(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
09 August 2018 (09.08.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/141978 A1

(51) International Patent Classification:
C12R 1/89 (2006.01) *C12N 1/12* (2006.01)
C12P 23/00 (2006.01) *C07C 403/24* (2006.01)

(21) International Application Number:
PCT/EP2018/052840

(22) International Filing Date:
05 February 2018 (05.02.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1701855.7 03 February 2017 (03.02.2017) GB
1702413.4 14 February 2017 (14.02.2017) GB

(71) Applicants: **THE MARINE BIOLOGICAL ASSOCIATION OF THE UNITED KINGDOM** [GB/GB]; The Laboratory, Citadel Hill, Plymouth Devon PL1 2PB (GB). **UNIVERSITY OF GREENWICH** [GB/GB]; Old Royal Navy College, Park Row, Greenwich Kent SE10 9LS (GB).

(72) Inventors: **SCHROEDER, Declan**; The Marine Biological Association of the United Kingdom, The Laboratory, Citadel Hill, Plymouth Devon PL1 2PB (GB). **HARVEY, Patricia**; University of Greenwich, Medway Campus, Faculty of Engineering and Science, Central Avenue, Gillingham, Chatham Kent ME4 4TB (GB). **XU, Yanan**; University of Greenwich, Medway Campus, Facuty of Engineering and Science, Central Avenue, Gillingham, Chatham Kent ME4 4TB (GB).

(74) Agent: **BRYERS LLP** et al.; 7 Gay Street, Bath, Bath and North East Somerset, BA1 2PH (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,

(54) Title: ALGAL STRAINS

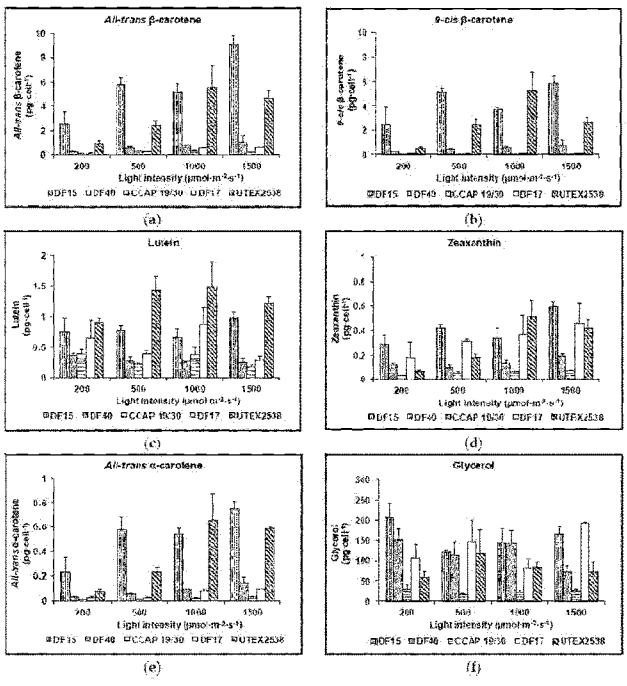

Figure 15

(57) Abstract: The present invention provides novel algal strains, including halotolerant hyper-accumulating carotenogenic strains of *Dunaliella salina*.

[Continued on next page]

WO 2018/141978 A1

MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with information concerning incorporation by reference of missing parts and/or elements (Rule 20.6)*

ALGAL STRAINS

Field of the Invention

The present invention relates to algal strains, compositions or taxonomy and in particular but not exclusively to algal compositions or clades for the production of compounds and in particular carotenes, such as β-carotene.

Background of the Invention

Algae are well known, for example there are over twenty strains of algae assigned to the genus, Dunaliella, have been isolated from various hypersaline environments. Dunaliella is one of the richest sources of natural β-carotene.

Dunaliella salina is a type of halophile green micro-algae. Halophiles are organisms living and thriving in environments of high salinity. D. salina is known for its high concentration of carotenoids that provide protection against the intense light normally measured in salt evaporation ponds. Due to its carotenoid accumulation, D. salina has various applications in health and nutrition.

Dunaliella salina is a rich source of natural orange-, yellow- or red-pigmented carotenoids. Their antioxidant and pro-retinoic acid activity may protect humans from compromised immune response, premature ageing, cancers, cardiovascular disease and arthritis.

Duniella salina consists of a species complex made up of a diverse range of isolates collected from natural and man-made hypersaline environments. Strains are routinely erroneously assigned to and within this complex.

Dunaliella (Chlorophyceae, Dunaliellales) is a genus of algae with immense economic potential owing to its production of an array of exploitable compounds, including β-carotene, glycerol and phytosterols. Microalgal species within this genus include halophilic and halotolerant strains, with D. salina able to tolerate NaCl saturation of approximately 5.5 M NaCl. Dunaliella lacks a polysaccharide cell wall and has a flexible cell membrane capable of rapidly changing shape in response to osmotic stress. Cell size is also highly variable with differences in cell size related to growth conditions, e.g. nutrients, pH and salt concentration and can also vary within the same culture at the same stage of growth, for example D. salina exhibits 5-29 μm cell length, 3.8-20.3 μm width).

Within the family Dunaliellaceae are four sections of Dunaliella; section Tertiolectae which are oligo-euhaline and do not accumulate carotenes and grow at an optimum salinity of <6% NaCl; section

Dunaliella which are halophilic species that accumulate carotenes; section Virides which are hyperhaline, always green and radially symmetrical; and Peirceinae which are hyperhaline, always green but cells are bilaterally symmetrical.

Presently, within the section Dunaliella are three accepted species, D. salina, D. parva and D. pseudosalina as well as D. bardawil, which is also considered as D. salina in some studies. D. salina is particularly polymorphic, with cells pigmented green or red, dependent on the amount of β-carotene accumulated. Under high stress conditions, members of this species can accumulate >5% β-carotene dry weight. D. bardawil was originally reported to accumulate considerably larger amounts of β-carotene compared with D. salina, accumulating the β-carotene in membrane-free globules in the interthylakoid spaces of the chloroplast.

D. parva typically accumulates up to 5% β-carotene dry weight and D. pseudosalina cells can be pigmented yellow/orange due to the accumulation of canthoxanthin as their main carotenoid pigment, as opposed to β-carotene.

The section Virides is a large group, encompassing 11 species. However, four of these species (D. baas-beckingii, D. media, D. ruineniana and D. gracilis) have been described based on a single field collection so are essentially provisional names. Variations in cell shape and size, flagella length and stigma are criteria often used to delineate these species from each other, e.g. D. bioculata has 2 stigmata whereas the other species have 1. Members of the Virides are often difficult to classify purely based on morphology, with D. viridis and D. minuta being particularly difficult to resolve due to their similar cell size and the ambiguous characteristic of cell form, defining these two from each other.

Genotyping is now considered imperative for the accurate classification of Dunaliella species, with the ribosomal markers 18S rRNA and ITS2 at the forefront of phylogenetic analyses; however, other commonly used markers for Dunaliella sp. and, indeed other marine algae. Furthermore, by exploiting the sporadic occurrence of the group I introns within the 18S rDNA, the 18S has proven useful as a size indicator of different Dunaliella species. The absence of introns in the 18S rDNA in the Tertiolectae, 1 intron in D. salina, 2 introns in D. bardawil and 1 intron for D. viridis (that differs to the aforementioned species), have been used to define these species.

The hypervariable regions, ITS 1, 5.8S rRNA and ITS 2, have been frequently employed by molecular studies of this important algal group; interestingly, some studies have favoured to use each spacer sequence separately in phylogenetic analysis. Assunção et al. 2012 undertook a comprehensive analysis of 3 Dunaliella species using ITS 2, identifying 5 main clades, viridis, salina I, salina II, tertiolecta and pseudosalina, with the latter clade being somewhat ambiguous due to problems with taxonomy.

ITS markers are often favoured for identification as there are many copies in the genome making it easy to amplify, it is biparentally inherited and insertions/deletions within the sequence are common meaning there is good variability between species. Other nuclear markers used for the analysis of Dunaliella from hypersaline environments include the large subunit rDNA and small subunit rDNA and these have shown good potential for taxonomic resolution.

Both rbcL and tufA are plastid genes, with tufA encoding elongation factor Tu, and has become more frequently used for molecular studies of algae. Availability of sequences, however, is limited with only two sequences for Dunaliella sp. available in Genbank, making phylogenetic analysis less conclusive compared to other markers. Moniz et al., 2014 used the tufA gene in their analysis of the order Prasiolales (Chlorophyta) finding that there was good agreement between the phylogenies of tufA, rbcL and psaB. Presently, tufA as a marker for Dunaliella taxonomy has not been thoroughly examined and its potential not fully realised. rbcL analysis has typically not highlighted intraspecific variation to the same degree as other markers such as the ITS regions.

Suitable molecular tools are needed for accurate identification of species as this will aid in accurately identifying those isolates that will be economically valuable, e.g. strains of D. salina that produce high levels of 9-cis-β- carotene that is economically more valuable than its isomer all-trans-β-carotene, and to further understand the molecular evolution of this important group. This study set out to genetically characterise a range of Dunaliella isolates collected from a range of geographical provinces including Israel, Spain, South Africa, Italy and Namibia. We sought to employ a suite of molecular tools to provide a comprehensive analysis of different markers and their suitability for application to the genus Dunaliella.

Summary of the Invention

Accession numbers:

CCAP 19/40	Dunaliella salina PLY_DF-40
CCAP 19/41	Dunaliella salina rubeus PLY_DF-15

Dunaliella strains from Eilat, Israel and Monzon, Spain, were isolated and characterised by the Marine Biological Association, UK (https://www.mba.ac.uk/). These strains are now available by application to the MBA Culture Collection.

Dunaliella salina strain DF15 isolated from Eilat, has a significantly higher cellular content and higher productivity of carotenoids than many other hyper-accumulating carotenogenic strains such as D. bardawil UTEX 2538.

Dunaliella salina strain DF40 isolated from Eilat is very similar to D. bardawil and grows in very highly saline water associated with crystallizing salt ponds.

Some aspects of the present invention relate to halotolerant hyper-accumulating carotenogenic strains of Dunaliella.

An aspect of the present invention provides a new composition or taxonomy containing algae from the sections Dunaliella and Virides.

A further aspect provides an algal composition or taxonomy comprising, consisting of or including D. rubeus, D. salina aureus and D. velox.

A further aspect provides a composition or taxonomy containing algae from the genus Dunaliella in a body of aqueous nutrient solution, in which the algae is from a clade other than the D. bardawil and salina salina clades.

The composition or taxonomy may be provided in a salt solution.

The salt solution may contain less than 6% NaCl.

A further aspect provides a method of strain selection or determination for the production of β-carotene by growing alga of the genus Dunaliella and extracting the β-carotene produced thereby.

The alga may be from a clade other than the previously described D. bardawil and D. salina salina clades.

The alga may be selected from the group comprising: (a) D. salina aureus SA3, SA4, T68, T41, T36, and T37 (b) D. rubeus DF15 (c) D. velox SA5, SA6.

An aspect of the present invention provides a method of producing β-carotene by growing alga of the genus Dunaliella and extracting the β-carotene produced thereby.

The alga may be from a clade other than the previously described D. bardawil and D. salina salina clades.

The alga may be selected from the group comprising: (a) D. salina aureus SA3, SA4, T68, T41, T36, and T37 (b) D. rubeus DF15 (c) D. velox SA5, SA6.

A further aspect provides an algae bio-refinery, for example a microalgae-based biorefinery, comprising or including one or more strains of algae described herein for the production of one or more (useful) compounds.

Some aspects of the present invention relate to an algae biorefinery, based on biomass from the halotolerant microalga Dunaliella salina.

The biorefinery elements may be integrated and optimised using sustainability assessments. The biorefinery elements include:

1. Biomass: New strains cultivated in lakes, raceways and photobioreactors.
2. Bioprocessing: Key biomass processing technologies are applied to the biomass.

Some aspects include innovative spiral plate technology for dynamic settling:
Ultramembrane filtration
Supercritical CO2 techniques
High performance counter-current chromatography

In the biorefinery of the present invention several extraction methods may be used (supercritical CO_2, high-performance counter-current chromatography, membrane technology, hydrophobic interaction resins and ion exchange resins). Each of these technological pathways is suitable for efficiently recovery of a specific class of products.

After harvesting, the collected algal biomass, usually a liquid suspension, should be processed immediately. If it has to be shipped for processing it should be dried to avoid degradation of biomass and costs of transporting water. Two different methods include:

1. Freeze-drying or lyophilisation, a dehydration process which involves freezing the material and then reducing the surrounding pressure to allow the frozen water to sublimate (from solid to gas).
2. Spray-drying. Here, the algal suspension is rapidly dried with a hot gas.

The present invention utilizes three harvesting systems for Dunaliella biomass, depending on the end-products required:

1. Stacked disk clarifier centrifuges: Dunaliella cells are enriched in β-carotene and lack of glycerol and other water soluble compounds.
2. Evodos dynamic settling machines: Dunaliella cells are rich in β-carotene with predominantly intact cells.
3. Membrane filtration: Dunaliella cells enriched in β-carotene retain on membrane, while solutes as well as bacteria and viruses pass through.

Aspects and embodiments of the present invention may comprise, consist of or include: (i) process of isolating and characterizing a new microorganism; (ii) new microorganism as produced by a defined process; (iii) new microorganism per se; and (iv) process of cultivation or otherwise using a known or new microorganism to: (a) a form of multiplied microorganism itself, for example biomass, and (b) a by-product of microbial growth, for example a useful industrial product.

Referring now to the drawings, aspects and embodiments of the present invention are further described. One of ordinary skill in the art will appreciate the many possible applications and variations of the present invention based on the following examples of possible embodiments of the present invention.

The example embodiments are described in sufficient detail to enable those of ordinary skill in the art to embody and implement the systems and processes herein described. It is important to understand that embodiments can be provided in many alternate forms and should not be construed as limited to the examples set forth herein.

Accordingly, while embodiments can be modified in various ways and take on various alternative forms, specific embodiments thereof are shown in the drawings and described in detail below as examples. There is no intent to limit to the particular forms disclosed. On the contrary, all modifications, equivalents, and alternatives falling within the scope of the appended claims should be included. Elements of the example embodiments are consistently denoted by the same reference numerals throughout the drawings and detailed description where appropriate.

Unless otherwise defined, all terms (including technical and scientific terms) used herein are to be interpreted as is customary in the art. It will be further understood that terms in common usage should also be interpreted as is customary in the relevant art and not in an idealised or overly formal sense unless expressly so defined herein.

Brief description of Figures and Tables

Figure 1: Images of the different Dunaliella strains isolated in this study (a) D. salina bardawil DF40 (b) D. salina bardawil DF45 (c) D. salina salina DF17 (d) D. salina salina T41 (e) D. salina salina T37 (f) D. salina rubeus DF15 (g) D. velox SA6 (h) D. tertiolecta DF44 (i) D. viridis SA2 (j) D. bioculata DF48 (k) D. minuta T34 (l) D. minuta T75. Scale bar is equivalent to 25µm.

Figure 2: Neighbour-joining tree of Dunaliella strains isolated during this study and sequences from Genbank based on a 477bp alignment of the LSU gene. Bootstrap values were retrieved from 1000 replicates and those >70% are indicated at the nodes for neighbour-joining, maximum likelihood and Bayesian respectively. Strain names followed by an asterisk indicate proposed taxonomic changes, where D. bardawil UTEX 2538 and ATCC 30861 has been proposed as D. salina (Borowitzka & Siva, 2007); D.

bioculata UTEX 199 has been proposed as D. tertiolecta (Assuncao et al., 2012); D. salina UTEX 200 has been proposed as D. viridis (Borowitzka & Siva, 2007); D. peircei UTEX 2192 has been proposed as D. maritima (Borowitzka & Siva, 2007) and D. viridis (Assuncao et al., 2012).

Figure 3: Neighbour-joining tree of Dunaliella strains isolated during this study and sequences from Genbank based on a 521bp alignment of the rbcL gene. Bootstrap values were retrieved from 1000 replicates and those >70% are indicated at the nodes for neighbour-joining, maximum likelihood and Bayesian respectively. Strain names followed by an asterisk indicate proposed taxonomic changes, where D. bardawil UTEX 2538 and ATCC 30861 has been proposed as D. salina (Borowitzka & Siva, 2007); D. bioculata UTEX 199 has been proposed as D. tertiolecta (Assuncao et al., 2012); D. salina UTEX 200 has been proposed as D. viridis (Gonzalez et al., 2001; Borowitzka & Siva, 2007).

Figure 4: Neighbour-joining tree of Dunaliella strains isolated during this study and sequences from Genbank based on a 614bp alignment of the tufA gene. Bootstrap values were retrieved from 1000 replicates and those >70% are indicated at the nodes for neighbour-joining, maximum likelihood and Bayesian respectively. Strain names followed by an asterisk indicate proposed taxonomic changes, where D. bardawil UTEX 2538 and ATCC 30861 has been proposed as D. salina (Borowitzka & Siva, 2007).

Figure 5: Neighbour-joining tree of Dunaliella strains isolated during this study and sequences from Genbank based on an 530bp alignment of the ITS1, 5.8S, ITS2. Bootstrap values were retrieved from 1000 replicates and those >70% are indicated at the nodes for neighbour-joining, maximum likelihood and Bayesian respectively. Strain names followed by an asterisk indicate proposed taxonomic changes, where D. bardawil UTEX 2538 and ATCC 30861 has been proposed as D. salina (Borowitzka & Siva, 2007); D. bioculata UTEX 199 has been proposed as D. tertiolecta (Assuncao et al., 2012); D. peircei CCAP 19/2 has been proposed as D. tertiolecta (Gonzalez et al., 2001); D. parva SAG 19-1 has been proposed as D. maritima (Borowitzka & Siva, 2007) and D. viridis (Assuncao et al., 2012); D. salina CCAP 19/3 has been proposed as D. viridis (Borowitzka & Siva, 2007); Dunaliella sp. ABRIINW M1/1 has been proposed as D. viridis (Assuncao et al., 2012).

Figure 6: Neighbour-joining tree of Dunaliella strains isolated during this study and sequences from Genbank based on an alignment of a concatenation of the ITS-LSU-rbcL-tufA sequences used to produce Fig. 2-Fig. 5. Bootstrap values were retrieved from 1000 replicates and those >70% are indicated at the nodes for neighbour-joining, maximum likelihood and Bayesian respectively. Images show the morphology of representative strains and the scale bar is equivalent to 25µm. Strain names followed by an asterisk indicate proposed taxonomic changes, where D. bardawil UTEX 2538 and ATCC 30861 has been proposed as D. salina (Borowitzka & Siva, 2007).

Figure 7: CLUSTALW alignment of the V9 SSU sequences generated in this study and sequences from Genbank. Dots 517 indicate identical nucleotides and letters indicate nucleotide substitutions.

Figure 8: Maximum likelihood phylogenetic tree based on ITS1-ITS2, rbcL, 28S rDNA & 18S rDNA sequences.

Figure 9: Growth curves for DF15 (red line) and DF17 (green line) where arrows indicate timings of RNA extractions for transcriptome sequencing. Inserts: Images of DF15 (red box) and DF17 (green box).

Figure 10: Microscopy observation of Dunaliella cells and photographs of stationary phase cultures of CCAP 19/30, UTEX 2538, DF17, DF40 and DF15 grown under a light intensity of 100~200 µmol $\cdot m^{-2} \cdot s^{-1}$ at 20°C. (**a**) Microscopy photographs taken through a light microscope (Nikon Eclipse Ti-U) with a magnification of 600x; (**b**) Differential interference contrast (DIC) microscopy photographs taken through a confocal microscope (ZEISS LSM 880) with a magnification of 630x. (**c**) Photographs of the cultures obtained for each Dunaliella strain grown under identical conditions.

Figure 11. Growth curves for the five Dunaliella strains: (**a**) CCAP 19/30; (**b**) DF15; (**c**) DF17; (**d**) DF40; (**e**) UTEX 2538 each grown under four identical light intensities of 200, 500, 1000 and 1500 µmol m^{-2} s^{-1} at a light/dark cycle of 12 h light and 12 h dark (LD200, LD500, LD1000 and LD1500); (**f**) specific growth rates of each strain grown under the four light intensities. Each culture condition was set up in triplicate.

Figure 12. Photosynthesis (a) and respiration (b) of the five Dunaliella strains cultivated under four light intensities of 200, 500, 1000, and 1500 µmol $\cdot m^{-2} \cdot s^{-1}$. Samples were taken at the mid log phase and all culture conditions were repeated at least in triplicates.

Figure 13. Cellular content of total chlorophyll (**a**) and total carotenoids (**b**) of the five Dunaliella strains grown under four light intensities of 200, 500, 1000, and 1500 µmol m^{-2} s^{-1}. Samples were taken at the mid log phase and all culture conditions were repeated at least in triplicates.

Figure 14. HPLC chromatograms of MTBE/ethanol extracts of the five Dunaliella strains cultivated under 1500 µmol $\cdot m^{-2} \cdot s^{-1}$. The major peaks shown are: (1) lutein, (2) zeaxanthin, (3) *all-trans* α-carotene, (4) *all-trans* β-carotene and (5) *9-cis* β-carotene.

Figure 15. Cellular contents of (**a**) all-trans β-carotene, (**b**) 9-cis β-carotene, (**c**) lutein, (**d**) zeaxanthin, (**e**) all-trans α-carotene and (**f**) glycerol in the five Dunaliella strains cultivated under four light intensities of 200, 500, 1000, and 1500 µmol m^{-2} s^{-1}. Samples were taken at the mid log phase and all culture conditions were repeated at least in triplicate.

Figure 16. Cluster dendograms of *all-trans* β-carotene, glycerol, lutein, zeaxanthin, all trans α-carotene, photosynthesis, respiration, total carotenoids, and total chlorophyll for all five *Dunaliella* strains cultivated at four light intensities. (**a**) CCAP 19/30; (**b**) DF15; (**c**) DF17; (**d**) DF40 and (**e**) UTEX 2538.

