performed including all publically available ITS2
521 sequences of *Coelastrella* strains from the NCBI database. Our results showed that our
522 six novel strains fell into one clade together with other *Coelastrella* strains SAG2123 and
523 SAG217-5 (Figure 5).
524



525

526 Figure 5: Un-rooted ITS2 sequence tree including our novel strains DOE0088, DOE0155,
527 DOE0202, DOE0369, EN1234, EN1235, and also known strains of the Scenedesmaceae.
528 Bootstrap values are based on 500 replicates and mapped to the internal branches. Branch
529 lengths are drawn proportional to inferred nucleotide substitutions.
530

531



532 Figure 6: Light microscopic photos of cells of *Coelastrella* sp. strain DOE0202 (A-H). A)
533 Several autosporangia containing daughter cells. B) Side view of vegetative, oval,
534 coccoid cells with one cup-shaped chloroplast, which contains one pyrenoid. C) Cells
535 stained with iodine show a ring of starch indicated by black color on the outside of the
536 pyrenoid. Note the Scenedesmus-like coenobia of smaller cells that were released from
537 the autosporangia. D) Larger sized vegetative cells with the hyaline cell wall. E)
538 Daughter cells just released from an autosporangium. The hyaline cell wall of the mother
539 cell is visible in the background. F) A variety of vegetative cells in different growth
540 stages. G) An autosporangium is visible on the lower left side with cells that are in the
541 carotenogenic process. H) Center view of two vegetative carotenogenic cells where
542 protruberances are visible on the cell wall. I) Shown are two SEM images where striated
543 cells ridges are visible along the oval cells. In addition, some connections are visible
544 between ridges. The black bar represents 10 µm for the light microscopic and the electron
545 microscopic images. The SEM images were taken by Dr. Cooke at New Mexico State
546 University.

547

548 Based on microscopic investigation, we found that the new *Coelastrella* strains had a

549 broad range of coccoid cell morphologies. As one example, images of the strain

550 DOE0202 are presented in Figure 6. Flagellate cells were never observed regardless of
551 growth conditions. Reproduction occurs through division of autospores, which release
552 two to eight daughter cells by rupture of the parental cell wall. Depending on the growth
553 conditions, cells were either solitary or found in coenobia. Only sometimes, during the
554 exponential growth phase in batch cultures, coenobia were found that looked similar to
555 coenobia of *Scenedesmus* species (Figure 6C). Solitary cells were round to ellipsoid and
556 3-10 μm long. When elipsoid, cells often had polar thickenings. Cells contained one
557 chloroplast with one pyrenoid. In growing cultures, vegetative cells were green. Resting
558 cells from cultures in the stationary phase became first brown and then orange due to
559 accumulation of orange colored oil bodies in the cytosol (Figure 6). Regardless of growth
560 conditions, cell walls were hyaline. By light microscopy only sometimes cell walls had
561 ridges appearing as protruberances (Figure 6H), but additional electron microscopy
562 confirmed that all cells had ridges on their cell walls (Figure 6I). Nevertheless, there was
563 variability in the appearance of the ridges. Such variability in the cell wall appearance
564 and ridges is in agreement with the description of *Coelastrella* sp. strain SAG 217-5 [33],
565 which according to our recovered ITS2 phylogenetic trees (Figures 4 & 5) is a close
566 relative to our novel strains. Our novel strains cluster not only very closely together in the
567 phylogenetic analysis, but they also have identical cell morphologies and physiological
568 characteristics. Therefore, we assume that all our strains (DOE0088, DOE0155,
569 DOE0202, DOE0369, EN1234, and EN1235) belong to the same species. Currently, not
570 all species within the genus of *Coelastrella* are well defined and our strains fall into the
571 *sensu lato* group [93]. Consequently, we refrain from providing species names for our
572 novel strains, which will have to wait until the *Coelastrella* genus undergoes a closer

573 taxonomic reinvestigation.

574 Our newly isolated *Coelastrella* strains originated from different freshwater habitats
575 from a variety of locations within the United States (Table 1), indicating that this
576 *Coelastrella* species has a broad distribution within the United States. In their original
577 habitats such as birdbaths, fountains, and temporary waterbodies (roadside ditches), the
578 *Coelastrella* strains were often found co-existing with strains of the green alga
579 *Haematococcus pluvialis*. Similar to *H. pluvialis*, our new isolates were also
580 carotenogenic and formed spores that survived desiccation. To the best of our knowledge,
581 this is the first description of *Coelastrella* strains isolated from the United States. Thus,
582 all these *Coelastrella* isolates represent novel strains, possibly of a new species. Of note
583 is that all of these novel *Coelastrella* strains performed not only well in the laboratory,
584 but strain DOE0202 also tested positively in small raceway-type ponds [46], indicating
585 that – similar to *H. pluvialis* – the new *Coelastrella* strains might be employed as new
586 platform strains for biofuels and/or bioproducts generation.

587 Several of the well performing strains (DOE1418, DOE1357, DOE1051) appear to be
588 of the *Desmodesmus* genus. This was confirmed by the strains cell morphology (Figure
589 SM-2) as oval cells also exhibit the characteristic spines at the ends of their coenobia
590 [80]. For the genus of *Desmodesmus*, detailed species level studies comparing
591 morphology and ITS2 rDNA phylogenies are lacking to date [80], except for an account
592 on four closely related species bearing lateral spines [39]. Based on us finding multiple
593 well performing strains belonging to the *Desmodesmus* genus, we hypothesize that
594 species of this genus may be excellent platforms for biofuel production. However,
595 proving this hypothesis would require further large-scale testing in mass culture, which is

beyond the scope of this work.

The strain DOE0259 falls into one clade with *Ankistrodesmus* sp. SP2-15 (Figure 7), although quite some evolutionary distance exists between the two. Currently, only three other *Ankistrodemus* strains besides DOE0259 have their ITS2 region sequences in the NCBI database. All three strains showed similar far distances from our newly isolated strain, so only one was included in the tree shown in Figure 4. Extending a search in the NCBI database with the 3' end of the 18S, ITS1, 5.8S, and ITS2 sequences from DOE0259 (Genebank Accession Number KT274017) also identified this strain as an *Ankistrodesmus*, with it sharing the highest percent identity (99%) with *Ankistrodesmus* RS-2012, although with low query coverage (59%). Lack of exact matches for the 18S gene, the ITS1 spacer, the 5.8S gene, or the ITS1 spacer for strain DOE0259 suggests that it may represent a new species. Unfortunately, the genus of *Ankistrodesmus* is polyphyletic [94, 95, 96] as there has been difficulty discerning the different crescent-shaped morphological features of species, which are highly variable and challenging to identify by light microscopy. Figure 7 shows images of cells of the strain DOE0259. The crescent-shaped cells were always solitary. The basic cell morphology is coccoid with cells being elongated sickle-shaped having rounded ends (Figure 7). Due to the known uncertainty regarding the taxonomic position of species within the genus *Ankistrodesmus* and the family of the Selenastraceae in general [94, 95], any further classification of strain DOE0259 was beyond the scope of this manuscript.

