Molecular and phylogenetic analysis reveals new diversity of *Dunaliella salina* from hypersaline environments

Andrea Highfield¹, Angela Ward¹, Richard Pipe¹ and Declan C. Schroeder¹,²,³

¹The Marine Biological Association of the United Kingdom, The Laboratory, Citadel Hill, Plymouth PL1 2PB, UK; ²School of Biological Sciences, University of Reading, Reading RG6 6LA, UK and ³Veterinary Population Medicine, College of Veterinary Medicine, University of Minnesota, St Paul, MN 55108, USA

**Abstract**

Twelve hyper-β carotene-producing strains of algae assigned to the genus *Dunaliella salina* have been isolated from various hypersaline environments in Israel, South Africa, Namibia and Spain. Intron-sizing of the SSU rDNA and phylogenetic analysis of these isolates were undertaken using four commonly employed markers for genotyping, LSU rDNA, ITS, *rbcL* and *tufA* and their application to the study of *Dunaliella* evaluated. Novel isolates have been identified and phylogenetic analyses have shown the need for clarification on the taxonomy of *Dunaliella salina*. We propose the division of *D. salina* into four sub-clades as defined by a robust phylogeny based on the concatenation of four genes. This study further demonstrates the considerable genetic diversity within *D. salina* and the potential of genetic analyses for aiding in the selection of prospective economically important strains.

**Introduction**

*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 phytoesters (reviewed by Avron & Ben-Amotz, 1992). Species within this genus include halophilic and halotolerant strains and are frequently encountered in hypersaline environments. *Dunaliella* lacks a rigid cell wall and has a flexible cell membrane capable of rapidly changing shape in response to osmotic stress (Oliveira et al., 1980) and this can confound morphological identification. Cell size is highly variable with differences being related to growth conditions, e.g. nutrients, light intensity and salt concentration (Borowitzka & Brown, 1974), and can also vary within the same culture, for example *D. salina* (Teodoresco, 1905) can exhibit 5–29 µm cell length and 3.8–20.3 µm width (Borowitzka & Siva, 2007).

Early observations and descriptions of the genus *Dunaliella* laid the foundations of a well characterised group (Teodoresco, 1905; Lerche, 1937; Massjuk, 1973) with the family Dunaliellaceae comprising of four sections of *Dunaliella*, as described by Massjuk (1973); section Tertiolectae which are oligo-euhaline, 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 *Petriceinae* 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*. The latter has been debated as to whether it should actually be deemed to be a separate species, or is in fact *D. salina* considering its characteristics match the species description originally specified by Teodoresco (1905) (Borowitzka & Siva, 2007; González et al., 2009). *Dunaliella salina* has been found to be able to tolerate salt saturation of approximately 5.5 M NaCl (reviewed by Avron & Ben-Amotz, 1992). It is a particularly polymorphic species, with cell physiology related to abiotic factors. Cells can be pigmented green or red depending on the amount of β-carotene accumulated and under high stress conditions, members of this species can accumulate >5% β-carotene dry weight (Ben-Amotz & Avron, 1983). *D. bardawil* was isolated from a salt pond near Bardawil, Israel in 1976 and was reported to accumulate considerably larger amounts of β-carotene compared to *D. salina*, accumulating β-carotene in membrane-free globules in the interbilyakoid spaces of the chloroplast (Ben-Amotz et al., 1982). However, it is now unclear if the strain of *D. salina* used to compare to *D. bardawil* in this study was a good representative of *D. salina*, as subsequent reports have indicated that *D. bardawil* UTEX (Culture Collection of Algae at the University of Texas at Austin) 2538 and *D. salina* UTEX 1644 actually have a similar carotenoid: chlorophyll ratio (Jahnke, 1999). Morphologically it has proven difficult to resolve differences between strains of *D. salina*, however, four formae have previously been described in the literature. *D. salina* spp. *salina* fo. *sibirica*; a species with cells that are broader in the median or anterior region, *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 µm³ 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 (Massjuk & Radchênkô, 1973). 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.

Genotyping is now considered imperative in aiding the classification of Dunaliella species, with the ribosomal markers at the forefront of phylogenetic analyses (e.g. González et al., 2001; Olmos et al., 2009; Buchheim et al., 2010; Assunção et al., 2012), however, other commonly used markers for Dunaliella sp. and, indeed other marine algae, include rbcL (Preetha et al., 2012) and tufA (reviewed by Leliäert et al., 2014). Furthermore, by exploiting the sporadic occurrence of the group I introns within the 18S nuclear SSU rDNA (Wilcox et al., 1992), the SSU marker has been used as a size indicator of different Dunaliella species (Olmos-Soto et al., 2002). The absence of introns in the SSU rDNA in the Tertiolectae, one intron in D. salina, two introns in D. bardawil and one intron for D. viridis (that differs to the aforementioned species), have been used as an aid in identifying these species. Moreover, Olmos et al. (2009) also used the method as a way to distinguish hyper-producers of β-carotene from other Dunaliella sp. and to discriminate between D. salina var. Teod and D. salina/bardawil.

