Algal Research 12 (2015) 545–549

Contents lists available at ScienceDirect

Algal Research

journal homepage: www.elsevier.com/locate/algal

Short communication

Informatics for improved algal taxonomic classification and research: A case study of UTEX 2341

Brendan T. Higgins [a], David Nobles [b], Yan Ma [c], William Wikoff [c], Tobias Kind [c], Oliver Fiehn [c], Jerry Brand [b], Jean S. VanderGheynst [a,*]

[a] Biological and Agricultural Engineering, University of California, Davis, CA 95616, United States
[b] UTEX The Culture Collection of Algae, University of Texas, Austin, TX 78712, United States
[c] University of California Davis, Genome Center, Davis, CA 95616, United States

ARTICLE INFO

Article history:
Received 19 May 2015
Received in revised form 22 August 2015
Accepted 31 October 2015
Available online 12 November 2015

Keywords:
UTEX 2341
Chlorella minutissima
Auxenochlorella protothecoides

ABSTRACT

Algal research has the potential to address numerous industrial and environmental challenges. Despite this potential, informatics data for algae are relatively modest limiting the development of algal biotechnology. This study reports on the re-classification of the green algae strain UTEX 2341 as a case study demonstrating the need for improved public informatics data for algae. Observations in the literature from lipid and pigment research led to uncertainty surrounding the identity of UTEX 2341, originally classified as *Chlorella minutissima*. In this study, lipid and conserved-gene sequencing data were combined to establish the identity of UTEX 2341 as *Auxenochlorella protothecoides*. Unambiguous identification of strains that are maintained over long periods of time will require sufficient allocation of resources to culture collections. Furthermore, development of repositories for biochemical and physiological data, along with DNA sequence information, are needed to facilitate sharing of strain data between research institutions and culture collections.

1. Introduction

Algal research has tremendous potential to address global challenges related to food and renewable energy production [1], climate change [2–4], and water quality [5,6]. Microalgae have high photosynthetic efficiency and are the fastest-growing primary producers on Earth [1]. Many strains can also grow in extreme environments unsuitable for terrestrial plant cultivation, such as brackish water and waste streams [7], reducing competition for land and water required for food production. Furthermore, numerous studies have reported accumulation of products in algae that have applications in nutrition and medicine and that could serve as precursors for commodity fuels and chemicals [8–10]. Despite this industrial potential, algal bioinformatics data are limited [11].

Algae represent the most evolutionarily and biochemically diverse group of organisms on earth, yet the NCBI database contains less than 150 completed and ongoing eukaryotic algal genome projects compared to 168 completed assemblies for *Saccharomyces cerevisiae* and 3005 completed assemblies for *Escherichia coli* [12]. This modest amount of genomic data limits the development of algal biotechnology and makes taxonomic classification of algae challenging. Full genome sequencing should be a priority, but a concerted effort will take years. Moreover, full genome sequencing is unlikely to be practical for ongoing strain quality control regimens employed by labs and culture collections. Instead, a combination of conserved gene sequencing [13], lipidomics, and metabolomics analyses can help fill the bioinformatics void, reduce the prevalence of taxonomic misclassification, and form the core of strain quality control protocols.

The majority of microalgal strains available from major culture collections were isolated and accessioned prior to the routine use of DNA sequencing for strain characterization. These strains were assigned taxonomic identities based on morphological and/or biochemical traits which did not necessarily reflect phylogeny. Morphological and biochemical characteristics alone are inadequate to elucidate the evolutionary relationships of microalgae. DNA sequence comparisons are also necessary for phylogenetic analyses. Indeed, taxonomic revisions within microalgal groups have become commonplace since the use of DNA sequences for identification has become widespread [14,15]. For example, the genus *Chlorella* historically contained small, green, spherical, asexual, unicellular microalgae. However, comparative sequence analyses using rDNA genes demonstrated that these characteristics had limited phylogenetic information and that the genus *Chlorella* was polyphyletic [16]. The number of species in this group has been greatly reduced to include only three species that are considered the "true" *Chlorella* [17]. Difficulty in obtaining definitive identifications of microalgae has practical implications for quality assurance in culture collections. Algal culture collections often lack the resources to create DNA barcodes for their existing holdings and to utilize DNA sequencing

