Genomic adaptations of the green alga *Dunaliella salina* to life under high salinity.

Jürgen E.W. Polle[1,2,3*]
Sara Calhoun[3]
Zaid McKie-Krisberg[1,4]
Simon Prochnik[3,5]
Peter Neofotis[6]
Won C. Yim[7]
Leyla T. Hathwaik[7]
Jerry Jenkins[3,8]
Henrik Molina[9]
Jakob Bunkenborg[10]
Igor Grigoriev[3,11]
Kerrie Barry[3]
Jeremy Schmutz[3,8]
EonSeon Jin[12]
John C. Cushman[7]
Jon K. Magnusson[13]

[1]Department of Biology, Brooklyn College of the City University of New York, Brooklyn, NY 11210, USA
[2]The Graduate Center of the City University of New York, New York, NY 10016 USA
[3]U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4]Current address: Department of Information Services and Technology, SUNY Downstate Health Sciences University, Brooklyn, NY 11203, USA
[5]Current address: MBP Titan LLC, South San Francisco, CA 94080, USA
[6]Current address: U.S. Department of Energy – Plant Research Laboratory, Michigan State University, East Lansing, MI, 48824 USA
[7]University of Nevada, Department of Biochemistry and Molecular Biology, Reno, Nevada, USA
[8]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA
[9]The Proteomics Resource Center, The Rockefeller University, New York, New York, USA
[10]Alphalyse A/S, Odense, Denmark
[11]Department of Plant and Microbial Biology, University of California - Berkeley, 111 Koshland Hall, Berkeley, CA 94720, USA
[12]Department of Life Science, Hanyang University, Research Institute for Natural Sciences, Seoul, Republic of Korea
[13]Pacific Northwest National Laboratory, Richland, Washington, USA

* Corresponding author: Dr. Jürgen E.W. Polle, jpolle@brooklyn.cuny.edu

Abstract

Life in high salinity environments poses challenges to cells in a variety of ways: maintenance of ion homeostasis and nutrient acquisition, often while concomitantly enduring saturating irradiances. *Dunaliella salina* has an exceptional ability to thrive even in saturated brine solutions. This ability has made it a model organism for studying responses to abiotic stress factors. Here we describe the occurrence of unique gene families, expansion of gene families, or gene losses that might be linked to osmoadaptive strategies. We discovered multiple unique genes coding for several of the homologous superfamily of the Ser-Thr-rich glycosyl-phosphatidyl-inositol-anchored membrane family and of the glycolipid 2-alpha-mannosyltransferase family, suggesting that such components on the cell surface are essential to life in high salt. Gene expansion was found in families that participate in sensing of abiotic stress and signal transduction in plants. One example is the patched family of the Sonic Hedgehog receptor proteins, supporting a previous hypothesis that plasma membrane sterols are important for sensing changes in salinities in *D. salina*. We also investigated genome-based capabilities regarding glycerol metabolism and present an extensive map for core carbon metabolism. We postulate that a second broader glycerol cycle exists that also connects to photorespiration, thus extending the previously described glycerol cycle. Further genome-based analysis of isoprenoid and carotenoid metabolism revealed duplications of genes for 1-deoxy-D-xylulose-5-phosphate synthase (DXS) and phytoene synthase (PSY), with the second gene copy of each enzyme being clustered together. Moreover, we identified two genes predicted to code for a prokaryotic-type phytoene desaturase (CRTI), indicating that *D. salina* may have eukaryotic and prokaryotic elements comprising its carotenoid biosynthesis pathways. In brief, our genomic data provide the basis for further gene discoveries regarding sensing abiotic stress, the metabolism of this halophilic alga, and its potential in biotechnological applications.

1. Introduction

Species of the genus *Dunaliella* are the dominant primary biomass producer for many hyper-saline lagoons and lakes [1,2,3]. The unicellular green alga *Dunaliella salina* TEODORESCO [4] is the type species of the genus *Dunaliella* [4,5,6,7] and often found to dominate salt-saturated brines.

To live in hyper-saline waters and to survive in habitats where rapid changes in salinities can occur, organisms have to adapt to maintain ion homeostasis and acquire nutrients efficiently [8,9]. Different hypersaline lakes and saltern crystallizer ponds vary greatly in their nutrient content, sometimes even with no detected nitrate and/or low phosphate concentrations [10]. Moreover, different carbonate salts can form during evaporation and with increasing salinities resulting in depletion of inorganic carbon from the water bodies as the salts precipitate [11,12]. In hypersaline sodium chloride solutions the solubility of iron is mainly determined by pH and temperature, with the solubility being extremely low and most iron present in unavailable colloidal form [13]. Consequently, algae in hypersaline environments have to actively scavenge and take up iron. Several known adaptations for *Dunaliella* to life in habitats that result in high and often varying salinities include the following:

1. Instead of a rigid cell wall, cells are surrounded by an elastic pericellular matrix [4,7,14], which allows instant cell volume adjustments either by expansion when the extracellular salinity suddenly drops or by rapid shrinking when the salt concentration around the cell increases.
2. Cells use glycerol as the main osmotic counter molecule [2,15,16,17,18,19]. The mechanistic aspects of hyper- and/or hypo-osmotic changes in cell size, shape, and metabolism of *Dunaliella* species have been studied intensely [20,21,22,23,24]. Cells not only evolved to grow in a saline/hypersaline environment, but also can adjust within minutes to severe alterations in salinity that result in hypertonic or hypotonic osmotic shock.
3. Cells have special ion transporters in the plasma membrane to acquire iron [8,22,25,26].
4. Cells possess unique carbonic anhydrases involved in carbon concentration [25,27,28].

The species *D. salina* and its close relative *D. bardawil* are model organisms not only for studies regarding adaptions to high salt, but also for general stress metabolism [29]. For example, cells of *D. salina* and *D. bardawil* respond to environmental abiotic stresses by accumulation of secondary β-carotene in plastidic globules to levels of about 8% of the dry weight [30]. Because of their carotenogenic capabilities, both species have been used commercially already for decades for β-carotene production [31,32,33,34] and both species are models for stress-induced β-carotene accumulation [32,35,36].

For the first time, the genome of *D. salina* [37] provided us with the opportunity to mine for and identify novel genes that may be involved in osmoadaptation mechanisms. Our data may serve as a basis for comparison and validation of other studies. Here we report on genomic adaptations to life in high salt and on the annotation of a number of genes coding for components involved in environmental stress response, with an emphasis on signaling, transport mechanisms, glycerol metabolism, and carotenoid metabolism.

2. Materials and Methods

2.1 Cultivation

The alga *D. salina* strain CCAP 19/18 was obtained from the Culture Collection of Algae and Protozoa (CCAP) (Argyll, Scotland). Based on analysis of morphological characters, physiological capability for carotenogenesis under abiotic stress conditions, and molecular rDNA marker sequence analysis, the strain CCAP19/18 is a genuine strain of the species *D. salina* [38]. Cells were grown in 1.0 M NaCl artificial sea water (ASW) medium (see the ASW recipe at www.ccap.ac.uk) under a photon flux density of 150 µmol m^{-2} s^{-1} (16 h light (26°C)/ 8 h dark (20°C) cycle with bubbling air. Alternatively, cells were cultivated at different salinities in artificial seawater medium according to Pick et al. [39]. To obtain broad coverage for the transcriptome, in addition to taking samples during the different phases of growth of phototrophic batch cultures, samples were taken as time courses following transitions from regular growth in batch culture to anaerobiosis, salt stress, light stress, and nutrient stress (low nitrogen, phosphorus, sulphur, and iron conditions). Samples were taken at 0h, 1h, 2h, 3h, 6h, 24h, 48h following transition of growth conditions. In addition, samples were obtained from short-term osmotic shock experiments (0h, 15 min, 30 min, 60 min, 90 min).

2.2 RNA Extraction, Library Preparation, and Sequencing

To improve gene models, we used RNA sequencing information. For this purpose, total RNA was isolated from cells using either the QuickPrep^{TM} mRNA Purification Kit (Millipore-Sigma, St. Louis, MO) or using a standard Trizol protocol [40]. RNA integrity and purity was determined by gel electrophoresis and by spectrometry. Total RNA was provided to JGI for 454 Sequencing using standard library and sequencing protocols. In addition, biomass from short term osmotic shock experiments was sent to Macrogen Corporation (Seoul, Republic of Korea) for total RNA extraction, standard library preparation, and Illumina HiSeq 2000 platform sequencing (Paired-end, 100 bp).

To generate sequencing data on the Roche 454 Titanium platform, double-stranded cDNA samples were fragmented via sonication to 400-800 bp. These fragments were end polished and ligated to a set of Y-shape adaptor. This process is called “454 Rapid Library Construction”. The 454 library fragments were then clonally amplified in bulk by capturing them through hybridization on microparticle beads. Resulting beads were covered with millions of copies of a single DNA fragment (range 400-800 bp) where each bead contains a different clonally amplified library fragment. Beads were then subjected to emulsion-based Polymerase Chain Reaction. After amplification, the beads were recovered from the emulsions and were loaded into the wells of a PicoTiterPlate device (PTP) such that wells contained single DNA beads. The PTP was then inserted into the 454 Genome Sequencer FLX-Titanium instrument for sequencing where sequencing reagents were sequentially flowed over the plate and the sequence of the DNA fragments was determined.