The variable regions, ITS 1, 5.8S rRNA and ITS 2, have been frequently employed by molecular studies of this important algal group (e.g. González et al., 2001) with some studies focusing on the ITS 2 spacer sequence (Assunção et al., 2012). Assunção et al. (2012) undertook a comprehensive analysis of Dunaliella species using ITS 2, and identified three main clades within the section Dunaliella; salina I, salina II, and pseudosalina. Following on from this, however, Assunção et al. (2013) allied the pseudosalina clade as putative D. viridis and furthermore, in agreement, Borowitza & Siva (2007), had also previously re-classified members of the pseudosalina clade based on morphology. Notably, they reassigned D. salina CCAP 19/3 as D. viridis, D. parva SAG 19-1 as D. maritima and D. salina UTEX 200 as D. viridis making the pseudosalina clade indeterminate. Assunção et al. (2013) concluded that many of the taxonomic assignments of Dunaliella strains based on morphological and physiological measurements were flawed and sequence analysis of the ITS 2 was required to affirm the taxonomic affiliation of numerous strains. Moreover, it appeared that some strains sequenced were different to what they were originally designated, potentially as a result of cross-contamination. ITS markers are often favoured as there are many copies in the genome making it easy to amplify 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 (Buchheim et al., 2010) and these have shown good potential for taxonomic resolution.

In addition to sequencing and phylogeny of the ITS 2, analysis of compensatory base changes (CBC) in the rRNA secondary structure of this region has also been shown to be useful in identifying organisms that are distinct species. A minimum of one CBC (with 93% confidence in plants and fungi) between two organisms can designate them as separate species (Muller et al., 2007). Notably, however, the absence of CBCs does not necessarily mean that they are the same species. Assunção et al. (2012) used this method in their investigations of Dunaliella sp. but found its application to the delineation of the group limited due to a lack of CBCs identified in isolates known to be distinct species.

Both rbcL and tufA are plastid genes, with tufA encoding elongation factor Tu, and has become more frequently used for molecular studies of algae (Cook et al., 2011; Moniz et al., 2014). Availability of sequences, however, is limited 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 phylogeny of tufA, rbcL and psbA. 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. Preetha et al. (2012) reported that rbcL gene phylogeny of Dunaliella did not show the same level of heterogeneity as the ITS, yet, the novel isolates analysed in their study showed similar groupings in both phylogenetic trees.

In addition to these useful barcoding markers for clonal isolates which can also be used for environmental analysis, the V9 variable region of the SSU rRNA, has frequently been employed for next generation sequencing of environmental samples (Amaral-Zettler et al., 2009; de Vargas et al., 2015) with an extensive database of sequences now available. The application of V9 to Dunaliella sp. taxonomy and its ability to resolve different species is therefore of importance to validate with regards to metagenomic diversity. To ensure this region can resolve the diversity of this group means that metagenomic analysis of hypersaline environments using this marker will accurately resolve the Dunaliella community diversity.

Suitable molecular markers 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 investigate a range of new Dunaliella isolates collected from a range of geographical provinces including Israel, Spain, South Africa and Namibia. These were compared to reference strains previously collected from Chile, Australia, Mexico, Norway and Israel. 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.

Materials and methods
Dunaliella isolation and culture

Water samples (50 mL) were collected from hypersaline locations detailed in Table 1 and were transported back to the laboratory. Dunaliella sp. were identified via light microscopy and isolated using single cell-picking with a micropipette and dilution techniques according to Anderson & Kawachi (2005). Established cultures were maintained in 30 mL F/2 media (Guillard & Ryther, 1962) with the addition of 50 g L$^{-1}$ sea salts (Sigma) (1.45 M NaCl), at 25°C continuous light of 100 μmol m$^{-2}$ s$^{-1}$. Cultures were regularly sub-cultured into fresh media every 2–3 weeks and cell imaging was undertaken 2 weeks post sub-culture using a DMi8 live cell imaging system (Leica). All isolates have been deposited in the Marine Biological Association culture collection.