617

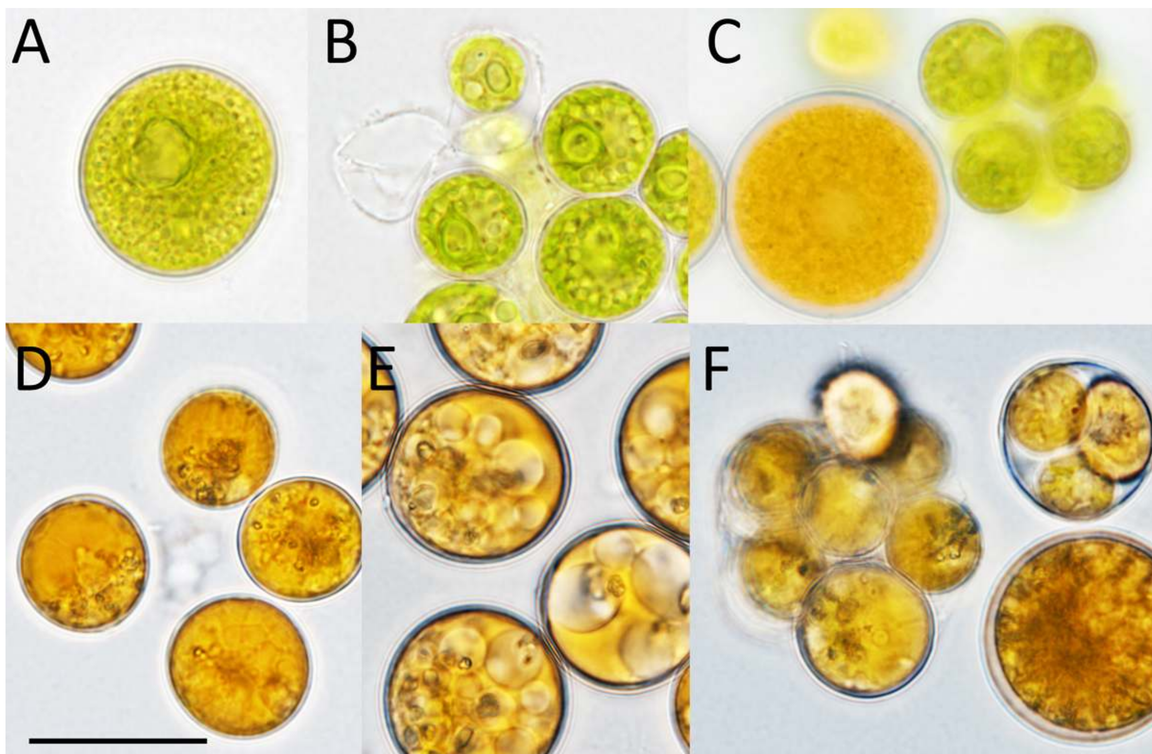
618 Figure 7: Shown are images of cells of the novel *Ankistrodesmus* strain DOE0259. Some
619 cells have oil bodies that are visible as blue roundish structures in the cytosol. The black
620 bar represents 10 µm. A) Two cells accumulating numerous oil bodies in the cytosol. B)
621 Images of several cells illustrating the plasticity in cell morphology. C) Single cell. D)
622 Two cells in the early division process. E) Two cells released from an autosporangium. F)
623 Four daughter cells still in the autosporangium. G) Four cells released from the
624 autosporangium.
625

626 *3.4.2. Chlamydomonadales (Class of the Chlorophyceae)*

627 Two taxa, strain DOE0101 and strain EN1423, fell into the Chlamydomonadales
628 order (Figure 4). Strain EN1423 was isolated from hyper-saline soil from a roadside ditch
629 close to the Laguna Del Perro in New Mexico, USA (Table 1). Originally the strain was
630 maintained in a 0.5 M NaCl containing artificial *Dunaliella* medium [97] and also tested
631 for biomass productivity in bubbling columns. However, in the saline medium large cell
632 clumps developed that stuck to the glass vessel, thus preventing accurate productivity

633 determination. Later the strain was transferred to BG11 freshwater medium [51] and
634 found to grow well in that medium.
635 As the ITS2 sequence alone could not provide an accurate species determination for
636 EN1423, the rDNA sequence containing the 3' end of the 18S gene, ITS1, 5.8S gene,
637 ITS2, and the 5' end of the 28S gene (Accession Number KT274016) was used for a
638 BLASTn search into the NCBI database. This search resulted in a hit with 100% identity
639 for our partial 18S gene sequence to the species *Borodinellopsis texensis* (NCBI
640 Accession number KM020129). In addition to the molecular marker sequence, the cell
641 morphology and physiology of strain EN1423 was investigated and found to match the
642 description of *B. texensis* (Figure SM-3) [98]. To our knowledge, as based on its lipid
643 over-accumulation characteristics, this is the first report of this species having potential
644 as a platform strain for biofuel generation.
645 While strain DOE0101 may not have been the very best biomass producer in the
646 laboratory (see section 3.2. above), this strain was highly carotenogenic (see section 3.3.
647 above) and it was successfully tested in small raceway-type ponds [46]. Classification of
648 strain DOE0101 to the species or genus level remained challenging. From our genome
649 sequences (unpublished) of the strain DOE0101, we located the 18S gene (NCBI
650 Accession number AJ249515), the 28S gene (NCBI Accession number KC145458), the
651 psaB gene (NCBI Accession number JN63055), and the rbcL sequence (NCBI Accession
652 number KC145509). For each of these four molecular markers, searches within the NCBI
653 database showed the best hits to be *Dysmorphococcus globosus* (SM 1), which is in the
654 Chlamydomonodales order, but it has no ITS2 sequence available. Although the sequence
655 identity was high, for DOE0101's psaB and rbcL genes, the query coverage with *D.*

656 *globosus* was only 85 and 86 %, respectively, indicating that strain DOE0101 is not a
657 close relative. Also, in contrast to the flagellate *D.* globosus, cells of strain DOE0101
658 were always coccoid (Figure 8). Therefore, though the ITS2 sequence-structure
659 phylogeny places DOE0101 within the Chlamydomonadales, even use of further
660 molecular marker sequence could not improve classification, because not enough
661 sequences are available from public databases.
662



663

664 Figure 8: Shown are images of cells of the strain DOE0101 indicating the variety of cell
665 types. The bar indicates 10 µm. A) Vegetative, coccoid cell from an actively growing
666 culure in the beginning stages of division of the pyrenoid. B) Several cells each
667 containing one cup-shaped chloroplast and one pyrenoid per chloroplast. C) One large
668 cell in the beginning phase of carotenogenesis and one autosporangium. D) Single cells
669 with orange oil bodies. E) Orange cells with large oil bodies and unknown white
670 appearing structures. F) Two autosporangia originating from orange cells.
671