In addition, transcriptomics data were obtained using Sanger sequencing. The 2 kb and 8 kb insert-size libraries were made from oligo(dT) primed cDNA and cloned into the pCMV-SPORT6 vector (Invitrogen Corp., Carlsbad, CA). PolyA RNA was primed with an oligo dT primer (5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTTT-3'), ligated to a SalI adapter (5'-TCGACCCACGCGTCCG and 5'- CGGACGCGTGGG) and digested with NotI. cDNA was size selected using 1.1% agarose gel electrophoresis, then ligated into NotI and SalI-digested pCMV-SPORT6 vector.

2.3 Genome annotation and sequence analysis

The genome assembly of *D. salina* [37] was annotated using the US Department of Energy Joint Genome Institute (JGI) annotation pipeline, which combines several gene predictors [41,42]. Gene annotations are available at Phytozome (https://phytozome.jgi.doe.gov). Predicted proteins were functionally annotated using SignalP [43] for signal sequences, TMHMM [44] for transmembrane domains, InterProScan [45] for protein domains, protein alignments to the National Center for Biotechnology Information (NCBI) non-redundant (NR) protein set, SwissProt [46], the Kyoto Encyclopedia of Genes and Genomes database (KEGG) to retrieve EC numbers [47], and the eukaryotic clusters of orthologs (KOG) to retrieve function descriptions [48]. Hits from InterPro and SwissProt were used to map Gene Ontology terms [49]. Complementary, the genome along with functional annotations is available at PhycoCosm (https://phycocosm.jgi.doe.gov).

2.4 Comparative genomics.

A comparative genomics approach was applied with the aim to identify the number of unique, over- and under-represented genes within the genome of *D. salina*. For unique genes the analysis was performed in PhycoCosm. The expansion and contraction of gene families were estimated using the Computational Analysis of gene Family Evolution (CAFE) software package version 4.2.1, which provides a statistical foundation for evolutionary inferences [50,51]. The input of gene families for CAFE was based on the counts of PFAM gene families in the six genomes listed in Supplemental Table 2. Gene families with the largest variance (i.e., gene families with counts that differed by more than 60 across these genomes) were removed when estimating the birth-death parameter to ensure convergence. Expansions or contractions of gene families with a p-value <= 0.01 were considered significant, indicating that the observed gene family counts deviated from the expected model for gene gain and loss.

For comparison the closely related green algae *Chlamydomonas reinhardtii* [52], *Volvox carteri* [53], *Gonium pectorale* [54], *Chromochloris zofingiensis* [55], and *Monoraphidium neglectum* [56] were used.

2.5 Metabolic Network Reconstruction

For reconstruction of the core carbon metabolic network each reaction was subjected to manually curation using information on reactions and enzymes from the literature and multiple public databases. The KEGG database was used to create the scaffold for the network, which is presented in a style similar to the KEGG pathways. KEGG annotations from the *D. salina* Phytozome and PhycoCosm portals were then used for metabolic pathway mapping. In addition, we performed tBLASTn searches of the *D. salina* genome using proteins from *C. reinhardtii* and from the model plant *Arabidopsis thaliana* (https://www.arabidopsis.org/). Further, we checked on predictions for genes coding for enzymes belonging to larger protein super-families using BLASTp searches into the NR database.

3. Results and Discussion

3.1 Genome Expansion

As still little is known about eukaryotic adaptations regarding life in hypersaline environments, we sequenced and annotated the genome of *D. salina* [37] to identify the features that could be associated with adaptations to high salt. In brief, the 343.7 Mbp (including 34% repeat content) *D. salina* genome is two- to three-fold larger than those of *V. carteri* (131.16 Mbp with 17% repeat content) or *C. reinhardtii* (111.1 Mbp with 12% repeats). The genome of *D. salina* is inflated, because 53% of the genome is contained in introns (Suppl. File 1), whereas in *C. reinhardtii* and in *V. carteri* the introns only make up about 30% of the genomes (Suppl. File 1). With 16,697 loci coding for 18,801 protein-coding transcripts (including iso-forms) [37], *D. salina* has a gene content similar to *C. reinhardtii* [52] and *V. carteri* [53] with 16,403 and 13,076 genes, respectively. The genome of *D. salina* is reasonably complete because both 97.6% CEGMA [57] and 89% BUSCO [58] coverages are high and comparable to that of other related green algal genomes (https://phycocosm.jgi.doe.gov; Suppl. Figure 1), which enables accurate metabolic reconstruction.

A comparison of gene structures of *D. salina* with those found in its close relatives within the Chlorophycean algae showed that on average genes in *D. salina* are about twice as long (11,502 nt, Dusalv1.0) as those in *C. reinhardtii* (5,482 nt, Chlrev5.6) or *V. carteri* (6,264 nt, Vocalv2.1). As the average exon length is shorter in *D. salina* (227 nt) than in *C. reinhardtii* (366 nt) or *V. carteri* (387 nt), but the number of exons per gene in the algae is similar, the longer gene length is due to the average introns of genes in *D. salina* being much longer (1,434 nt, average intron length) as compared to *C. reinhardtii* (268 nt) or *V. carteri* (416 nt). Additional comparison of the 5' exon-intron and the 3' intron-exon splice boundaries did not reveal any specific features for *D. salina* (Suppl. Figure 2). We can only speculate that the intron length expansion has evolutionary implications, but any further analysis of the introns in *D. salina* would be beyond the scope of this manuscript.

In the context of the gene structures, we also compared the codon usage among *D. salina*, *C. reinhardtii*, and *V. carteri* (Suppl. File 1). There exist some differences in codon preferences for the amino acids proline, threonine, and alanine. The frequency of the base adenine in the third codon position is higher for these three amino acids and as well overall higher in *D. salina* than in the other two related algae (Suppl. File 1).

PhycoCosm annotation revealed 10,958 PFAM descriptions (58% of the genes) and 6,058 genes (38%) could be affiliated with GO terms. As KEGG annotations for metabolic pathways are conservative in PhycoCosm, for metabolic pathway analysis, we also checked the Phytozome annotations and we performed NCBI BLASTp searches. Our use of NCBI BLASTp searches was limited and mostly restricted to instances where genes were predicted to code for members of larger protein super-families.

3.2 Unique protein families

Due to *D. salina* being a halophilic organism, which can even grow in saturated brines, one might have expected the presence of genes coding for proteins that confer properties that represent adaptations to life in high salt. *D. salina* is known to have a unique duplicated form of the carbonic anhydrase, which is absent from other green algal species [25,27,28]. Further, specific transferrin proteins residing in the plasma membrane had been identified that allow iron acquisition [8,22,25]. Based on the hypothesis that the genome of *D. salina* contains genes coding for unique proteins, we mined the genome based on a comparison of PFAM domain annotations. We compared *D. salina* against its close relatives *C. reinhardtii* [52], *V. carteri* [53], *G. pectoral* [54], as well as *C. zofingiensis* [55], and *M. neglectum* [56]. Such comparative analysis feature is provided in the PhycoCosm portal.

Many PFAM domains appeared to be unique in *D. salina* as compared to other green algae (Suppl. Table 1). However, here we discuss only a few gene families that may be relevant for adaptations to life in high salt. The first example is the fourteen genes coding for proteins containing the PFAM domain PF00191. PF00191 refers to a group of proteins collectively called "annexins", which form a homologous superfamily. Although these proteins are well studied in animals and humans [59], not much is known about annexins in plants and algae except that they have diverse functions in plants [60,61]. Many different types of annexins exist, which bind calcium, thus performing calcium-dependent functions and making annexins responsive to calcium as a signal. Calcium signaling is a major component in plant cells stress responses [61] and has been postulated to play a role in stress response signaling pathways in *D. salina* [29,63], in particular in response to changes in extracellular salinities [64]. Our finding will allow future targeted research into the role of calcium signaling with regards to molecular stress responses in *D. salina*.

A second example of a unique gene family in *D. salina* is that coding for proteins containing the PFAM domain PF10342, with six members found in *D. salina* as compared to its relatives within the Chlorophyceae. These proteins are glycosylphosphatidylinositol-anchored in the plasma membrane and are essential for plants at the interface between the cell wall and outer surface of the plasma membrane [65], often participating in signal transduction [66]. Again, we identified a unique gene family, which encodes for proteins possibly important for signal perception and transduction at the plasma membrane level.

A third example is the gene family represented by two genes coding for proteins including the PFAM domain PF01793 found in glycolipid 2-alpha-mannosyltransferases [EC 2.4.1.131]. This enzyme transfers an alpha-D-mannosyl residue from GDP-mannose to lipid-linked oligosaccharides, which function as intermediates in glycoprotein biosynthesis. This may be a function that aids in synthesis of the elastic periplasmic coat that surrounds cells of *D. salina*.

In summary, we discovered a variety of unique genes in *D. salina* that are not present in other closely related Chlorophycean algae. Some of these genes code for proteins where the predicted function may be relevant for adaptations to abiotic stress such as high salt.