DNA extraction, PCR and sequencing

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 which was 50 μL. PCR was carried out using a suite of primers (Table 2) in a Corbett Thermocycler (Corbett Research). PCR reactions were typically carried out in 50 μL volumes containing 2 μL DNA, 25 pmol each primer, 1 × reaction buffer, 2.5 mM MgCl$_2$, 0.0025 mM dNTPs, 1 Unit Gotaq polymerase (Promega) unless otherwise stated (Table 2). PCR reactions proceeded with an initial denaturation at 95°C for 5 min, followed by 35 cycles.
| MBA culture collection reference | Habitat               | Geographic origin         | Date collected (month/year) | Cell length (μm) (mean and range) | Cell width (μm) (mean and range) | Original microscopic identification | Cell colour at 25°C, constant light, 1.5 M NaCl |
|--------------------------------|-----------------------|---------------------------|-----------------------------|-----------------------------------|----------------------------------|-----------------------------------|------------------------------------------|
| DF15                           | Saltpan               | Eilat, Israel             | 03/2014                     | 18.2 (14.4–21.5)                  | 13.8 (10.4–18.7)                 | Dunaliella salina                   | Red                                      |
| DF17                           | Raceway               | Eilat, Israel             | 03/2014                     | 12.7 (10.9–15.5)                  | 10.3 (8.7–12.9)                  | Dunaliella salina                   | Yellow                                  |
| DF40                           | Raceway               | Monzon, Spain             | 05/2015                     | 14.4 (13.7–15.1)                  | 13.3 (12.0–14.6)                 | Dunaliella salina                   | Red                                      |
| DF41a                          | Raceway               | Monzon, Spain             | 11/2015                     | N/A                               | N/A                              | Dunaliella salina                   | Red                                      |
| DF45                           | Saltpan               | Monzon, Spain             | 11/2015                     | 13.6 (11.5–17.2)                  | 10.8 (7.9–13.2)                  | Dunaliella salina                   | Red                                      |
| SA3                            | Saltpan               | Velddrif, South Africa    | 02/2012                     | 13.9 (12.9–15.1)                  | 11.7 (10.4–12.9)                 | Dunaliella salina                   | Orange                                  |
| SA4                            | Saltpan               | Velddrif, South Africa    | 02/2012                     | 15.1 (12.9–17.2)                  | 12.2 (10.4–14.2)                 | Dunaliella salina                   | Orange                                  |
| T36                            | Saltpan               | Swakopmund, Namibia       | 06/2010                     | 18.0 (15.9–20.7)                  | 12.8 (9.8–16.1)                  | Dunaliella salina                   | Orange                                  |
| T37                            | Saltpan               | Swakopmund, Namibia       | 06/2010                     | 16.1 (11.9–19.3)                  | 12.7 (8.8–16.0)                  | 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, South Africa | 04/2013                   | 15.7 (13.3–18.1)                  | 15.2 (12.9–17.9)                 | Dunaliella salina                   | Orange                                  |
| T74a                           | Saltpan               | Velddrif, South Africa    | 04/2013                     | N/A                               | N/A                              | Dunaliella salina                   | N/A                                      |
| UTEX 999                       | Fjord                 | Oslo, Norway              | 1938                        | 9.4 (8.2–11.2)                    | 7.6 (6.2–8.7)                    | Dunaliella tertiolecta              | Green                                   |
| CCAP 19/30                     | Bardawil lagoon       | North Sinai, Israel       | 1978                        | 11.0 (9.9–13.8)                   | 7.9 (6.1–9.2)                    | Dunaliella bardawil now tertiolecta | Green                                   |
| UTEX 2538                      |                       |                           | 1982                        | 14.33 (10.4–18.4)                 | 10.51 (8.3–13.2)                 | Dunaliella bardawil                | Red                                     |
| ATCC 30861                     |                       |                           | 1982                        | 21.27 (17.4–23.0)                 | 18.78 (15.2–22.6)                | Dunaliella bardawil                | Red                                     |
| MUR 8                         | Hutt lagoon           | Australia                 | Unknown                     | 7.9 (6.9–9.0)                     | 6.4 (5.5–7.3)                    | Dunaliella salina                   | Green                                   |
| MUR 22                        | Hutt lagoon           | Australia                 | Unknown                     | 14.0 (12.5–15.6)                  | 11.2 (10–12.4)                   | Dunaliella salina                   | Orange                                  |
| CONC 001                       | Coastal saline pond   | Laguna la Rinconada, Chile| 1990                       | N/A                               | N/A                              | Dunaliella salina                   | N/A                                      |
| CONC 003, 004, 005, 006, 007, 008, 009 | Andean saline depression | Salar de Atacama, Chile | 1990 | N/A | N/A | Dunaliella salina | N/A |
| CCAP 19/18                     | Hypersaline Hutt lagoon | Australia                 | 1982                        | 14.1 (10.2–15.4)                  | 11.1 (9.8–15)                    | Dunaliella salina                   | Red                                     |
| CCAP 19/25                     | Unknown               | Unknown                   | Unknown                     | 14.0 (13.4–14.4)                  | 11.8 (11.1–12.3)                 | Dunaliella salina                   | Orange                                  |
| UTEX 1644                      | Point Colorado Salinas | Baja California, Mexico   | 1967                        | 15.5 (13.6–16.6)                  | 12.1 (11.1–14.0)                 | Dunaliella salina                   | Orange                                  |

1UTEX, Culture collection of Algae at the University of Texas, Austin, USA; 2CCAP, Culture collection of Algae and Protozoa, UK; 3ATCC, American type culture collection, VA, USA; 4MUR, Murdoch University Algal collection, Australia; 5CONC, Culture collection of Microalgae, Universidad de Concepcion, Concepcion, Chile.