* Corresponding author at: Department of Biological and Agricultural Engineering, One Shields Ave., Davis, CA 95616, United States.
E-mail address: jsvander@ucdavis.edu (J.S. VanderGheynst).

http://dx.doi.org/10.1016/j.algal.2015.10.016
2211-9264/

as a method of quality control to confirm the identity of strains at the time of accession. Moreover, maintenance of metabolically active cultures over long periods of time by serial transfer inevitably introduces the risk of cultures being compromised by contaminants or mislabeling. Culture collections have quality control mechanisms in place to identify and correct problems with compromised cultures. However, on rare occasions, these mechanisms fail, resulting in the loss of a culture or the distribution of a contaminated or misidentified culture. This is true even under the most rigorous operational standards such as those employed in the maintenance of mammalian tissue culture. For example, 360 known cross-contaminated or misidentified human cell lines have been identified and many of these have been used for published research [18]. This further underscores the importance of bioinformatics data for rapid identification and correction of such errors.

The green algae strain UTEX 2341 offers a compelling case study demonstrating the need for improved algal data. Like many other green algae, UTEX 2341 has properties that make it a prime candidate for the production of food, fuels, and bioproducts. Research has shown that it can synthesize significant quantities of oil, particularly the omega3 fatty acid eicosapentaenoic acid (EPA) [19]. For this reason, this strain has attracted attention as feed for shellfish [20] and was included in a Solazyme patent for oil production [21]. However, uncertainty surrounding the identity of this strain, originally classified as *Chlorella minutissima*, has arisen from lipid and pigment research [20,22]. We suggest that this issue has not been resolved, due to a lack of sequence, metabolomics, and lipidomics data. High-throughput sequencing and metabolic profiling capabilities are now available to enable large-scale population of public databases with algal informatics data. This would allow rapid cross-referencing of algal genetic and metabolic data, facilitate identification of promising new strains, and assist with correction of misidentified algae.

Without public informatics data, strain mutations and misidentification can take years to resolve, during which time researchers continue to publish results with these strains. For example, UTEX 2341 has been employed in at least 124 publications based on a Google Scholar search and this figure does not include publications that fail to reference the strain by name. We now provide evidence, based on its fatty acid profile and its 18S sequence, that UTEX 2341 is *Auxenochlorella protothecoides* rather than *C. minutissima* as currently annotated.

Early reports using UTEX 2341 reported the presence of long chain polyunsaturated fatty acids, predominantly EPA [19,23] (Table 1). Gladu et al. described the strain as a "dull yellow-green color" [20] which contrasted with the bright green UTEX 2341 strain that we obtained from UTEX in 2006. The fatty acid profile of another *C. minutissima* (UTEX 2219) also exhibited significant amounts of long chain PUFAs, including EPA [19]. However, a publication by Li et al. in 2011 showed that UTEX 2341 contained predominantly C18 fatty acids with no long chain PUFAs [24]. Li et al. noted this discrepancy with earlier studies but suggested that fatty acid composition is sensitive to growth medium salinity: Seto et al. [23] observed increased EPA production under higher salinity. Because Li et al. cultivated UTEX 2341 on freshwater medium while Seto et al. grew it in saline water, they concluded that the fatty acid profile shifted due to the salinity difference. We have now cultivated the current strain from UTEX on freshwater medium and artificial seawater, and have not observed significant differences in either growth characteristics or fatty acid composition between the two culture conditions. Moreover, our observed fatty acid profile closely matches that of Li et al. for cultures grown in either freshwater or saltwater media, with no detection of long chain PUFAs (Table 1). This suggests that differences in salinity cannot explain the discrepancy in the fatty acid profile of UTEX 2341 in 2011 compared to profiles observed in the 1980s and 1990s.