3.3 Protein family expansion

Based on assignments to PFAM domains, our CAFE analysis revealed a variety of genes that are over-represented in the genome of *D. salina* (Suppl. Table 2). Standing out in this analysis is the gene family coding for transferrins and transferrin-like proteins, because transferrins have previously been discovered in the plasma membrane of *D. salina* cells [25,67,68] and have been shown to participate in iron uptake [67]. Transferrin-like proteins that do not bind iron and may interact with bicarbonate have also been identified for *D. salina* [68]. Proteins of the transferrin superfamily are widespread in the animal kingdom [69], but have rarely been found in the plant kingdom [70]. Here, we verified not only the presence of genes, but also provide evidence that this gene family is expanded in *D. salina* as compared to other green algae and higher plant species. Our result regarding the transferrins confirms the validity of our PFAM analysis approach.

Similar to the unique mannosyltransferases mentioned in section 3.2 above, another expansion is the gene family coding for proteins containing the PF05637 domain representing the galactosyl transferase GMA12/MNN10 family (Suppl. Tab. 2). Such galactosyl transferases, which may function in glycosylation of extracellular proteins, possibly participate in synthesis of the pericellular matrix, which is essential to cells surviving in salinity conditions that require rapid volume changes (hyper- or hypoosmotic shock).

A third example for an over-represented gene family in *D. salina* as compared to other closely related green algae codes for proteins containing the PF02460 domain (patched family, Suppl. Tab. 2) present in the Sonic Hedgehog receptor family (InterPro entry IPR003392). This protein family is involved in sterol binding and sensing in animals and humans [71]. Although nothing is known about Sonic Hedgehog signaling in plants, the plasma membrane of *Dunaliella* cells reportedly contains sterols [72] and that these sterols appear to be essential for sensing osmotic changes in *Dunaliella* [73]. Consequently, discovery of this gene family presents an opportunity to further study the initial steps of osmosensing and down-stream signal transduction events in *D. salina*.

In summary, we identified several different gene families that are over-represented in the genome of *D. salina*. As the examples presented and discussed above demonstrate, several of these gene families code for proteins potentially involved in adaptations to life in high salinities, thus opening new avenues of research into stress biology including signal perception and transduction in *D. salina*.

3.4 Under-represented and potentially lost genes

Our PFAM analysis indicated that a variety of gene families were either under-represented in, or lost from, the genome of *D. salina* (Suppl. Tab. 2). One example for gene loss is the DOMON (named after dopamine β-monooxygenase N-terminal) domain PF03351, which is found in proteins that assist in the scavenging of extracellular heme or in sugar recognition. Loss of genes coding for such proteins may be explained by the absence of free heme in the hypersaline environments in which *D. salina* thrives.

In brief, some of the genes that are underrepresented or could not be identified in the genome of *D. salina* might represent adaptation to conditions of high salinities. Nevertheless, reduction in numbers or the absence of such genes is not necessarily proof for a specific role of those proteins. Our results clearly point towards some new research avenues in unraveling the mechanisms of salt adaptation.

3.5 Ion channels and ion transporters

In addition to the above presented results and discussion, Table 1 provides a brief comparative summary of genes predicted to code for ion channels and transporters in *D. salina* and related Chlorophyceae. Ion homeostasis has been studied in *Dunaliella* [8,22,74], but as ion relations within cells are complex and it was not the aim of this genomic study to provide specific cellular localization and function information about the different ion channels and transporters, we only present the list of genes. This screen for PFAMs is aimed to provide an overview of what types of channels/transporters are present in, or potentially absent, from *D. salina*, thus aiding future targeted work in this area. Overall, we identified several classes of voltage-gated, ligand-gated, and potentially mechanically-gated ion channels. Each of these classes contains potassium channels, which are important in plant osmotic stress responses [75].

Table 1: List of gene counts by PFAM annotation for a variety of ion channels/transporters identified for *D. salina* and compared to related Chlorophytes.

Annotation	Dunsal1	Chlre5.5	Gonpec1	Volcav2.1	Chrzof v5.2.3.1	Monneg1	Annotation Description
PF07885	4	9	3	7	3	5	Voltage-gated potassium channels (VGKCs)
PF03595	0	1	1	0	3	1	Voltage-dependent anion channel
PF08016	16	17	11	9	4	2	Polycystin cation channel, voltage-gated
PF00654	7	6	6	6	6	7	Voltage-gated chloride channel
PF01007	4	3	3	3	3	3	(Ligand-gated) Inward rectifier potassium channel transmembrane domain
PF00060	10	5	4	2	7	6	Ionotropic glutamate receptors (iGluRs) are ligand-gated ion channels
PF02931	0	4	4	1	1	0	Neurotransmitter-gated ion-channel ligand binding domain
PF02932	0	3	4	2	1	1	Neurotransmitter-gated ion-channel transmembrane region
PF01062	8	8	7	6	6	21	Calcium-activated Bestrophin, RFP-TM, chloride channel, plasmamembrane
PF02714	2	11	6	5	9	7	Calcium-dependent channel, 7TM region, putative phosphate
PF03493	0	2	1	1	2	2	Calcium-activated BK potassium channel alpha subunit
PF04547	0	1	0	0	2	1	Calcium-activated chloride channel
PF12166	1	1	1	0	1	2	Piezo non-specific cation channel, R-Ras-binding domain
PF01741	0	0	0	0	0	3	Large-conductance mechanosensitive channel, MscL
PF00924	5	8	6	6	5	4	Mechanosensitive ion channel, inward rectifier

							potassium channel transmembrane domain
PF01036	6	8	7	2	1	1	Bacteriorhodopsin-like protein
PF05197	1	1	1	0	0	0	TRIC channel
PF06011	2	7	6	2	2	0	Transient receptor potential (TRP) ion channel
PF01699	9	6	7	7	8	8	Sodium/calcium exchanger protein
PF02705	0	3	2	1	5	2	K^+ potassium transporter

3.6 Acclimation to changes in extracellular salinity in D. salina

Sensing and signal transduction in response to changes in extracellular salinity in cells of *D. salina* is hypothesized to be similar to what is known from fungi [8,36], where MAPK Signaling through the HOG pathway is essential [76]. However, to date, nothing is known about the details of this signaling cascade in *Dunaliella*. Overall, we discovered seven genes that code for predicted MAPKs, which represent four different classes. Nevertheless, a homology tBLASTn search using the yeast HOG1 protein into the *D. salina* genome did not result in unambiguous identification of a specific potential HOG1 homologue. Additionally, BLASTp searches in the NCBI NR database using translated sequences from the predicted genes coding for MAPK proteins only returned general hits showing the protein kinase family domains typically found in MAPKs.

To potentially pinpoint genes coding for the components of the Sln1 and Sho1 branches of the HOG pathway [76], we discovered candidate genes coding for four MAPK Kinases (MAPKK) as well as for 10 MAPK kinase kinases (MAPKKK). Again, we could not assign specific function for any of the coded proteins based on homology and homology searches as the NCBI NR database only returned hits to general domains signifying MAPKs in general. Nevertheless, our result with the ten genes coding for predicted MAPKKK proteins is low when compared to the reported 33 MAPKKK for another *Dunaliella* species [77]. This may mean that we did not discover the full set of kinases involved in MAPK signaling. Alternatively, reduced numbers for this specific set of protein kinases may be part of the overall reduction in the cellular set of protein kinases in *D. salina*, which is 30-50% reduced, as revealed by a comparison of the most prevalent PFAM domains of protein kinases in other green algae. For example, *C. reinhardtii* has 538 genes and *V. carteri* has 348 genes coding for proteins with protein kinase domains (IPR000719). In contrast, in the genome of *D. salina* only 246 genes were discovered. This low number might be due to imprecise gene identification, but our CEGMA and BUSCO analyses demonstrated good gene coverage. Clearly more experimental research and genomic mining for further potentially missed genes coding for genes such as the MAP Kinase families is required to unravel the functions of the products of the discovered genes in *D. salina*.

3.7 Reconstruction of a core carbon metabolic network

As glycerol metabolism is part of the core carbon metabolism of cells, to approach the question on anabolism and catabolism of glycerol, a metabolic network was reconstructed for the core carbon metabolism in *D. salina* based on genomic analysis and is presented in Fig. 1. The metabolic network contains glycolysis/gluconeogenesis, the Calvin-Benson-Bassham cycle (CBB-Cycle), starch and sucrose metabolism, photorespiration, the Krebs cycle, and some interfacing connections to isoprenoid metabolism, lipid and amino acid metabolism. The different genes coding for enzymes included in this map are listed in Suppl. File 2. Part of this metabolic network had recently been published [78], but the version we present here is focused on core carbon metabolism and contains additional reactions that had not been considered previously. Additional reactions include sucrose as well as trehalose metabolism and reactions of dicarboxylic acid metabolism.