*Isolates lost from culture.
of denaturation at 95°C for 30 s, annealing at 54°C for 45 s and extension at 72°C for 1 min unless otherwise stated (Table 2). PCR reactions had a final extension step of 72°C for 5 min. PCR products were either sequenced directly using the respective primers (Source Bioscience, Cambridge) 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). Purified PCR product (1 μL) was ligated into the pCR2.1 TA cloning vector (Invitrogen) and transformed according to manufacturer’s instructions. PCR amplicons from colony PCR with M13 primers were sequenced (Source Bioscience, Cambridge).

Sequences were manually verified for quality using Chromas (Technelysium Pty Ltd). Multiple sequence alignments were constructed in BioEdit 7.0 (Hall, 1999) using ClustalW. Phylogenetic analysis based on neighbour-joining and maximum likelihood was undertaken using MEGA 6 (Tamura et al., 2013). Bootstrap values were retrieved from 1000 replicates. Accession numbers for sequences generated in this study are detailed in Supplementary Table 1.

### Results

#### Microscopy

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* (Figure 1). These cultures (DF40, DF41, DF45, DF17, DF15, T36, T27, T41, T68, SA4 and SA3) ranged in mean cell length of 12.6–18.2 μm and cell width of 10.2–15.1 μm (Table 1). Cell shape was variable within cultures and stigma was not easily microscopically identified due to the presence of refractile granules.

### 18S intron-sizing method and sequence analysis

As anticipated, according to Olmos *et al.* (2009), *D. tertiolecta* had no introns (Table 3). We also identified two different groups of *D. salina*, as reported in Olmos *et al.* (2009), those with two introns and those with one. Further to this, however, some isolates that were assigned as *D. salina* microscopically were found to have no introns, i.e. they had the same MA1/2 18S amplicon size as *D. tertiolecta*. In this study, we PCR-amplified from two of the same isolates as Olmos and co-workers, with differing results. They reported *D. bardawil* UTEX 2538 to have two introns whereas we sized it to have one. Moreover, we sized *D. bardawil* ATCC (American Type Culture Collection) 30861, which is thought to be the same strain as UTEX 2538, and found it to have no introns. Strain ATCC 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 UTEX collection as strain 2538 by R. Adams between 1980 and 1982. The culture collection legacy of this strain continues as two other culture collections host this strain as the same as the original *D. bardawil* isolated by Ben-Amotz and Avron. SAG (Culture Collection of Algae and Protozoa) at Göettingen University (Göttingen, Germany) 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) identifying it as strain 19/30. Olmos *et al.* (2009) reported the CCAP 19/30 sequence in GenBank, that was deposited by ‘Herve’ in 2006, to have one intron which is in agreement with our result for ATCC 30861, however, both our 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 different strains and for those strains that had introns all of them shared the same sequence in GenBank, that was deposited by ‘Herve’ in 2006, to have one intron which is in agreement with our result for ATCC 30861, however, both our study and Olmos *et al.* (2009) reported CCAP 19/30 to have no introns.

### Table 2. Primers used in this study

| Primer name | Target | Sequence (5’-3’) | Size (bp) | Anneal temp (°C) | PCR reaction | Reference |
|-------------|--------|-----------------|-----------|------------------|--------------|-----------|
| AB1F        | ITS    | AATCTATCAATTACACACACCG | 700       | 52               | 20 pmol primers; 1.5 mM MgCl₂; 1 min extension | Hejazi *et al.* (2010) |
| AB2R        |        | TTTCACTTCGTGATATTAGG  |           |                  |              |           |
| MA1F        | 18S Dunaliellaceae | CCGGATCTCAGCAGTAGCTATGCCTGCTC | 1700-2500 | 52               | Extension time 3 minutes; 30 pmol primers; 1.5 mM MgCl₂ | Olmos *et al.* (2009) |
| MA2R        |        | CGGAATTCTCTGAGGTTTACC |           |                  |              |           |
| rbclL       | rbcl Chlorophyta | CGTACGATATGAAACCAATAT | 700       | 54               | As text | Nozaki *et al.* (1995) |
| rbClR       |        | AAAGATTTAATCAGGGCA |           |                  |              |           |
| tufA_F      | tufA Chlamydomonales | GAAGARAAAGCWCGYGGATTAC | 750       | 54               | As text | This study |
| tufA_R      |        | CCATACCCGTTGACATCTTT |           |                  |              |           |
| 1391F       | Universal SSU rDNA (V9) | GTACACACGCGGCTGC | 168       | 54               | 25 μl reaction; 10 pmol primer; 1 mM MgCl₂; 0.5 U taq; 20 s annealing and extension | Lane (1991); Medlin *et al.* (1988) |
| EukB        |        | TGTACCTTGTCAGGTTCACCACCTAC |           |                  |              |           |
| D1R_F       | Universal LSU rDNA | ACCGCCGTAAATTAAAGCATA | 760       | 60               | 20 pmol primers; 1.5 mM MgCl₂; 1 min extension | Scholin *et al.* (1994) |
| D2C_R       |        | GCTTGGTCCGTTGTTTCAAGA |           |                  |              |           |
DF15 amplicons shared, there was 98.5% similarity and the large amplicon intron and small amplicon intron shared 89.6% and 88.8% identity with *D. bardawil* UTEX 2538, respectively. BLASTn analysis showed these introns to be most similar to *Dunaliella* sp. ST13 (>98%). The second intron in the larger product was unique, sharing 99.23% identity also with *Dunaliella* sp. strain ST13.