672 In summary, while *Borodinellopsis texensis* strain EN1423 and the
673 Chlamydomonadales strain DOE0101 are in the same order as the species *H. pluvialis*, in
674 contrast to flagellate green cells of *H. pluvialis*, non-stressed green cells of the novel
675 strains were always coccoid under our growth conditions. Also in contrast to *H. pluvialis*,
676 which is well-known for application in mass production of the brick-red carotenoid
677 astaxanthin, cells of DOE0101 appear orange under stress (Figure 3 & 8) with the orange
678 pigments in DOE0101 having been identified as precursors of astaxanthin (Figure 3). As
679 the strains EN1423 and DOE0101 accumulate secondary carotenoids, both strains may
680 find further use in future applications for carotenoid production.
681
682 *3.4.3. Chlorellales (Class of the Trebouxiophyceae)*
683 Green algae with the spherical morphotype named *Chlorella* belonging to the class of
684 the Trebouxiophyceae have traditionally been used as model organisms for studies of
685 photosynthesis and biotechnological applications for decades [99]. As shown in Figure 4,
686 many of our top performing strains clustered within the class of the Trebouxiophyceae.
687 However, the taxonomic status of many of the new strains within the order of the
688 Chlorellales was not too well resolved when the rDNA ITS2 marker was used (Figure 4).
689 The best support for identification was obtained for strain DOE1135, which appears to be
690 closely related to *Chlorella vulgaris*. DOE0314 falls with *Dicloster acutus*, though a
691 significant distance exists between the two. The most internal branches of the other new
692 strains were also not well supported, but in the terminal parts of the phylogenetic tree
693 they clearly nested into highly supported clades. As the phylogeny based only on the
694 rDNA ITS 2 spacer region was not providing sufficient information regarding the

695 taxonomic position at the genus or even species level, an additional molecular

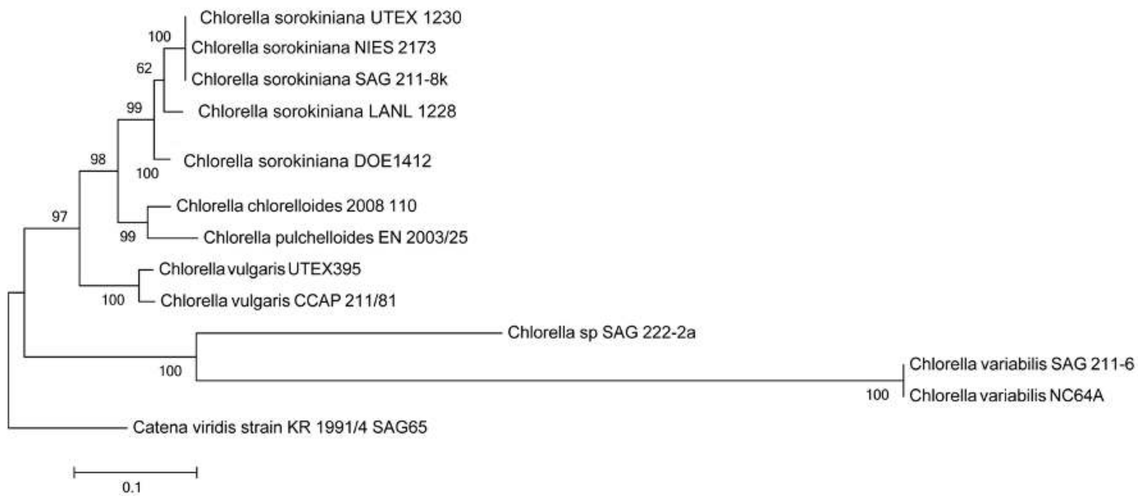
696 comparison was performed using the rDNA region including the 3' end of the 18s gene,

697 the ITS1 region, the 5.8s gene, the ITS2 region, and 5' end of the 28s gene [100]. The

698 resulting phylogenetic tree (Figure 9) allowed classification of the strain DOE1412 as *C.*

699 *sorokiniana*.

700



701

702 Figure 9: Phylogenetic tree based on the rDNA region including the 3' end of the 18s
703 gene, the ITS1 region, the 5.8s gene, the ITS2 region, and 5' end of the 28s gene. The
704 sequences for the strains *Chlorella sorokiniana* UTEX 1230 and *Chlorella sorokiniana*
705 LANL 1228 were kindly provided by Dr. Starkenburg.
706

707 In the past few years, the classification of coccoid taxa within the Chlorellaceae

708 including the genus of *Chlorella* has receive a lot of attention and many species were re-

709 classified [96, 99, 100, 101, 102]. Molecular and morphological characters together are

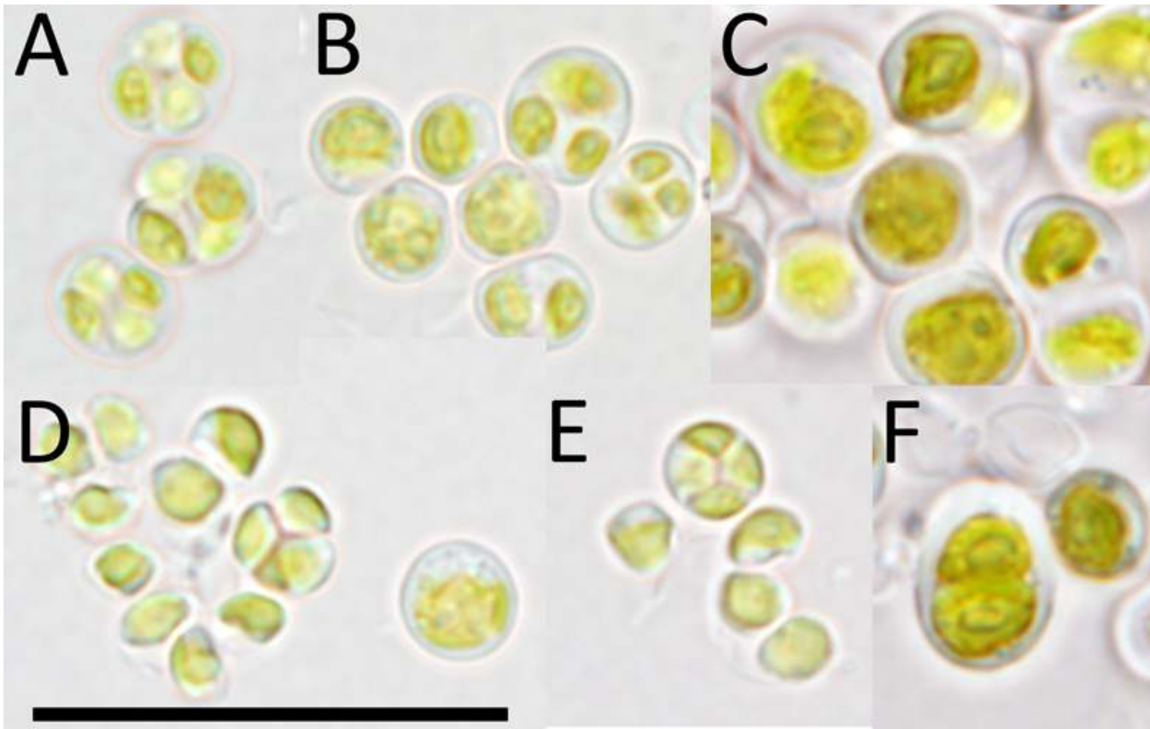
710 used for delineation of coccoid species [96, 100]. *C. sorokiniana* strain DOE1412 is of

711 unique interest, because of its high productivity, and is the subject of research in several

712 major laboratories. As a reference, we included images of cells of in Figure 10 below.