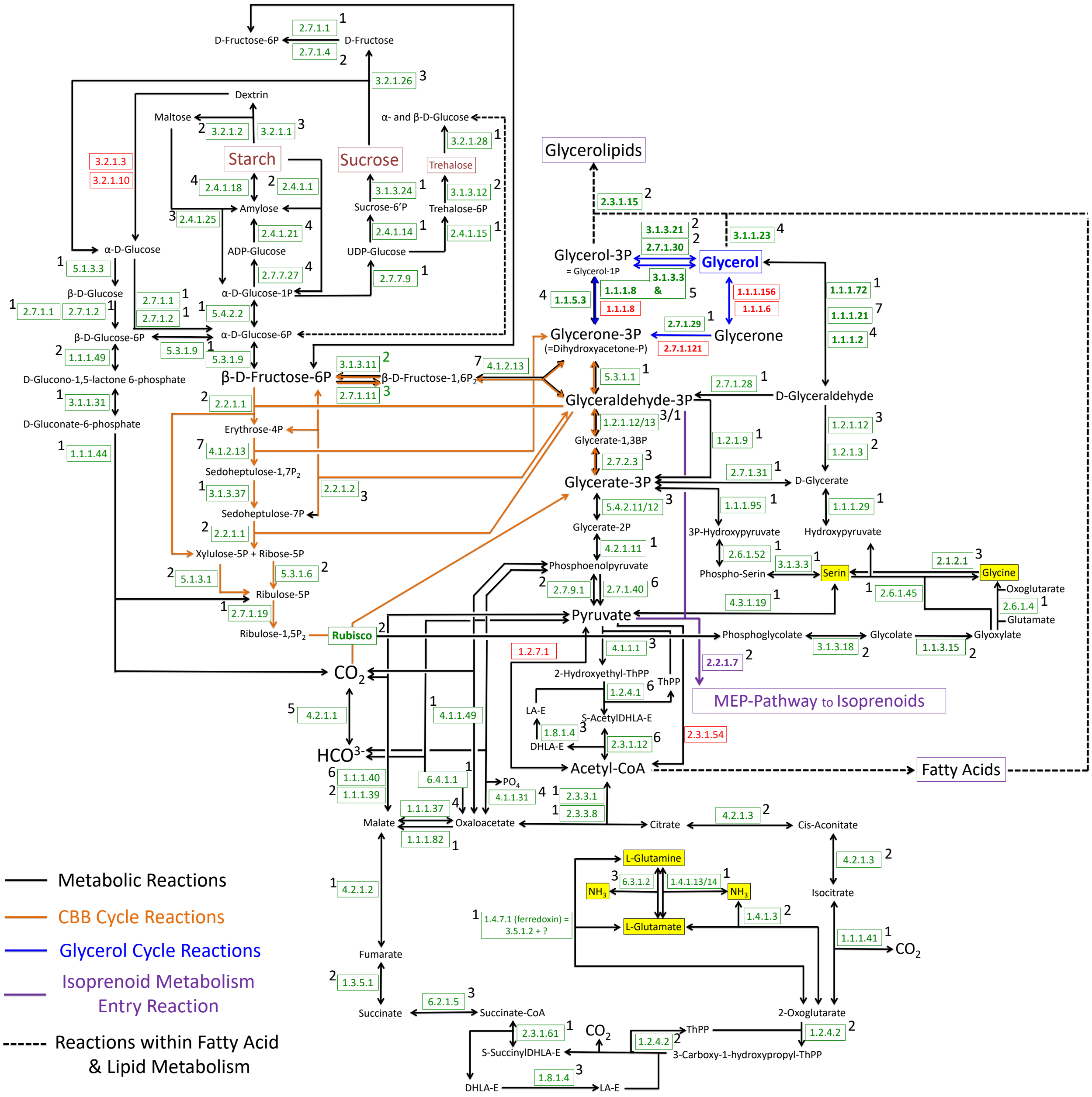

Figure 1: Reconstruction of the core carbon metabolism including glycerol metabolism (purple lines) based on functional annotation of the *D. salina* genome. The metabolic map shown does not include information on the cellular localization of the enzymes. Identified genes for enzymes are shown as EC

numbers in green boxes. The box around EC1.1.1.8 and EC 3.1.3.21 indicates gene-fusion products containing the glycerol-3-phosphate dehydrogenase and the glycerol-3-phosphate phosphatase function. Next to the boxes with EC numbers, an additional number indicates the number of genes coding for that enzyme. Calvin-Benson-Bassham Cycle reactions are shown in light brown and the connection to isoprenoid metabolism is shown in purple lines. Connections to lipid metabolism are shown in dotted lines.

Various species of the genus *Dunaliella* were used over the past decades to elucidate the response of *Dunaliella* cells to hypo-osmotic (decrease in salinity) and hyper-osmotic (increase in salinity) changes in the medium [22,36,79,80]. Glycerol was found to be the main osmolyte [15,81,82]. Maintaining osmotic balance in response to changes in the external salt concentration depends on the intracellular synthesis or degradation of glycerol [15,83]. Early on, a glycerol cycle was proposed to exist in *Dunaliella* [17,23,84].

Central to the known glycerol cycle is the generation of glycerol-3-phosphate from glycerone-3-phosphate (aka, Dihydroxyacetonephosphate, DHAP) either provided through glycolytic reactions [85] stemming from fructose-1,6-bisphosphate or through the pentose phosphate pathway with the intermediate glycerate-3-phosphate [18,21]. This reversible reaction is catalyzed by the glyceraldehyde-3-phosphate dehydrogenase (GPDH), which exists as two classes encoded by different genes [86]:

1) The nicotinamide-adenine dinucleotide (NAD+)-dependent enzyme [EC 1.1.1.8] is encoded by genes that only result in a single function GPDH, or is coded in fusion genes that produce bi-functional enzymes including the [EC 1.1.1.8] GPDH function together with the HAD-superfamily hydrolase domain performing the [EC 3.1.3.3] phosphatase function. This previously predicted bi-functional enzyme [87,88] had indeed recently been demonstrated to convert glycerone-3-phosphate directly into glycerol [89]. We identified two genes (Dusal.1194s00002.1, Dusal.0486s00010.1) that code for this bi-functional enzyme (Suppl. File 2). At a third locus, the gene Dusal.0246s00004.1 coding for an enzyme containing the [EC 3.1.3.3] domain together with the N-terminal domain of [EC 1.1.1.8] is followed by gene Dusal.0246s00004.1, which only codes for the C-terminal domain [EC 1.1.1.8]. As proper curation of the two genes is impossible due to a gap in sequence between both genes, for metabolic network construction purposes, we decided to treat both gene models as coding for one bi-functional enzyme. A similar situation was encountered for a fourth locus where the gene Dusal.0832s00001.1 annotated as the [EC 3.1.3.3] is directly followed by the gene Dusal.0832s00002.1 annotated as the [EC 1.1.1.8] GPDH. As both genes are separated by a gap in the genome sequence, we also decided to treat both genes as coding for one bi-functional enzyme. In addition, a tBLASTn search using the protein predicted from gene Dusal.1194s00002.1 revealed the presence of a fifth gene, which was not annotated, but is predicted to code for a bi-functional enzyme. In summary, we identified five putative genes coding for bi-functional GPDH, but we did not find any gene coding for a single function GPDH. The bi-functional GDPH isoforms had also been identified in *D. viridis* [90] and in the freshwater alga *C. reinhardtii* [91,92]. Moreover, the activity of the bi-functional GPDH enzymes from *C. reinhardtii* had been established in yeast for isoforms 2 and 3 [93]. Additionally, isoform 2 function had been demonstrated *in vitro* and the gene was upregulated in response to salt stress in *C. reinhardtii* [92]. Consequently, we conclude that these bi-

functional enzymes present in *D. salina* are not specific adaptations to life in hypersaline conditions.

Our finding of five genes coding for GDPH is somewhat similar to that recently reported by Wu et al. [86]. However, in contrast to our results, they identified one gene coding for a single-function GPDH and four genes coding for bi-functional GPDH. Although Wu et al. [86] reported their results under the species name *D. salina*, the strain they used, FACH435, belongs to a different *Dunaliella* species. This may explain the differences in our results. In any case, lack of a gene coding for a single-function GPDH [EC 1.1.1.8] may not be crucial for glycerol-3-phosphate biosynthesis, because it had been shown for the bi-functional enzyme isoform GPD2 in *C. reinhardtii* that its two enzymatic functions may be modulated by posttranslational modifications [92] to suit the cell's metabolic demands. For our metabolic network reconstruction, we did not consider cellular compartments, but there is evidence that the bi-functional GPDH isoforms are localized either to the cytosol or the chloroplast [86,92].

2) The inner mitochondrial membrane bound, ubiquinone-dependent enzyme [EC 1.1.5.3] genes code only for single-function GPDHs. We identified four genes coding for this enzyme, which is higher than the two genes reported by Wu et al. [86].

Independently from the bi-functional GPDH, dephosphorylation of glycerol-3-phosphate by the enzyme glycerol-3-phosphate phosphatase [EC 3.1.3.21] will produce glycerol. This phosphatase function had been reported for cell extracts of *D. salina* [94] and we identified one gene coding for this enzyme (Dusal.0430s00018.1). For a second gene (Dusal.0132s00008.1), the functional annotation is ambiguous (Suppl. File 2).