Phylogenetic analysis

Strains CCAP 19/30 and MUR 8 were omitted from the phylogenetic analysis as morphological analysis showed that they did not turn red under stress. Furthermore, BLAST analysis of the sequences generated from these two strains and also CONC 001 showed them to be more similar to *D. tertiolecta* (data not shown) and hence confirmed their exclusion from the *D. salina* phylogeny. GenBank holds a vast number of sequences originating from *Dunaliella* species which are available to include in phylogenetic analysis. Because of some of the problems reported with strain identification and/or cross contamination we chose to only include key strains that we obtained from culture collections or DNA where the culture was not available.

The large-subunit (LSU) rDNA phylogeny formed a strongly supported clade of five out of seven of the *D. salina* strains isolated from Salar de Atacama, Chile, in 1990, however, other subclades were weakly supported, owing to only a small number of nucleotide substitutions (Figure 2). The clade containing *D. bardawil* UTEX 2538 clustered separately from *D. salina* UTEX 1644 owing to 6 bp substitutions across 477 bp of the alignment.

The *rbcL* phylogeny (Figure 3) divided *D. salina* broadly into two separate clades, corresponding to the clustering patterns described by Assunção *et al.* (2012) (salina I and salina II). The *rbcL* phylogeny further sub-divides the Chilean clade identified in the LSU phylogeny, but with the addition of CONC 009 which clustered separately in the LSU phylogeny.

---

**Fig. 1.** Images of a representative set of *Dunaliella salina* strains isolated in this study compared with strains imaged from culture collections. All isolates were cultured in identical conditions and imaging for each isolate was undertaken 2 weeks following sub-culturing. Numerals in brackets after the strain name indicates the *D. salina* sub-group. (a) *D. salina* DF40 (I), (b) *D. salina* DF45 (I), (c) *D. salina* CCAP 19/18 (II), (d) *D. salina* MUR22 (II), (e) *D. salina* DF17 (II), (f) *D. salina* CCAP 19/25 (II), (g) *D. salina* UTEX 1644 (II), (h) *D. salina* T41 (III), (i) *D. salina* T37 (III), (j) *D. salina* DF15 (IV). Scale bar is equivalent to 25 μm.
The tufA phylogeny (Figure 4) grouped all the new isolates and previously characterized *D. salina* strains in one clade with all the sequences >99% similar to each other.

The phylogenetic tree based on ITS1 + ITS2 (Figure 5) provided much more resolution compared with the other trees and was robust, supported by high bootstrap values. In order to complement the phylogeny of González et al. (2009) who identified three main *D. salina* clades, we included the sequence for *D. salina* AC144 in this analysis. This strain was the sole member of one of the clades identified by González et al. (2009) and therefore was of interest to examine its relationship with the strains from this study. Two salina groups were identified in this phylogeny, according to Assunção et al. (2012), and these were clearly separated by long branch lengths with salina I which encompasses *D. bardawil* UTEX 2538 and salina II which encompasses *D. salina* UTEX 1644. The Chilean strains were separated into two groups, matching the LSU phylogeny. 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. In the LSU phylogeny (Figure 2) DF15 clustered with the *D. salina* UTEX 1644 (salina II) group but with long branch lengths separating them. However in the *rbcL* (Figure 3) phylogeny, this strain groups with *D. bardawil* UTEX 2538 (salina I). The ITS phylogenetic tree places DF15 intermediary to the clades classified as salina I or II by Assunção et al. (2012). BLASTn analysis of the DF15 ITS sequence identified *Dunaliella* sp. ST13 as most similar yet with only 98.67% identity. The sequence for *D. salina* AC144 weakly clustered with the top half of the tree with low bootstrap values.

**Table 3.** Results from intron sizing and sequence alignment of introns

| Isolate name   | # Introns | Intron 1 group | Intron 2 group |
|----------------|-----------|----------------|----------------|
| 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         | –              | –              |
| T4             | 0         | –              | –              |
| SA3            | 0         | –              | –              |
| SA4            | 0         | –              | –              |
| T36            | 0         | –              | –              |
| T37            | 0         | –              | –              |
| T41            | 0         | –              | –              |
| T68            | 0         | –              | –              |
| DF15           | 1 and 2   | B              | Q              |
| *D. tertiolecta* UTEX 999 | 0   | –              | –              |
The concatenated tree of the 4 markers, ITS-LSU-rbcL-tufA (Figure 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. Four sub-clades of Dunaliella salina could be identified, and these were supported by significant bootstrap values. Within each of the subclades the strains shared 99% identity, with all the strains morphologically identified as D. salina sharing at least 97% identity.