713

714



715 Figure 10: Shown are photos of cells of *Chlorella sorokiniana* strain DOE1412. The
716 black bar represents 10 µm. A) Autosporangia. B) Some autosporangia and several cells
717 in the early stages of cell division. C) Several larger cells with the pyrenoid clearly
718 visible. D) Several autosporangia and one larger cell. One autosporangium and an opened
719 autosporangium with released daughter cells. F) Larger cells with the pyrenoids clearly
720 visible within the chloroplast.
721

722 Although strain DOE1412 is not a representative of a new species and concern about
723 redundancies regarding already existing strains in culture collections may exist, the first
724 analysis of a draft genome of strain DOE1412 revealed that its genome contains
725 significant differences to two other *C. sorokiniana* strains UTEX1230 and LANL1228
726 (Starkenburg, personal communication). Therefore, it can be concluded that the new
727 isolate strain DOE1412 greatly expands the germ plasma of the species *C. sorokiniana*.
728 Comparative genomic work is ongoing and the analysis of the results will be published
729 separately.

730 In addition to determining its biomass productivity, the fatty acid profile for *C.*

731 *sorokiniana* DOE1412 was obtained and analyzed for a culture grown in a
732 photobioreactor (Figure SM-6). Cells of *C. sorokiniana* had C16:0 and C18:1 fatty acids
733 as the main constituents, which is in line with a previous report about *C. sorokiniana*
734 strains having long-straight chain alkanes and fatty alcohols being major compounds of
735 extracts [103].

736 Within the Chlorellales clade, *C. sorokiniana* DOE01412 proved to be one of the best
737 growers of the first strains isolated during the NAABB project. It is of note that *C.*
738 *sorokiniana* strain DOE1412 was found to grow well in LED-lighted 800L indoor
739 raceway ponds [104]. When tested outdoors, it grew very well, significantly better then
740 even gold standard *N. salina*. In 23,000 L NAABB outdoor raceway ponds with paddle
741 wheels, it had a maximum productivity of 30g/m^2/day, tolerating temperatures from 40°-
742 110°F, withstanding a range of salinities, and accumulating about 25% lipids. The strain
743 also faired well in economical models, and oil from it was successfully converted to jet
744 fuel. Consequently, *C. sorokiniana* strain DOE1412 was brought forward as a top
745 producer, and its discovery was cited as a major deliverable for the NAABB project [46].

746 Like *Scenedesmus*, *Chlorella* strains have also been of interest as biofuel feedstocks.
747 For example, *C. sorokiniana* is regarded as one of the most promising algae feedstocks,
748 because the algae can grow in autotrophic, heterotrophic, and mixotrophic conditions
749 [105]. Our new strains offer an expansion of the genetic resources available to create
750 even more productive algae from this clade.

751

752 **4. Conclusions**

753

754 This work characterized and classified the top algal strains to come out of a multi-
755 year screening effort to find algal strains suitable for further development as biofuel
756 feedstocks. Thirty of the best performing strains were deposited with the UTEX algal
757 culture collection. The project successfully isolated a variety of strains, which showed
758 potential as future platform strains. Several novel strains belonging to the *Coelastrella*
759 genus were found, which are reported on for the first time for the United States. In
760 addition, one new strain designated as DOE0101 was from the Chlamydomonadales.
761 Both, the Chlamydomonadales strain DOE0101 and all the *Coelastrella* strains,
762 accumulate carotenoids and may become platform strains for future carotenoid
763 production. That these novel strains appear previously uncharacterized in a biofuel
764 context despite their high growth rates and carotenoid-rich profiles demonstrates the
765 untapped diversity of the green microalgae. In addition, several new strains of species
766 already known to be fast growers were isolated, thus validating our overall phyco-
767 prospecting approach. These strains of known species such as *S. obliquus*, now offer an
768 expansion of the genomic resources available for these clades. Lastly, the discovery of *C.*
769 *sorokiniana* strain DOE1412 is particularly noteworthy as it is currently one of the most
770 promising algal species being developed for biofuels applications. Understanding the
771 evolution of the highly productive strains isolated in this study and characterizing their
772 metabolic and growth traits is a key step in bringing forth cultivars of microalgae for
773 biofuel, and high-value compound, production.
774
775 **Contributions**

J.P. designed and oversaw the entire isolation and screening project. P.N., J.P., and A.H. conceived and wrote the paper. A.H., W.C, and F.J. conducted the screening, P.N., J.P., S.T., and F.J., conducted the lipid analysis, and P.N., J.P., K.S., and Q.W. conducted the phylogenetic analysis. Light microscopy for many strains was performed by A.G. and the Transmission Electron Micrograph for *Coelastrella* strain DOE0202 was provided by O.H. The fatty acid profile for *C. sorokiniana* strain DOE1412 was provided by S.T. Together, J.P. (jpolle@brooklyn.cuny.edu) and P.N. (pneofotis@gc.cuny.edu) declare the integrity of this work as a whole.

Acknowledgements

The authors thank the US Department of Energy for funding under the grant #DE-EE0003129 and #DE-EE0003046-28302B. The authors also gratefully acknowledge support from Airforce Office of Scientific Research under grants number grant #FA9550-08-1-0170 and #FA9550-08-1-0403. The authors greatly appreciate support by their collaborators within the National Alliance for Advanced Biofuels and Bioproducts, specifically Dr. J. Olivares and Dr. R. Sayre. The authors would also like to thank Dr. J. Nishiura for his aid in florescent microscopy as well as for his discussions. The authors would also like to thank the members of the boyscout troop 1949 in Katy, TX. The authors thank Dr. Mahendra-Perumal and Ms. S. Registe for their technical support on strain isolation and screening. We thank Ms. K. Laje for technical assistance with sample preparation for the SEM analysis.