Lipid profiling by thin layer chromatography performed by Haigh et al. [25] in 1996 showed that UTEX 2341 synthesized large quantities of diacylglyeryltrimethylhomoserine (DGTS) under certain growth conditions. This study also showed that EPA accounted for 92.9% of the fatty acids found in DGTS and that DGTS was a major source of EPA in UTEX 2341. Like Vazhappilly et al. [19], Haigh et al. cultivated UTEX 2341 on saltwater medium. Our own lipid profiling of intracellular lipids from UTEX 2341 by LC-MS revealed no intracellular DGTS regardless of culture medium salinity (Table 2). Interestingly, the current UTEX 2341 strain was found to secrete DGTS when grown on complex solid medium as documented by Kind et al. in 2012 [22]. To understand if this discrepancy is due to culture conditions, we repeated the experiment carried out by Kind et al. and found that both our 2006 strain and the current UTEX strain do in fact secrete DGTS when grown on complex solid medium (Table 2). Nevertheless, the LC-MS data revealed that acyl chains associated with DGTS were predominantly medium chain fatty acids, not EPA as reported by Haigh et al. These lipid results

Table 1
Fatty acid composition of algal lipids based on FAME analysis.

	Chlorella minutissima						*Auxenochlorella protothecoides*		
Strain	UTEX 2341					UTEX 2219	Isolate	UMN280	CCAP 211/10C
Study	This study		Li et al. 2011 [24]	Vazhappily and Chen, 1998 [19]	Seto et al. 1984[a] [23]	Vazhappily and Chen, 1998 [19]	Pasaribu et al. 2014 [27]	Zhou et al. 2012 [28]	This study
Medium	N8-NH4	Artificial seawater	N8Y	Artificial seawater	Mann & Myers	Proteose	Artificial seawater	Municipal wastewater	N8-NH4
C12:0	–	–	–	–	–	–	–	0.2	–
C14:0	–	–	1.3	0.7	11.9	11.4	–	–	–
C14:1	–	–	–	4.7	–	2.7	–	–	–
C16:0	12.3 (0.5)	11.8 (0.1)	14.9	11.9	13.4	30.5	21.4	27.9	17.3 (<0.1)
C16:1	–	–	1.0	15	21.2	26.8	2.1	0.2	–
C16:2	–	–	–	–	2.1	–	–	0.6	–
C16:3	–	–	–	–	–	–	–	0.6	–
C18:0	1.2 (0.2)	1.3 (0.2)	1.5	7.8	–	4.9	1.5	4.6	2.6 (<0.1)
C18:1	0.6 (0.5)	0.8 (0.1)	1.7	12	1.4	3.3	16.1	–	1.4 (<0.1)
C18:2	22.0 (0.2)	26.4 (0.1)	22.9	6.3	1.7	0.3	18.2	28.8	33.5 (<0.1)
C18:3	63.9 (1.3)	59.7 (0.1)	54.7	4.6	–	10.3	29.0	36.9	45.1 (<0.1)
C20:4	–	–	–	5.6	3.4	2.6	–	–	–
C20:5	–	–	–	31.3	44.7	3.3	–	–	–
C22:5	–	–	–	0.1	–	3.9	0.6	–	–

For data from this study, values are averages followed by (s.d.) based on n = 3 biological replicates for each experimental treatment. All samples were derivatized simultaneously and GC injection was randomized.

[a] The strain used by Seto et al. was deposited in the UTEX collection as 2341.

Table 2
Number of distinct compounds measured in each lipid class in UTEX 2341 based on LC-MS/MS analysis.