V9 (SSU rDNA) amplicon sequencing was undertaken for DF15, DF17, D. bardawil UTEX 2538, D. salina CONC 008, D. salina CCAP 19/18, D. salina UTEX 1644, D. salina CCAP 19/25, D. salina CONC 003, D. salina UTEX 1644, D. salina CCAP 19/18 and D. tertiolecta UTEX 999 only (Figure 7). Nucleic acid alignment showed that Dunaliella of the tentative clades, salina I, and tertiolecta were identical, however salina II, was different. DF15 was identical to salina I and D. tertiolecta.

**Discussion**

The intron sizing method previously revealed useful information regarding the history of Dunaliella spp. in culture collections. Whilst providing some data indicating the taxonomic affiliation of particular strains we found the method was not sufficiently robust. The fact that a group of D. salina have no introns (the same as D. tertiolecta) negates the intron-sizing as a tool for separating these strains from each other as both an isolate of D. tertiolecta and D. salina would produce the same size amplicon, despite these two being very different species. Selection based on biochemistry could however easily separate these two species prior to intron-sizing if this method was preferred. Olmos et al. (2009) concluded that cross-contamination of cultures had occurred which resulted in differences in amplicon sizes in what was thought to be the same strain. Our study showed similar discrepancies with identical strains D. bardawil, strain UTEX 2538 and strain ATCC 30861 from separate culture collections having a different number of introns. This could indicate cross-contamination in its culturing history, primary non-clonal cultures or could indicate intron loss in one of the strains in the different culture collections. Indeed in our phylogenetic analyses the two aforementioned strains consistently group together with identical tufA and rbcL sequences, ITS sequences differing by 2 bp over a 553 bp multiple sequence alignment and LSU sequences differing by 5 bp substitutions over 477 bp. Considering the original isolate was collected in 1976, genetic divergence in separate culture collections is conceivable although without more information this cannot be determined. Certainly, with advances in cryopreservation one would be able to track such events more closely by resurrecting and comparing with regularly sub-cultured algae.

Our study confirms that, based on the sequencing analysis of four marker genes, the strains CCAP 19/30, MUR 8 and CONC 001 that we acquired in this study are not D. salina. This is not to say that other cultures of these strains around the world are not true representatives of the original designation of these strains, however, but because of our observations they were not included in the analyses here. These results are a cautionary tale for anyone working on algal cultures, with genotyping of the
strains recommended as an integral step to ascertain the taxonomic affiliation of the strains one is studying.

The sequences for DF15 nuclear SSU rDNA were shown to be unique, with two different versions amplified. Whilst there is a high copy number of the rDNA, concerted evolution typically moves towards resulting identical sequences of this gene within the genome so it is surprising to get two different versions of this gene in DF15. Since its clonality was verified through multiple rounds of single cell picking, intragenomic sequence variation is a possible explanation. Certainly, Alverson & Kolnick (2005) reported intragenomic nucleotide polymorphisms within the SSU rDNA of the diatom *Skeletonema*, 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 this was indeed the case one would anticipate identifying this heterogeneity in more strains.

Comparing the four markers analysed in this study shows that *tufA* is the least useful in resolving intraspecific diversity within *D. salina*. However, the *tufA* analysis has been of use in that it supports the designation of all of these strains to the species *D. salina*. González et al. (2009) discussed the high intraspecific diversity within *D. salina* and the potential existence of more than one biological species, yet these data support the existence of only one. The LSU and *rbcL* phylogenies show more resolution than *tufA*, providing some information on the relationships between the different strains, but are less robust than the ITS. The phylogenetic tree based on the ITS regions gives the most information, with subdivisions strongly supported by high bootstrap values. The concatenated tree of the four genes further reinforces the ITS region phylogeny with a clearer delineation of different phylogroups. Including intron presence/absence data and sequences can provide further supporting information on some subdivisions, such as the separate DF15 clade.

As a marker on its own, the V9 region of SSU rDNA does not offer enough variability for accurate taxonomy for the *D. salina* species. The V9 region lends itself to next generation sequencing methodologies, due to its heterogeneity and short length, however, the fact that not all the groups can be resolved with this marker raises questions on its suitability for analysis of hypersaline microbial communities, as a significant portion of *D. salina* diversity will be missed. The V2–V4 region of the SSU rDNA was found to have the best phylogenetic resolution compared with the other hypervariable regions in dinoflagellates (Ki, 2012), however, this was not tested here.