References

799 [1] Markou, G., & Nerantziz, E. 2013. Microalgae for high-value compounds and
800 biofuels production: A review with focus on cultivation under stress
801 conditions. *Biotechnology Advances* **31**:1532-42.
802 [2] National Research Council of the National Academies 2012. Sustainable
803 Development of Algal Biofuels in the United States*.* The National Academies
804 Press, Washington, D.C., USA.
805 [3] Wijffels, R.H., & Barbosa, M.J. 2010. An Outlook on Microalgal Biofuels. *Science*
806 **329**:796-799.
807 [4] Wilkie, A. C., Edmundson, S. J. & Duncan, J. G. 2011. Indigenous algae for local
808 bioresource production: Phycoprospecting. *Energy for Sustainable*
809 *Development* **15**:365-71.
810 [5] Barclay, W. & Apt, K. 2013. Strategies for Bioprospecting Microalgae for Potential
811 Commercial Applications. *In*: Richmond, R. & Hu, Q. [Eds.] *Handbook of*
812 *Microalgal Culture: Applied Phycology and Biotechnology.* Wiley-Blackwell, pp.
813 69-79.
814 [6] Georgianna, D. R. & Mayfield, S. P. 2012. Exploiting diversity and synthetic
815 biology for the production of algal biofuels. *Nature* **488**:329-35.
816 [7] Mutanda, T., Ramesh, D., Karthikeyan, S., Kumari, S., Anandraj, A. & Bux, F. 2011.
817 Bioprospecting for hyper-lipid producing microalgal strains for sustainable
818 biofuel production. *Bioresource Technology* **102**:57-70.
819 [8] Sheehan, J., Dunahay, T., Benemann, J. & Roessler, P. 1998. A look back at the US
820 Department of Energy's aquatic species program - Biodiesel from Algae. *Report*
821 *NREL/TP-580-24190.* National Renewable Energy Laboratory, Golden, CO, USA.
822 [9] U.S. Department of Energy 2010. U.S. DOE Roadmap: National Algal Biofuels
823 Technology Roadmap. Fishman, D., Majumdar, R., Morello, J., Pate, R. & Y., Y.
824 [Eds.]. Office of Energy Efficiency and Renewable Energy Biomass Program.
825 [10] Borowitzka, M. A. 2013. High-value products from microalgae-their
826 development and commercialisation. *Journal of Applied Phycology* **25**:743-56.
827 [11] Mulders, K.J.M., Lamers, P.P., Martens, D.E., & Wijffels, R.H. 2014. Phototrophic
828 pigment production with microalgae: Biological constraints and opportunties.
829 *Journal of Phycology* **50**:229-242.
830 [12] Olaizola, M. 2003. Commercial development of microalgal biotechnology: from
831 the test tube to the marketplace. *Biomolecular Engineering* **20**:459-66.
832 [13] Wijffels, R. H., Kruse, O. & Hellingwerf, K. J. 2013. Potential of industrial
833 biotechnology with cyanobacteria and eukaryotic microalgae. *Current Opinion*
834 *in Biotechnology* **24**:405-13.
835 [14] Scaife M.A., Nguyen G.T.D.T., Rico J., Lambert D., Helliwell K.E., Smith A.G. 2015.
836 Establishing *Chlamydomonas reinhardtii* as an industrial biotechnology host.
837 *Plant Journal* **82** (3):532-546.
838 [15] Gimpel, J. A., Specht, E. A., Georgianna, D. R., & Mayfield, S. P. 2013. Advances in
839 microalgae engineering and synthetic biology applications for biofuel
840 production. *Current Opinion in Chemical Biology* **17**:489-95.
841 [16] Del Campo, J. A., Garcia-Conzalez, M. & Guerrero, M. G. 2007. Outdoor
842 cultivation of microalgae for fuel production. *Applied Energy* **88**:3524-31.
843 [17] Pulz, O. & Gross, W. 2004. Valuable products from biotechnology of microalgae.
844 *Applied Microbiology and Biotechnology* **65**:635-48.

845 [18] Vanthoor-Koopmans, M., Wijffels, R.H., Barbosa, M.J., Eppink, M.H.M. 2013.
846 Biorefinery of microalgae for food and fuel. *Bioresource Technology* **135**:142-
847 149.
848 [19] Gershwin, M. E. & Belay, A. B. 2008. *Spirulina in human nutrition and health.*
849 CRC Press, Boca Raton, USA.
850 [20] Guedes, A. C., Amaro, H. M. & Malcata, F. X. 2011. Microalgae as sources of high
851 added-value compounds - a brief review of recent work. *Biotechnology*
852 *Progress* **27**:597-613.
853 [21] Guerin, M., Huntley, M. E. & Olaizola, M. 2003. *Haematococcus* astaxanthin:
854 applications for human health and nutrition. *Trends in Biotechnology* **21**:210-
855 16.
856 [22] Yuan J.-P., Peng J., Yin K., Wang J.-H. 2011. Potential health-promoting effects of
857 astaxanthin: A high-value carotenoid mostly from microalgae. *Molecular*
858 *Nutrition and Food Research* **55** (1):150-165..
859 [23] Borowitzka, L. J. 1991. Development of western biotechnology algal beta-
860 carotene plant. *Bioresource Technology* **38**:251-52.
861 [24] Ben-Amotz, A., Polle, J. E. W. & Rao, D. V. S. 2009. *The Alga Dunaliella:*
862 *Biodiversity, Physiology, Genomics and Biotechnology* Science Publishers
863 Enfield, New Hampshire, USA
864 [25] Elliott, L. G., Feehan, C., Laurens, L. M. L., Pienkos, P. T., Darzins, A. & Posewitz,
865 M. C. 2012. Establishment of a bioenergy-focused microalgal culture collection.
866 *Algal Research* **1**:102-13.
867 [26] Ratha, S. K. & Prasanna, R. 2012. Bioprospecting microalgae as potential
868 sources of "Green Energy" - challenges and perspectives. *Applied Biochemistry*
869 *and Microbiology* **48**:109-25.
870 [27] Ratha, S. K., Prasanna, R., Gupta, V., Dhar, D. W. & Saxena, A. K. 2012.
871 Bioprospecting and indexing the microalgal diversity of different ecological
872 habitats of India. *World Journal of Microbiology & Biotechnology* **28**:1657-67.
873 [28] Draaisma, R. B., Wijffels, R. H., Slegers, P. M., Brentner, L. B., Roy, A. & Barbosa,
874 M. J. 2013. Food commodities from microalgae. *Current Opinion in*
875 *Biotechnology* **24**:169-77.
876 [29] Clerck, O.D., Guiry, M.D., Leliaert, F., Samyn, Y., Verbruggen, H. 2013. Algal
877 Taxonomy: A Road to Nowhere? *Journal of Phycology* **49**:215-225.
878 [30] John, D. M. 1994. Biodiversity and conservation; algal perspective. *The*
879 *Phycologist* **38**:3-21.
880 [31] Norton, T. A., Melkonian, M. & Andersen, R. A. 1996. Algal biodiversity.
881 *Phycologia* **35**:308-26.
882 [32] Guiry, M.D., & Guiry, G.M. 2012. AlgaeBase. World-wide electronic publication.
883 Available at http://www.algaebase.org.
884 [33] Hegewald, E., Wolf, M., Keller, A., Friedl, T. & Krienitz, L. 2010. ITS2 sequence-
885 structure phylogeny in the Scenedesmaceae with special reference to
886 *Coelastrum* (Chlorophyta, Chlorophyceae), including the new genera
887 *Comasiella* and *Pectinodesmus*. *Phycologia* **49**:325-35.
888 [34] Hegewald, E., Bock, C. & Krienitz, L. 2013. A phylogenetic study on
889 Scenedesmaceae with the description of a new species of *Pectinodesmus* and