Lipids	Positive mode			Negative mode		
	N8-NH4	Artificial seawater	Solid ATCC #5, secrete	N8-NH4	Artificial seawater	Solid ATCC #5, secrete
MGDG	23	23	0	9	9	1
DGDG	15	15	12	2	2	7
SQDG	0	0	0	14	14	12
DGTS	0	0	18	0	0	0
lysoDGTS	0	0	3	0	0	0
PC	43	43	16	11	11	4
lysoPC	4	4	0	3	3	0
PE	13	13	1	9	9	0
lysoPE	1	1	0	1	1	0
PG	0	0	0	1	1	5
PS	3	3	0	0	0	1
PI	0	0	0	2	2	2
DG	9	9	1	0	0	0
TG	71	71	18	0	0	0
Total	182	182	69	52	52	32

Lipids shown had a LipidBlast [35] PepSearch dot product of >600. Intracellular lipids were analyzed for autotrophic UTEX 2341 cultures grown on liquid N8-NH_4 or artificial seawater medium. Secreted lipids were analyzed for UTEX 2341 grown on complex solid medium (ATCC #5 sporulating agar). Annotation was performed by MS-DIAL software (v1.85) [36].

indicate the possibility of a strain mix-up, culture mislabeling, or else a mutation that occurred at UTEX. This analysis also underscores the importance of using consistent culture conditions when employing metabolite analysis as part of strain identity confirmation.

To confirm the identity of our UTEX 2341 strain, we obtained a partial sequence (1668 bp) of its 18S gene by PCR followed by Sanger sequencing. When trimmed, this sequence was an exact match to a 1257 bp segment of the 18S gene in *A. protothecoides* (UTEX 250). We also sequenced a 1593 bp segment of the 18S gene from (*Auxeno*)*chlorella protothecoides* (CCAP 211/10C) and observed an exact match between this partial sequence and that of UTEX 2341. We also sequenced the 18S gene from the UTEX 2341 strain currently maintained at UTEX and observed an exact sequence match to our strain from 2006. When placed within a phylogenetic tree based on 18S partial sequences of ~1 kb, we found that UTEX 2341 fell within the region occupied by *A. protothecoides*, and exhibited marked distance from other *C. minutissima* species (Fig. 1). For example, the 18S gene from *C. minutissima* SAG 1.8 shared only 90.6% identity with that of UTEX 2341 over a 1668 bp overlapping segment. Our phylogenetic tree also suggests potential misclassification of a *C. protothecoides* species (strain 124) that clusters with two *C. minutissima* strains. Huss et al. sequenced strain 124′s 18S gene and concluded that this strain was not correctly annotated as *C. protothecoides* but that it fell within the *Chlorella* umbrella [26]. Thus, their phylogenetic tree based on 18S sequences also suggests significant misclassification of green algae. We also sequenced UTEX 2341′s internal transcribed spacer (ITS2, KT368819) between the 5.8S and 28S ribosomal rRNA genes and confirmed that this matched closely with that of another *A. protothecoides* strain (UTEX 31, KJ676121.1).

Partial 23S sequences (~500 bp) from a variety of *Chlorella* species also showed that UTEX 2341 clusters with *A. protothecoides* [21]. While this concurs with our classification of UTEX 2341 as *A. protothecoides*, this small segment is highly conserved at the genus level and therefore could not be used for conclusive identification of UTEX 2341.

Several previous analyses of *A. protothecoides*' fatty acid profile [27,28] match quite well with our fatty acid analysis of UTEX 2341 (Table 1). They contain predominantly C18 fatty acids with minimal long chain PUFAs. The fatty acid composition data, together with the partial sequences of 18S and ITS2 genes suggest that UTEX 2341 should be reclassified as *A. protothecoides*. That said, the lipid and fatty acid profiles from studies conducted in the 1980s and 1990s suggest that UTEX 2341 may have originally been a eustigmatophyte [20] but no genetic sequencing data was published to provide verification. It is possible that a strain switch occurred prior to 2006, accounting for the discrepancies in UTEX 2341′s lipid composition and physical appearance between previous and more recent publications.