Figure 17. Principle component analysis of 11 traits (*all-trans* β-carotene, *9-cis* β-carotene, glycerol, lutein, zeaxanthin, *all-trans* α-carotene, photosynthesis, respiration, total carotenoids, total chlorophyll, and specific growth rate) for all five *Dunaliella* strains cultivated at four light intensities.

Table 1: Dunaliella sp. isolates collected during this study.

Table 2: Primers used in this study.

Table 3: Results from intron sizing and sequence alignments of introns.

Table 4: Gene annotation and relative expression analysis in biochemical pathway genes involved in carotenogenesis for strain DF15. Gene expression normalised to cell number (Figure 9) and data is shown in linear RPKM values. Grey= 0-10 and blue = 10-100 (no to low level expression); yellow=100-400 and orange=400-800 (medium levels of expression); brown=800-1200, red=1200-1600 and dark red=1600+(high levels of expression).

Table 5: Gene annotation and relative expression analysis in biochemical pathway genes involved in carotenogenesis for strain DF17. Gene expression normalised to cell number (Figure 9) and data is shown in linear RPKM values. Grey= 0-10 and blue = 10-100 (no to low level expression); yellow=100-400 and orange=400-800 (medium levels of expression); brown=800-1200, red=1200-1600 and dark red=1600+(high levels of expression).

Table 6. Chlorophyll and carotenoid content of the five Dunaliella strains at early stationary phase; cultures were maintained in the incubator at (20 ± 2) °C with 12/12 light/dark with light intensity of ~50 μmol photons m^{-2} s^{-1} for ~4 weeks.

Table 7. Two-way ANOVA analysis of the responses of all examined variables (photosynthesis, respiration, doubling time, all-trans β-carotene, 9-cis β-carotene, glycerol, lutein, zeaxanthin, all-trans _-carotene, total carotenoids, total chlorophyll) to strain and light intensity and their interaction (Light intensity*Strain). The values of all observations were transformed by taking log function, square root function or reciprocal to fit linear models. Df: degrees of freedom; F values:variation between sample means; P values: significance levels and a star (*) indicates $P \leq 0.05$, two stars (**) $P \leq 0.01$ and three stars (***) $P \leq 0.001$.

Detailed description

Microscopy observation

Microscopy clearly identified cells that had accumulated β-carotene owing to the orange-red colouration and hence those affiliated to the section Dunaliella, species D. salina (Fig. 1). These cultures (DF40, DF41, DF45, DF17, DF15, T36, T27, T41, T68, SA4 and SA3) ranged in mean cell length of 12.65-18.22 µm and cell width of 10.26-15.15 µm. Cell shape was variable and stigma was not easily identifiable due to the presence of refractile granules. The green cultures of SA1, SA2, DF48, SA7, T32, T34, T76, proposed to belong within the broad taxonomic section 'Virides', were much smaller ranging from a mean of 6.27-7.49 µm length; 3.2-4.9 µm width. Whilst T77 had similar features to the others, it was much smaller (mean cell length 2.33µm; width 2.04). The stigma was clearly visible in these cells. In the case of DF48, SA7, T32, T34, T75 and T76, the anterior of the cell was colourless and free of chloroplast with the pyr enoid often clearly defined in the basal cell, furthermore, DF48, SA7, T32, had the presence of 2 stigmata in the anterior part of the cell, a feature described in D. bioculata, a member of the Virides. SA5 and SA6 were slightly larger than the aforementioned green strains (7.61-8.13 µm length; 5.49-6.41 µm width). Cells were oval/pyriform and the stigma could be identified.

18S Intron-sizing method and sequence analysis

As anticipated, according to Olmos et al. 2009, isolates identified microscopically as D. tertiolecta had no introns (Table 3). We also identified two different strains of D. salina, as reported in Olmos et al. (2009), those with 2 introns and those with one. Further to this however, isolates that were assigned as D. salina microscopically were found to have no introns, i.e. the same MA1/2 amplicon size as D. tertiolecta. In this study, we amplified from two of the same isolates as Olmos and co-workers, with differing results. They reported D. bardawil UTEX 2538 to have 2 introns whereas we sized it to have 1. Moreover, we sized D. bardawil ATCC 30861 to have 0 introns which is intriguing as this strain and UTEX 2538 are meant to be the same culture. Strain ATCC (American Type Culture Collection) 30861 was isolated by Ben-Amotz and Avron in 1976 from a salt pond near Bardawil, Israel, and this strain was then deposited in the Culture Collection of Algae at the University of Texas at Austin (UTEX) by R. Adams between 1980 and 1982. So, not only do we have contradicting results regarding D. bardawil UTEX 2538 18S sizing we also are raising the issue as to whether it is the same strain as ATCC 30861. The culture collection legacy of this strain continues as three other culture collections host this strain as the same to the original D. bardawil isolated by Amotz and Avron. DCCBC (Dunaliella Culture Collection at Brooklyn College) received the strain from UTEX in 2002, SAG (Culture Collection of Algae at Goettingen University) received the strain from Avron via Prof Thompson in 1988, re-naming it D. salina SAG 42.88 in 2001, and CCAP (Culture Collection of Algae and Protozoa) received the strain from SAG in 1996 (Muller, 2005). Olmos et al., 2009 reported the CCAP 19/30 sequence in genbank, deposited by Herve in 2006, to have 1 intron.

Both this study and Olmos et al., 2009 reported CCAP 19/30 to have no introns. In this study we also sequenced the 18S amplicons for the majority of isolates, detecting 5 different sequences for each of the introns (Table 3).

DF15 was the only strain to produce two different sized-amplicons with the intron sizing method, leading us to suspect contamination of the culture. To ensure this was not a result of contamination, further rounds of single-cell picking from the culture was undertaken to ensure clonality and 6 more cells were picked from this culture, washed and grown up in higher salt media (150 g/L) to select for D. salina. The 18S PCR was repeated with all the newly isolated clones with the PCR consistently generating 2 amplicons. Both PCR products were sequenced, with the smaller product found to contain one intron and the larger product had two introns. When the sequences were compared, nucleotide substitutions were detected across the whole sequence (both within the intron and exon regions).

Phylogenetic analysis

The large-subunit phylogeny distinguished three virides clades supported by high bootstrap values (Fig. 2) with SA1 and SA2 clustered with a known species of D. viridis. The LSU phylogeny created a large clade of strains assigned as D. salina and D. tertiolecta with weakly supported subclades, owing to only a small number of nucleotide substitutions between strains of D. tertiolecta and D. salina, e.g. D. tertiolecta UTEX 199 has 2bp difference across the 477bp alignment with D. salina DF41. Within this large group however, SA5 and SA6 were separated and clustered with D. salina UTEX 200 (now viridis) and D. peircei UTEX 2192 (AKA SAG19-1) (now viridis or maritima) which have been assigned within the 'pseudosalina' clade by Assunção et al. (2012) or a viridis clade by Assunção et al. (2013). The SA5/SA6 clade is supported high bootstrap values.

The rbcL phylogeny (Fig. 3) also grouped strains assigned as D. tertiolecta with D. salina, however there were 2 separate clades of D. salina, which follow the clustering patterns described by Assunção et al. (2012) (salina I and salina II). The rbcL phylogeny is in agreement with the LSU phylogeny in that it forms strongly supported clade of SA5 and SA6 and three Virides clades. Interestingly, however, T32, clusters differently within the LSU phylogeny (with T76, T34 and T77), compared to the rbcL phylogeny where it is identical to SA7 and DF48. The tufA phylogeny (Fig. 4) strongly supports the clustering of D. tertiolecta in a separate clade to the D. salina clade which also encompasses the strains, SA5 and SA6.

The Virides form 3 strongly supported subdivisions separated by long branch lengths with the distance between T76 and DF48 equating to approximately 31 nucleotide substitutions between the two subclades (comparing 614bp sequence). The phylogenetic tree based on ITS1, 5.8S and ITS2 (Fig. 5) provided much more resolution compared to the other trees and was a robust tree supported by high bootstrap values. Two salina groups were identified, according to Assunção et al. (2012) and these were clearly separated by long branch lengths. Due to ambiguous bases in the DF15 sequences, this amplicon was cloned and sequenced, hence the inclusion in the tree of four DF15 sequences. DF15 clones were

found to have nucleotide substitutions across the ITS region when sequenced with the three sequences identified found to share >99% identity. Whilst grouping within the salina clade, DF15 does not cluster within either of the sub clades classified as salina I or II by Assunção et al. (2012). SA5/6 formed a separate clade with tentatively characterised D. viridis, D. maritima and D. tertiolecta. In all four trees, SA1 and 2 clustered together forming a clade with D. viridis sequences from genbank (where available) that was separated by significant genetic distance and supported by high bootstrap values.

The concatenated tree of the 4 markers, ITS-LSU-rbcL-tufA (Fig. 6) was constructed to further resolve/affirm any clades/sub-clades to provide reliable information on the taxonomy of these strains in combination with microscopy. It is important to note that in this tree, only sequences that were generated in this study were used as it was difficult to confidently match sequences from genbank that were derived from the same culture, particularly due to mis- identification and cross contamination problems. Three sub-clades of the section Dunaliella were detected that morphologically were identified as D. salina, and these were supported by significant bootstrap values. Within each of the subclades the strains shared 99% identity, with all the strains morphological identified as D. salina sharing at least 97% identity. D. tertiolecta strains (section Tertiolectae) shared 99% identity and 97% identity with members of the section Dunaliella, with SA5/6 sharing 97% and 98% identity with members of section Dunaliella and Tertiolectae respectively. A well delineated Virides section could be identified supported by high probability, with members of this section sharing 91-93% identity with members of the Tertiolectae and Dunaliella. The three clades identified within the Virides showed greater divergence with SA1/2 (D. viridis) sharing 91% with the DF48 clade (D. bioculata) and 92% identity with the T75 clade (D. minuta).

V9 amplicon sequencing was undertaken for DF15, DF17, DF40, D. bardawil UTEX 2538, Dunaliella sp. CCAP19/30, and D. tertiolecta UTEX 199 only, due to problems with sequence quality. The V9 primers used were found to not only amplify Dunaliella DNA but also bacterial DNA (cultures were non-axenic) and hence PCR products required cloning prior to sequencing to provide good quality sequences. Sequences amplified from the Dunaliella sp. CCAP19/30 culture included Dunaliella sp. and Halomonas sp. bacteria, and DF40 sequenced clones (6) were all Marinobacter sequences, suggesting this was preferentially amplified by the V9 primers. However, multiple sequence alignment (Fig. 7) of the sequences generated and those from Genbank showed that the members of the tentative clades, salina I, and tertiolecta were identical, however salina II, and viridis were different. DF15 was identical to salina I and tertiolecta species.

The intron sizing method revealed useful information regarding the history of Dunaliella spp. in culture collections. Olmos et al. (2009) concluded cross contamination of cultures had occurred resulting in differences in amplicon sizes in what was thought to be the same culture. Our study confirms that, based on sequencing analysis of 4 marker genes, the current CCAP 19/30 is in fact a strain of D. tertiolecta, not D. salina. With regards to the supposedly same cultures of D. bardawil, strain UTEX 2538 had 1intron and strain ATCC 30861 had 0 introns indicating they are not the same. The tufA and

rbcL sequences for these 2 strains are identical but the ITS sequences differ by 2bp over the 553bp in the multiple sequence alignment and the LSU sequence has 5 bp substitutions between them. Based on these clear differences in the ribosomal genes, we conclude that these two strains are not the same. This could be a result of cross-contamination, or alternatively if the original culture was not clonal, different culturing conditions could have selected for different strains. These results are a cautionary tale for anyone working on algal cultures, with genotyping of the strains an integral step to ascertain the taxonomic affiliation of the strains one is studying. The fact that a group of D. salina have no introns (the same as tertiolecta) negates the intron-sizing as a tool for separating these strains from each other as both an isolate of tertiolecta and salina would produce the same size amplicon despite these two being very different species.

Whilst there is a high copy number of the rDNA, concerted evolution typically results in identical sequences of this gene within the genome so it is surprising to get two different versions of this gene in DF15. Intragenomic sequence variation however is conceivable in the case of DF15, since its clonality has been verified. Certainly, Alverson and Kolnick (2005) reported intragenomic nucleotide polymorphisms within the 18S rDNA of the diatom Skelatonema, however the differences identified were random single nucleotide polymorphisms rather than whole introns. We propose that DF15 represents a novel lineage and as such concerted evolution has not progressed to completion as seen in the other Dunaliella species. Alternatively, it is possible that the rate of variation is exceeding concerted evolution with the hypersaline harsh environment selecting for heterogeneity. However, if indeed this was the case we would predict to see this heterogeneity in more strains.

Comparing the four markers analysed in this study shows that the LSU and rbcL are least useful in resolving the Dunaliellaceae family with the section Dunaliella and Tertiolectae forming weakly supported subclades with low boot strap values (<70%) yet they do cluster as expected and provide some information on the relationships between the different strains. Both these markers identify the two groups of D. salina as identified by Assunção et al. (2012). These markers also separate SA5 and SA6 from the other strains and form a well-supported clade of D. viridis strains. tufA, which to our knowledge has never been used for phylogenetic analysis of Dunaliella, was able to separate Viridies, Tertiolectae and Dunaliella with high reliability however the two D. salina groups and SA5/6 clusters are less clear. The phylogenetic tree based on the ITS regions gave the most information, with subdivisions strongly supported by high bootstrap values. The concatenated tree further reinforces the ITS regions with a clearer delineation of different phylogroups and adding the information on the intron presence/absence and sequence can provide further information on the subdivisions.

As a marker on its own, the V9 region does not offer enough variability for accurate taxonomy for the Dunaliella species. The V9 region lends itself to next generation sequencing methodologies, due to its heterogeneity and short length, however, the fact that 3-4 groups cannot be resolved using this marker raises questions on its suitability for analysis of hypersaline microbial communities as a significant portion

of Dunaliella diversity will be missed. The V2-V4 region of the 18S was found to have best phylogenetic resolution compared to the other hypervariable regions in dinoflagellates (Ki 2012), however, this was not tested here.

The section Dunaliella contains hyperhaline species that accumulate carotenes and includes the species, D. salina, D. pseudosalina and D. parva. Two clusters of D. salina have been previously identified, e.g. salina I and salina II based on the ITS2 sequences (Assunção et al. 2012). Strains isolated during this study fall into both these groups which can be distinguished from each other based on rbcL, LSU and ITS. One distinctive D. salina strain DF15 (high β-carotene producer), however, was significantly different and does not delineate with either groups. This strain is somewhat of an enigma with the different markers used clustering it (weakly) with different groups, e.g. DF15 clusters with DF17 and D. bardawil UTEX2538 in the LSU and rbcL phylogenies, respectively. Morphologically it has proven difficult to resolve differences between D. salina strains, however, 4 formae have previously been described in the literature. D. salina spp. salina fo. sibirica; a freshwater species with cells that are broader in the middle or anterior region (Massjuk and Radcênko 1973), D. salina spp. salina fo. oblonga; cylindrical cells of 7-28 µm length, 5-13 µm width, D. salina spp. salina fo. magna; cells ovoid with a cell volume >1000 μm^3 and cell length 7.5-29 µm, cell width 7.5-21 µm, and D. salina spp. salina fo. salina; which is similar to magna but with smaller cell dimensions. These formae have primarily been based on cell shape which can be an ambiguous criterion to use as cell shape can vary considerably based on differing culture conditions therefore we feel it unreliable to use further as a taxonomic criterion. Whilst it still may be valid to use these names to describe D. salina, with genetic data now available, we believe a revision in taxonomy is demanded. There is definitive D. salina taxa, i.e. hyper β carotene producers, with sequence information providing robust segregation into 3 clades. We propose that members of the previously identified salina clade II (Assunção et al. 2012) to be referred to as D. salina. This group contains the original isolate of what we know as D. salina, isolated in 1967 by Loeblich from Baja California and is deposited within the UTEX culture collection as #1644. However, this species should be subdivided into D. salina salina and D. salina aureus based on the phylogenetic analysis that sub-divides this D. salina clade into two. What has been identified as the salina I clade should be referred to as D. bardawil as it refers to the first isolate within this clade, D. bardawil, isolated in 1976. Although, there is some dispute over the existence of D. bardawil due to conflicting morphological and physiological data, the genetic data supports the delineation D. bardawil from D. salina. Whilst some studies have reported similarities between these species (Jahnke et al. 1999; Borowitzka and Siva 2007), other studies using genetically verified strains of D. salina and D. bardawil have identified physiological differences, e.g. Gomez et al. (2003) reported that D. salina CONC007 (verified as D. salina by Assunção et al. 2012) had the highest carotenoid to chlorophyll ratio in high salinities, whereas D. bardawil showed the opposite trend. Clearly, morphological and physiological properties can vary within a species so the genetic data is required to robustly classify the different isolates. Finally, we propose, the new clade that contains DF15 as the sole member as D. rubeus (rubeus - latin for red and the first

name for the giant in Harry Potter, Rubeus Hagrid) owing to its large size under high stress and intense red colouration. All of these strains should be classified under the section Dunaliella.

SA5 and 6 cluster with sequences from the tentatively designated pseudosalina by Assunção et al. (2012) and viridis IV clade as defined by Assunção et al. (2013). D. pseudosalina cells are classified as green to orange in colour with wide cylindrical, oval shaped cells exhibiting radial symmetry. Cells range from 11-23 µm length; 6-16 µm width with flagella 1.5 x cell length and a large distinctive elongated stigma (Massyuk 1973). D. parva can turn orange/red under high stress typically accumulating up to 5% β carotene dry weight (less than D. salina) and have a distinct stigma (Massyuk 1973). SA5 and SA6 are smaller cells that resemble those of the viridis clade rather than D. pseudosalina and furthermore do not accumulate β carotene as one would expect in D. parva. However, the phylogenetic analysis clearly separates them from the other D. viridis sequences available, assigned as viridis clade I by Assunção et al. (2013), with significant genetic distance. Furthermore, SA5 and 6 have a unique intron further supporting they are a distinct clade. Sequence analysis of the ITS region showed most homology (only 1 bp across 573bp sequence) to a strain isolated from India (MBTB-CMFRI-S118) with morphological similarities to our strains and interestingly the authors of this paper concluded that this strain was a Dunaliella salina allied strain (Preetha et al. 2012). Not having a type pseudosalina strain that we can genotype means we cannot ascertain whether these strains are actually this species genetically either or a new strain as the morphological data is not robust enough on its own. Therefore, we have assigned these as Dunaliella velox due to its fast swimming nature as it appears to be unique to other characterised strains of Dunaliella.

Section virides

The phylogenetic analysis of all genes (and the concatenated genes) defines three distinct groups of strains that have microscopically been affiliated to the section virides, as defined by Massjuk (1973), that are hypersaline species with radial symmetry and do not accumulate carotenoids. SA1 and SA2 cluster with other D. viridis strain sequences where available (Figs. 2, 3 and 5) and hence we have assigned these to this genus. A further 2 clades that are morphologically similar to the viridis group, including strains SA7, T32, DF48, T75, T34, T76 and T77, share a similar cell size and shape, yet are genetically distinct. Based on the descriptions of other species within the virides we can discount these clades as D. baas-beckingii, D. media, D. ruineniana, D. gracilis, D. carpatica, D. granulata, and D. terricola as the cell size and shape descriptions are incongruent with the morphology of these strains. Morphologically, the cells match the descriptions of D. viridis and D. minuta, with DF48, SA7 and T32 showing similarities to the description of D. bioculata with regards to cell shape and size and the presence of 2 stigmata. As the D. viridis sequences form a separate cluster, separated by considerable genetic distance this is discounted as the taxonomic affiliation of these strains. Based on the morphology and genetic analysis, we propose the following strains, T75, T34, T76, and T77 that share 99% identity for the concatenated genes, to be D. minuta with their sequences for LSU, rbcL and ITS clustering with Dunaliella sp. CCMP

367, which has a similar morphology to these strains (NCMA website). Unfortunately, the only available D. minuta in a culture collection is that of CCAP 19/5 which has now been re-classified as D. tertiolecta, so there is no type species for this 'minuta' group. We therefore propose Dunaliella minuta CCMP 367, isolated by Dodson from San Diego Bay, California, as the type species for D. minuta. Of note, T77 is significantly smaller than the other cells in this clade and is more spherical in shape and fits with the description of D. minutissima yet genetically is not distinct. Therefore, we propose that D. minutissima is actually another morphological variation of D. minuta. This minuta clade shares 96% identity with the DF48, SA7 and T32 clade but is separated by high bootstrap support (interestingly T32 groups with minuta for rbcL only). Whilst the cell size and shape are similar, in the case of DF48, SA7 and T32, they appear to have 2 eye spots and furthermore the intron classification separates the groups, indicating they are genetically distinct. The presence of 2 eyespots would automatically draw one to classify these strains as D. bioculata. Genetic data has made this difficult to verify as the original D. bioculata CCAP 19/4 AKA UTEX 199, described by Butcher, 1959, was re-assigned within the tertiolecta clade based on genetic analysis of the ITS2 (Assunção et al., 2012). This was following the re-identification, by Borowitzka and Siva (2007), who concluded the culture was now D. tertiolecta. Therefore, we propose these sequences to be truly representative of D. bioculata and represent the type species for this clade. Other strains examined by Borowitzka and Siva (2007), MUR26 originally identified as D. salina and MUR56 originally identified as D. viridis appear to actually be strains of D. bioculata as well, however sequences are not available for these as yet.

Based on the alignments and phylogenies, genetic data points to collapsing the Dunaliella and Tertiolectae into one section with D. tertiolecta showing high genetic similarity to representatives of the section Dunaliella, particularly when one compares how divergent the Section Virides is. Furthermore, other species classified within the Tertiolectae, including D. maritima, D. quartolecta, D. polymorpha and D. primolecta have not been shown to have differing sequences, in fact D. quartolecta CCAP 19/8, which was confirmed morphologically as this species by Borowitzka and Siva (2007) but has the same sequence as strains of D. tertiolecta (Assuncao et al. 2012). Morphological data used to delineate the Tertiolectae (Butcher 1959) has been criticised as unreliable and it is conceivable that the other members of this group are purely morphological variants of the same species, D. tertiolecta. We propose that the section Tertiolectae be absorbed within the section Dunaliella and this should also include the strains classified as D. velox.