890 the new genera *Verrucodesmus* and *Chodatodesmus* (Chorophyta,
891 Chlorophyceae). *Fottea* **13**:149-64.
892 [35] An, S. S., Friedl, T. & Hegewald, E. 1999. Phylogenetic relationships of
893 *Scenedesmus* and *Scenedesmus*-like coccoid green algae as inferred from ITS-2
894 rDNA sequence comparisons. *Plant Biology* **1**:418-28.
895 [36] van Hannen, E. J., Fink, P. & Lurling, M. 2002. A revised secondary structure
896 model for the internal transcribed spacer 2 of the green algae *Scenedesmus* and
897 *Desmodesmus* and its implication for the phylogeny of these algae. *European*
898 *Journal of Phycology* **37**:203-08.
899 [37] Coleman, A. W. 2003. ITS2 is a double-edged tool for eukaryote evolutionary
900 comparisons. *Trends in Genetics* **19**:370-75.
901 [38] Coleman, A. W. 2009. Is there a molecular key to the level of "biological species"
902 in eukaryotes? A DNA guide. *Molecular Phylogenetics and Evolution* **50**:197-
903 203.
904 [39] Hegewald, E., Schmidt, A., Braband, A. & Tsarenko, P. 2005. Revision of the
905 *Desmodesmus* (Sphaeropleales, Scenedesmaceae) species with lateral spines. 2.
906 The multi-spined to spineless taxa. *Archiv fur Hydrobiologie/Supp* ***157***
907 Algological Studies:1-38.
908 [40] Jeon, S. L. & Hegewald, E. 2006. A revision of the species *Desmodesmus*
909 *perforatus* and D tropicus (Scenedesmaceae, Chlorophyceae, Chlorophyta).
910 *Phycologia* **45**:567-84.
911 [41] Schultz, J., Muller, T., Achtziger, M., Seibel, P. N., Dandekar, T. & Wolf, M. 2006.
912 The internal transcribed spacer 2 database - a web server for (not only) low
913 level phylogenetic analyses. *Nucleic Acids Research* **34**:W704-W07.
914 [42] Schultz, J. & Wolf, M. 2009. ITS2 sequence-structure analysis in phylogenetics: A
915 how-to manual for molecular systematics. *Molecular Phylogenetics and*
916 *Evolution* **52**:520-23.
917 [43] Coleman, A.W. 2007. Pan-eukaryote ITS2 homologies revealed by RNA
918 secondary structure. *Nuclei Acids Research* **25**:3322-3329.
919 [44] Mai, J.C. and Coleman, A.W. 1997. The internal transcribed spacer 2 exhibits a
920 common secondary structure in green algae and flowering plants. *Journal of*
921 *Molecular Evolution* **44**:258-271.
922 [45] Schultz, J., Maisel, S., Gerlach, D., Muller, T. & Wolf, M. 2005. A common core of
923 secondary structure of the internal transcribed spacer 2 (ITS2) throughout the
924 Eukaryota. *RNA - a Publication of the RNA Society* **11**:361-64.
925 [46] NAABB Final Report 2014. National Alliance for Advanced Biofuels and
926 Bioproducts Synopsis (NAABB) Final Report.
927 http://www.energy.gov/eere/bioenergy/downloads/national-alliance-
928 advanced-biofuels-and-bioproducts-synopsis-naabb-final.
929 [47] Neofotis, P., Huang, A., Chang, W. H., Joseph, F. & Polle, J. E. W. 2015. Microalgae
930 strain isolation, screening, and identification for biofuels and high value
931 products *In*: Slocombe, S. P. & Benemann, J. R. [Eds.] *Micro-Algal Production for*
932 *Biomass and High-Value Products.* CRC Press, Taylor and Francis, LLC.
933 [48] Andersen, R. A. & Kawachi, M. 2005. Chapter 6: Traditional microalgae isolation
934 and techniques *In*: Anderson, R. A. [Ed.] *Algae Culturing Techniques* Elsevier
935 Academic Press & The Phycological Society of America Amsterdam

936 [49] Sieracki, M. E., Poulton, N. & Crosbie, N. 2005. Automated isolation techniques
937 for microalgae *In*: Anderson, R. A. [Ed.] *Algal Culturing Techniques* Elsevier
938 Academic Press Amsterdam.
939 [50] Dolezel, J., Greilhuber, J., Suda, J. 2007. *Flow Cytometry with Plant Cells: Analysis*
940 *of Genes, Chromosomes, and Genomes.* Wiley-VCH, Germany.
941 [51] Andersen, R. A. 2005. *Algal Culturing Techniques.* Elsevier Academic Press,
942 Amsterdam.
943 [52] Bigogno, C., Khozin-Goldberg, I., Boussiba, S., Vonshak, A. & Cohen, Z. 2002.
944 Lipid and fatty acid composition of the green oleaginous alga *Pariotochloris*
945 *incises* the richest plant source of arachidonic acid. *Phytochemistry* **60**:497-503.
946 [53] Bligh, E. G. & Dyer, W. J. 1959. A rapid method of total lipid extraction and
947 purification *Canadian Journal of Biochemistry and Physiology* **37**:911-17.
948 [54] Unkefer, C. J., Starkenburg, S. R., Twary, S. N., Park, M. S., Chain, P. S. & Unkefer,
949 P. J. 2011. The National Alliance for Advanced Biofuels and Bioproducts
950 (NAABB). *Abstracts of Papers of the American Chemical Society* **242**.
951 [55] Quinn, J. C., Yates, T., Douglas, N., Weyer, K., Butler, J., Bradley, T. H. & Lammers,
952 P. J. 2012. Nannochloropsis production metrics in a scalable outdoor
953 photobioreactor for commercial applications. *Bioresource Technology*
954 **117**:164-71.
955 [56] White, T. J., Bruns, T., Lee, S. J. & Taylor, J. 1990. Amplification and direct
956 sequencing of fungal ribosomal RNA genes for phylogenetics. *In*: Innis, M. A.,
957 Gelfand, D. H., Sninsky, J. J. & White, T. J. [Eds.] *PCR Protocols.* Academic Press,
958 San Diego, pp. 315-24.
959 [57] Keller, A., Schleicher, T., Schultz, J., Muller, T., Dandekar, T. & Wolf, M. 2009.
960 5.8S-28S rRNA interaction and HMM-based ITS2 annotation. *Gene* **430**:50-57.
961 [58] Kessler, E., Schafer, M., hummer, C., Kloboucek, A. & Huss, V. A. R. 1997.
962 Physiological, biochemical, and molecular characters for the taxonomy of the
963 subgenera of *Scenedesmus* (Chlorococcales, Chlorophyta). *Botanica Acta*
964 **110**:244-50.
965 [59] Worden, A. Z., Lee, J. H., Mock, T., Rouze, P., Simmons, M. P., Aerts, A. L., Allen, A.
966 E., Cuvelier, M. L., Derelle, E., Everett, M. V., Foulon, E., Grimwood, J., Gundlach,
967 H., Henrissat, B., Napoli, C., McDonald, S. M., Parker, M. S., Rombauts, S.,
968 Salamov, A., Von Dassow, P., Badger, J. H., Coutinho, P. M., Demir, E., Dubchak, I.,
969 Gentemann, C., Eikrem, W., Gready, J. E., John, U., Lanier, W., Lindquist, E. A.,
970 Lucas, S., Mayer, K. F. X., Moreau, H., Not, F., Otillar, R., Panaud, O., Pangilinan, J.,
971 Paulsen, I., Piegu, B., Poliakov, A., Robbens, S., Schmutz, J., Toulza, E., Wyss, T.,
972 Zelensky, A., Zhou, K., Armbrust, E. V., Bhattacharya, D., Goodenough, U. W., Van
973 de Peer, Y. & Grigoriev, I. V. 2009. Green Evolution and Dynamic Adaptations
974 Revealed by Genomes of the Marine Picoeukaryotes *Micromonas*. *Science*
975 **324**:268-72.
976 [60] Seibel, P. N., Muller, T., Dandekar, T., Schultz, J. & Wolf, M. 2006. 4SALE - A tool
977 for synchronous RNA sequence and secondary structure alignment and editing.
978 *BMC Bioinformatics* **7**:498.
979 [61] Seibel, P. N., Muller, T., Dandekar, T. & Wolf, M. 2008. Synchronous visual
980 analysis and editing of RNA sequence and secondary structure alignments
981 using 4SALE. *BMC Research Notes* **1**:91.