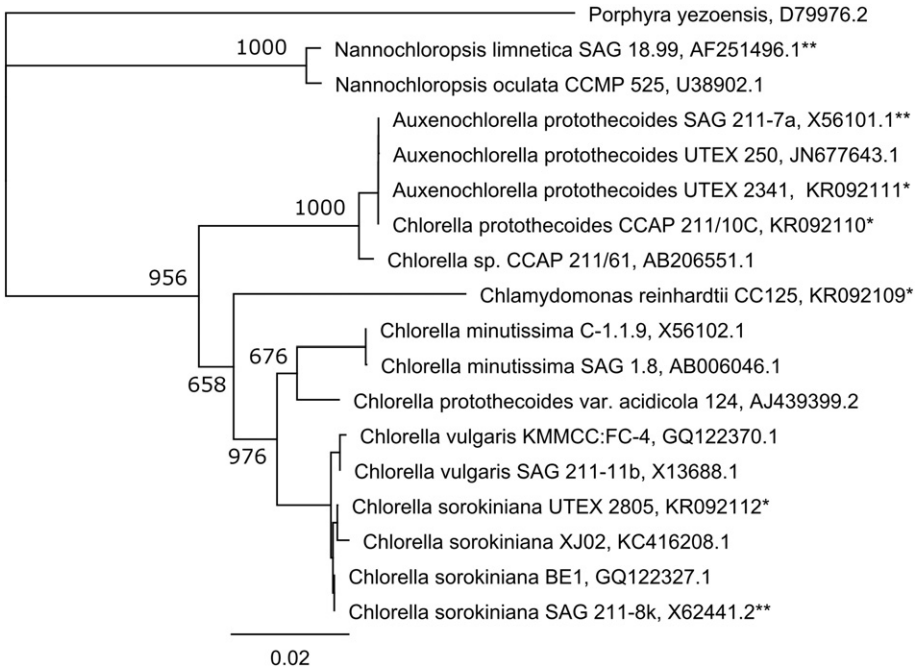

Fig. 1. Phylogenetic tree based on partial 18S sequences of ~1000 bp. GenBank Accession numbers follow strain names and bootstrap values are shown for internal nodes based on 1000 replications. *Partial sequences that were obtained in this work. The 18S gene for *Chlorella sorokiniana* (UTEX 2805) was included for method validation and resulted in an exact match to an existing sequence (AM423162.1). The 18S gene for *Chlamydomonas reinhardtii* (CC125) exactly matched those of strains CC621 and SAG 11-32a (JX888472.1 and AB511835.1). **Type strain of species.

This case study combines lipid and conserved-gene sequencing data to establish the identity of UTEX 2341 and demonstrates that more extensive application of bioinformatics to algae is needed. Taxonomic assignments using ribosomal RNA can help clarify relationships among algae, but are most applicable to genus-level identification due to highly conserved sequences among strains of the same species such as those of *A. protothecoides*. Strain level distinctions require comparison across multiple genes and would be greatly facilitated by complete genome sequencing. Even as genomes come online, we believe that periodic conserved gene sequencing and metabolite analysis can serve as powerful quality control methods for long-term strain maintenance.

In addition to providing accurate strain comparisons and identification, bioinformatics data can facilitate basic and applied algal research. Given the industrial potential of algae such as UTEX 2341 for biofuel, bioproducts, and food production, it is imperative that culture collections have sufficient resources for unambiguous identification of strains that are maintained over long periods of time. Also the development of repositories for biochemical and physiological data, along with DNA-sequence information, are needed to facilitate sharing of strain data between research institutions and culture collections. These publicly accessible data are critical to fully realize the potential of algae for industrial applications and address environmental challenges facing the planet.