In combination with morphological analysis the suite of genes used for molecular analysis has permitted the separation of D. salina isolates collected from South Africa, Israel, Namibia, Spain and Italy into three subspecies and identified three clusters within the section virides. Certainly, it appears to be beneficial to use several gene markers for phylogeny in order to generate a robust tree, however at the very least we are in agreement with other studies that the ITS regions are the most appropriate for resolving different clades. However, we propose that ITS2 should be used with ITS1 and 5.8S to provide even more genetic information. In accordance with Assunção et al. (2012) we have identified 'type strains' for

several Dunaliella taxa. Furthermore, strains deposited in culture collections require updating to reflect new information on their taxonomic affiliation as well as Genbank entries. Based on our conflicting genetic data regarding D. bardawil ATCC 30861/UTEX 2538/CCAP 19/30, which is supposedly the same strain, we strongly recommend that when a strain is acquired from a culture collection basic genetic screening (ITS) is warranted to confirm the identity of the strain and thus adding integrity to the scientific study.

Example 1

50ml water samples were collected from hypersaline environments and transported back to the laboratory. Dunaliella sp. were isolated using single cell picking with a micropipette and dilution techniques according to Andersen and Kawachi, 2005 (Table 1). Cultures were maintained in 30 mL F/2 media (Guillard and Ryther 1962) with the addition of 50 gL-1 sea salts (Sigma) (1.45 M NaCl), at 25°C continuous light of 100 photumols-1. Microscopy and cell measurements were undertaken 2 weeks post subculture into fresh media using a DMi8 live cell imaging system (Leica).

DNA was extracted from 10 mL late exponential cultures using the DNeasy blood and tissue kit (Qiagen) according to the manufacturer's instructions with the exception of the elution volume that was 50 ul. PCR was carried out using a suite of primers (Table 2) in a Corbett Thermocycler. PCR reactions were typically carried out in 50µl volumes containing 2µl DNA, 25pmol each primer, 1 x buffer 2.5mM MgCl2, 0.0025mM dNTPs, 1 Unit Gotaq polymerase (Promega) unless otherwise stated (Table 2). PCR reactions proceeded with an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of denaturation at 95°C for 30 seconds, annealing at 54°C for 45 seconds and extension at 72°C for 1 minute unless otherwise stated (Table 2). PCR reactions had a final extension step of 72°C for 5 minutes. PCR products were either sequenced directly using the respective primers (source bioscience) or in some cases cloning was necessary to ensure a single sequence was obtained. In these instances, PCR products were electrophoresed on a 1.2% (w/v) agarose gel in 1 x TAE and purified using the Zymoclean gel purification kit (Cambridge Biosciences). 1µl purified PCR product was ligated into the pCR2.1 TA cloning vector (Invitrogen) and transformed according to manufacturer's instructions.

Sequences were manually verified for quality using Chromas (Technelysium Pty Ltd). Multiple sequence alignments were constructed in BioEdit 7.0 (Hall 1999) using ClustalW with other available sequences from Genbank. Phylogenetic analysis based on neighbour-joining and maximum likelihood was undertaken using MEGA 6 (Tamura et al. 2013) and Bayesian using Geneious (Kearse et al. 2012). Bootstrap values were retrieved from 1000 replicates.

Genetic analyses reveal that five halotolerant Dunaliella strains (DF15, DF17, DF40, DF44 & DF48, Table 1) could be clustered into three main clades *Dunaliella salina* (DF15, DF17 & DF40), *D. tertiolecta* (DF44) and *Dunaliella minuta* (DF48) (Figure 8). In addition, the three *D. salina* strains clustered into three

distinct clades. DF17 and DF40 grouped with *D. salina* previously isolated from similar environments. However, DF15 was genetically distinct.

Example 2

We focussed our attention on the two *D. salina* strains DF15 and DF17 by performing controlled salt stress experiments. Triplicate cultures of both strains were grown in 1.8 M (final conc) seasalt-ErdSchrieber media under 175 µmol s^{-1} m^{-1} light in constant light at 25 °C. Cell counts were taken over a 30 day period and two time points selected (arrows, Figure 9) for total RNA extraction. In addition, at each of the second time points a salt stress exposure of 2.5 M sea salt for 2 hours was carried out. Total RNA was extraction from these salt stress exposures as well.

Four NEXTERA clone libraries for each strain were carried out on one of the culture replicates. Next-generation Illumina 2x300bp HiSeq RNAseq was carried out on these samples. A combined total of 170 million reads were achieved for both DF15 and DF17.

The raw sequence data obtained from the cDNA libraries were pooled and subjected to filtering and trimming of adaptors for cDNA synthesis, primers, poly (A/T) tails and potential contaminating vector sequences. Following the sequence trimming, the reads for each of the DF strains were assembled together using SeqMan NGen v 14 (DNASTAR). A total 86 million reads survived the QC filtering, assembled into 50,922 transcripts of an average length of 1,445kb with 32,383 having a size >1kb. Thirty eight percent (19,327) of the 50,922 transcripts returned a BLAST match, which in turn resulted in 6,674 unique gene assignments. Upon closer examination of the MEP/DOXP and GGPP pathways, which provides the precursors (isopentenyl diphosphate and dimethylallyl diphosphate) and geranylgeranyl diphosphate (GGPP), respectively, for the carotenogenic pathway, we found that both strains DF15 (Table 4) and DF17 (Table 5) remain largely unaffected with an increase in cell biomass (t1 vs t2), harvesting (t2 vs control) or salt stress (t2 vs salt stress).

We also note that certain genes occurred as single copies in both genomes (e.g. CMS or PDS), while others appear to have multiple orthologs (e.g. GGPPS). Moreover, different orthologs are preferentially expressed between strains. Most notable are the genes that encode for ZDS and lycopene cyclase that produces the compounds lycopene and beta-or alpha carotene, respectively (Tables 4 & 5). The specific orthologs in each strain will have different kinetic properties, which in turn will result in the different efficiencies in the overall production of carotene. This would therefore explain the different pigmentation as observed for DF15 (red) vs DF17 (yellow). In addition, the different lycopene cyclase othologs are likely to encode cyclases that result in either beta-carotene (red) or alpha-carotene-Lutein (yellow) accumulation.

Example 3

D. salina UTEX 2538 was obtained from the Culture Collection of Algae at The University of Texas at Austin (UTEX, Austin, TX, USA) and D. salina CCAP 19/30 was obtained from the Culture Collection of Algae and Protozoa at Scottish Marine Institute (CCAP, Scotland, UK). D-Factory strains DF15 and DF17 were isolated from a salt pond in Eilat, Israel, and DF40 was isolated from a salt pond in Monzon, Spain. The new isolates were identified as strains of or closely related to Dunaliella salina (bardawil) by The Marine Biological Association (MBA, Plymouth, Devon, UK) and are now deposited at the MBA culture collection (www.mba.ac.uk/culture-collection/). Algae were cultured in Modified Johnsons Medium (Borowitzka 1988) containing 10 mM $NaHCO_3$ with the pH value adjusted to 7.5 with 10 mM Tris-buffer, and 1.5 M NaCl, which has been tested as the optimal salinity for cell growth of the strains. Cultures were maintained in a temperature-controlled growth chamber at 20±2 °C with illumination provided under a 12 h light, 12 h dark cycle (12/12 LD) by white light emitting diode (LED) lights with a light intensity of ~200 µmol photons m^{-2} s^{-1}.

For algal cultivation, small stock cultures were grown to mid-log phase at 25°C in an incubator and diluted 1 in 50 (v/v) as inoculum for larger cultures in each experiment. Erlenmeyer flasks containing 500 mL culture each were maintained at 25°C in an ALGEM Environmental Modeling Labscale Photobioreactor (Algenuity, Bedfordshire, UK) with strictly controlled conditions of light, temperature and mixing level. Under 12/12 LD conditions, cell growth under a range of light intensities (200, 500, 1000, and 1500 µmol m^{-2} s^{-1}) of white LED light were compared. Each growth condition was set up at least in triplicate. Cell growth was monitored automatically in the bioreactor by recording the value obtained for light scatter at 725 nm in optical density (OD) units.

Cultures were mixed at 100 rpm for 10 min every hour before measuring the OD. Cell concentration was determined by counting the cell number in culture broth using a haemocytometer after fixing the cells with 2% formalin. The maximum specific growth rate of all cultures was calculated to compare cell growth under different conditions.

The Eclipse Ti-U inverted research microscope (Nikon, Tokyo, Japan) with a Nikon Digital Sight DS-Fi1 camera system was used to take brightfield microscopy photographs of cells of each Dunaliella strain. The objective lens used was Nikon Splan Fluor ELWD 60x/0.7 and the ocular lens was Nikon CFI 10x/22. The NIS-Elements Advanced Research Microscope Imaging Software (Nikon, Tokyo, Japan) was used to acquire the photos. Differential interference contrast (DIC) microscopy photographs were also obtained using a confocal microscope system ZEISS LSM 880 (Carl Zeiss Microscopy GmbH, Jena, Germany). The ZEISS Plan Apochromat 63x/1.4 oil DIC objective lens and the Carl Zeiss PI 10x/23 ocular lens were used. Images were acquired and analyzed through the ZEN 2.1 LSM software (Carl Zeiss Microscopy GmbH).

Algae grown under different light intensities were harvested during mid log phase of growth at the end of the light period. Pigments were extracted from the biomass harvested from 1 mL samples of the

cultures using 1 mL of 80% (v/v) acetone. The absorbance of the acetone extract after clarification at the centrifuge was measured at 480 nm for total carotenoids using an ultraviolet (UV)/Vis spectrophotometer. The content of total carotenoids was calculated according to Strickland & Parsons (1972). Chlorophyll a, b, and total chlorophyll were evaluated by measuring absorbance of the acetone extract at 664 nm and 647 nm and calculated according to Porra et al (1989).

The compositions of pigments extracted from different strains were analyzed using high performance liquid chromatography (HPLC) with diode array detection (DAD) (Agilent Technologies 1200 series, Agilent, Santa Clara, CA, USA). Carotenoid standards of all-trans α-carotene, all-trans β-carotene and zeaxanthin were obtained from Sigma-Aldrich, Inc. (Merck KGaA, Darmstadt, Germany). Lutein and 9-cis β-carotene were obtained from Dynamic Extractions (Tredegar, UK). Carotenoids and chlorophylls were extracted from freshly harvested cells using methyl tert-butyl ether (MTBE) and Methanol (MeOH) (20:80) as extraction solvent. 15 mL of algal culture was centrifuged at 3000 g at 18 °C for 5 min and the pellet was extracted with 10 mL MTBE-MeOH (20:80) and sonicated for 20 s. The sample was clarified by centrifugation at 3000 g at 18 °C for 5 min, then 1-2 mL of the supernatant was filtered through 0.45 µm syringe filter into amber HPLC vials. It was then analyzed using a YMC30 250 x 4.9 mm I.D S-5µ HPLC column with DAD at 25 °C , and isocratic elution with 80% methanol: 20% MTBE, flow rate of 1 ml min^{-1}, pressure of 90 bar. The quantities of 9-cis and all-trans β-carotene, all-trans β-carotene, lutein, and zeaxanthin in the biomass were determined from the corresponding standard curves. Glycerol, known to be regulated by salinity, was determined according to the procedures described in a Xu et al (2016).

Cells were harvested during the exponential phase and $NaHCO_3$ was added to a final concentration of 10 mM 5 min before the start of each measurement. The rates of net O_2 evolution and dark respiration were measured as described by Brindley et al. (2010) at 25°C using a Clark-type electrode (Hansatech Instruments Ltd, Norfolk, UK). O_2 evolution was induced with 1500 µmol photons m^{-2} s^{-1} actinic light. After initial 30 min of dark adaption, O_2 evolution was measured for 5 min followed by dark respiration for 20 min. The average net rate of photosynthesis was then determined from the oxygen concentration gradient recorded over 5 min, dO_2/dt. Dark respiration was determined by following the same procedure, except that oxygen uptake was calculated from data recorded during the last 5 min of the 20 min experiment. Sodium dithionite was used to calibrate the oxygen electrode.

The data generated in this study was analyzed in R (Rstudio, Boston, MA, USA). A two-way analysis of variance (ANOVA) analysis was performed to study the relationships of a series of variables measured with two factors in this work: strain and light intensity. The two-way ANOVA tests three omnibus effects: the main effect of strain or light intensity, and the interaction effect between these two factors. Correlation analysis was used to evaluate the association between each pair of the variables and the Pearson correlation method was chosen to measure the linear dependence between two variables. In

correlation analysis, a correlation coefficient (the Pearson Product Moment correlation coefficient) was estimated for each pair of the variables studied. Whether or not an observed correlation is statistically significant or not was evaluated by P values (significant when $P \leq 0.05$). Hierarchical cluster analysis is based on the strength of the correlations and the distance in the clustering dendrogram reflects the dissimilarity among these parameters. Traits examined with strong correlations are grouped as a cluster. A principle component analysis was carried out using the whole data set to reveal the relatedness between the examined traits.

Growth curves for the five strains cultivated under the same conditions of different light intensities of 200, 500, 1000, and 1500 μmol m^{-2} s^{-1} are shown in Figure 11, from which the maximum specific growth rate was calculated for each growth condition. Generally, these strains grew at a faster rate under higher light intensities. This is clearly shown for CCAP 19/30 and DF17. All strains showed the slowest growth rates under 200 μmol m^{-2} s^{-1} light intensity. In DF15 and UTEX 2538, when increasing the light intensity from 1000 μmol m^{-2} s^{-1} to 1500 μmol m^{-2} s^{-1}, no further improvement in cell growth rate was observed. It is likely that the optimal light intensity for fastest growth of DF15 or UTEX 2538 is around 1000 μmol m^{-2} s^{-1}, while 1500 μmol m^{-2} s^{-1} or higher is optimal for the other three strains under the specific growth conditions used (white LED lights under 25 °C with 10 min mixing at 100 rpm every hour). DF15 had the slowest growth rate and CCAP 19/30 the fastest.

Figure 12a shows that as the light intensity increased, the rate of photosynthesis decreased for DF17, DF40 and CCAP 19/30, indicating that these three strains are susceptible to photoinhibition. However, DF15 and UTEX 2538 did not exhibit photoinhibition with increase in light intensity, suggesting that these two strains have a more robust photoprotection mechanism. Figure 12b shows that the dark respiration rate patterns were similar for DF17, DF40, and CCAP 19/30. These three strains showed a slight decrease or no change in dark respiration rate with the increase in light intensity. DF15 and UTEX 2538 had a similar pattern to each other and their respiration rate increased slightly with increase in light intensity. From statistical analysis using two-way ANOVA, both strain difference and light intensity were significant factors affecting photosynthesis; less significant was the interaction between them. However, light intensity showed no significant impact on dark respiration, but strain played a major role in the observed differences in dark respiration (Table 7).

Cellular contents of total chlorophyll and total carotenoids were determined for the five *Dunaliella* strains grown under the four light intensities (200, 500, 1000, and 1500 μmol m^{-2} s^{-1}) using UV/Vis spectrometry (Figure 13). Generally, the cellular content of total chlorophyll decreased while total carotenoids increased with the increase in light intensity for all five *Dunaliella* strains. Statistical analysis showed that strain difference significantly affected total carotenoids and total chlorophyll content, although total carotenoids and total chlorophyll content also responded significantly to light intensity (Table 7).

HPLC-DAD was used to quantify the contents of major carotenoids, namely lutein, zeaxanthin, *all-trans* β-carotene, *9-cis* β-carotene, and *all-trans* α-carotene, in each strain acclimated in response to four light intensities, to understand the effect of light in carotenoid metabolism. Figure 14 shows HPLC chromatograms of the pigment extracts from the five *Dunaliella* strains grown under the light intensity of 1500 μmol m^{-2} s^{-1}. It is clear that CCAP 19/30 does not accumulate β-carotene even under high light intensity. DF15, DF40 and UTEX 2538 have a similar pigment profile and β-carotene dominates the carotenoid composition. DF17 produced a higher relative amount of zeaxanthin under high light stress compared with the other strains, indicating the important role of zeaxanthin in DF17 for photoprotection.

The major difference between the strains was their ability to accumulate β-carotene. As shown in Figure 15a, b, the contents of all-trans and 9-cis β-carotene increased with increasing light intensity in all five strains apart from UTEX 2538, which produced the highest cellular amount of all-trans (5.6 ± 1.8 pg $cell^{-1}$) or 9-cis β-carotene (5.3 ± 1.5 pg $cell^{-1}$) under 1000 μmol m^{-2} s^{-1}. Under the highest light intensity studied (1500 μmol m^{-2} s^{-1}), the cellular contents of all-trans β-carotene and 9-cis β-carotene were 9.0 ± 0.7 and 5.9 ± 0.6 pg $cell^{-1}$ in DF15; 1.1 ± 0.5 and 0.8 ± 0.4 pg $cell^{-1}$ in DF40; and 0.6 ± 0.0 and 0.1 ± 0.0 pg $cell^{-1}$ in DF17. In CCAP 19/30, the highest all trans β-carotene content (0.3 ± 0.0 pg $cell^{-1}$) was obtained at 1000 μmol m^{-2} s^{-1}, and only a very small amount of 9-cis β-carotene was detected at all light intensities (~0.01 pg $cell^{-1}$). All five strains achieved the highest all-trans β-carotene productivity at 1500 μmol m^{-2} s^{-1} (3.2 ± 0.0, 3.5 ± 0.0, 1.3 ± 0.0, 2.6 ± 0.0 and 2.9 ± 0.0 mg L^{-1} day^{-1} for CCAP 19/30, DF15, DF17, DF40 and UTEX 2538 respectively), and also the highest 9-cis β-carotene productivity at 1500 μmol m^{-2} s^{-1} except that UTEX 2538 has the highest yield of 9-cis β-carotene at 1000 μmol m^{-2} s^{-1} (0.2 ± 0.0, 2.3 ± 0.0, 0.2 ± 0.0, 2.0 ± 0.0 and 2.2 ± 0.0 mg L^{-1} day^{-1} for CCAP 19/30, DF15, DF17, DF40 and UTEX 2538 respectively). From the two-way ANOVA analysis, the cellular contents of all-trans or 9-cis β-carotene were found to vary significantly among strains and under different light intensities (Table 7). CCAP 19/30, DF17 and DF40 had similar responses to increasing light with a mild β-carotene accumulation, while DF15 and UTEX 2538 significantly increased β-carotene content with increasing light (Figure 15a, b). DF15 and UTEX 2538 have significantly higher cellular contents of all-trans- or 9-cis β-carotene than the other three strains and DF15 contains a higher cellular content of β-carotene than UTEX 2538 under most of the light conditions. UTEX 2538, already known to be a massive carotene-accumulating strain, had faster growth rates than DF15 under all light intensities examined, as shown in Figure 11f. On the other hand, DF15 accumulated a high carotene content even under the lowest light intensity tested here. In Dunaliella, variation in β-carotene content has been reported to correlate with the integral irradiance received during a division cycle and to be a specific mechanism of photoprotection, which may explain why DF15 has a higher cellular content of β-carotene than UTEX 2538. DF15 has the advantage of accumulating a large amount of β-carotene even without light stress (Figure 15), and also highest productivity of both all-trans and 9-cis β-carotene under light stress,

therefore has great potential for the commercial production of β-carotene with less light energy input required.

The cellular content of lutein in the five Dunaliella strains grown under various light intensities is shown in Figure 15c. All strains accumulated considerably different amounts of lutein and the response to increasing light intensities varied among different strains. Lutein increased with light intensity from 200 μmol m^{-2} s^{-1} to 1000 μmol m^{-2} s^{-1} and then decreased when light increased to 1500 μmol m^{-2} s^{-1} in UTEX 2538. In DF15, lutein content did not change with light intensity from 200 μmol m^{-2} s^{-1} to 1000 μmol m^{-2} s^{-1} and only increased from 1000 to 1500 μmol m^{-2} s^{-1}. Both DF15 and UTEX 2538 accumulated significantly larger amounts of lutein under high light compared with the other strains. DF17 had the highest lutein content at 1000 μmol m^{-2} s^{-1}, and the lowest at 1500 μmol m^{-2} s^{-1}. Two-way ANOVA shows the cellular content of lutein is significantly affected by both the strain and light intensity.

Figure 15d shows that zeaxanthin content in all strains increased with light intensity. DF15 accumulated the highest amount of zeaxanthin, followed by DF17, UTEX 2538, DF40 and CCAP 19/30 accumulated the lowest amount. Two-way ANOVA analysis (Table 7) shows that the factors of strain and light intensity determined the accumulation of zeaxanthin. Zeaxanthin accumulation was significantly different among strains and at different light intensities. Among the different strains, DF17 and UTEX 2538 had similar responses in terms of zeaxanthin accumulation.

The cellular content of all-trans α-carotene of the five strains grown under different light intensities is shown in Figure 15e and the cellular content of glycerol is shown in Figure 15f. The content of all-trans α-carotene in DF15 or UTEX 2538 was much higher than that in the other three Strains. α-carotene is the precursor of lutein but surprisingly α-carotene did not respond to light stress in the same way as lutein. All-trans α-carotene increased with the light intensity in all strains examined, and its response to increasing light intensity was very similar to the pattern of accumulation obtained for all-trans and 9-cis β-carotene.