There is definitive *D. salina* taxa defined by its morphology and physiology with sequence information from our study demonstrating a robust segregation into four clades. We support the two sub-clades designated by Assunção et al. (2012), *salina* I and II, and identify two further groups. *Salina* II, as previously identified (Assunção et al., 2012), comprises 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. The *salina* I clade contains originally classified *D. bardawil*...
strains including the first isolate within this clade, *D. bardawil* UTEX 2538 isolated in 1976 and also *D. salina* CCAP 19/18, which has recently been proposed as a reference strain for the species by Polle et al. (2020). We propose two other sub-clades within *D. salina*, the salina III group which encompasses strains T36, T37, T41, T68, SA3 and SA4 isolated in this study. These form a robust group and therefore are considered distinct from the salina II group. Finally, we propose a new salina IV sub-clade that contains DF15 as the sole member. The ITS phylogeny potentially also suggests the *D. salina* AC144 (as analysed by González et al., 2009) is a unique strain, however due to unavailability of this strain, comprehensive genotyping could not be undertaken to corroborate this.

Strain DF15 is a distinctive strain of *D. salina* as, significantly, it does not delineate with either group. This strain is somewhat of an enigma with the different markers used clustering it (weakly) with different groups, e.g. DF15 clusters with strain DF17 for LSU but with strain UTEX 2538 in the rbCL phylogenies. Interestingly, biochemical analyses of this strain found it to be the highest producer of 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 at four different light intensities and showed that DF17 and DF40 clustered closely whereas our phylogenetic analysis separates them into different clades. This highlights the ambiguities between genetic data and biochemical data, however, we believe that these data are supportive of DF15 belonging to a unique sub-clade of *D. salina*. Upon identifying other strains that can be assigned to this clade we can confirm that it is a group characterized by the production of exceptional quantities of carotenoids. Genetic information from this strain and the identification of its sub-clade could be used to screen environments for this strain which have the potential to produce similarly high quantities. Due to ongoing revisions and changes in *D. salina* taxonomy we have refrained from the designation of new sub-species.

In conclusion this study has revealed the further diversity of *D. salina* isolates collected from South Africa, Israel, Namibia, Spain and in particular the novel strain DF15. In combination with morphological analysis the suite of genes used for molecular analysis has permitted the separation of *D. salina* into four sub-clades. 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. The classification of *Dunaliella* has proven to be a challenge over many decades and we recommend that the *Dunaliella* community needs to come together to tackle the challenge in unison with the vision to facilitating the selection of important strains for key uses.

**Supplementary material.** The supplementary material for this article can be found at [https://doi.org/10.1017/S0025315420001319](https://doi.org/10.1017/S0025315420001319)

**Acknowledgements.** We thank Patricia Gomez for supplying DNA for Chilean *Dunaliella* sp. and Prof. W. Wilson for helpful comments that improved the manuscript. We would also like to thank Dr Maya Pfaff, Department of Environmental Affairs, Forestry and Fisheries South Africa for assisting Prof. Declan Schroeder in the sampling of the South African salt pans and Sophia Rhodes for providing her facilities for purposes of sample processing prior to transport.

**Financial support.** This research was supported by the D-factory and the EU KBBE.2013.3.2-02 programme (D-Facory: 368 613870) and the ValgoOrize consortium, subsidised by the Interreg 2 Seas programme 2014–2020, co-financed by the European Fund for Regional Development in the framework of subsidy contract nr. 2005-17.

**References**

Alverson AJ and Kolnick L (2005) Intragenomic nucleotide polymorphism among small subunit (18S) rDNA paralogs in the diatom genus *Skeletomena* (Bacillariophyta). *Journal of Phycology* 41, 1248–1257.

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

Anderson RA and Kawachi M (2005) Traditional microalgae isolation techniques. In Anderson RA (ed.), *Algal Culturing Techniques*. London: Academic Press.

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

Assunção P, Jaén-Molina R, Caujapé-Castells J, Wolf M, Buheim MA, de la Jara A, Freijanes K, Carmona L and Mendoza H (2013) Phylogenetic analysis of ITS2 sequences suggests the taxonomic re-structuring of *Dunaliella viridis* (Chlorophyceae, Dunaliellales). *Physiological Research* 61, 81–88.

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

Ben-Amotz A and Avron M (1983) On those factors which determine the massive β-carotene accumulation in the halotolerant alga *Dunaliella bardawil*. *Plant Physiology* 72, 593–597.
Ben-Amotz A, Katz A and Avron M (1982) Accumulation of β carotene in halo-tolerant algae: purification and characterisation of β-carotene-rich globules from Dunaliella bardawil (Chlorophyceae). Journal of Phycology 18, 529–537.

Borowitzka LJ and Brown AD (1974) The salt relations of marine and halo-philic species of the unicellular green algae, Dunaliella. Archives of Microbiology 96, 37–52.

Borowitzka MA and Siva CJ (2007) The taxonomy of the genus Dunaliella (Chlorophyta, Dunaliellales) with emphasis on the marine and haloophilic species. Journal of Applied Phycology 19, 567–590.

Buchheim MA, Kirkwood AE, Buchheim JA, Verghese B and Henley WJ (2010) Hypersaline soil supports a diverse community of Dunaliella (Chlorophyceae). Journal of Phycology 46, 1038–1047.