2. Methods

C. minutissima (UTEX 2341), *A. protothecoides* (CCAP 211/10C), and *Chlamydomonas reinhardtii* (CC125) were cultured in N8-NH_4 freshwater medium [29] unless otherwise noted. *Chlorella sorokiniana* (UTEX 2805) was cultured in N8 medium [30]. Saltwater studies were conducted using artificial seawater based on UTEX artificial seawater with modifications: 18 g/L NaCl, 764 mg/L NH_4Cl, 50 mg/L KH_2PO_4, 13 mg/L $CaCl_2 \cdot 2H_2O$, 10 mg/L FeNa EDTA, 1250 mg/L $MgSO_4$, 3 g/L Tris, 500 mg/L KCl, 1 ml micronutrient solution. The micronutrient solution contained 3.58 g/L $Al_2(SO_4)_3 \cdot 18H_2O$, 12.98 g/L $MnCl_2 \cdot 4H_2O$, 1.83 $CuSO_4 \cdot 5H_2O$, and 3.2 g/L $ZnSO_4 \cdot 7H_2O$. Algae for lipid analysis (UTEX 2341 and CCAP 211/10C) were cultivated at 28 °C in 300 ml hybridization tubes under T5 lamps (10,000 lx, 16:8 light–dark cycle) as described previously [29]. Algae grown for DNA extraction were cultivated in bottles under the same illumination as hybridization tubes. Algae were harvested and lipids were extracted using the Folch method as described previously [29]. Fatty acid methyl esters and GCMS were performed using methods described previously [31]. The LC-MS/MS was performed using a published method with some modifications [32]: the gradient was shortened to 15 min and three collision energies were applied in the MS/MS level (20, 40, 60 V CID).

DNA was extracted from freeze-dried algae using the Fast DNA Spin kit (MP Biomedicals). PCR reactions using the forward primer GCCAGT AGTCATATGCTTGTCTC, and the reverse primers CATCTAAGGGCATCAC AGACCTGT and TCCGCAGGTTCACCTACGGA were used to amplify partial segments of the 18S rRNA gene. These primers were based on universal fungal primers with modifications [33]. PCR was performed on a StepOnePlus qPCR instrument (Applied Biosystems) using *Taq* polymerase (Takara). The PCR cycle was 95 °C for 25 s, 53 °C for 30 s, and 72 °C for 90 or 105 s (for the longer amplicon) over 40 cycles. Gel electrophoresis was used to confirm correct amplification and the PCR product was cleaned using a PCR clean-up kit (QIAGEN). Sanger sequencing was performed using both forward and reverse primers at the UC Davis DNA Sequencing Facility. Partial sequences were aligned and trimmed using the Geneious software package to obtain a ~1700 bp 18S sequence. Geneious was also used to compare 18S sequences to those found in the NCBI database and to construct a phylogenetic tree using the Temura-Nei genetic distance model and the neighbor-joining tree building method. *Auxenochlorella* strains were included in the tree to show similarity with UTEX 2341 while several "true *Chlorella*" species (*Chlorella vulgaris* and *Chlorella sorokinana*) and *C. minutissima* strains were included to show contrasts. *Nannochloropsis* species were included because UTEX 2341 had previously been proposed as a member of this genus. The red algae *Porphyra yezoensis* was included as an outgroup. Strains were selected largely based on public availability of nearly complete 18S sequences. PHYLIP was used to perform 1000 bootstrap replicates using the neighbor-joining method [34].

Acknowledgments

This research was supported by NSF grants CBET-1438211 and MCB-1139644 and National Institute of Food and Agriculture project CA-D-BAE-2228-RR.