Whilst the accumulated data permit elucidation of strain differences for carotenoid production, they also provided the opportunity to explore the use of statistical analysis to provide new insights into carotenoid metabolism coupled to the interdependent metabolic functions of photosynthesis and respiration. This was possible with the large set of data generated across five strains and four light intensities combined with tools of ANOVA analysis, correlation analysis, and principal component analysis used in this study. With the quantitative data obtained for the five Dunaliella strains, statistical analysis was used as a tool in order to assess the strength of the correlations among the carotenoids and other cell growth parameters and examine the differences among the five strains. A correlation and clustering analysis was performed on the growth, photosynthesis and pigment data presented, to all five strains grown under four light conditions. The analysis was performed for each strain using all variables

examined in this study (all-trans β-carotene, 9-cis β-carotene, glycerol, lutein, zeaxanthin, all-trans α-carotene, photosynthesis, respiration, total carotenoids, total chlorophyll, and specific growth rate). Among them, glycerol is known to maintain osmotic balance in Dunaliella strains and as expected, the cellular content of glycerol would not respond to changes in light intensity, as shown in Figure 15f. Glycerol content therefore was used to index the analysis.

The clustering dendrogram of the examined traits for each strain is shown in Figure 16 and depicts graphically several features of note amongst the strains. First, it shows that the individual carotenoids of all-trans β-carotene, 9-cis β-carotene, zeaxanthin and all trans α-carotene in the four D. salina strains are strongly correlated with each other but significantly not with lutein, except in CCAP 19/30. From this, it is clear that there is greater similarity between the four D. salina strains (DF15, DF17, DF40, and UTEX 2538) than with the CCAP19/30 strain. Second, the correlation analysis shows that accumulation of carotenoids is positively correlated with photosynthesis over all light intensities for the D. salina strains (also shown in Figure 17), signifying a role for carotenoids in photoprotection. Third, lutein is not correlated closely with the other carotenoids, but correlates more strongly with photosynthesis and respiration. This result suggests an important and not hitherto identified role for lutein in coordinated control of the cellular functions of photosynthesis and respiration in response to changes in light conditions, which is moreover broadly conserved in Dunaliella strains. Glycerol, which was not expected to change with light intensity, is weakly correlation with the different carotenoids in the Dunaliella strains as anticipated, but also correlates more closely with either photosynthesis or respiration.

A principle component analysis was performed with all strains growing at all tested conditions as shown in Figure 17. The examined 11 traits can be roughly grouped into four groups as shown in the graph, where all-trans α-carotene, all-trans β-carotene, 9-cis β-carotene and zeaxanthin were clustered closely, lutein, respiration and total chlorophyll were found in a second cluster, glycerol and photosynthesis were closely correlated, and the specific growth rate stands separately. The formation of two separate clusters of the carotenoids indicates two functionally distinct mechanisms for coordinated adaptation to changes in light conditions, broadly conserved between DF15, DF40, CCAP 19/30, DF17 and UTEX 2538. More importantly, it shows that DF17 and DF40 performed similarly under the tested environmental conditions; that DF15 is closely related to UTEX 2538, and that CCAP 19/30 is different compared to all the other strains.

It is noteworthy that the statistical analysis based on the data obtained from the biochemical characterization suggests a grouping of the five strains into three different groups: (1) DF15 and UTEX 2538; (2) DF17 and DF40; and (3) CCAP 19/30 as shown in Figure 17. However, genetic using the approaches of bar coding shows a higher similarity between DF40 and UTEX 2538, and therefore groups the five strains into four different groups: (1) DF40 and UTEX 2538, (2) DF17, (3) DF15, and (4)

CCAP 19/30 as shown in the phylogenetic tree (Figure 4-6). This indicates the complicity in Dunaliella by using a single classification method, and the importance of strain selection for the commercial production of Dunaliella biomass and natural β-carotene.

Although illustrative embodiments of the invention have been disclosed in detail herein, with reference to the accompanying drawings, it is understood that the invention is not limited to the precise embodiments shown and that various changes and modifications can be effected therein by one skilled in the art without departing from the scope of the invention. Furthermore, although individual embodiments of the invention may have been described, the invention is intended to include combinations of those embodiments.

References

Alverson AJ, Kolnick L (2005) Intragenomic nucleotide polymorphism among small subunit (18S) rDNA paralogs in 377 the diatom genus Skeletonema (Bacillariophyta). J Phycol 41: 1248-1257

Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using 379 massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS ONE 4: e6372

Assunção P, Jaén-Molina R, Caujapé-Castells J, de la Jara A, Carmona L, Freijanes K, Mendoza H (2012) 382 Molecular taxonomy of Dunaliella (Chlorophyceae), with a special focus on D. salina: ITS2 sequences revisited with as 383 extensive geographical sampling. Aquat Biosys 8: 2.

Assunção P, Jaén-Molina R, Caujapé-Castells J, Wolf M, Bucheim MA, de la Jara A, Freijanes K, Carmona L, 386 Mendoza H (2013) Phylogenetic analysis of ITS2 sequences suggests the taxonomic re-structuring of Dunaliella viridis 387 (Chlorophyceae, Dunaliellales). Phycol Res 61: 81-88

Avron M, Ben-Amotz A (1992) Eds. Dunaliella: Physiology, Biochemistry, and Biotechnology. Boca Raton: CRC Press.

Ben-Amotz A, Avron M (1983) On those factors which determine the massive β-carotene accumulation in the 397 halotolerant alga Dunaliella bardawil. Pl Physiol 72:593–597

Borowitzka LJ, Brown AD (1974) The salt relations of marine and halophilic species of the unicellular green algae, Dunaliella. Arch Microbiol 96: 37–52.

Borowitzka, M.A. (1988) Algal growth media and sources of cultures. In Microalgal Biotechnology; Borowitzka, M.A., Borowitzka, L.J., Eds.; Cambridge University Press: Cambridge, UK; pp. 456–465.

Borowitzka MA, Siva CJ (2007) The taxonomy of the genus Dunaliella (Chlorophyta, Dunaliellales) with emphasis on 395 the marine and halophilic species. J Appl Phycol. 19: 567-590.

Brindley, C.; Acién, F.G.; Fernández Sevilla, J.M. The oxygen evolution methodology affects photosynthetic rate measurements of microalgae in well-defined light regimes. Biotechnol. Bioeng. 2010, 106, 228–237.

Bucheim MA, Kirkwood AE, Bucheim JA, Verghese B, Henley WJ (2010) Hypersaline soil supports a diverse community of Dunaliella (Chlorophyceae). J Phycol 46: 1038-1047.

Cook SS, Whittock L, Wright SW, Hallegraeff GM (2011) Photosynthetic pigment and genetic differences between 403 two Southern Ocean morphotypes of Emiliania huxleyi (Haptophyta). J Phycol. 47: 615-626.

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C, Le Bescot N, Probert I, Carmichael M, Poulain J, Romac S, Colin S, Aury JM, Bittner L, Chaffron S, Dunthorn M, Engelen S, Flegontova 407 O, Guidi L, Horák A, Jaillon O, Lima-Mendez G, Lukeš J, Malviya S, Morard R, Mulot M, Scalco E, Siano R, Vincent F, Zingone A, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Acinas SG, Bork P, Bowler C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Not F, Ogata H, Pesant S, Raes J, Sieracki ME, Speich S, Stemmann L, 410 Sunagawa S, Weissenbach J, Wincker P, Karsenti E (2015) Eukaryotic plankton diversity in the sunlit ocean. Science 411 348: 6237

Guillard RRL, Ryther JH (1962) Studies of marine planktonic diatoms. I. Cyclotella nana (Hustedt) and Detonula 414 confervaceae (Cleve) Gran. Can J Microbiol 8: 229-239

Gómez PI, Barriga A, Cifuentes AS, Gonzalez MA (2003) Effect of salinity on the quantity and quality of carotenoids 417 accumulated by Dunaliella salina (strain CONC-007) and Dunaliella bardawil (strain
5 ATCC 30861) Chlorophyta. Biol 418 Res 36: 185-192.

González MA, Coleman AW, Gómez PI, Montoya R (2001) Phylogenetic relationship among various strains of 421 Dunaliella (Chlorophyceae) based on nuclear ITS rDNA sequences. J Phycol. 37: 604–611.

Hall, TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 423 95/98/NT. Nucl Acid S. 41:95-98
10 Hejazi MA, Barzegari A, Gharajeh NH, Hejazi MS (2010) Introduction of a novel 18S rDNA gene arrangement along with distinct ITS region in the saline water microalga Dunaliella. Saline Systems. 6:4

Jahnke LS. Massive carotenoid accumulation in Dunaliella bardawil induced by ultraviolet-A radiation. J Photochem Photobiol B: Biol. 1999; 48:68–74. doi: 10.1016/S1011-1344(99)00012-3.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S,
15 Duran C, Thierer T, Ashton B, Mentjies P, Drummond A (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28: 1647-1649.

Ki, JS (2012) 18S rRNA using a large dataset for marker considerations. J Appl Phycol. 24: 1035-1043.

Lane DJ (1991) 16S/23S sequencing. In: Nucleic Acid Technologies in Bacterial Systematic. Stackebrandt
20 E, 433 Goodfellow M (Eds) pp. 115–175, Wiley, NY.

Leliaert F, Verbruggen H, Vanormelingen P, Steen F, Lopez-Bautista JM, Zuccarello GC, De Clerck O (2014) 435 DNA-based species delimitation in algae. Eur J Phycol. 49: 179-196

Massjuk NP (1973) Morphology, taxonomy, ecology and geographic distribution of the genus Dunaliella Teod. and prospects for its potential utilisation. Naukova Dumka, Kiev [original in Russian]
25 Massyuk NP, Radchenko MI (1973) New taxa of the genus Dunaliella Teod. III. Ukr Bot Zh 30:470–471

Medlin LK, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-442 like rRNA-coding regions. Gene. 71 491–499.

Moniz MBJ, Guiry M, Rindi F (2014) tufA phylogeny and species boundaries in the green algal order Prasiolales 444 (Trebouxiophyceae, Chlorophyta). Phycologia. 53: 396-406.
30 Muller J (2005) Genetic fingerprints of microalgal culture strains: amplified fragment length polymorphism (AFLP) for 447 investigations below the species level. PhD thesis. Universität zu Göttingen.

Nozaki H, Itoh M, Sano R, Uchida H, Watanabe MM, Kuroiwa T (1995) Phylogenetic relationships within colonial 449 volvocales (Chlorophyta) inferred from rbcL gene sequence data. J Phycol. 31: 970-979.
35 Olmos J, Paniagua J, Contreras R (2000) Molecular identification of Dunaliella sp. utilizing the 18S rDNA gene. Lett 451 Appl Microbiol. 30: 80–84.

Olmos SJ, Ochoa L, Paniagua-Michel J, Contreras R (2009) DNA fingerprinting differentiation between β- carotene hyperproducer strains of Dunaliella from around the world. Saline Systems. 5:5.

Olmos-Soto J, Paniagua-Michel J, Contreras PR, Trujillo L (2002) Molecular identification of β-carotene hyper-456 producing strains of Dunaliella from saline environments using species-specific oligonucleotides. Biotech Lett. 24: 365–369.

Porra, R.J.; Thompson, W.A.; Kriedemann, P.E. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: Verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim. Biophys. Acta Bioenerg. 1989, 975, 384–394.

Preetha K, John L, Subin CS, Vijayan KK (2012) Phenotypic and genetic characterisation of Dunaliella (Chlorophyta) 460 from Indian salinas and their diversity. Aquat Biosys. 8: 27. 461

Scholin CA, Herzog M, Sogin M, Anderson DM (1994) Identification of group- and strain-specific genetic markers for 463 globally distributed Alexandrium (Dinophyceae). II. Sequence analysis of a fragment of the LSU rRNA gene. J Phycol. 464 30: 999–1011.

Strickland, J.; Parsons, T.R. A Practical Handbook of Seawater Analysis, 2nd ed.; Fisheries Research Board of Canada: Ottawa, Canada, 1972; pp. 185–190.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis 466 Version 6.0. Mol Biol Evol. 30: 2725-2729 467

Wilcox LW, Lewis LA, Fuerst PA, Floyd GL (1992) Group I introns within the nuclear-encoded small-subunit rRNA 468 gene of three green algae. Mol Biol Evol. 9: 1103-1118.

Xu, Y., Ibrahim, I. M. & Harvey, P. J. The influence of photoperiod and light intensity on the growth and photosynthesis of Dunaliella salina (chlorophyta) CCAP 19/30. Plant Physiol. Biochem. 106, 305–315 (2016).

Table 1: *Dunaliella* sp. isolates collected during this study

MBA culture collection	Habitat	Geographical origin	Month and Year Collected	Cell length (µm) Mean/Range	Cell width (µm) Mean/Range	Original microscopic identification	Cell colour at 25°C, constant light, 1.5M NaCl
DF15	Saltpan	Eilat, Israel	03/2014	18.22 (14.4-21.5)	13.82 (10.4-18.7)	*Dunaliella salina*	Red
DF17	Raceway	Eilat, Israel	03/2014	12.65 (10.9-15.5)	10.26 (8.67-12.9)	*Dunaliella salina*	Yellow
DF40	Raceway	Monzon, Spain	05/2015	14.37 (13.7-15.1)	13.28 (12.0-14.6)	*Dunaliella salina*	Red
DF41*	Raceway	Monzon, Spain	11/2015	N/A	N/A	*Dunaliella salina*	Red
DF45	Saltpan	Monzon, Spain	11/2015	13.63 (11.5-17.2)	10.84 (7.9-13.2)	*Dunaliella salina*	Red
DF44	Saltpan	Monzon, Spain	11/2015	8.56 (7.52-9.31)	7.63 (6.49-8.60)	*Dunaliella tertiolecta*	Green
DF48	Saltpan	Monzon, Spain	1/2015	7.49 (6.02-9.01)	4.90 (4.1-6.48)	*Dunaliella* sp. (Viridis)	Green
SA1	Saltpan	Velddrif, R.S.A.	02/2012	6.38 (4.47-7.56)	3.53 (2.85-4.12)	*Dunaliella* sp. (Viridis)	Green
SA2	Saltpan	Velddrif, R.S.A.	02/2012	7.04 (6.21-8.52)	4.06 (3.12-4.93)	*Dunaliella* sp. (Viridis)	Green
SA3	Saltpan	Velddrif, R.S.A.	02/2012	13.90 (12.9-15.1)	11.68 (10.4-12.9)	*Dunaliella salina*	Orange
SA4	Saltpan	Velddrif, R.S.A.	02/2012	15.09 (12.9-17.2)	12.19 (10.4-14.2)	*Dunaliella salina*	Orange
SA5	Saltpan	Velddrif, R.S.A.	02/2012	7.61 (6.99-8.49)	5.49 (4.76-6.43)	*Dunaliella* sp.	Green
SA6	Saltpan	Velddrif, R.S.A.	02/2012	8.13 (7.44-9.12)	6.41 (5.60-7.60)	*Dunaliella* sp.	Green
SA7	Saltpan	Velddrif, R.S.A.	02/2012	7.22 (5.78-8.48)	3.82 (3.40-4.20)	*Dunaliella* sp. (Viridis)	Green
T32	Saltpan	Swakopmund, Namibia	06/2010	6.41 (5.13-7.84)	3.35 (2.94-3.96)	*Dunaliella* sp. (Viridis)	Green
T34	Saltpan	Swakopmund, Namibia	06/2010	6.27 (5.56-7.31)	3.92 (3.20-4.46)	*Dunaliella* sp. (Viridis)	Green
T36	Saltpan	Swakopmund, Namibia	06/2010	17.98 (15.9-20.7)	12.83 (9.8-16.1)	*Dunaliella salina*	Orange
T37	Saltpan	Swakopmund, Namibia	06/2010	16.06 (11.9-19.3)	12.68 (8.81-16.0)	*Dunaliella salina*	Orange
T41	Saltpan	Swakopmund, Namibia	06/2010	15.3 (13.4-17.9)	12.2 (10.0-13.5)	*Dunaliella salina*	Orange
T68	Saltpan	Port Elizabeth, R.S.A.	04/2013	15.72 (13.3-18.1)	15.15 (12.9-17.9)	*Dunaliella salina*	Orange
T74*	Saltpan	Velddrif, R.S.A.	04/2013	N/A	N/A	*Dunaliella salina*	N/A
T75	Saltpan	Port Elizabeth, R.S.A.	04/2013	6.87 (5.63-7.95)	3.20 (2.80-3.96)	*Dunaliella* sp. (Viridis)	Green
T76	Saltpan	Trapani, Italy	12/2012	6.55 (5.97-7.25)	3.28 (2.36-3.65)	*Dunaliella* sp. (Viridis)	Green
T77	Saltpan	Trapani, Italy	12/2012	2.33 (2.11-2.81)	2.04 (1.51-2.55)	*Dunaliella* sp. (Viridis)	Green
UTEX 999	Fjord	Oslo, Norway	1938	9.44 (8.21-11.20)	7.59 (6.17-8.70)	*Dunaliella tertiolecta**	Green
CCAP 19/30	Bardawil lagoon	North Sinai, Israel	1978	10.99 (9.9-13.8)	7.85 (6.05-9.22)	*Dunaliella bardawil* now *tertiolecta**	Green
UTEX 2538				21.27 (17.4-23.7)	18.78 (15.2-22.6)	*Dunaliella bardawil*	Red
ATCC 30861				14.33 (10.4-18.4)	10.51 (8.3-13.2)	*Dunaliella bardawil*	Red

*Isolates lost from culture

Table 2: Primers used in this study

Primer name	Target	Sequence (5'-3')	Size (bp)	Anneal Temp (°C)	PCR reaction	Reference
AB1F	ITS (Universal)	AATCTATCAATAACCACACCG	700	52	20pmol primer; 1.5mM $MgCl_2$; 1min extension	Hejazi et al., 2010
AB2R		TTTCATTCGCCATTACTAAGG				
MA1F	18S Dunaliellaceae	CGGGATCCGTAGTCATATGCTTGT CTC	1700-2500	52	Extension time 3 minutes; 30 pmol primers; 1.5mM $MgCl_2$	Olmos et al. 2009
MA2R		CGGAATTCCTTCTGCAGGTTCACC				
rbcLF	rbcL Chlorophyta	CGTGACAAATTAAACAAATATGG	700	54	As text	Nozaki et al., 1995
rbcLR		AAGATTTCAACTAAAGCTGGCA				
tufA_F	tufA Chlamydomonales	GAAGARAAAGCWCGYGGTATTAC	750	54	As text	This study
tufA_R		CCATACCRCGTTCRAIDICTT				
1391F	Universal Eukaryotic 18S (V9)	GTACACACCGCCCGTC	168	54	25µl reaction; 10pmol primer; 1mM $MgCl_2$; 0.5U taq; 20 sec anneal extend	Lane et al., 1991; Medlin et al., 1988
EukB		TGATCCTTCTGCAGGTTCACCTAC				
D1R_F	28S Universal	ACCCGCTGAATTTAAGCATA	760	60	20pmol primers; 1.5mM $MgCl_2$; 1 min extension	Scholin et al., 1994
D2C_R		GCTTGGTCCGTGTTTCAAGA				

Table 3: Results from intron sizing and sequence alignment of introns

Isolate name	# Introns	Intron 1 group	Intron group 2
DF45	2	A	P
DF41	2	A	P
DF40	2	A	P
D. bardawil ATCC 30861	0	-	-
D. bardawil UTEX 2538	1	A	-
DF17	0	-	-
T74	0	-	-
SA3	0	-	-
SA4	0	-	-
T36	0	-	-
T37	0	-	-
T41	0	-	-
T68	0	-	-
DF28	1	F	-
DF15	1 and 2	B	Q
SA5	1	N/A	R
SA6	1	N/A	R
D. tertiolecta UTEX 199	0	-	-
DF44	0	-	-
Dunaliella sp. CCAP 19/30	0	-	-
SA1	1	N/A	S
SA2	1	N/A	S
AW1(DF48)	2	E	U
T76	1	N/A	T
T77	1	N/A	
T34	2	D	T
T75	2	C	T
SA7	2	E	U
T32	2	E	U