Regarding the generation of glycerol from glycerol-3-phosphate originally only the above-mentioned glycerol-3-phosphate phosphatase [EC 3.1.3.21] was believed to be responsible, whereas the glycerol kinase [EC 2.7.1.30] was thought to perform the opposite reaction involved in removal of glycerol under hypo-osmotic conditions [83]. We identified one gene coding for a glycerol kinase (Suppl. File 2). Alternatively, removal of glycerol was proposed to occur by conversion to glycerone via the dihydroxyacetone reductase [EC 1.1.1.156] or the glycerol dehydrogenase [EC 1.1.1.6] followed by phosphorylation to glycerone-3-phosphate by the glycerone kinase [EC 2.7.1.29], thus connecting glycerol back to glycolysis. Overall, the annotation based on homology of proteins with assignments to specific dehydrogenases such as the dihydroxyacetone reductase [EC 1.1.1.156] and the glycerol dehydrogenase [EC 1.1.1.6] proved to be very difficult due to the existence of a relatively large, non-specific enzyme family belonging to the oxidoreductases. Therefore, we could not specifically identify any gene predicted to code for the dihydroxyacetone reductase [EC 1.1.1.156] or the glycerol dehydrogenase [EC 1.1.1.6]. Both EC numbers are shown in red color in Fig. 1. Nevertheless, we were able to identify a number of potential alcohol dehydrogenases [EC 1.1.1.-] that could reversibly connect glycerol to glyceraldehyde and through aldehyde dehydrogenases [EC 1.2.1.-] to glycerate (Fig. 1), which is essential in the photorespiratory pathway. In addition, we identified one gene coding for a triokinase (EC 2.7.1.28/2.7.1.29), which could phosphorylate glycerone or glyceraldehyde. Moreover, a glycerate-3-phosphate kinase [EC 2.7.1.31] could phosphorylate glycerate. Taken these findings together, we propose that glycerol could be connected to glycolysis/gluconeogenesis in more ways than previously proposed, likely providing a wider glycerol cycle in *D. salina*.

In brief, discovery of a variety of genes coding for enzymes of the superfamily of alcohol dehydrogenases [EC 1.1.1.-] and aldehyde dehydrogenases [EC 1.2.1-], with difficulties in homology-based assignments of functions, demonstrates the necessity of further biochemical research to validate the reactions of the current predicted functions. Only then will we be able to verify the novel proposed second glycerol cycle that includes glyceraldehyde and glycerate.

While reconstructing the network for core carbon metabolism, we noticed that several crucial enzymes were only represented with one copy in the genome. An example for an enzyme essential for the lower glycolysis pathway with only one gene copy is the enolase [EC 4.2.1.1]. The enolase might play a role in salt stress responses in *D. salina*, possibly by restricting carbon flow through the lower glycolysis pathway and consequently forcing carbon flux into reactions towards glycerol [95,96]. In this context, we had previously proposed that the cellular location of enolase is crucial for cellular carbon partitioning in green algae [97]. In summary, the enolase is an enzyme that was shown to impact carbon flux in cells and its localization as well as its functions should be investigated in more detail.

A second enzyme with only one gene is triose phosphate isomerase (TPI) [EC 5.3.1.1]. This finding is in contrast to many other green algae and higher plants, which have two genes coding for one cytoplasmic and one plastidic isoform. The TPI is at the cross-roads of triose phosphate metabolism participating in glycolysis/gluconeogenesis, the CBB Cycle, and in providing the precursors for terpenoid backbone metabolism. As TPI is essential for proper carbon allocation in the CBB cycle, it has to be localized in the chloroplast in *D. salina*. The TPI had also been reported to be upregulated at the protein level in response to salt stress [84]. In conjunction with the recent discovery that TPI in animals is regulated through phosphoenolpyruvate by competitive inhibition [98], TPI is a prime target for future research regarding carbon partitioning in *D. salina* and green algae in general.

Another part of the metabolic network revealed the presence of genes coding for enzymes involved in sucrose and trehalose metabolism. For example, we identified a gene for the UTP-glucose-1-phosphate uridylyltransferase [EC 2.7.7.9]. Further genes coding for all enzymes involved in synthesis of sucrose and trehalose were found (see Fig. 1 and Suppl. File 2). Currently, it is unknown what the exact function(s) of sucrose in green algae is (are), but it was observed in *D. tertiolecta* that cells accumulated considerable amounts of sucrose with increasing temperature [99,100]. As the interactions of glycerol and sucrose metabolism have not been studied in *D. salina*, our results provide the groundwork for future investigations.

Notably, our reconstruction of the network for core carbon metabolism revealed the lack of a gene coding for the pyruvate ferredoxin/flavodoxin oxidoreductase [EC 1.2.7.1]. In addition, we did not find any gene(s) coding for the pyruvate formate lyase [EC 2.3.1.54]. Both enzymes participate in pyruvate metabolism and provide alternative reactions for cells to "bypass" pyruvate dehydrogenase (PDH) [EC 1.2.4.1]. As the PDH produces acetyl-CoA, its function is essential for provision of the precursor for fatty acid biosynthesis and thus lipid biosynthesis. Lack of the availability of alternatives for pyruvate to acetyl-CoA conversion potentially makes the PDH the bottleneck for carbon flux into fatty acids and lipids in *D. salina*. We hypothesize that accumulation of β-carotene under stress conditions may be driven by reduced PDH activity, thus driving carbon through pyruvate (and glyceraldehyde-3-phosphate) into the isoprenoid biosynthesis pathway (Fig. 1) [101].

In summary, reconstruction of the core carbon metabolic network allowed us to pinpoint a variety of core carbon network components that represent potential bottlenecks (enolase and PDH), potential regulatory elements such as the TPI, or currently understudied osmolytes such as sucrose and trehalose.

3.8 Isoprenoid Metabolism

D. salina is unique within the green algae by accumulating β-carotene in globuli within the chloroplast in response to abiotic stress such as nutrient deficiency or increased irradiance. Due to this capability, *D. salina* is not only a model organism for carotenoid metabolism [32,36], but also for commercial production of β-carotene [35]. Carotenoids are made up of eight isoprene units. As shown in Fig. 1, the entry reaction into isoprenoid metabolism is the condensation of pyruvate with glyceraldehyde-3-phosphate to form, through a series of steps, isopentenyl diphosphate and its isomer dimethylallyl phosphate (Fig. 2). Below, we describe in more detail first the terpenoid backbone biosynthesis and then carotenoid metabolism.

3.8.1 Terpenoid Backbone Biosynthesis

To dissect isoprenoid metabolism, we followed the KEGG descriptions to map reactions constituting the terpenoid backbone biosynthesis. The metabolic map presented in Fig. 2 begins with the entry reaction condensing glyceraldehyde-3P and pyruvate to 1-deoxy-D-xylulose 5-phosphate. In green algae, this entry-step reaction and the following seven reactions, generally referred to as the methyl-erythritol-phosphate (MEP) pathway, localize to the chloroplast [102].

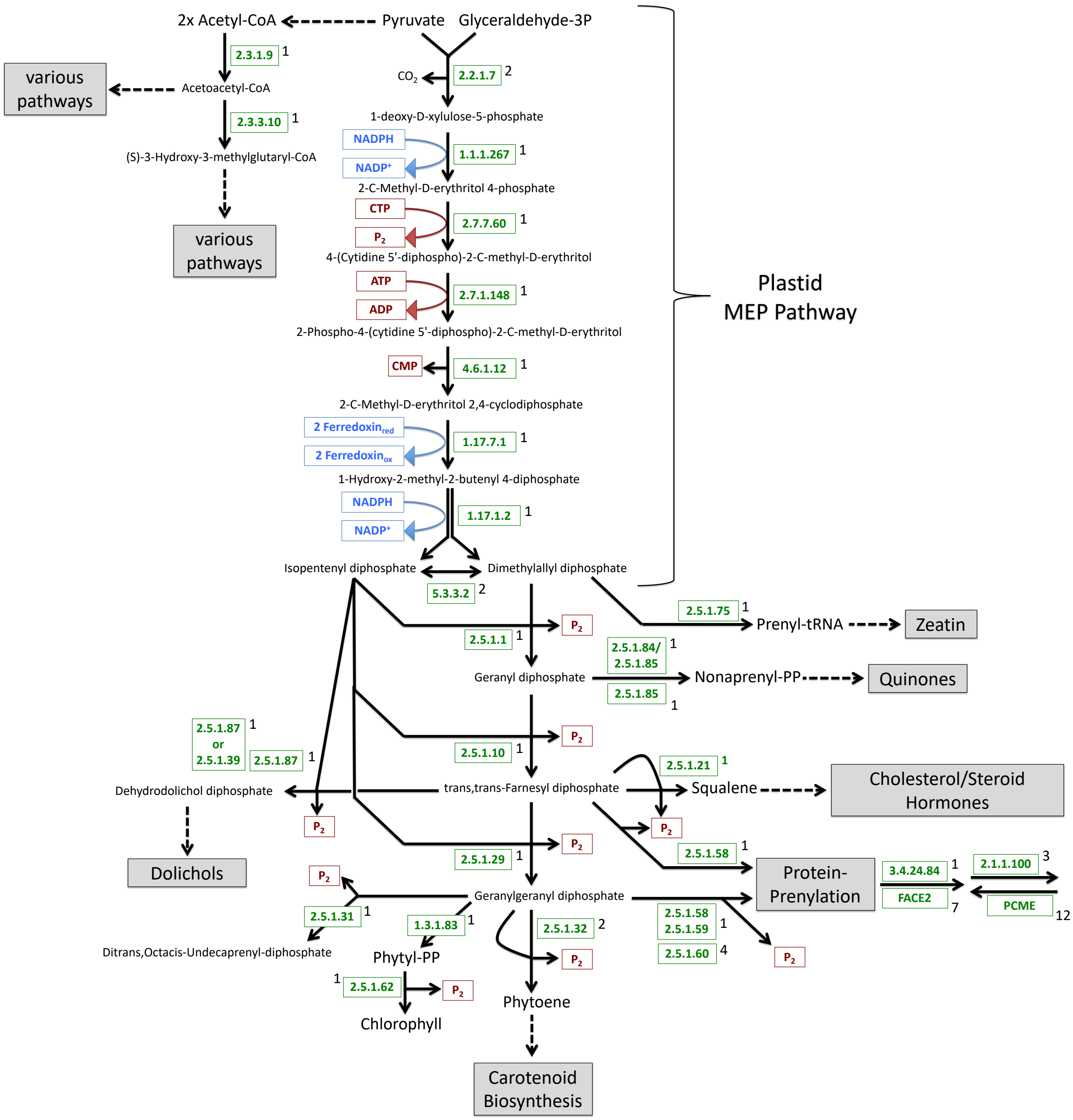

Figure 2: Metabolic map for terpenoid backbone biosynthesis according to the genes that could be functionally annotated in the genome of *D. salina*. This map includes the plastidic MEP pathway and the different prenyl transferases that provide entry-step reactions into a variety of other pathways that are requiring isoprene subunits.