References

[1] J. Sheehan, T. Dunahay, J. Benemann, P. Roessler, A look back at the U.S. Department of Energy's Aquatic Species Program—Biodiesel from Algae, National Renewable Energy Laboratory, Golden, CO, 1998.
[2] L. Cheng, L. Zhang, H. Chen, C. Gao, Carbon dioxide removal from air by microalgae cultured in a membrane-photobioreactor, Sep. Purif. Technol. 50 (2006) 324–329.
[3] J.E. Keffer, G.T. Kleinheinz, Use of *Chlorella vulgaris* for CO_2 mitigation in a photobioreactor, J. Ind. Microbiol. Biotechnol. 29 (2002) 275–280.
[4] M.G. de Morais, J.A. Costa, Biofixation of carbon dioxide by *Spirulina sp.* and *Scenedesmus obliquus* cultivated in a three-stage serial tubular photobioreactor, J. Biotechnol. 129 (2007) 439–445.
[5] F.B. Green, L.S. Bernstone, T.J. Lundquist, W.J. Oswald, Advanced integrated wastewater pond system for nitrogen removal, Water Sci. Technol. 33 (1996) 207–217.
[6] R.J. Craggs, W.H. Adey, B.K. Jessup, W.J. Oswald, A controlled stream mesocosm for tertiary treatment of sewage, Ecol. Eng. 6 (1996) 149–169.
[7] A.D. Brown, L.J. Borowitzka, Halotolerance of *Dunaliella*, in: M. Levandowsky, S.H. Hutner (Eds.), Biochemistry and Physiology of Protozoa, Academic Press, Inc., New York 1979, pp. 139–190.
[8] T.M. Mata, A.A. Martins, N.S. Caetano, Microalgae for biodiesel production and other applications: a review, Renew. Sust. Energ. Rev. 14 (2010) 217–232.
[9] A.F. Talebi, S.K. Mohtashami, M. Tabatabaei, M. Tohidfar, A. Bagheri, M. Zeinalabedini, H. Hadavand Mirzaei, M. Mirzajanzadeh, S. Malekzadeh Shafaroudi, S. Bakhtiari, Fatty acids profiling: a selective criterion for screening microalgae strains for biodiesel production, Algal Res. 2 (2013) 258–267.
[10] O. Pulz, W. Gross, Valuable products from biotechnology of microalgae, Appl. Microbiol. Biotechnol. Adv. 65 (2004) 635–648.
[11] L. Tirichine, C. Bowler, Decoding algal genomes: tracing back the history of photosynthetic life on Earth, Plant J. 66 (2011) 45–57.
[12] NCBI, Nucleotide, NIH, 2015.
[13] T.A. Isenbarger, C.E. Carr, S.S. Johnson, M. Finney, G.M. Church, W. Gilbert, M.T. Zuber, G. Ruvkun, The most conserved genome segments for life detection on Earth and other planets, Orig. Life Evol. Biosph. 38 (2008) 517–533.
[14] C. Lemieux, C. Otis, M. Turmel, Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class *Trebouxiophyceae*, BMC Evol. Biol. 14 (2014) 211.
[15] K. Fucikova, F. Leliaert, E.D. Cooper, P. Skaloud, S. D'Hondt, O. De Clerck, F. Gurgel, L.A. Lewis, P.O. Lewis, J. Lopez-Bautista, C.F. Delwiche, H. Verbruggen, New phylogenetic hypotheses for the core *Chlorophyta* based on chloroplast sequence data, Front. Ecol. Evol. 2 (2014).
[16] V.A.R. Huss, C. Frank, E.C. Hartmann, M. Hirmer, A. Kloboucek, B.M. Seidel, P. Wenzeler, E. Kessler, Biochemical taxonomy and molecular phylogeny of the genus *Chlorella* sensu lato (*Chlorophyta*), J. Phycol. 35 (1999) 587–598.
[17] W. Luo, S. Pflugmacher, T. Proschold, N. Walz, L. Krienitz, Genotype versus phenotype variability in *Chlorella* and *Micractinium* (*Chlorophyta*, *Trebouxiophyceae*), Protist 157 (2006) 315–333.
[18] A. Capes-Davis, G. Theodosopoulos, I. Atkin, H.G. Drexler, A. Kohara, R.A. MacLeod, J.R. Masters, Y. Nakamura, Y.A. Reid, R.R. Reddel, R.I. Freshney, Check your cultures! A list of cross-contaminated or misidentified cell lines, Int. J. Cancer 127 (2010) 1–8.
[19] R. Vazhappilly, F. Chen, Eicosapentaenoic acid and docosahexaenoic acid production potential of microalgae and their heterotrophic growth, J. Am. Oil Chem. Soc. 75 (1998) 393–397.
[20] P.K. Gladu, G.W. Patterson, G.H. Wikfors, B.C. Smith, Sterol, fatty acid, and pigment characteristics of UTEX 2341, a marine eustigmatophyte identified previously as *Chlorella minutissima*, J. Phycol. 31 (1995) 774–777.
[21] A.G. Day, G. Brooks, S. Franklin, in: U.S.P. Office (Ed.), Chemicals produced from oil-bearing microbial biomass, Solazyme Inc., United States, 2012.
[22] T. Kind, J.K. Meissen, D. Yang, F. Nocito, A. Vaniya, Y.-S. Cheng, J. VanderGheynst, O. Fiehn, Qualitative analysis of algal secretions with multiple mass spectrometric platforms, J. Chromatogr. A 1244 (2012) 139–147.
[23] A. Seto, H.L. Wang, C.W. Hesseltine, Culture conditions affect eicosapentaenoic acid content of *Chlorella minutissima*, J. Am. Oil Chem. Soc. 61 (1984) 892–894.
[24] Z. Li, H. Yuan, J. Yang, B. Li, Optimization of the biomass production of oil algae *Chlorella minutissima* UTEX2341, Bioresour. Technol. 102 (2011) 9128–9134.
[25] W.G. Haigh, T.F. Yoder, L. Ericson, T. Pratum, R.R. Winget, The characterisation and cyclic production of a highly unsaturated homoserine lipid in *Chlorella minutissima*, Biochim. Biophys. Acta (BBA) 1299 (1996) 183–190.