Table 4

Pathway	Gene Name	Cluster	Acronyms or synonyms	Compound(s)	EC	DF15t1	DF15t2	DF15t2-control	DF15t2-salt stress
Non-mevalonate pathway for terpenoid synthesis (MEP/DOXP pathway) that occurs in chloroplasts	1-deoxy-D-xylulose-5-phosphate_synthase	cl_857208_2 cl_535498_12	DXS or DOXP synthase	Pyruvate + D-glyceraldehyde 3-phosphate ⇩ 1-deoxy-D-xylulose 5-phosphate	2.2.1.7	181.1165 [illegible]	105.172 [illegible]	[illegible]	[illegible]
	1-deoxy-D-xylulose_5-phosphate_reductoisomerase	DXR1_XM_00169: cl_45366_21 cl_45366_7	DXR or DOXP synthase	⇩	1.1.1.267	237.462 100.967 421.751	243.372 105.35 593.643	200.194 98.865 445.32	161.105 47.416 457.441
	4-diphosphocytidyl-2C-methyl-D-erythritol_synthase	cl_175958_14	CMS	2-C-methyl-D-erythritol 4-phosphate ⇩ diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	2.7.7.60	[illegible]	[illegible]	[illegible]	[illegible]
	4-diphosphocytidyl-2-C-methyl-D-erythritol_kinase	cl_187509_1 cl_187509_2 cl_708851_2 cl_708851_6 cl_639608_18	CMK	⇩	2.7.1.148	[illegible] [illegible] [illegible] 0.1205 0.8335	[illegible] [illegible] [illegible] 0 2.33	[illegible] [illegible] [illegible] 0.183 1.631	[illegible] [illegible] [illegible] 0.716 0.944
	2-C-methyl-D-erythritol_2,4-cyclodiphosphate_synthase	cl_762165_19 cl_762165_1 cl_762165_17	MCS	2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol ⇩ 2-C-methyl-D-erythritol 2,4-cyclodiphosphate	4.6.1.12	5.0385 148.5085 0.821	[illegible] 160.574 0.347	[illegible] 182.126 0.54	9.607 165.207 1.407
	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase or 1-hydroxy-2-methyl-2-(E)-butenyl_4-diphosphate_synthase	cl_459175_35 cl_459175_21	HDS (ferredoxin)	⇩ (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate	1.17.7.1	258.6705 253.91	189.35 [illegible]	104.842 279.465	[illegible] 314.937
	4-Hydroxy-3-methylbut-2-enyl diphosphate reductase	cl_665162_3 cl_665162_6 cl_665162_10 cl_665162_26 cl_462991_1 cl_462991_3	HDR or HMBPP reductase	⇩ isopentenyl diphosphate + dimethylallyl diphosphate	1.17.7.4	[illegible] 0.006 0.006 0.006 288.167 [illegible]	[illegible] 0.012 0.012 0.012 305.715 261.577	[illegible] 0.012 0.012 0.012 [illegible] [illegible]	[illegible] 0.012 0.012 0.012 [illegible] [illegible]
geranylgeranyl diphosphate synthesis in cytoplasm, ER or mitochondrion	Dimethylallyltranstransferase.	cl_531037_8 cl_531037_9	FPP synthase or Farnesyl pyrophosphate synthase	Dimethylallyl diphosphate + isopentenyl diphosphate+ (2E,6E)-farnesyl diphosphate ⇩ geranylgeranyl diphosphate (GGPP)	2.5.1.1	n/a n/a	n/a n/a	n/a n/a	n/a n/a
	Geranylgeranyl_diphosphate_synthase	cl_717241_11 cl_717241_12 cl_717241_24 cl_717241_69 cl_717241_8 cl_717241_18 cl_717241_34 cl_186223_14 cl_732062_4 cl_732062_5 cl_732062_6	GGPPS or GGPP synthase	⇩ geranylgeranyl diphosphate (GGPP)	2.5.1.10 or 2.5.1.29	[illegible] [illegible] 0 [illegible] 114.6445 112.522 140.7125 1.5 1.106 [illegible] [illegible]	[illegible] [illegible] 0 [illegible] 172.245 140.995 198.373 1.581 0.012 [illegible] [illegible]	[illegible] [illegible] 0 [illegible] [illegible] [illegible] [illegible] 4.566 0.081 [illegible] [illegible]	[illegible] [illegible] 0.16 [illegible] [illegible] [illegible] [illegible] 1.831 0.21 [illegible] [illegible]
Carotegenesis	15-cis-phytoene synthase	cl_705362_8 cl_705362_5	PSY	geranylgeranyl diphosphate (GGPP) ⇩ 15-cis-phytoene	2.5.1.32	140.4215	131.519	130.485	157.602
	Phytoene_dehydrogenase,_phytoenedesaturase	cl_844762_52	PDS	⇩ 9,15,9'-tri-cis-zeta-carotene	1.3.5.5	[illegible]	[illegible]	[illegible]	[illegible]
	Zeta-carotene_desaturase	cl_813223_22 cl_813223_24 cl_813223_26 cl_813223_29 cl_367788_5 cl_415028_1	ZDS	⇩		0.1815 0 0.128 174.7525 2.941 [illegible]	0 0.46 0.751 [illegible] 5.756 [illegible]	0 0.179 0.195 [illegible] 0 [illegible]	0 0.233 0.254 [illegible] 0 [illegible]
	Lycopene_cyclase	cl_1010176_1 cl_1010176_2 cl_178530_36 cl_107156_8	beta- or epsilon cyclase	7,9,7',9'-tetra-cis-lycopene ⇩	5.5.1.19 or 5.5.1.18	[illegible] 8.1865 [illegible] 155.947	[illegible] 12.065 [illegible] 116.242	[illegible] [illegible] [illegible] [illegible]	[illegible] [illegible] [illegible] [illegible]
	Beta-carotene_hydroxylase_chloroplast	cl_335275_7	B1 or CHY1	**Beta-carotene** (gamma and beta-end group) or alpha-cartoene(delta and alpha- end group) ⇩ zeaxanthin	1.14.13.129	[illegible]	[illegible]	[illegible]	[illegible]
	Zeaxanthin_epoxidase	cl_244459_3 cl_244459_4	ZEP	⇩ violaxanthin	1.14.13.90	[illegible] 3.26	[illegible] 6.872	[illegible] 9.085	[illegible] 5.25
Control	Elongation_factor_EF-1_alpha_subunit	XM_011400722.1				[illegible]	[illegible]	[illegible]	[illegible]

Table 5

Pathway	Gene Name	Cluster	Acronyms or synonyms	Compound(s)	EC	DF17t1	DF17t2	DF17t2-control	DF17t2-salt stress
				Pyruvate + D-glyceraldehyde 3-phosphate					
Non-mevalonate pathway for terpenoid synthesis (MEP/DOXP pathway) that occurs in chloroplasts	1-deoxy-D-xylulose-5-phosphate_synthase	cl_857208_2	DXS or DOXP synthase		2.2.1.7	314.301	303.138	199.635	171.115
		cl_535498_12				[illegible]	[illegible]	[illegible]	[illegible]
				1-deoxy-D-xylulose 5-phosphate					
	1-deoxy-D-xylulose_5-phosphate_reductoisomerase	DXR1_XM_00169:	DXR or DOXP synthase		1.1.1.267	130.976	[illegible]	123.574	102.087
		cl_45366_21				[illegible]	[illegible]	[illegible]	[illegible]
		cl_45366_7				175.214	122.779	147.148	149.656
				2-C-methyl-D-erythritol 4-phosphate					
	4-diphosphocytidyl-2C-methyl-D-erythritol_synthase	cl_175958_14	CMS		2.7.7.60	129.107	116.404	104.604	[illegible]
				diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol					
	4-diphosphocytidyl-2-C-methyl-D-erythritol_kinase	cl_187509_1	CMK		2.7.1.148	0	0	0	0
		cl_187509_2				[illegible]	[illegible]	[illegible]	[illegible]
		cl_708851_2				60.237	[illegible]	[illegible]	[illegible]
		cl_708851_6				[illegible]	[illegible]	[illegible]	[illegible]
		cl_639608_18				[illegible]	[illegible]	[illegible]	[illegible]
				2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol					
	2-C-methyl-D-erythritol_2,4-cyclodiphosphate_synthase	cl_762165_19	MCS		4.6.1.12	[illegible]	[illegible]	[illegible]	[illegible]
		cl_762165_1				0	0	0	0
		cl_762165_17				136.736	113.176	111.863	[illegible]
				2-C-methyl-D-erythritol 2,4-cyclodiphosphate					
	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase or 1-hydroxy-2-methyl-2-(E)-butenyl_4-diphosphate_synthase	cl_459175_35	HDS (ferredoxin)		1.17.7.1	[illegible]	[illegible]	[illegible]	[illegible]
		cl_459175_21				282.219	342.864	271.645	279.146
				(E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate					
	4-Hydroxy-3-methylbut-2-enyl diphosphate reductase	cl_665162_3	HDR or HMBPP reductase		1.17.7.4	0.218	0.012	0.452	0.251
		cl_665162_6				185.35	259.4	178.91	167.5
		cl_665162_10				112.094	[illegible]	[illegible]	[illegible]
		cl_665162_26				156.027	145.885	110.859	[illegible]
		cl_462991_1				[illegible]	31.195	[illegible]	[illegible]
		cl_462991_3				117.436	112.28	[illegible]	[illegible]
				isopentenyl diphosphate + dimethylallyl diphosphate					
				Dimethylallyl diphosphate + isopentenyl diphosphate+ (2E,6E)-farnesyl diphosphate					
geranylgeranyl diphosphate synthesis in cytoplasm, ER or mitochondrion	Dimethylallyltranstransferase.	cl_531037_8	FPP synthase or Farnesyl pyrophosphate synthase		2.5.1.1	n/a	n/a	n/a	n/a
		cl_531037_9				n/a	n/a	n/a	n/a
	Geranylgeranyl_diphosphate_synthase	cl_717241_11	GGPPS or GGPP synthase		2.5.1.10 or 2.5.1.29	[illegible]	[illegible]	[illegible]	[illegible]
		cl_717241_12				1.171	0.455	1.7	[illegible]
		cl_717241_24				256.763	159.221	196.423	180.783
		cl_717241_69				704.668	[illegible]	617.091	[illegible]
		cl_717241_8				1.294	0.377	0.805	1.045
		cl_717241_18				4.463	2.38	3.135	3.986
		cl_717241_34				185.407	[illegible]	[illegible]	[illegible]
		cl_186223_14				[illegible]	[illegible]	[illegible]	[illegible]
		cl_732062_4				[illegible]	[illegible]	[illegible]	[illegible]
		cl_732062_5				[illegible]	[illegible]	[illegible]	[illegible]
		cl_732062_6				0.842	1.33	0.982	0.971
				geranylgeranyl diphosphate (GGPP)					
				geranylgeranyl diphosphate (GGPP)					
Carotegenesis	15-cis-phytoene synthase	cl_705362_8	PSY		2.5.1.32	105.43	156.026	267.542	239.11
		cl_705362_5							
				15-cis-phytoene					
	Phytoene_dehydrogenase,_phytoenedesaturase	cl_844762_52	PDS		1.3.5.5	106.716	[illegible]	111.573	103.344
				9,15,9'-tri-cis-zeta-carotene					
	Zeta-carotene_desaturase	cl_813223_22	ZDS			[illegible]	[illegible]	[illegible]	[illegible]
		cl_813223_24				123.241	95.921	106.454	100.813
		cl_813223_26				[illegible]	[illegible]	[illegible]	[illegible]
		cl_813223_29				[illegible]	[illegible]	[illegible]	[illegible]
		cl_367788_5				[illegible]	[illegible]	[illegible]	[illegible]
		cl_415028_1				137.07	132.311	144.936	141.332
				7,9,7',9'-tetra-cis-lycopene					
	Lycopene_cyclase	cl_1010176_1	beta- or epsilon cyclase		5.5.1.19 or 5.5.1.18	0.012	0.012	0.012	0.012
		cl_1010176_2				0.012	0.012	0.012	0.012
		cl_178530_36				126.405	155.06	117.403	105.904
		cl_107156_8				3.653	3.33	2.131	2.897
				Beta-carotene (gamma and beta-end group) or alpha-cartoene(delta and alpha- end group)					
	Beta-carotene_hydroxylase_chloroplast	cl_335275_7	B1 or CHY1		1.14.13.129	[illegible]	[illegible]	[illegible]	[illegible]
				zeaxanthin					
	Zeaxanthin_epoxidase	cl_244459_3	ZEP		1.14.13.90	[illegible]	[illegible]	[illegible]	[illegible]
		cl_244459_4				9.425	7.227	[illegible]	[illegible]
				violaxanthin					
Control	Elongation_factor_EF-1_alpha_subunit	XM_011400722.1				[illegible]	[illegible]	[illegible]	[illegible]

SUBSTITUTE SHEET (RULE 26)

Table 6

Strains	chls (pg/cell)	caro (pg/cell)	caro / chls	% total carotenoids of dry weight
DF15	3.86	20.51	5.32	3.09
DF17	2.56	4.22	1.65	1.13
DF40	4.71	5.50	1.17	0.81
CCAP19/30	1.30	0.56	0.44	0.28
UTEX2538	10.59	9.07	0.86	1.40

Response	Light Intensity				Strain				Light Intensity*Strain			
	Df	*F* Values	*P* Values	Significance Level	Df	*F* Values	*P* Values	Significance Level	Df	*F* Values	*P* Values	Significance Level
Photosynthesis	3	8.1825	0.0002	***	4	71.2528	<2.2e−16	***	12	2.7966	0.0073	**
Respiration	3	1.7925	0.1641		4	52.7992	1.96e−15	***	12	2.4328	0.0176	*
Total carotenoids	3	2.9403	0.0446	*	4	693.560	<2.2e−16	***	12	7.9749	2.52e−07	***
Total chlorophyll	3	36.529	1.55e−11	***	4	161.782	<2.2e−16	***	12	10.285	8.41e−09	***
All-trans β-carotene	3	88.922	<2.2e−16	***	4	474.255	<2.2e−16	***	12	3.6878	0.0009	***
9-cis β-carotene	3	28.119	6.02e−10	***	4	730.574	<2.2e−16	***	12	6.8407	1.67e−06	***
Lutein	3	7.3679	0.0005	***	4	118.762	<2.2e−16	***	12	6.4955	3.08e−06	***
Zeaxanthin	3	35.542	2.31e−11	***	4	83.0526	<2.2e−16	***	12	5.2669	3.13e−05	***
All-trans α-carotene	3	113.39	<2.2e−16	***	4	408.180	<2.2e−16	***	12	5.9987	7.64e−06	***
Glycerol	3	2.1170	0.1132		4	95.5589	<2.2e−16	***	12	5.0858	4.50e−05	***

(*) indicates $P \leq 0.05$, (**) $P \leq 0.01$, (***) $P \leq 0.001$.

Table 7

CLAIMS

1. A composition or culture containing Dunaliella salina algae strain DF15, DF17 or DF40.

2. A composition or culture of algal strain CCAP accession number CCAP 19/40 for Dunaliella salina PLY_DF-40.

3. A composition or culture of algal strain CCAP accession number CCAP 19/41 for Dunaliella salina rubeus PLY_DF-15

4. A composition or culture as claimed in any preceding claim in a body of aqueous nutrient solution.

5. A composition or culture as claimed in any preceding claim, in which the composition is provided in a salt solution.

6. A composition or culture as claimed in claim 5, in which the salt solution contains less than 6% NaCl.

7. The use of an algae strain selected from the group consisting of Dunaliella salina strain DF15, DF17 or DF40 in a process for producing a commercially useful product.

8. The use of claim 7, in which the product is a carotenoid.

9. The use of claim 8, in which the product is β-carotene.

10. A method of producing β-carotene by growing alga selected from Dunaliella salina strain DF15, DF17 or DF40, and extracting β-carotene or isomers thereof produced thereby.

11. An algae biorefinery, based on biomass from Dunaliella salina algae strains DF15, DF17 or DF40.

12. β-carotene or isomers thereof produced by alga selected from Dunaliella salina strain DF15, DF17 or DF40.

13. Algal strain CCAP accession number CCAP 19/40.

14. Algal strain CCAP accession number CCAP 19/41.

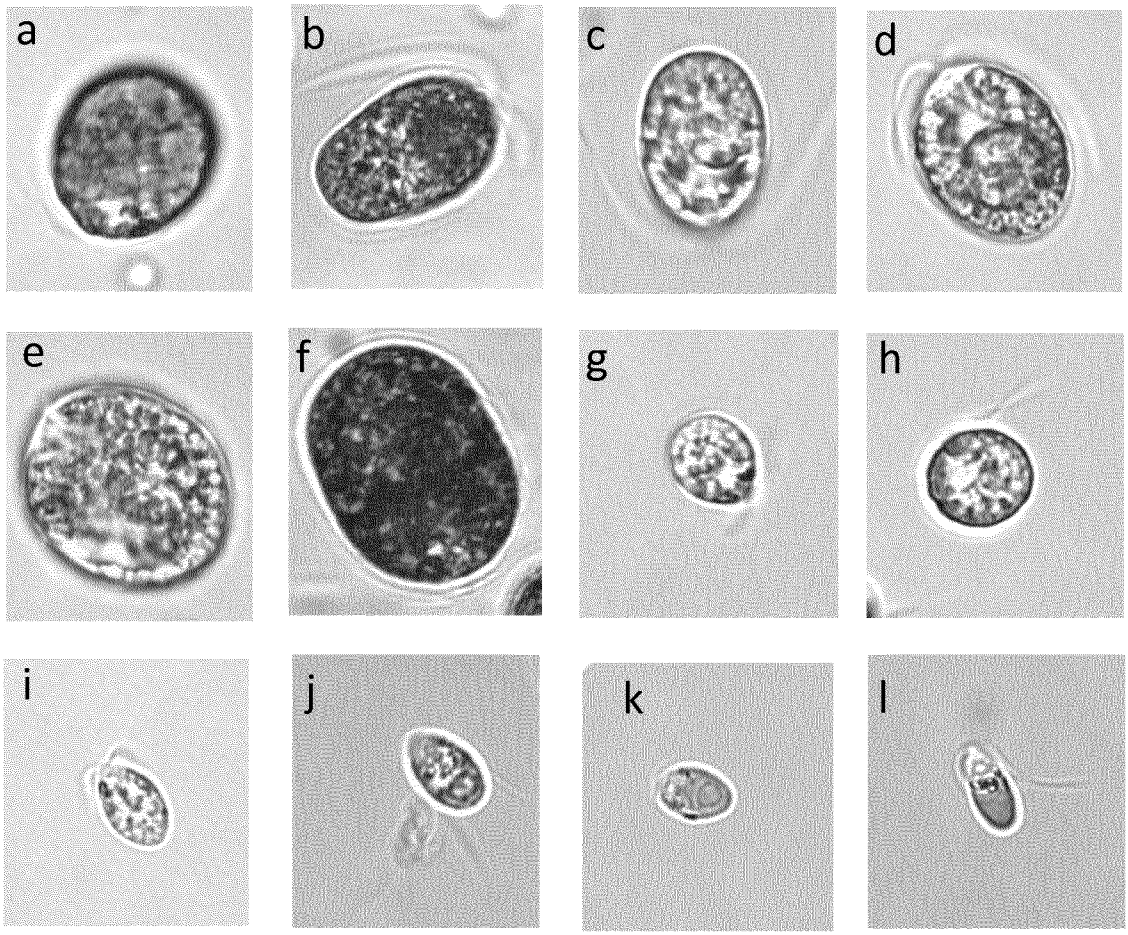

Figure 1

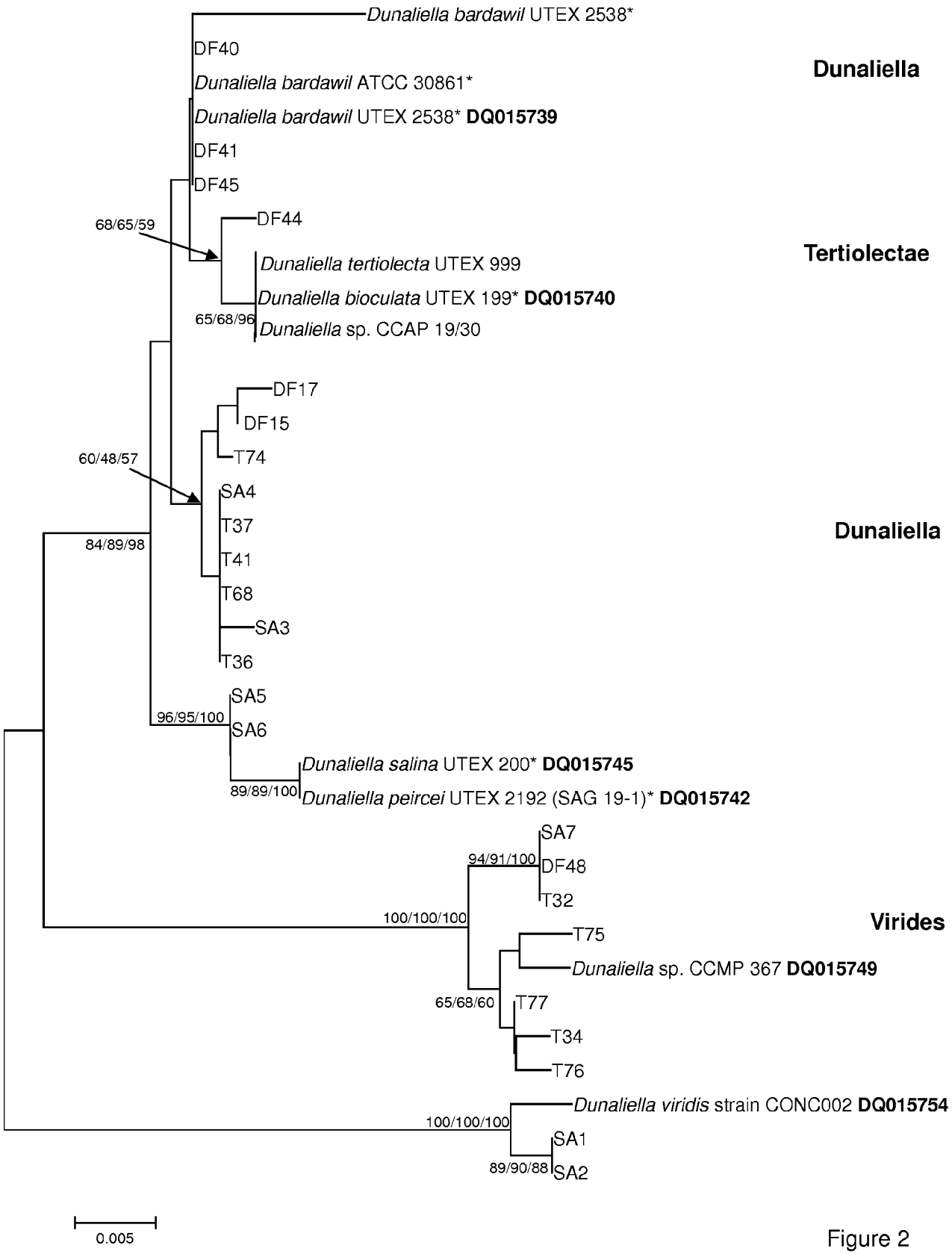
Dunaliella bardawil UTEX 2538*
DF40
Dunaliella bardawil ATCC 30861*
Dunaliella bardawil UTEX 2538* DQ015739
DF41
DF45
Dunaliella
68/65/59
DF44
Dunaliella tertiolecta UTEX 999
Dunaliella bioculata UTEX 199* DQ015740
65/68/96
Dunaliella sp. CCAP 19/30
Tertiolectae
DF17
DF15
60/48/57
T74
SA4
T37
84/89/98
T41
T68
SA3
T36
Dunaliella
SA5
96/95/100
SA6
Dunaliella salina UTEX 200* DQ015745
89/89/100
Dunaliella peircei UTEX 2192 (SAG 19-1)* DQ015742
SA7
94/91/100
DF48
T32
100/100/100
T75
Dunaliella sp. CCMP 367 DQ015749
Virides
65/68/60
T77
T34
T76
Dunaliella viridis strain CONC002 DQ015754
100/100/100
SA1
89/90/88
SA2
0.005