Our analysis of the genome regarding genes coding for enzymes participating in terpenoid backbone biosynthesis is summarized in Fig. 2. The numbers provided next to the green boxes with EC IDs indicate how many genes were identified and details on the genes are provided in Suppl. File 2. To put our results into context, we compared the numbers of genes identified in *D. salina* to its relatives *C. reinhardtii* and *Haematococcus pluvialis* [103] (Tab. 2). When exposed to abiotic stress conditions, zoospores of *D. salina* accumulate β-carotene in globuli in the plastid to a level of up to 8% of the cellular dry weight [30]. In contrast, vegetative cells of *C. reinhardtii* do not contain secondary carotenoids, but rather maturing zygotes accumulate secondary carotenoids in oil bodies located in the cytosol. Similar to *D. salina*, vegetative cells of *H. pluvialis* respond to abiotic stress by accumulation of secondary carotenoids of up to 4% of the cellular dry weight [104,105], but only in oil bodies located in the cytosol comparable to the situation in *C. reinhardtii* zygotes.

Table 2: Comparison of gene numbers identified in the genomes of the three related green algae *D. salina* (*D.s.*), *C. reinhardtii*, (*C.r.*) and *H. pluvialis* (*H.p.*). For *C. reinhardtii*, data were extracted from Phytozome. For *H. pluvialis*, data were extracted from Luo et al. [103]. The asterisk denotes that *H. pluvialis* is a diploid organism. Empty cells for *H. pluvialis* indicate that no information is available.

EC number	**Enzyme**	***D. s.***	***C. r.***	***H. p.****
2.2.1.7	1-deoxy-D-xylulose-5-phosphate synthase (dxs)	2	1	3
1.1.1.267	1-deoxy-D-xylulose-5-phosphate reductoisomerase (dxr)	1	1	3
2.7.7.60	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (ispD)	1	1	2
2.7.1.148	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase (ispE)	1	1	2
4.6.1.12	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (ispF)	1	1	2
1.17.7.1 (1.17.7.3)	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase (gcpE, ispG)	1	1	2
1.17.1.4	4-hydroxy-3-methylbut-2-enyl diphosphate reductase (lytB)	1	1	3
5.3.3.2	Isopentenyl-diphosphate Delta-isomerase (idi)	2	2	3
2.5.1.-	Prenyl transferase	1	none	
2.5.1.1.	Geranyl diphosphate synthase, GPS	1	none	2
2.5.1.1/2.5.1.10	Farnesyl diphosphate synthase (2E,6E?); FDPS	1	1	3
2.5.1.1/2.5.1.10/ 2.5.1.29	Geranylgeranyl pyrophosphate synthase	1	1	3
2.5.1.32	Phytoene synthase	2	1	1
2.5.1.32/2.5.1.21	Squalene/phytoene synthase	1	1	
1.3.5.5	15-cis-phytoene desaturase / plant-type phytoene desaturase; (PDS, crtP)	3	3	5
1.3.99.30	Bacterial-type phytoene desaturase, CrtI	2	0	
5.2.1.12	Zeta-carotene isomerase / 15-cis-zeta-carotene isomerase; Z-Iso	1	1	
1.3.5.6	Zeta-carotene desaturase / 9,9'-di-cis-zeta-carotene desaturase	1	1	3
5.2.1.13	Carotenoid isomerase, chloroplastic; CrtISO	3	3	
5.5.1.19	Lycopene beta-cyclase	1	3	3
5.5.1.18	Lycopene epsilon-cyclase (lcyE, crtL2)	2	1	
5.2.1.14	Beta-carotene isomerase;	3	2	
1.14.-.-	Carotene beta-ring hydroxylase?	1	none	3
1.14.-.-	Carotene beta-ring hydroxylase; Lut5	2	3	
1.14.14.158	Carotenoid epsilon hydroxylase; CYP97C1; LUT1;	1	1	
1.14.13.24	Beta-carotene hydroxylase; CrtZ	1	1	
1.14.14.1	Cytochrome P450 CYP4/CYP19/CYP26 subfamilies	1	2	

1.23.5.1	Violaxanthin de-epoxidase VDE, NPQ1	1	1	
1.14.13.90	Zeaxanthin epoxidase, ZEP	2	2	
1.14.99.63/1.14.99.64	Beta-carotene/zeaxanthin 4-ketolase; crtW, bkt	1	1	6

Regarding the MEP pathway, *D. salina* is similar to *C. reinhardtii*, with the only difference being two genes coding for 1-deoxy-D-xylulose-5-phosphate synthase (DXS) in *D. salina*. Considering that *H. pluvialis* is diploid, with three genes coding for the DXS, it is more similar to *D. salina*. This additional gene for the DXS might have been expected for *D. salina* and *H. pluvialis*, because both species induce the accumulation of carotenoids. In higher plants, DXS controls flux into the MEP pathway [106]. Often, multiple copies of *dxs* genes coding for different classes of DXS are found when differential regulation of isoprenoid biosynthesis occurs [106], either in different tissues (fruit ripening [107]) or in the context of stress responses (pine resins [108]). Based on this comparison to higher plant metabolism, we hypothesize that in *D. salina,* the two *dxs* genes are differentially regulated with one isoform for green cells and the second gene up-regulated specifically in response to abiotic stress.

One of the two *dxs* genes (Dusal.0031s00031.1) is clustered directly next to a gene coding for a phytoene synthase (PSY) (Dusal.0031s00032.1), but we cannot evaluate this finding at this time. Clustering of genes involved in isoprenoid biosynthesis had been discovered and such gene clustering may play a role in regulation of gene expression in higher plants [109]. To the best of our knowledge, this is the first report of a gene cluster for isoprenoid biosynthesis in a green alga. As both the DXS and the PSY are known entry-step enzymes for isoprenoid and carotenoid biosynthesis, co-regulation of both genes is likely, but its significance in regulation of secondary β-carotene accumulation in *D. salina* remains to be explored.

Taking the summary of prenyltransferase [102] for guidance, we identified at least one gene coding for all the different transferases that are deemed to be essential for green alga (Tab. 2 and Suppl. File 2). However, we also identified one gene predicted to code for a geranyl diphosphate synthase (GPS) [EC 2.5.1.1], which was not found in the genome of *C. reinhardtii*, but is present in the genome of *H. pluvialis*. Potentially, this additional predicted GPS may be involved in accumulation of secondary carotenoids by providing the required precursors.

In summary, genome-based evaluation of genes coding for enzymes participating in the terpenoid backbone metabolism revealed the DXS, PSY, and the GPS as targets for further investigations regarding regulation of stress-induced carotenoid accumulation.

3.8.2 Carotenoid Metabolism

Carotenoid biosynthesis begins with the PSY condensing two molecules of geranylgeranyl diphosphate. Fig. 3 shows the carotenoid biosynthesis pathway with genes identified mapped onto it. Details for the genes from *D. salina* are summarized in Suppl. File 2 and Table 2 compares the number of genes coding for enzymes of the core carotenoid biosynthesis pathway to *C. reinhardtii* and *H. pluvialis*. In general, the presence of genes coding for essential enzymes is conserved among these three related species of the Chlorophyceae. Nevertheless, there are some differences that may represent specific adaptations to stress in *D. salina*.