[26] V.A. Huss, C. Ciniglia, P. Cennamo, S. Cozzolino, G. Pinto, A. Pollio, Phylogenetic relationships and taxonomic position of *Chlorella*-like isolates from low pH environments (pH < 3.0), BMC Evol. Biol. 2 (2002) 13.
[27] B. Pasaribu, I.P. Lin, C.S. Chen, C.Y. Lu, P.L. Jiang, Nutrient limitation in *Auxenochlorella protothecoides* induces qualitative changes of fatty acid and expression of caleosin as a membrane protein associated with oil bodies, Biotechnol. Lett. 36 (2014) 175–180.
[28] W. Zhou, Y. Li, M. Min, B. Hu, H. Zhang, X. Ma, L. Li, Y. Cheng, P. Chen, R. Ruan, Growing wastewater-born microalga *Auxenochlorella protothecoides* UMN280 on concentrated municipal wastewater for simultaneous nutrient removal and energy feedstock production, Appl. Energy 98 (2012) 433–440.
[29] B. Higgins, J. VanderGheynst, Effects of *Escherichia coli* on mixotrophic growth of *Chlorella minutissima* and production of biofuel precursors, PLoS One 9 (2014), e96807.
[30] O.U. Tanadul, J.S. VanderGheynst, D.M. Beckles, A.L. Powell, J.M. Labavitch, The impact of elevated CO_2 concentration on the quality of algal starch as a potential biofuel feedstock, Biotech. Bioeng. 111 (2014) 1323–1331.
[31] B. Higgins, A. Thornton-Dunwoody, J.M. Labavitch, J.S. VanderGheynst, Microplate assay for quantitation of neutral lipids in extracts from microalgae, Anal. Biochem. 465 (2014) 81–89.
[32] Y. Ma, T. Kind, D. Yang, C. Leon, O. Fiehn, MS2Analyzer: a software for small molecule substructure annotations from accurate tandem mass spectra, Anal. Chem. 86 (2014) 10724–10731.
[33] T.D. Bruns, T.J. White, J.W. Taylor, Fungal molecular systematics, Annu. Rev. Ecol. Syst. 22 (1991) 525–564.
[34] J. Felsenstein, PHYLIP — Phylogeny Inference Package (Version 3.2), Cladistics 5 (1989) 164–166.
[35] T. Kind, K.H. Liu, Y. Lee do, B. DeFelice, J.K. Meissen, O. Fiehn, LipidBlast in silico tandem mass spectrometry database for lipid identification, Nat. Methods 10 (2013) 755–758.
[36] H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, M. Arita, MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, Nat. Methods 12 (2015) 523–526.