Figure 2

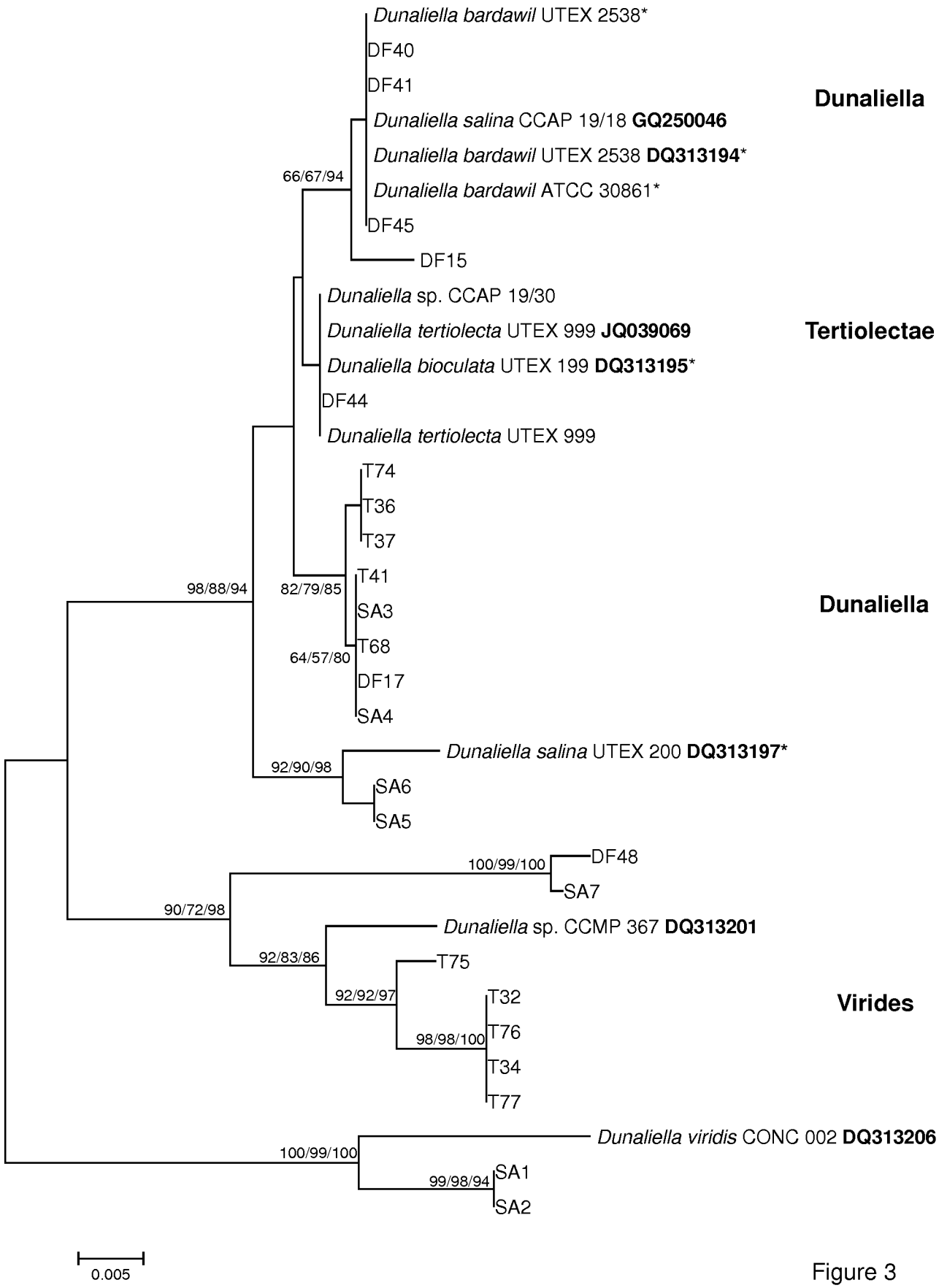
Dunaliella bardawil UTEX 2538*
DF40
DF41
Dunaliella salina CCAP 19/18 **GQ250046**
Dunaliella bardawil UTEX 2538 **DQ313194***
Dunaliella bardawil ATCC 30861*
DF45
DF15
Dunaliella sp. CCAP 19/30
Dunaliella tertiolecta UTEX 999 **JQ039069**
Dunaliella bioculata UTEX 199 **DQ313195***
DF44
Dunaliella tertiolecta UTEX 999
T74
T36
T37
T41
SA3
T68
DF17
SA4
Dunaliella salina UTEX 200 **DQ313197***
SA6
SA5
DF48
SA7
Dunaliella sp. CCMP 367 **DQ313201**
T75
T32
T76
T34
T77
Dunaliella viridis CONC 002 **DQ313206**
SA1
SA2
66/67/94
98/88/94
82/79/85
64/57/80
92/90/98
100/99/100
90/72/98
92/83/86
92/92/97
98/98/100
100/99/100
99/98/94
Dunaliella
Tertiolectae
Dunaliella
Virides
0.005

Figure 3

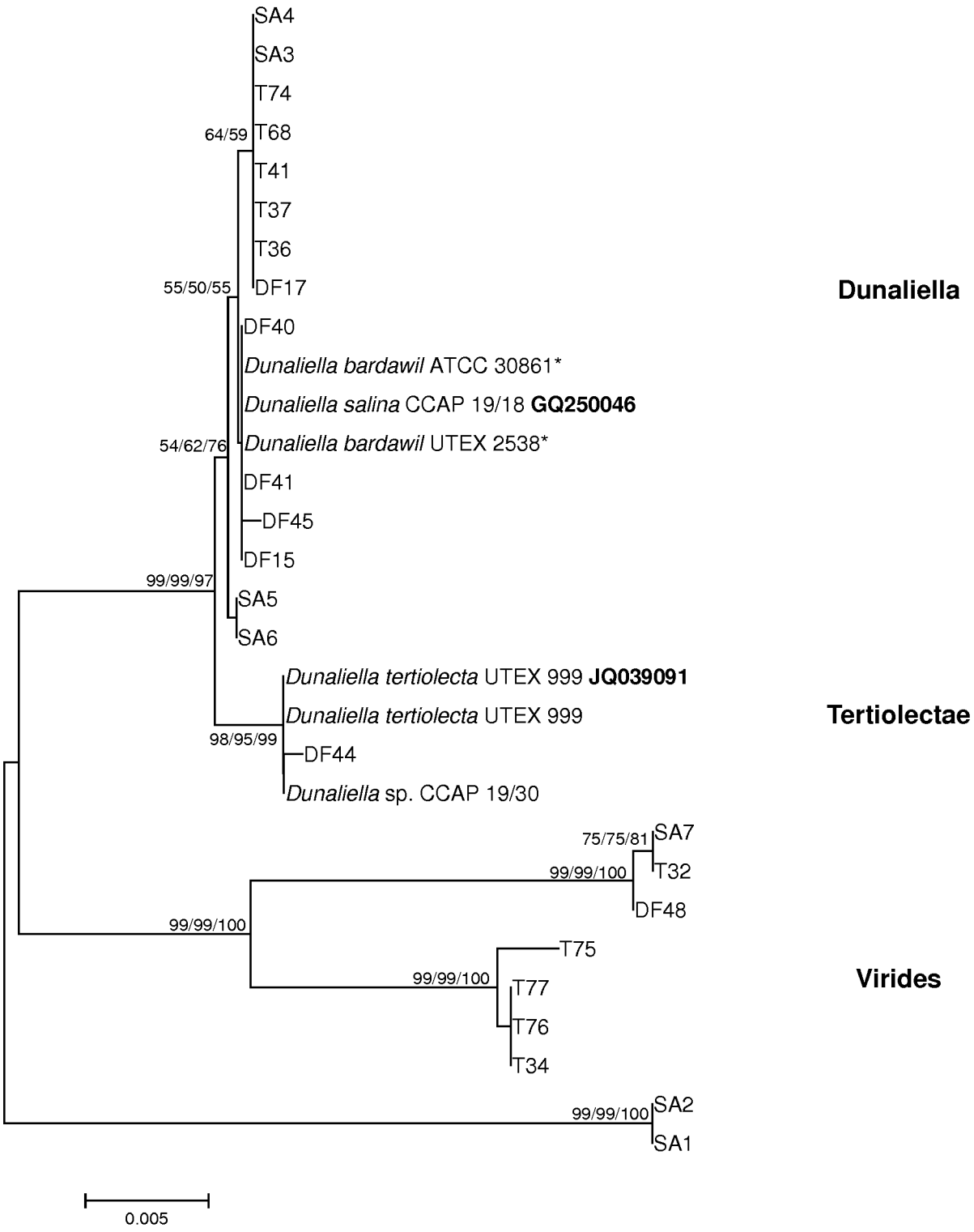
SA4
SA3
T74
64/59 T68
T41
T37
T36
55/50/55 DF17
DF40
Dunaliella bardawil ATCC 30861*
Dunaliella salina CCAP 19/18 GQ250046
54/62/76 Dunaliella bardawil UTEX 2538*
DF41
DF45
DF15
99/99/97
SA5
SA6
Dunaliella tertiolecta UTEX 999 JQ039091
Dunaliella tertiolecta UTEX 999
98/95/99
DF44
Dunaliella sp. CCAP 19/30
75/75/81 SA7
99/99/100 T32
DF48
99/99/100
T75
99/99/100
T77
T76
T34
99/99/100 SA2
SA1
0.005
Dunaliella
Tertiolectae
Virides

Figure 4

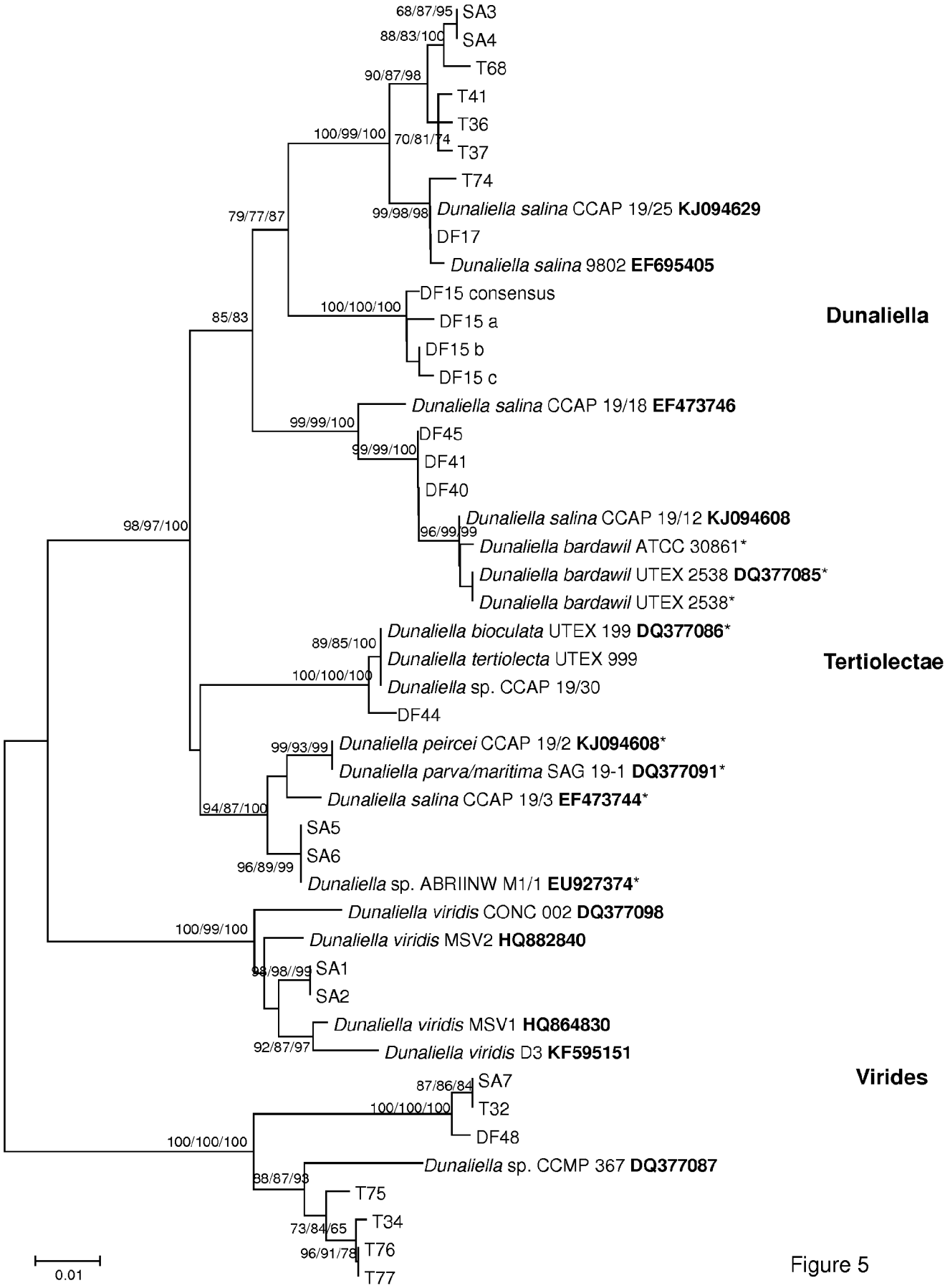
68/87/95
SA3
88/83/100
SA4
T68
90/87/98
T41
T36
100/99/100
70/81/74
T37
T74
99/98/98
Dunaliella salina CCAP 19/25 KJ094629
79/77/87
DF17
Dunaliella salina 9802 EF695405
DF15 consensus
100/100/100
85/83
DF15 a
Dunaliella
DF15 b
DF15 c
Dunaliella salina CCAP 19/18 EF473746
99/99/100
DF45
99/99/100
DF41
DF40
Dunaliella salina CCAP 19/12 KJ094608
98/97/100
96/99/99
Dunaliella bardawil ATCC 30861*
Dunaliella bardawil UTEX 2538 DQ377085*
Dunaliella bardawil UTEX 2538*
Dunaliella bioculata UTEX 199 DQ377086*
89/85/100
Dunaliella tertiolecta UTEX 999
Tertiolectae
100/100/100
Dunaliella sp. CCAP 19/30
DF44
99/93/99
Dunaliella peircei CCAP 19/2 KJ094608*
Dunaliella parva/maritima SAG 19-1 DQ377091*
Dunaliella salina CCAP 19/3 EF473744*
94/87/100
SA5
SA6
96/89/99
Dunaliella sp. ABRIINW M1/1 EU927374*
Dunaliella viridis CONC 002 DQ377098
100/99/100
Dunaliella viridis MSV2 HQ882840
98/98//99
SA1
SA2
Dunaliella viridis MSV1 HQ864830
92/87/97
Dunaliella viridis D3 KF595151
Virides
87/86/84
SA7
100/100/100
T32
DF48
100/100/100
Dunaliella sp. CCMP 367 DQ377087
88/87/93
T75
73/84/65
T34
96/91/78
T76
T77
0.01

Figure 5

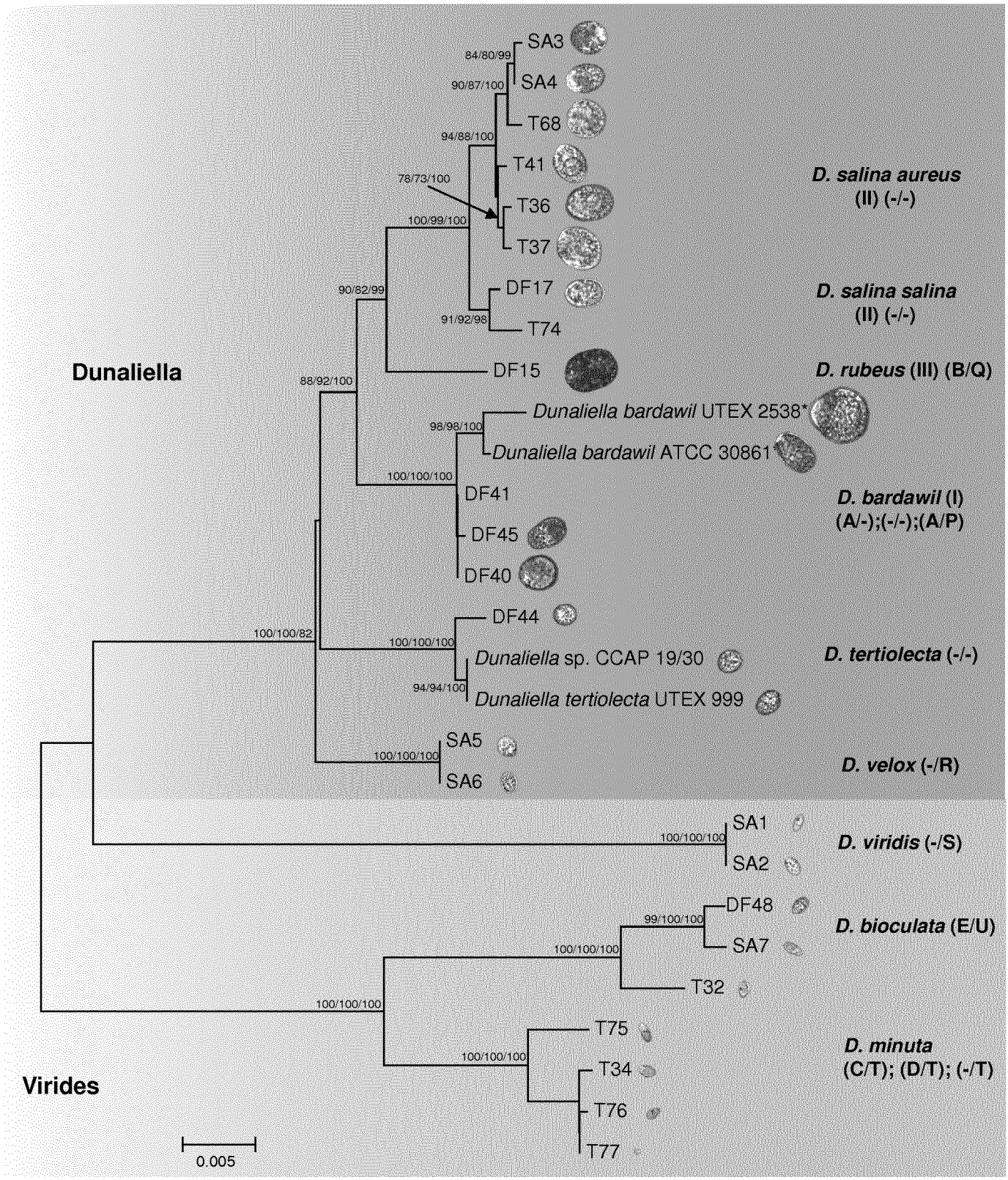
Dunaliella
Virides
SA3
SA4
T68
T41
T36
T37
DF17
T74
DF15
Dunaliella bardawil UTEX 2538*
Dunaliella bardawil ATCC 30861
DF41
DF45
DF40
DF44
Dunaliella sp. CCAP 19/30
Dunaliella tertiolecta UTEX 999
SA5
SA6
SA1
SA2
DF48
SA7
T32
T75
T34
T76
T77
84/80/99
90/87/100
94/88/100
78/73/100
100/99/100
90/82/99
91/92/98
88/92/100
98/98/100
100/100/100
100/100/82
94/94/100
0.005
D. salina aureus (II) (-/-)
D. salina salina (II) (-/-)
D. rubeus (III) (B/Q)
D. bardawil (I) (A/-);(-/-);(A/P)
D. tertiolecta (-/-)
D. velox (-/R)
D. viridis (-/S)
D. bioculata (E/U)
D. minuta (C/T); (D/T); (-/T)
25µm

Figure 6

10 20 30 40 50 60 70 80 90 100
AATGTTCTTCTCAGTACCAACAGAGGTTGCCCCCCATTGATACCGATCCAAACACTTCACCAGCACACCCAATCGGTAGGAGCGACGGGCGGTGTGTACA
DF15
D. bardawil UTEX 2538 DQ009777
D. bardawil UTEX 2538 (this study)
D. salina CCAP 19/30 DQ447648
D. salina 19/18 EF473745
D. pseudosalina isolate MAH KU641615
D. salina strain CCAP 19/12 KJ756842
D. salina strain CCAP 19/18 EF473745
D. salina strain SAG 19-3 EF473739
Dunaliella sp. SAS11133 KF054056
DF17
D. peircei strain UTEX LB 2192 DQ009778
D. terfiolecta UTEX 199 (this study)
D. terfiolecta strain CCAP 19/6B KJ756820
D. terfiolecta strain CCMP 1320 EF537907
D. bioculata strain UTEX LB 199 DQ009761
D. parva strain SAG 19-1 DQ009763
D. viridis CONC002 DQ009776
Dunaliella sp. CCMP 367 DQ009769