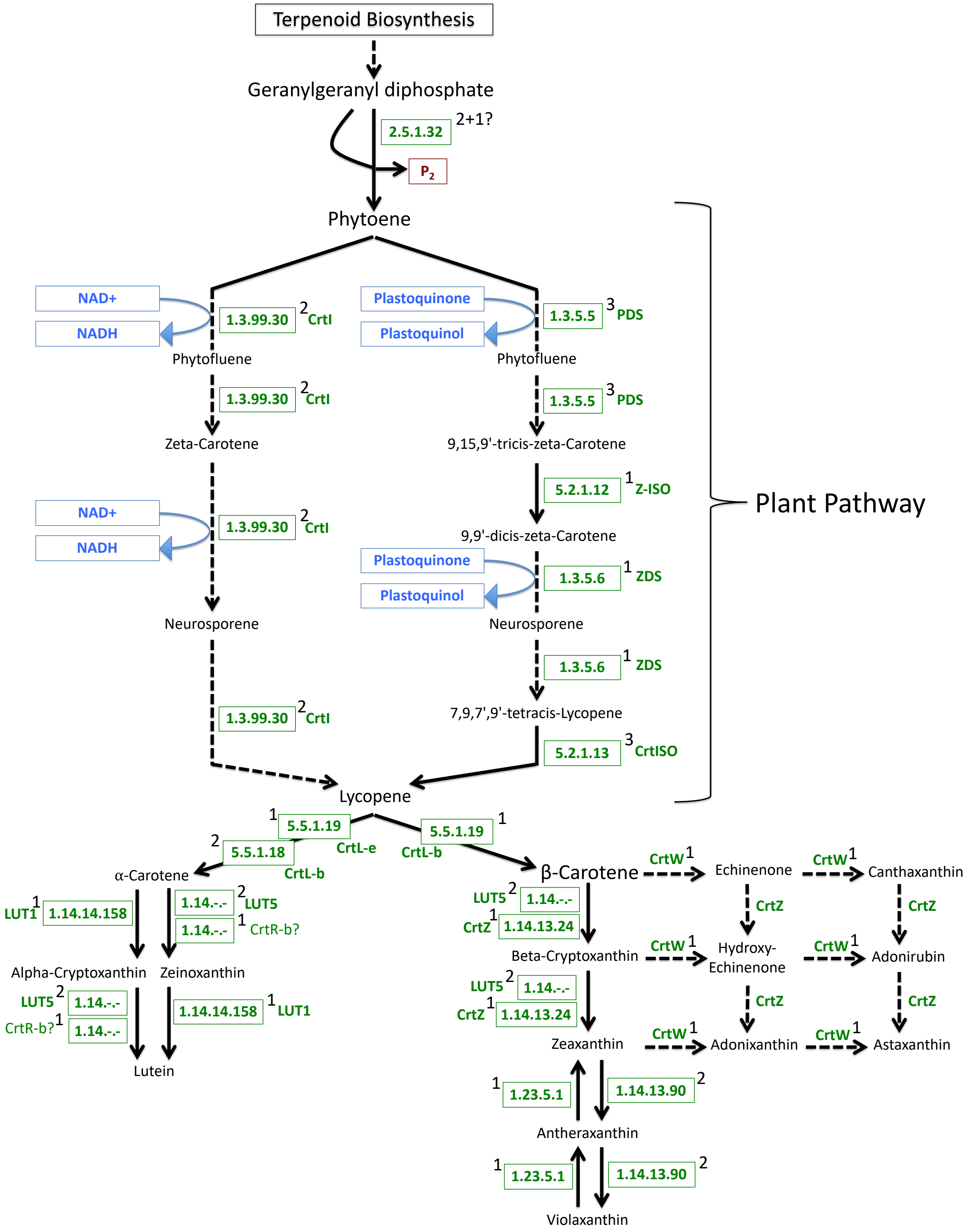

Figure 3: Genome-based carotenoid biosynthesis pathway identified for *D. salina*. Note the bacterial-type and the plant-type pathway from phytoene to lycopene. The three genes coding for carotene isomerases are not included in the presented pathway, but they are discussed in the text.

Regarding the entry-step into carotenoid biosynthesis in *D. salina*, we had previously discovered that two different genes code for two PSY isoforms [110]. This finding was confirmed and the existence of two *psy* genes is unique in *D. salina*, as compared to other green algae. Even for the carotenogenic *H. pluvialis,* only one gene had been reported, which is upregulated at the transcriptional level in response to abiotic stress [103,105]. Notably, in plants, PSY is one of the main control points in the pathway [111] and many higher plants that have carotenogenic tissues contain multiple genes for the PSY, which are differentially regulated during development, for example, in fruit ripening [111,112]. We hypothesize that differential expression of the *psy* genes are partially responsible for accumulation of β-carotene under abiotic stress conditions. Importantly, genomic information on the two *psy* genes should now allow testing and verification of this hypothesis.

In addition to the two *psy* genes, as indicated by the (+1) next to the box with [EC 5.2.1.32] in Fig. 3, we report on a third gene that is annotated as coding for a protein belonging to the squalene/phytoene synthase (SQS/PSY) superfamily (Suppl. File 2). Based on protein sequence homology, genes coding for this kind of SQS/PSY are also present in genomes of other green algae. Nevertheless, to date, no functional studies were reported on this protein leaving room for new discoveries.

In plants, other carotenoids are made from phytoene by sequential actions of dehydrogenases and isomerases resulting in synthesis of lycopene. In contrast to plants, some bacteria perform all these reactions by the use of one enzyme, phytoene desaturase (CRTI). As shown in Fig. 3, based on homology-based functional annotations, we discovered that *D. salina* has two genes coding for putative CRTI enzymes. Consequently, *D. salina* may have both the bacterial and plant pathways. This is a novel finding and raises new questions about carotenoid biosynthesis in this alga. For example, are the putative CRTI enzymes converting phytoene into lycopene? If true, then where in the cell would these enzymes be located and would they be participating in the stress-induced carotenogenesis process? If false, then what function would the putative CRTI enzymes play in carotenoid metabolism? These putative CRTI enzymes may be exploited in metabolic engineering not just in *D. salina*, but also for other green algae.

One topic not addressed in the carotenoid pathway map shown in Fig. 3 is that in *D. salina* many different forms of cis-β-carotene exist [113], although only trans-lycopene is made by the bacterial and plant pathways. Recently three genes were reported for 9-cis/all-trans β-carotene isomerases in *D. bardawil* [114]. Here we confirm the existence of these three genes also for *D. salina*.

A topic of interest for metabolic engineering is the production of astaxanthin in *Dunaliella* [115]. As shown in Fig. 3, astaxanthin is made by concerted action of hydroxylases (LUT5 or CRTZ) and of the β-carotene ketolase (BKT or CRTW). *H. pluvialis* is cultivated commercially for astaxanthin production, but *D. salina* is only known to accumulate β-carotene. Here we report for the first time that the genome of *D. salina* contains a gene coding for BKT. Nevertheless, we did not identify any assembled transcripts from RNA-Seq data coding for the BKT, although we had cultivated *D. salina* under many different stress conditions. In contrast, the expression of *bkt* genes in *H. pluvialis* is immediately induced when cells are exposed to stress [105]. The lack of induction of the *bkt* gene by abiotic stress could be one reason why vegetative cells of *D. salina* do not accumulate astaxanthin. This lack of *bkt* expression was recently also described for *C. reinhardtii* [116].

In summary, we report that carotenoid biosynthesis of *D. salina* has some aspects that are different from the related *C. reinhardtii* and *H. pluvialis*. In particular, discovery of a gene cluster containing one *dxs* and one *psy* gene as well as potential presence of genes coding for the bacterial type CRTI open new avenues of dissection of molecular regulation of carotenogenesis and concomitantly are of great interest for future metabolic engineering.

3.9 Plastid-localized Carotenoids Globules.

A specific feature of *D. salina* and *D. bardawil* is that stress-induced synthesis of secondary β-carotene is localized to lipid globuli within the chloroplast [30]. All other green algae accumulate secondary carotenoids in cytosolic oil bodies [117]. These β-carotene globuli from *D. salina* are similar to plant plastoglobuli [118,119] and are hypothesized to have originated from eyespot globuli [120]. Plant plastoglobuli are surrounded by a coat that is made up of special plastid-lipid-associated protein-fibrillins (PAP-fibrillins) [121]. We observed that the genome of *D. salina* contains 13 genes predicted to code for PAP-fibrillins (Suppl. File 2), which is consistent with a report that β-carotene globules from *D. bardawil* contain PAP-fibrillins [122].

4. Conclusions

Our analysis of the genome of *D. salina* has identified a variety of gene families that could provide further insights into osmoadaptations in this halophilic green alga. We found unique gene families coding for the Ser-Thr-rich glycosyl-phosphatidyl-inositol-anchored membrane proteins, which are likely plasma membrane bound. Further, mannosyltransferases (unique genes) and galactosyltransferases (expanded gene family) are implicated in synthesis of glycoproteins that are possibly part of the elastic pericellular matrix. Additionally, the transmembrane proteins of the Patched family are receptors transducing Hedgehog signals. Taken together, these findings suggest that modifications at the plasma membrane level play major roles in adaptations to life in hypersaline environments. This finding may not be unexpected, because the plasma membrane is the first contact point with the environment. Nevertheless, our results represent discoveries that will enable and advance innovative studies regarding dissection of the molecular mechanisms underlying osmoadaptations in this alga.

Importantly, several of the expanded and under-represented as well as potentially lost gene families we identified are predicted to code for proteins that may be involved in the molecular mechanisms of sensing and signal transduction in response to abiotic stress conditions. An example is the gene family of the annexins, which suggests an important role for calcium in signaling and signal transduction in stress response in *D. salina*. Moreover, the patched family proteins are known to play a role in regulation of development in animals. Its functions are unknown for *D. salina*, opening novel avenues for research into these kinds of proteins in algae.

The core carbon metabolic network together with the mapped terpenoid/carotenoid metabolism provide roadmaps for investigations of regulation of carbon partitioning into glycerol, carbohydrates, lipids and carotenoids. The enolase, TPI, DXS, PSY, and CRTI are potential novel targets for metabolic engineering in directing carbon flux for example towards carotenoids.

The annotated genome of *D. salina* will enable in-depth studies regarding ion homeostasis in this halophilic eukaryote. Our findings regarding ion channels and transporters may stimulate investigations into the molecular mechanisms of salinity tolerance.