Cook SS, Whittock I, Wright SW and Hallegraeff GM (2011) Photosynthetic pigment and genetic differences between two Southern Ocean morphotypes of Emiliania huxleyi (Haptophyta). Journal of Phycology 47, 615–626.

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berny C, Le Bescot N, Probert I, Carmichael M, Poulin J, Romac S, Colin S, Aury JM, Bittner L, Chaﬀron S, Dunthorn M, Engelen S, Flegontova O, Guidi I, 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, Pichéral M, Searson K, Kandels-Lewis S, Acinas SG, Bork P, Bowler C, Gorsky G, Grimley N, Hingamp P, Judecine D, Not F, Ogata H, Pesant S, Raes J, Sieracki ME, Speich S, Stemmann L, Sunagawa S, Weissbach J, Wincker P and Karsenté E (2015) Eukaryotic plankton diversity in the sunlit ocean. Science (New York, N.Y.) 348, 6237.

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

González MA, Gómez PI and Polle JEW (2009) Taxonomy and phylogeny of the genus Dunaliella. In Ben-Amotz A, Polle JEW and Subba Rao DV (eds), Biotechnology Letters 24, 1035–74.

Guillard RRL and Ryther JH (1962) Studies of marine planktonic diatoms. I. Cyclotella nana (Hustedt) and Detonula confervacea (Cleve) Gran. Canadian Journal of Microbiology 8, 229–239.

Hall TA (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41, 95–98.

Hejazi MA, Barzegari A, Gharajeh NH and Hejazi MS (2010) Introduction of a novel 18S rDNA gene arrangement along with distinct ITS region in the saline water microalg Dunaliella. Saline Systems 6, 4.

Jahnke LS (1999) Massive carotenoid accumulation in Dunaliella bardawil induced by ultraviolet-A radiation. Journal of Photochemistry and Photobiology B: Biology 48, 68–74.

KJ JS (2012) 18S rRNA using a large dataset for marker considerations. Journal of Applied Phycology 24, 1035–1043.

Lane DJI (1991) 16S/23S Sequencing. In Stackebrandt E and Goodfellow M (eds), Nucleic Acid Technologies in Bacterial Systematics. New York, NY: Wiley, pp. 115–175.

Leliaert F, Verbruggen H, Vanomringen P, Steen F, López-Bautista JM, Zuccarello GC and De Clerck O (2014) DNA-based species delimitation in algae. European Journal of Phycology 49, 179–196.

Leche W (1937) Untersuchungen uber Entwicklung und Fortpflanzung in der Gattung Dunaliella. Archiv fur Proteinkunde 88, 236–268.

Massjuk NP (1973) Morphology, Taxonomy, Ecology and Geographic Distribution of the Genus Dunaliella. Teod. and Prospects for its Potential Utilisation. Kiev: Naukova Dumka Press.

Massjuk NP and Radchenko MI (1973) New taxa of the genus Dunaliella Teod. III. Ukrainian Botanical Journal 30, 470–471.

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

Moniz MJ, Guiy M and Rindi F (2014) Tufa phylogeny and species boundaries in the green algal order Prasiolales (Trebouxiophyceae, Chlorophyta). Phycologia 53, 396–406.

Muller T (2005) Genetic fingerprints of microalgal culture strains: amplified fragment length polymorphism (AFLP) for investigations below the species level (PhD thesis). Universität zu Göttingen.

Müller T, Philippi N, Dandekar T, Schultz J and Wolf M (2007) Distinguishing species. RNA 13, 1469–1472.

Nozaki H, Itoh M, Sano R, Uchida H, Watanabe MM and Kuroiwa T (1995) Phylogenetic relationships within colonial volvocales (Chlorophyta) inferred from rbcL gene sequence data. Journal of Phycology 31, 970–979.

Olivaie L, Bisalputra T and Antia NJ (1980) Ultrastructural observation of the surface coat of Dunaliella tertiolecta from staining with cationic dyes and enzyme treatments. New Physiologist 85, 385–392.

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

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

Polle JEW, Jin ES and Ben-Amotz A (2020) The alga Dunaliella revisited: looking back and moving forward with model and production organisms. Algal Research 49, 101948.

Preetha K, John L, Subin CS and Vijayan KK (2012) Phenotypic and genetic characterisation of Dunaliella (Chlorophyta) from Indian salinas and their diversity. Aquatic Biosystems 8, 27.

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

Tamura K, Stecher G, Peterson D, Filipski D and Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution 30, 2725–2729.

Teodorescu EC (1905) Organisation et développement du Dunaliella, nouveau genre de Volvocacée-Polyblepharidée. Beithefe zum Botanischen Centralblatt 18, 215–232.

Wilcox LW, Lewis LA, Fuerst PA and Floyd GL (1992) Group I introns within the nuclear-encoded small-subunit rRNA gene of three green algae. Molecular Biology and Evolution 9, 1103–1118.

Xu Y, Ibrahim IM, Wosu CI, Ben-Amotz A and Harvey PJ (2018) Potential of new isolates of Dunaliella salina for natural β-carotene production. Biology 7, 14.