Figure 7

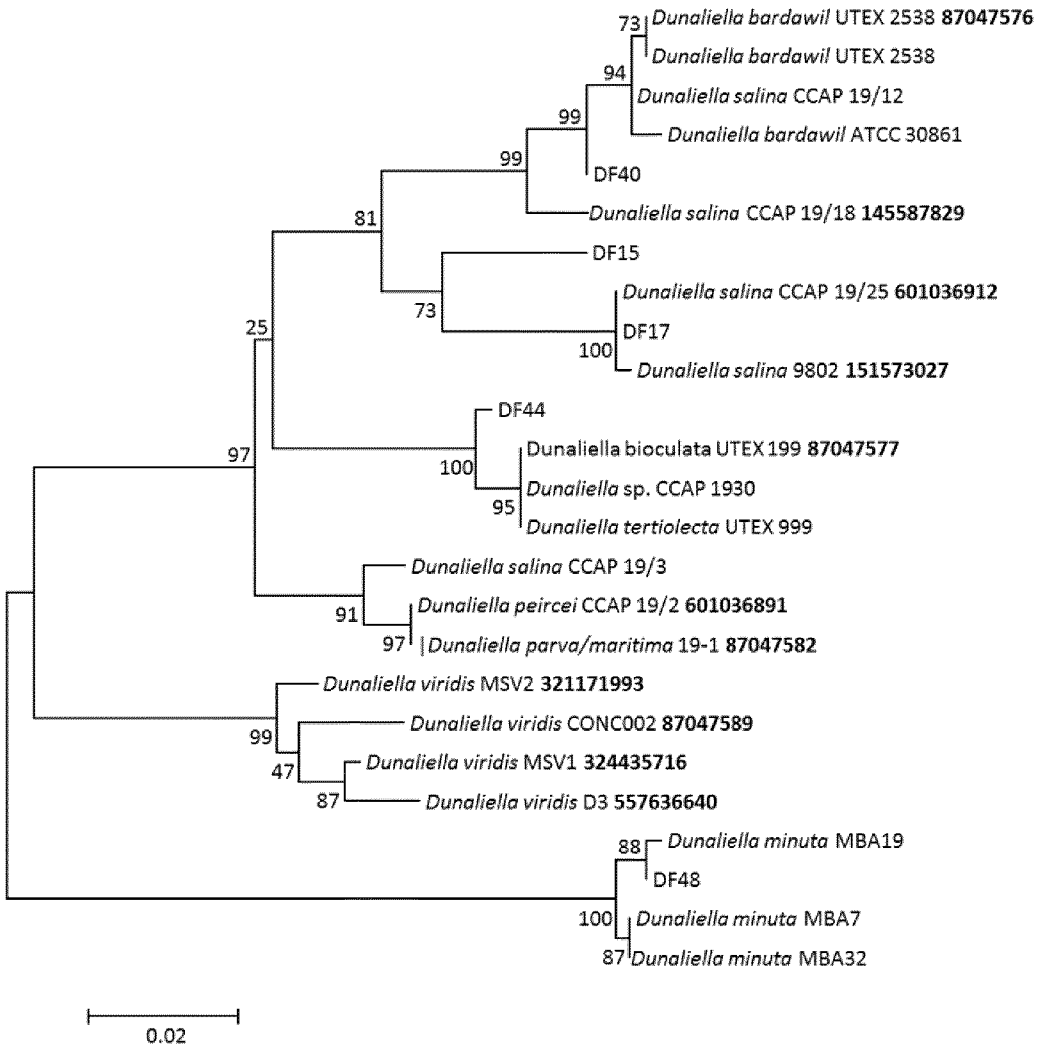
Dunaliella bardawil UTEX 2538 87047576
Dunaliella bardawil UTEX 2538
Dunaliella salina CCAP 19/12
Dunaliella bardawil ATCC 30861
DF40
Dunaliella salina CCAP 19/18 145587829
DF15
Dunaliella salina CCAP 19/25 601036912
DF17
Dunaliella salina 9802 151573027
DF44
Dunaliella bioculata UTEX 199 87047577
Dunaliella sp. CCAP 1930
Dunaliella tertiolecta UTEX 999
Dunaliella salina CCAP 19/3
Dunaliella peircei CCAP 19/2 601036891
Dunaliella parva/maritima 19-1 87047582
Dunaliella viridis MSV2 321171993
Dunaliella viridis CONC002 87047589
Dunaliella viridis MSV1 324435716
Dunaliella viridis D3 557636640
Dunaliella minuta MBA19
DF48
Dunaliella minuta MBA7
Dunaliella minuta MBA32
73
94
99
99
81
73
25
100
97
100
95
91
97
99
47
87
88
100
87
0.02

Figure 8

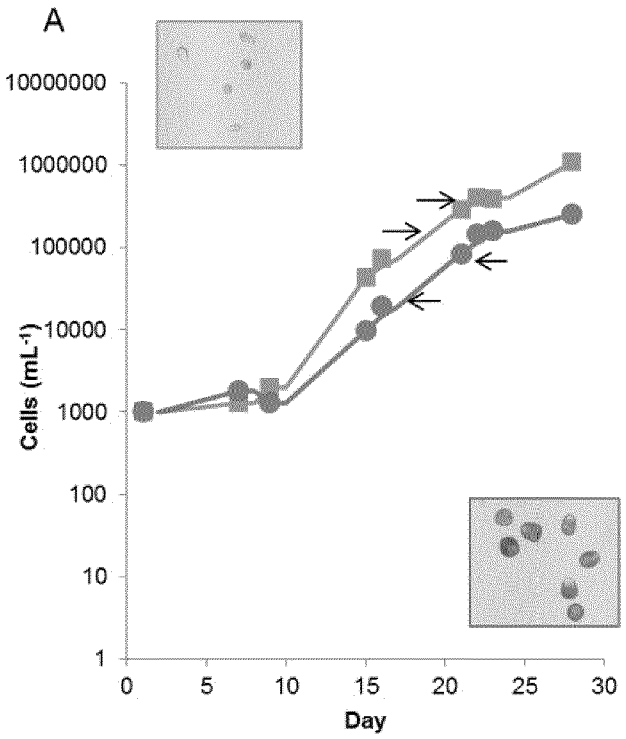
A
10000000
1000000
100000
10000
1000
100
10
1
Cells (mL-1)
0
5
10
15
20
25
30
Day

Figure 9

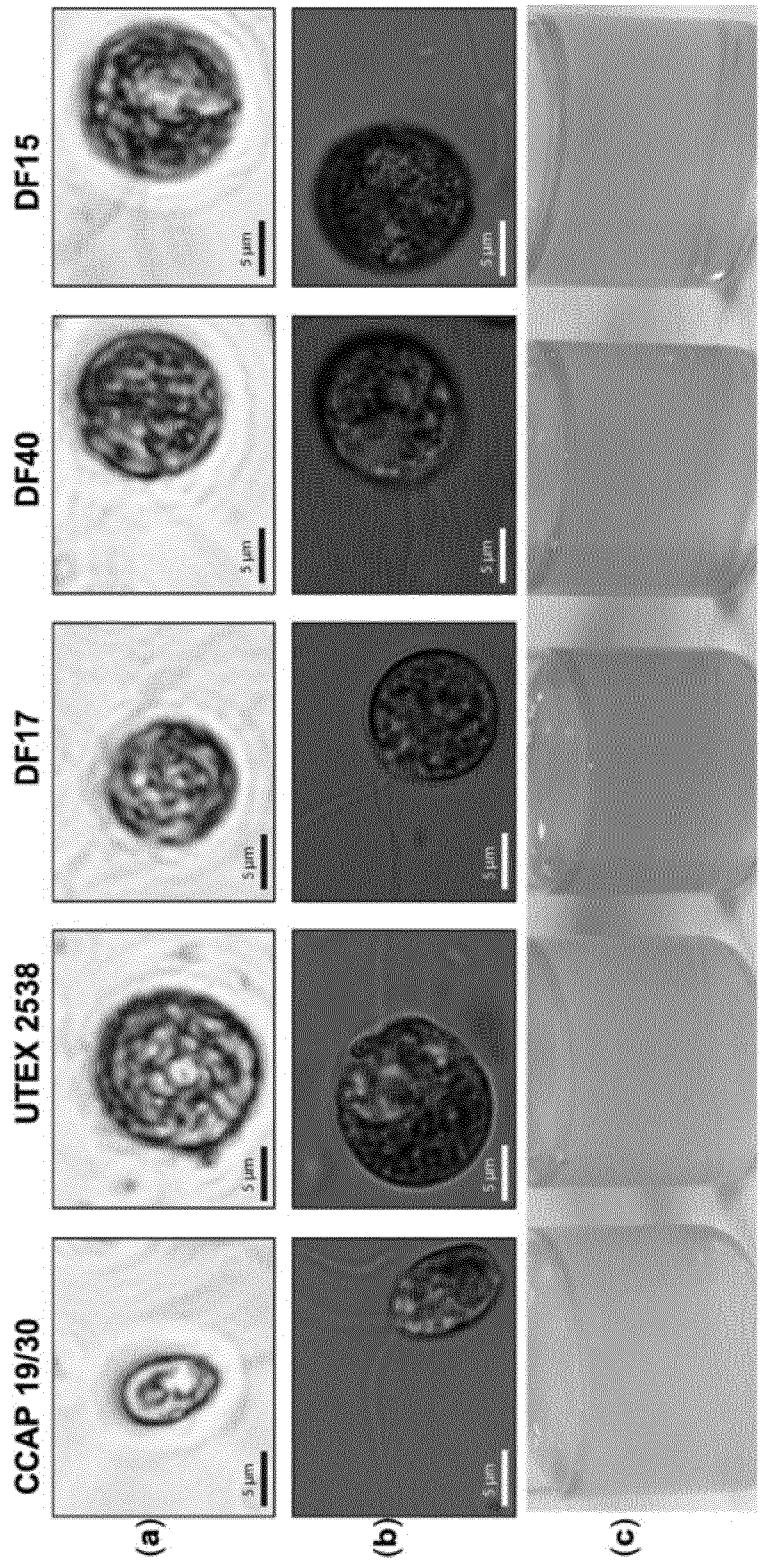

Figure 10

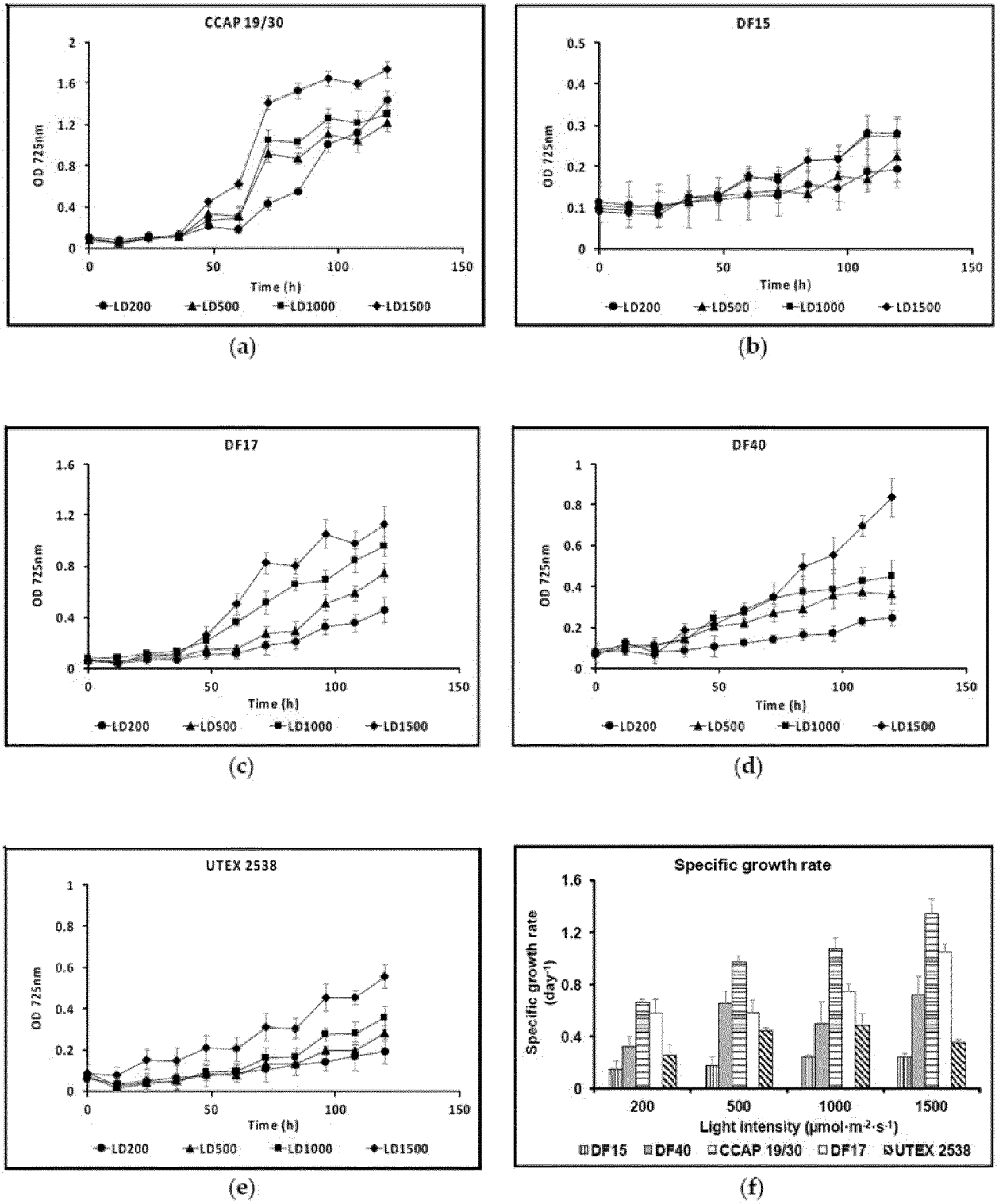
CCAP 19/30
OD 725nm
Time (h)
LD200
LD500
LD1000
LD1500
(a)
DF15
OD 725nm
Time (h)
LD200
LD500
LD1000
LD1500
(b)
DF17
OD 725nm
Time (h)
LD200
LD500
LD1000
LD1500
(c)
DF40
OD 725nm
Time (h)
LD200
LD500
LD1000
LD1500
(d)
UTEX 2538
OD 725nm
Time (h)
LD200
LD500
LD1000
LD1500
(e)
Specific growth rate
Specific growth rate (day-1)
Light intensity (µmol·m-2·s-1)
DF15
DF40
CCAP 19/30
DF17
UTEX 2538
(f)

Figure 11

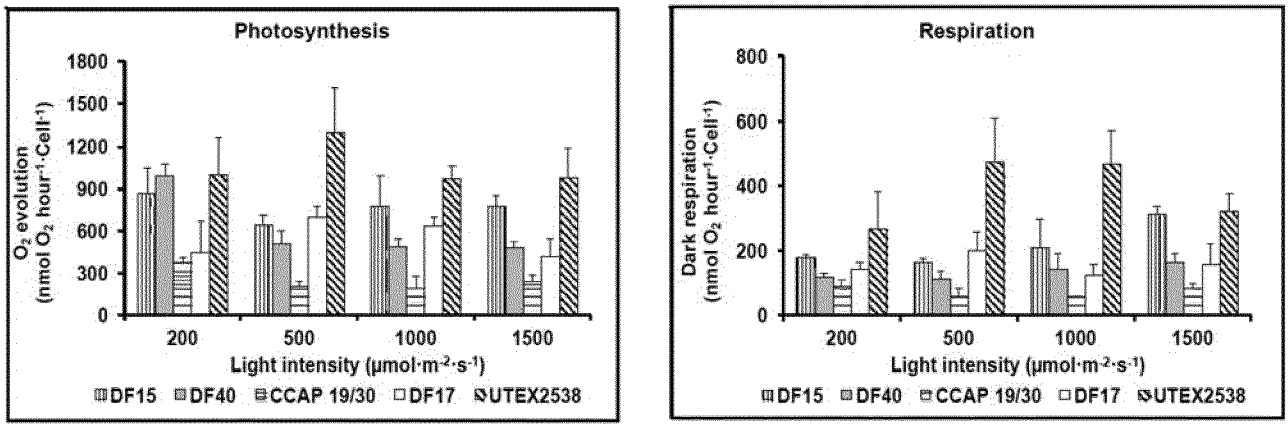

Figure 12

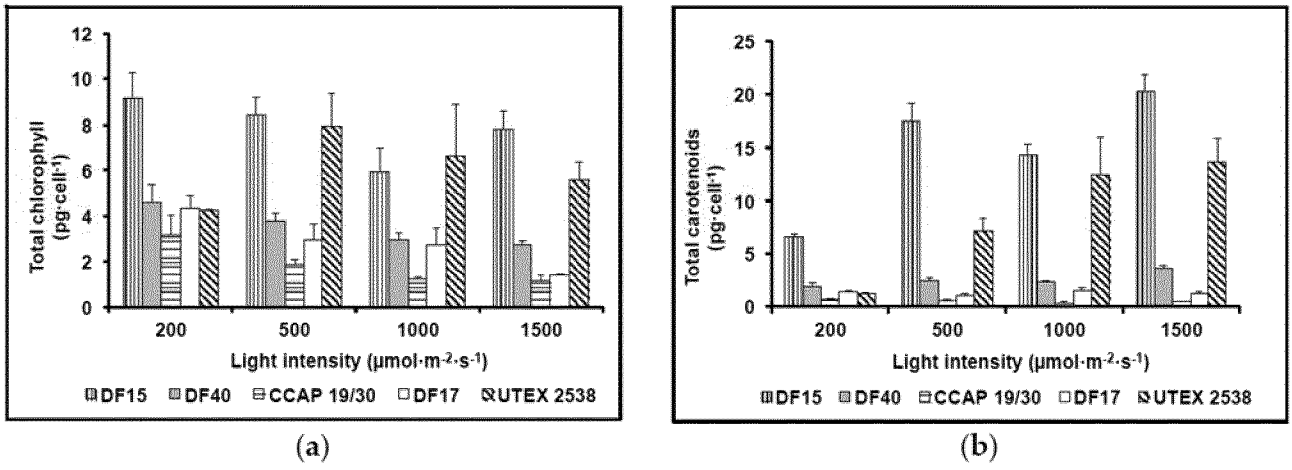

Figure 13

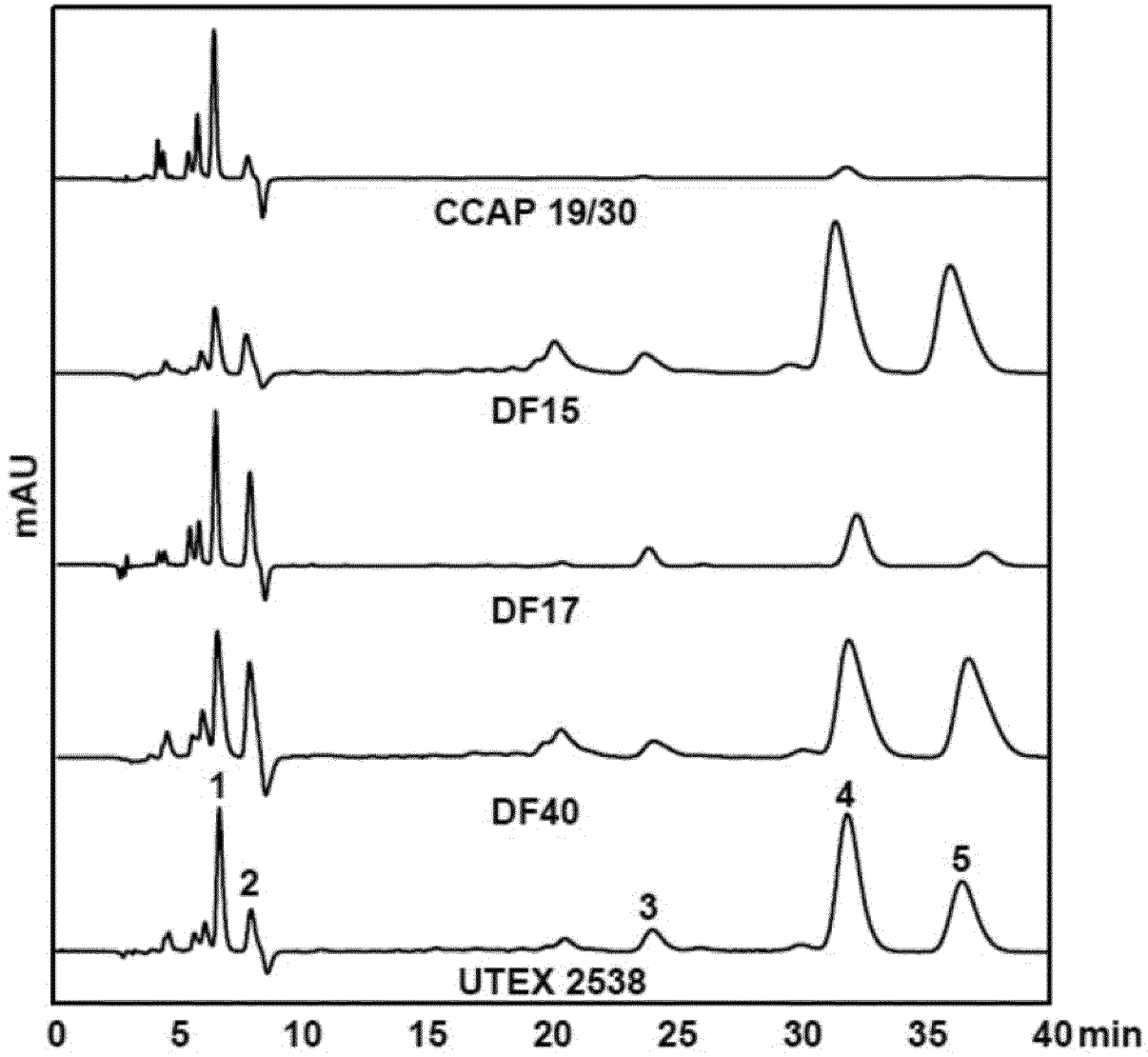
CCAP 19/30
DF15
DF17
DF40
UTEX 2538
mAU
1
2
3
4
5
0
5
10
15
20
25
30
35
40 min