982 [62] Friedrich, J., Dandekar, T., Wolf, M. & Muller, T. 2005. ProfDist: a tool for the
983 construction of large phylogenetic trees based on profile distances.
984 *Bioinformatics* **21**:2108-09.
985 [63] Wolf, M., Ruderisch, B., Dandekar, T., Schultz, J. & Muller, T. 2008. ProfDistS:
986 (profile-) distance based phylogeny on sequence-structure alignments.
987 *Bioinformatics* **24**:2401-02.
988 [64] Paradis, E. 2006. *Analysis of Phylogenetics and Evolution with R.* Springer, New
989 York, NY, USA
990 [65] Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. 2008.
991 GenBank. *Nucleic Acids Research* **33**:D25-D30.
992 [66] Wolf, M., Achtziger, M., Schultz, J., Dandekar, T. & Muller, T. 2005. Homology
993 modeling revealed more than 20,000 rRNA internal transcribed spacer 2
994 (ITS2) secondary structures. *RNA - A Publication of the RNA Society* **11**:1616-
995 23.
996 [67] Selig, C., Wolf, M., Muller, T., Dandekar, T. & Schultz, J. 2008. The ITS2 Database
997 II: homology modelling RNA structure for molecular systematics. *Nucleic Acids*
998 *Research* **36**:D377-D80.
999 [68] Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M. & Turner, D.
1000 H. 2004. Incorporating chemical modification constraints into a dynamic
1001 programming algorithm for prediction of RNA secondary structure.
1002 *Proceedings of the National Academy of Sciences of the United States of America*
1003 **101**:7287-92.
1004 [69] Zhang R., Patena W., Armbruster U., Gang S.S., Blum S.R., Jonikas M.C. 2014.
1005 High-throughput genotyping of green algal mutants reveals random
1006 distribution of mutagenic insertion sites and endonucleolytic cleavage of
1007 transforming DNA. *Plant Cell* **26** (4):1398-1409.
1008 [70] Griffiths, M. J. & Harrison, S. T. L. 2009. Lipid productivity as a key characteristic
1009 for choosing algal species for biodiesel production. *Journal of Applied Phycology*
1010 **21**:493-507.
1011 [71] Zhu, Y., Albrecht, K. O., Elliott, D. C., Hallen, R. T. & Jones, S. B. 2013.
1012 Development of hydrothermal liquefaction and upgrading technologies for
1013 lipid-extracted algae conversion to liquid fuels. *Algal Research* **2**:455-64.
1014 [72] Barreiro, D. L., Prins, W., Ronsse, F. & Brilman, W. 2013. Hydrothermal
1015 liquefaction (HTL) of microalgae for biofuel production: State of the art review
1016 and future prospects. *Biomass & Bioenergy* **53**:113-27.
1017 [73] Stephens, E., Ross, I. L. & Hankamer, B. 2013. Expanding the microalgal industry
1018 - continuing controversy or compelling case? *Current Opinion in Chemical*
1019 *Biology* **17**:444-52.
1020 [74] Elliott D.C., Biller P., Ross A.B., Schmidt A.J., Jones S.B. 2015. Hydrothermal
1021 liquefaction of biomass: Developments from batch to continuous process.
1022 *Bioresource Technology* **178**:147-156.
1023 [75] Boussiba S. 2000. Carotenogenesis in the green alga *Haematococcus pluvialis*:
1024 Cellular physiology and stress response. Physiologia Plantarum, **108** (2):111-
1025 117.