Figure 14

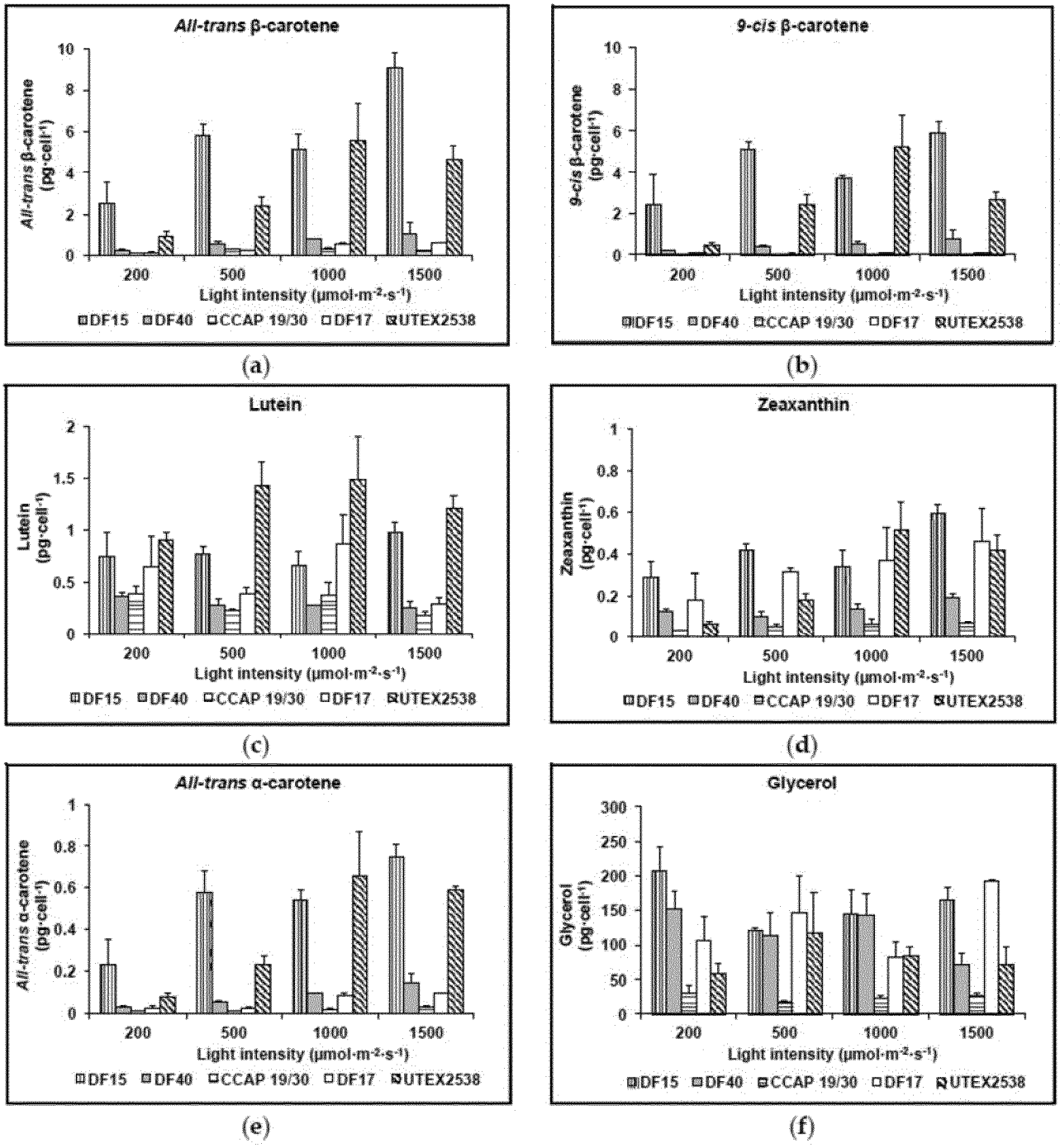
All-trans β-carotene
All-trans β-carotene (pg·cell⁻¹)
Light intensity (µmol·m⁻²·s⁻¹)
DF15 DF40 CCAP 19/30 DF17 UTEX2538
(a)
9-cis β-carotene
9-cis β-carotene (pg·cell⁻¹)
Light intensity (µmol·m⁻²·s⁻¹)
DF15 DF40 CCAP 19/30 DF17 UTEX2538
(b)
Lutein
Lutein (pg·cell⁻¹)
Light intensity (µmol·m⁻²·s⁻¹)
DF15 DF40 CCAP 19/30 DF17 UTEX2538
(c)
Zeaxanthin
Zeaxanthin (pg·cell⁻¹)
Light intensity (µmol·m⁻²·s⁻¹)
DF15 DF40 CCAP 19/30 DF17 UTEX2538
(d)
All-trans α-carotene
All-trans α-carotene (pg·cell⁻¹)
Light intensity (µmol·m⁻²·s⁻¹)
DF15 DF40 CCAP 19/30 DF17 UTEX2538
(e)
Glycerol
Glycerol (pg·cell⁻¹)
Light intensity (µmol·m⁻²·s⁻¹)
DF15 DF40 CCAP 19/30 DF17 UTEX2538
(f)

Figure 15

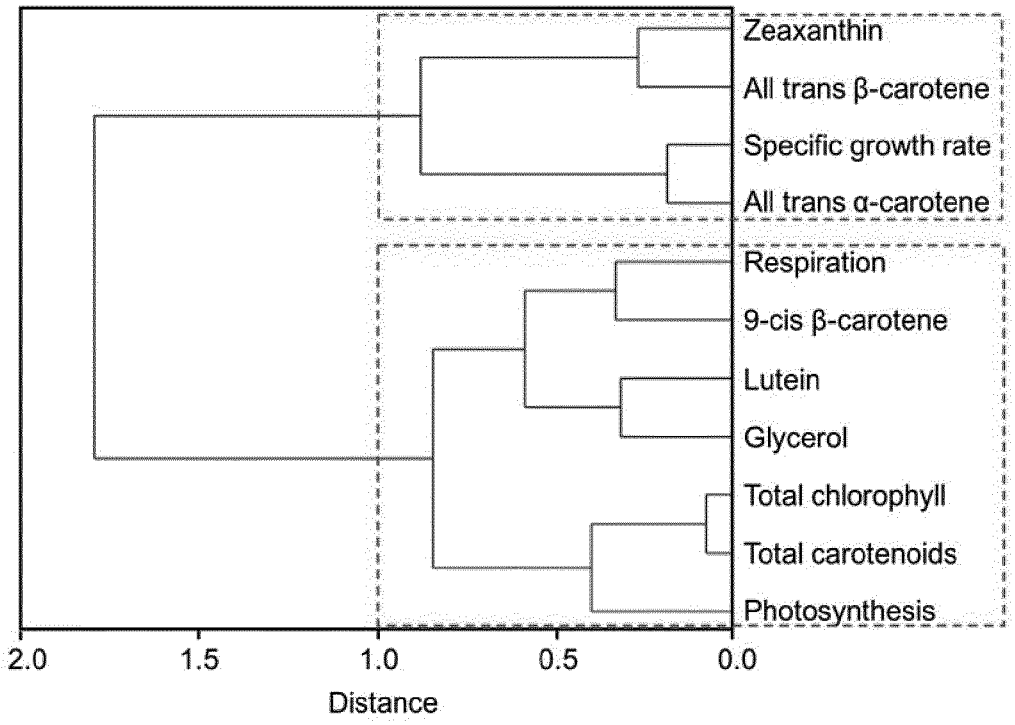

Figure 16a

CCAP 1930

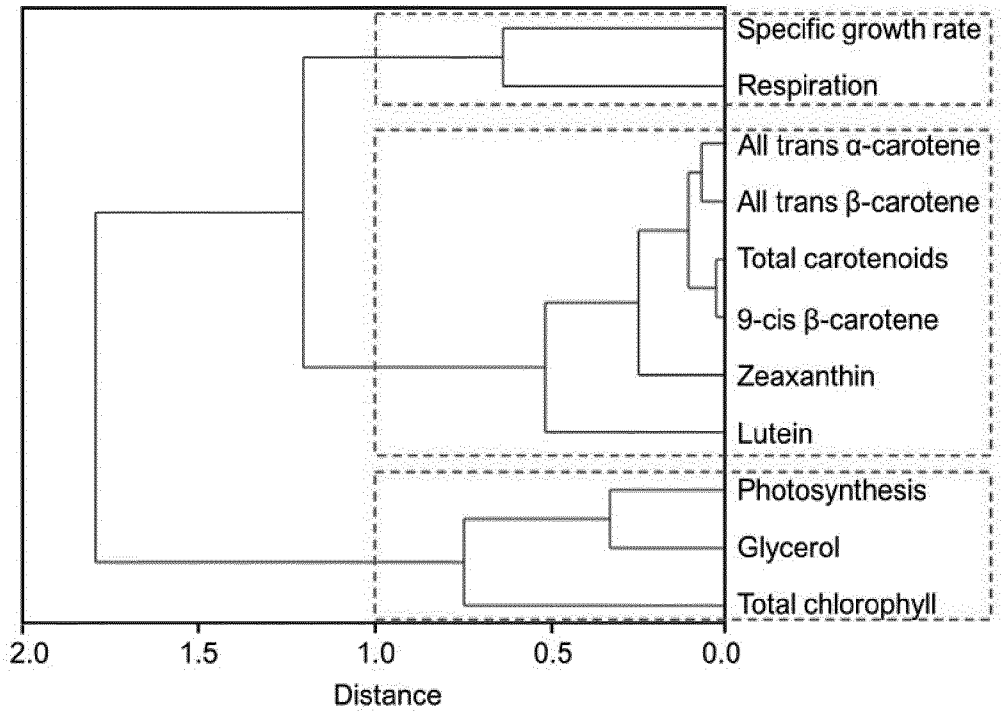

Figure 16b

DF15

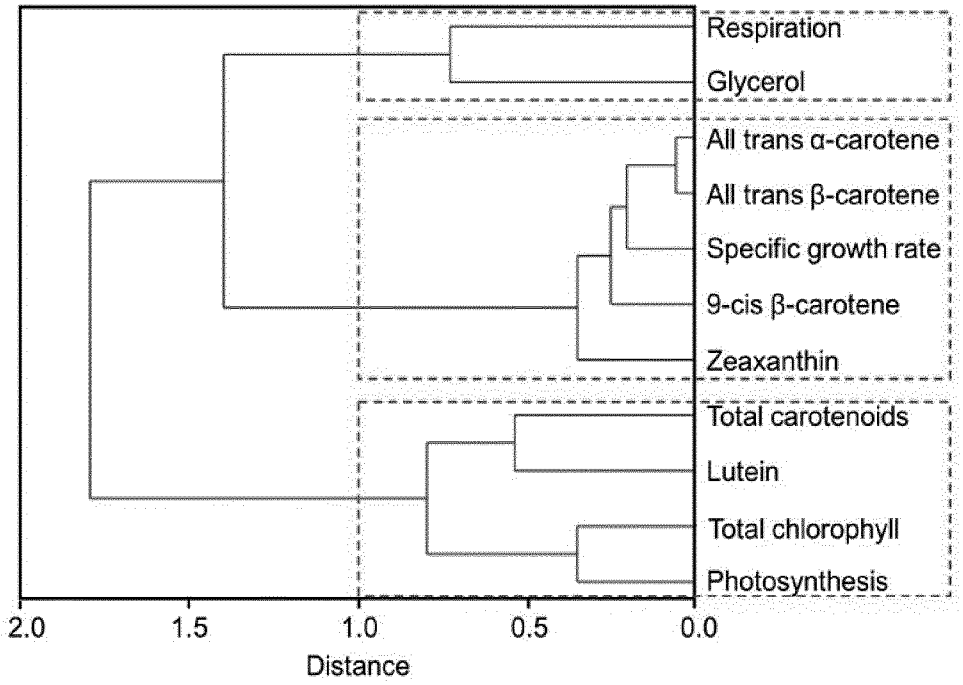

Figure 16c

DF17

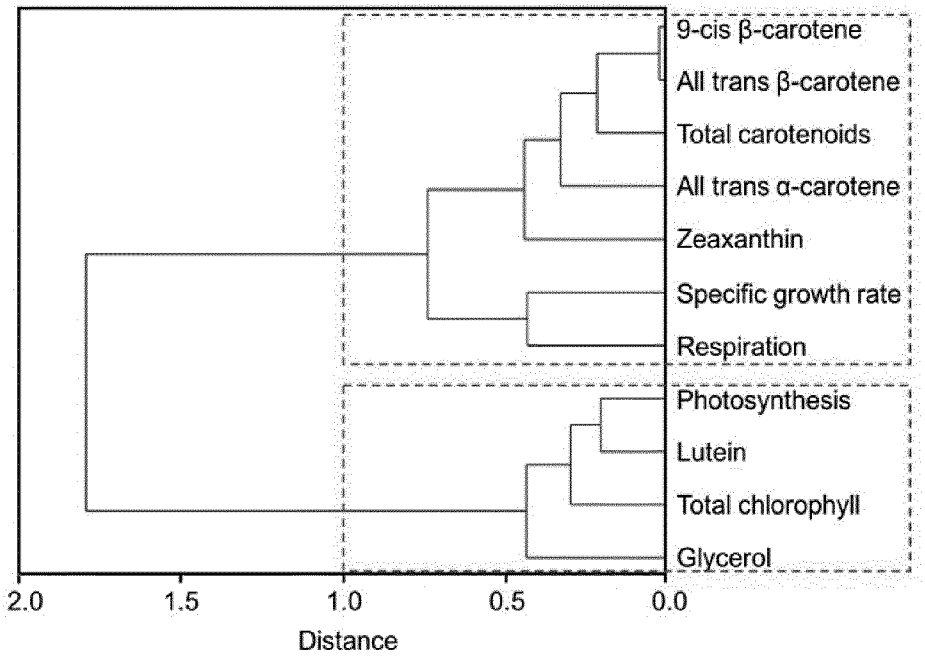

Figure 16d

DF40

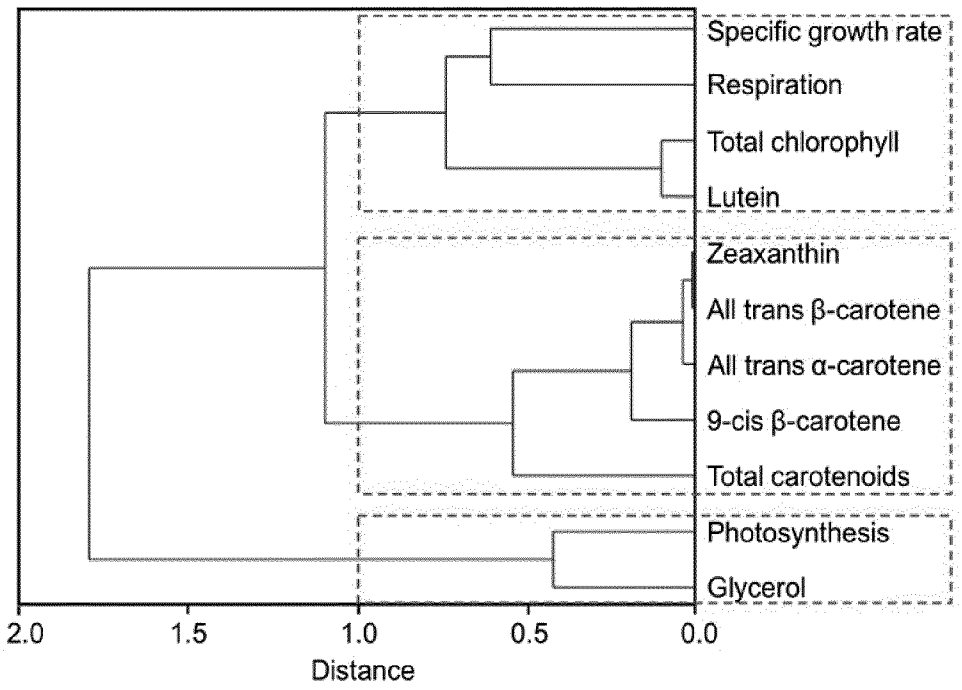

Figure 16e

UTEX 2538

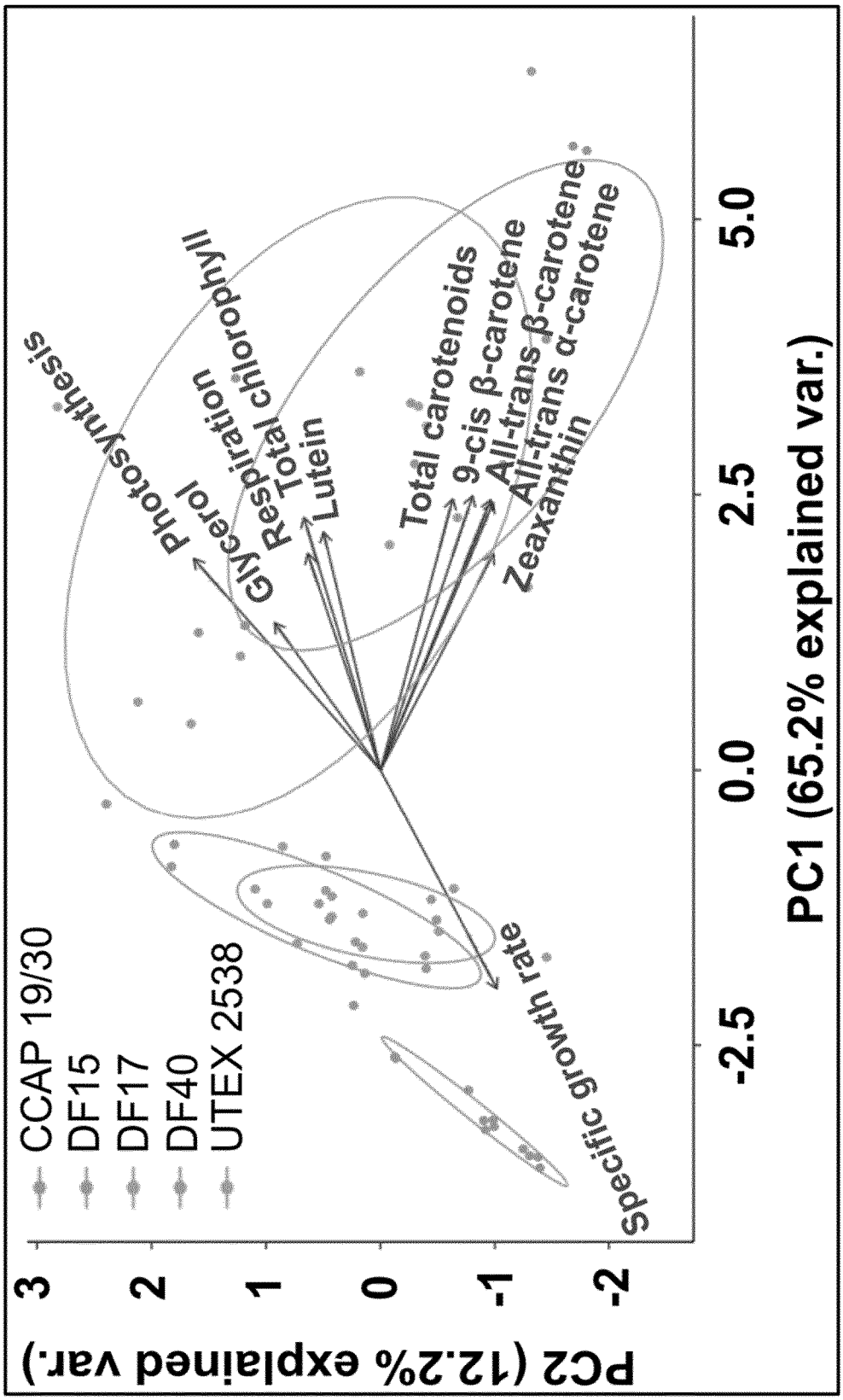
PC2 (12.2% explained var.)
3
2
1
0
-1
-2
-2.5
0.0
2.5
5.0
PC1 (65.2% explained var.)
CCAP 19/30
DF15
DF17
DF40
UTEX 2538
Photosynthesis
Phytol
Glycerol
Respiration
Total chlorophyll
Lutein
Total carotenoids
9-cis β-carotene
All-trans β-carotene
All-trans α-carotene
Zeaxanthin
Specific growth rate

Figure 17

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2018/052840

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12R1/89 C12P23/00 C12N1/12 C07C403/24
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12R C12P C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	OLMOS JORGE ET AL: "DNA fingerprinting differentiation between Î-carotene hyperproducer strains of Dunaliella from around the world", SALINE SYSTEMS, BIOMED CENTRAL, LONDON, GB, vol. 5, no. 1, 30 June 2009 (2009-06-30), page 5, XP021059921, ISSN: 1746-1448, DOI: 10.1186/1746-1448-5-5 the whole document page 6; table 1	12

-/--

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 May 2018	16/07/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Blanco Urgoiti, B

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	M. GARCÍA-GONZÁLEZ ET AL: "Conditions for open-air outdoor culture of Dunaliella salina in southern Spain", JOURNAL OF APPLIED PHYCOLOGY., vol. 15, no. 2/3, 1 March 2003 (2003-03-01), pages 177-184, XP055474343, NL ISSN: 0921-8971, DOI: 10.1023/A:1023892520443 the whole document -----	12
X	XU YANAN ET AL: "The influence of photoperiod and light intensity on the growth and photosynthesis ofDunaliella salina(chlorophyta) CCAP 19/30", PLANT PHYSIOLOGY AND BIOCHEMISTRY, GAUTHIER-VILLARS, PARIS, FR, vol. 106, 17 May 2016 (2016-05-17), pages 305-315, XP029673811, ISSN: 0981-9428, DOI: 10.1016/J.PLAPHY.2016.05.021 the whole document -----	12

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2018/052840

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

3, 14(completely); 1, 4-12(partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 3, 14(completely); 1, 4-12(partially)

   A composition or culture of Dunaliella salina algae strain DF15; use for producing a carotenoid; a method for producing beta-carotene; an algae biorefinery; beta-carotene;and algal strain thereof.

   ---

2. claims: 1, 4-12(all partially)

   A composition or culture of Dunaliella salina algae strain DF17; use for producing a carotenoid; a method for producing beta-carotene; an algae biorefinery; and beta-carotene thereof.

   ---

3. claims: 2, 13(completely); 1, 4-12(partially)

   A composition or culture of Dunaliella salina algae strain DF40; use for producing a carotenoid; a method for producing beta-carotene; an algae biorefinery; beta-carotene;and algal strain thereof.