1026 [76] Jin E., Lee C.-G., Polle J.E.W. 2006. Secondary carotenoid accumulation in
1027 *Haematococcus* (chlorophyceae): Biosynthesis, regulation, and biotechnology.
1028 Journal of Microbiology and Biotechnology **16** (6):821-831.
1029 [77] Byrdwell, W. C. 2005. [Ed.] Modern Methods for Lipid Analysis by Liquid
1030 Chromatography/ Mass Spectrometry and Related Techniques. AOCS
1031 Publishing.
1032 [78] Buchheim, M.A., Sutherland, D.M., Schleicher, T., Forster, F., Wolf, M. 2012.
1033 Phylogeny of Oedogoniales, Chaetophorales and Chaetopeltidales
1034 (Chlorophyceae): inferences from sequence-structure analysis of ITS2. *Annals*
1035 *of Botany* 109:109-116.
1036 [79] Hoshina, R. 2014. DNA analyses of a private collection of microbial green algae
1037 contribute to a better understanding of microbial diversity. *BMC Research*
1038 *Notes* **7** (1):592
1039 [80] Vanormelingen, P., Hegewald, E., Braband, A., Kitschke, M., Friedl, T., Sabbe, K. &
1040 Vyverman, W. 2007. The systematics of a small spineless *Desmodesmus* species,
1041 D-costato-granulatus (Sphaeropleales, Chlorophyceae), based on ITS2 rDNA
1042 sequence analyses and cell wall morphology. *Journal of Phycology* **43**:378-96.
1043 [81] Leliaert, F., Smith, D. R., Moreau, H., Herron, M. D., Verbruggen, H., Delwiche, C.
1044 F. & De Clerck, O. 2012. Phylogeny and Molecular Evolution of the Green Algae.
1045 *Critical Reviews in Plant Sciences* **31**:1-46.
1046 [82] Eliáš, M., Němcová, Y., Škaloud, P., Neustupa, J., Kaufnerová, V. & Šejnohová, L.
1047 2010. Hylodesmus singaporensis gen. et sp. nov., a new autosporic subaerial
1048 green alga(Scenedesmaceae, Chlorophyta) from Singapore. *Intern. J. System.*
1049 *Evol. Microbiol* **60**:1224-35.
1050 [83] Buchheim, M. A., Keller, A., Koetschan, C., Forster, F., Merget, B. & Wolf, M. 2011.
1051 Internal Transcribed Spacer 2 (nu ITS2 rRNA) Sequence-Structure
1052 Phylogenetics: Towards an Automated Reconstruction of the Green Algal Tree
1053 of Life. *Plos One* **6**(2):e16931.
1054 [84] Becker, E. W. 1984. Biotechnology and Exploitation of the Green Alga
1055 *Scenedesmus obliquus* in India *Biomass* **4**:1-19.
1056 [85] Grobbelaar, J. U., Soeder, C. J. & Stengel, E. 1990. Modeling algal productivity in
1057 large outdoor cultures and waste treatment systems. *Biomass* **21**:297-314.
1058 [86] Breuer G., Lamers P.P., Martens D.E., Draaisma R.B., Wijffels R.H. 2012. The
1059 impact of nitrogen starvation on the dynamics of triacylglycerol accumulation
1060 in nine microalgae strains. *Bioresource Technology* **124**:217-226.
1061 [87] Li, X., Hu, H. Y. & Zhang, Y. P. 2011. Growth and lipid accumulation properties of
1062 a freshwater microalga *Scenedesmus sp* under different cultivation
1063 temperature. *Bioresource Technology* **102**:3098-102.
1064 [88] Tang, D. H., Han, W., Li, P. L., Miao, X. L. & Zhong, J. J. 2011. CO(2) biofixation and
1065 fatty acid composition of *Scenedesmus obliquus* and *Chlorella pyrenoidosa* in
1066 response to different CO(2) levels. *Bioresource Technology* **102**:3071-76.
1067 [89] Rodolfi, L., Zittelli, G. C., Bassi, N., Padovani, G., Biondi, N., Bonini, G. & Tredici, M.
1068 R. 2009. Microalgae for Oil: Strain Selection, Induction of Lipid Synthesis and
1069 Outdoor Mass Cultivation in a Low-Cost Photobioreactor. *Biotechnology and*
1070 *Bioengineering* **102**:100-12.

1071 [90] Feng et al. 2014 Growth and lipid accumulation characteristics of *Scenedesmus*
1072 *obliquus* in semi-continuous cultivation outdoors for biodiesel feedstock
1073 production. *Bioresource Technology* **173**:406–414
1074 [91] Abou-Shanab, R. A. I., Hwang, J. H., Cho, Y., Min, B. & Jeon, B. H. 2011.
1075 Characterization of microalgal species isolated from fresh water bodies as a
1076 potential source for biodiesel production. *Applied Energy* **88**:3300-06.
1077 [92] Ho, S. H., Chen, C. Y., Yeh, K. L., Chen, W. M., Lin, C. Y. & Chang, J. S. 2010.
1078 Characterization of photosynthetic carbon dioxide fixation ability of
1079 indigenous *Scenedesmus obliquus* isolates. *Biochemical Engineering Journal*
1080 **53**:57-62.
1081 [93] Kaufnerová, V., Eliáš, M. 2013. The demise of the genus *Scotiellopsis* Vinatzer
1082 (Chlorophyta). *Nova Hedwigia* **97** (3-4):415-428.
1083 [94] Krienitz L., Ustinova I., Friedl T., Huss V.A.R. 2001. Traditional generic concepts
1084 versus 18S rRNA gene phylogeny in the green algal family Selenastraceae
1085 (Chlorophyceae, Chlorophyta). *Journal of Phycology* **37** (5):852-865.
1086 [95] Krienitz L., Bock C., Nozaki H., Wolf M. 2011. Ssu rRNA gene phylogeny of
1087 morphospecies affiliated to the bioassay alga "*Selenastrum capricornutum*"
1088 recovered the polyphyletic origin of crescent-shaped chlorophyta. *Journal of*
1089 *Phycology* **47** (4):880-893.
1090 [96] Krienitz L., Bock C. 2012. Present state of the systematics of planktonic coccoid
1091 green algae of inland waters. *Hydrobiologia* **698** (1):295-326.
1092 [97] Polle J.E.W., Struwe L., Jin E. 2008. Identification and characterization of a new
1093 strain of the unicellular green alga *Dunaliella salina* (Teod.) from Korea.
1094 *Journal of Microbiology and Biotechnology* **18** (5):821-827.
1095 [98] Dykstra R.F. 1971. *Borodinellopsis texensis*: Gen. et Sp. Nov. A new alga from the
1096 Texas gulf coast. In: Bruce C. Parker and R. Malcolm Brown, Jr. [Eds.]
1097 *Contributions in Phycology,* Students of Harold C. Bold, pp.1-8
1098 [99] Krienitz, L., Huss, V.A.R., Bock, C. 2015. *Chlorella*: 125 years of the green
1099 survivalist. *Trends in Plant Science* **20** (2):67-69.
1100 [100] Bock C., Krienitz L., Proschold T. 2011. Taxonomic reassessment of the genus
1101 *Chlorella* (Trebouxiophyceae) using molecular signatures (barcodes), including
1102 description of seven new species. *Fottea* **11** (2):293-312.
1103 [101] Lemieux C., Otis C., Turmel M. 2014. Chloroplast phylogenomic analysis
1104 resolves deep-level relationships within the green algal class
1105 Trebouxiophyceae. *BMC Evolutionary Biology* **14** (1):211.
1106 [102] Luo W., Proschold T., Bock C., Krienitz L. 2010. Generic concept in *Chlorella*-
1107 related coccoid green algae (Chlorophyta, Trebouxiophyceae). *Plant Biology* **12**
1108 (3): 545-553.
1109 [103] Yang, X. W., Dai, X., Guo, H., Geng, S. & Wang, G. Y. 2013. Petrodiesel-like
1110 straight chain alkane and fatty alcohol production by the Microalga *Chlorella*
1111 *sorokiniana*. *Bioresource Technology* **136**:126-30.
1112 [104] Huesemann, M. H., Van Wagenen, J., Miller, T., Chavis, A., Hobbs, S. & Crowe, B.
1113 2013. A screening model to predict microalgae biomass growth in
1114 photobioreactors and raceway ponds. *Biotechnology and Bioengineering*
1115 **110**:1583-94.

1116
1117
1118
1119
1120
1121
1122
1123
1124

[105] Kim, S., Park, J. E., Cho, Y. B. & Hwang, S. J. 2013. Growth rate, organic carbon and nutrient removal rates of *Chlorella sorokiniana* in autotrophic, heterotrophic and mixotrophic conditions. *Bioresource Technology* **144**:8-13.