In conclusion, our results provide targets not only for future research into adaptations to life in high salt in this model algal, but also for genes that may be applied to engineering of salt tolerant crop species [123].

CRediT authorship contribution statement

Juergen Polle: Conceptualization, Resources, Supervision, Formal Analysis, Visualization, Writing – Original Draft.
Sara Calhoun: Formal analysis, Visualization, Writing – Review & Editing.
Zaid McKie-Krisberg: Formal analysis, Visualization, Writing – Review & Editing.
Simon Prochnik: Investigation
Peter Neofotis: Formal analysis
Won C. Yim: Investigation
Leyla T. Hathwaik: Investigation
Jerry Jenkins: Investigation
Henrik Molina: Investigation, Formal analysis
Jakob Bunkenborg: Formal analysis
Igor Grigoriev: Supervision, Writing – Review & Editing.
Kerrie Barry: Supervision
Jeremy Schmutz: Supervision
EonSeon Jin: Supervision, Resources
John C. Cushman: Conceptualization, Supervision, Resources, Writing – Review & Editing
Jon R. Magnusson: Conceptualization, Supervision, Resources, Writing – Review & Editing

Acknowledgements

The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research was supported, in part, under the National Science Foundation Grants CNS-0958379, CNS-0855217, and ACI-1126113, the City University of New York High Performance Computing Center at the College of Staten Island, the Sun Grant Initiative, Western Regional Center (U.S. Department of Transportation), and the University of Nevada Agricultural Experiment Station (NEV-00342). We thank Dr. Duc Tran, Dr. Erika Lindquist, and Dr. Mark Buchheim for technical support.

Statement of informed consent

No conflicts, informed consent, human or animal rights applicable.

Declaration of author's agreement to authorship and submission

All authors have agreed to authorship and the submission of this manuscript for peer review.

Declaration of competing interest

There is no conflict of interest that could be perceived to influence the outcomes of the present research.

References

1. Oren, A. 2005. A hundred years of *Dunaliella* research: 1905-2005. Review Saline Systems, 1(2), pp. 1-14.
2. Ginzburg, M. 1987. *Dunaliella*: A green alga adapted to salt. Advances in Botanical Research 14, pp. 93-183.
3. Polle J.E.W., Tran D., and A. Ben-Amotz (2009) History, distribution, and habitats of algae of the genus *Dunaliella* Teodoresco (Chlorophyceae). Chapter 1, In: The Alga *Dunaliella*: Biodiversity, Physiology, Genomics, and Biotechnology. Edited by Ben-Amotz A., Polle J.E.W., Rao D.V.S. Enfield, New Hampshire, USA, pp. 1-13. ISBN 9781578085453.
4. Teodoresco, E.C. (1905) Organization et developpement du *Dunaliella*, nouveau genre de Volvocacee-Polyblepharidee. Beihefte Botanisches Zentralblatt. Band 18 Abt.1, pp. 215-232.
5. Hamburger, C. (1905) Zur Kenntnis der *Dunaliella salina* und einer Amöbe aus Salinenwasser von Cagliari. Archiv für Protistenkunde, 6, pp. 111-130.
6. Borowitzka M.A., Siva C. (2007) The taxonomy of the genus *Dunaliella* (Chlorophyta, *Dunaliellales*) with emphasis on the marine and halophilic species. Journal of Applied Phycology, 19(5), pp. 567-590.
7. Ben-Amotz A., Polle, J.E.W., Subba Rao D.V. (2009) The alga *Dunaliella*: biodiversity, physiology, genomics and biotechnology. Science Publishers, Enfield, NH, USA. ISBN 9781578085453.
8. Pick, U. (1998), *Dunaliella* - a model extremophilic alga. Israel Journal Plant Science, 46, pp. 131-140.
9. Gunde-Cimerman N., Plemenitas A., Oren A. (2018) Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. FEMS Microbiology Reviews, 42 (3), pp. 353-375.
10. Oren A. (2009) Saltern evaporation ponds as model systems for the study of primary production processes under hypersaline conditions. Aquatic Microbial Ecology, 56 (2-3), pp. 193-204.
11. Jones B.F., Naftz D.L., Spencer R.J., Oviatt C.G. (2009) Geochemical evolution of Great Salt Lake, Utah, USA. Aquatic Geochemistry, 15 (1-2), pp. 95-121.
12. Lazar B., Starinsky A., Katz A., Sass E., Ben-Yaakov S. (1983) The carbonate system in hypersaline solutions: Alkalinity and $CaCO_3$ solubility of evaporated seawater. Limnology and Oceanography, 28 (5), pp. 978-986.
13. Liu X., Millero F.J. (1999) The solubility of iron hydroxide in sodium chloride solutions. Geochimica et Cosmochimica Acta, 63 (19-20), pp. 3487-3497.

14. Oliveira L., Bisalputra T., Antia N.J. (1980) Ultrastructural observation of the surface coat of *Dunaliella tertiolecta* from staining with cationic dyes and enzyme treatments. New Phytologist, 85 (3), pp. 385-392.
15. Ben-Amotz A. and M. Avron (1973) The role of glycerol in the osmotic regulation of the halophilic alga *Dunaliella parva*. Plant Physiology 51, pp. 875-878.
16. Borowitzka L.J., Brown A.D. (1974) The salt relations of marine and halophilic species of the unicellular green alga *Dunaliella*. Archives of Microbiology, 96(1), pp. 37-52.
17. Wegman K. (1971) Osmotic regulation of photosvnthetic glycerol production in *Dunaliella*. Biochimica et Biophysica Acta, 234, pp. 317-323.
18. Chitlaru E., Pick U. (1991) Regulation of glycerol synthesis in response to osmotic changes in *Dunaliella*. Plant Physiology 96 (1), pp. 50–60.
19. Thompson Jr. G.A. (1994) Mechanisms of osmoregulation in the green alga *Dunaliella salina*. Journal of Experimental Zoology, 268 (2), pp. 127-132.
20. Pick U. (2002) Adaptation of the halotolerant alga *Dunaliella* to high salinity. In: Salinity, environment, plants, molecules. Eds. Läuchli A. & Lüettge U., Kluwer Acad. Pub./ Dordrecht, The Netherlands, pp. 97-112.
21. Goyal A. (2007) Osmoregulation in *Dunaliella*, Part II: Photosynthesis and starch contribute carbon for glycerol synthesis during a salt stress in *Dunaliella tertiolecta*. Plant Physiology and Biochemistry, 45 (9), pp. 705-710.
22. Chen H., Jiang J.-G. (2009) Osmotic responses of *Dunaliella* to the changes of salinity. Journal of Cellular Physiology, 219 (2), pp. 251-258.
23. Chen, H., Lu, Y., Jiang, J.-G. (2012) Comparative analysis on the key enzymes of the glycerol cycle metabolic pathway in *Dunaliella salina* under osmotic stresses. PLoS ONE 7 (6), art. no. e37578.
24. Oren A. (2017) Glycerol metabolism in hypersaline environments. Environmental Microbiology, 19 (3), pp. 851-863.
25. Katz A., Waridel P., Shevchenko A., Pick U. 2007. Salt-induced changes in the plasma membrane proteome of the halotolerant alga *Dunaliella salina* as revealed by blue-native gel electrophoresis and nano-LC-MS/MS analysis. Molecular and Cellular Proteomics, 6, pp. 1459-1472.
26. Katz A., Paz Y., Pick U. (2009) Salinity tolerance and iron deprivation resistance mechanisms revealed by proteomic analyzes in *Dunaliella salina*. Chapter 14 In: The Alga *Dunaliella*: Biodiversity, Physiology, Genomics, and Biotechnology. Edited by Ben-Amotz A., Polle J.E.W., Rao D.V.S. Enfield, New Hampshire, USA, pp. 341-358.
27. Fisher M., Gokhman I., Pick U., Zamir A. (1996) A salt-resistant plasma membrane carbonic anhydrase is induced by salt in *Dunaliella salina*. Journal of Biological Chemistry, 271 (30), pp. 17718-17723.
28. Jeon H., Jeong J., Baek K., McKie-Krisberg Z., Polle J.E.W., Jin E. (2016) Identification of the carbonic anhydrases from the unicellular green alga *Dunaliella salina* strain CCAP 19/18. Algal Research, 19, pp. 12-20.
29. Cowan A.K., Rose P.D., and L.G. Horne (1992) *Dunaliella salina*: A model system for studying the response of plant cells to stress. Journal of Experimental Botany, 43 (12), pp. 1535-1547.

30. Ben Amotz A, Katz A, Avron M (1982) Accumulation of β-carotene in halotolerant algae: purification and characterization of β-carotene-rich globules from *Dunaliella bardawil*. (Chloropyceae). Journal of Phycology, 18, pp. 529–537.
31. Lamers P.P., Janssen M., De Vos R.C.H., Bino R.J., Wijffels R.H. (2008) Exploring and exploiting carotenoid accumulation in *Dunaliella salina* for cell-factory applications. Trends in Biotechnology, 26 (11), pp. 631-638.
32. Ramos A.A., Polle J., Tran D., Cushman J.C., Jin ES., Varela J.C. (2011) The unicellular green alga *Dunaliella salina* Teod. as a model for abiotic stress tolerance: genetic advances and future perspectives. Algae, 26(1), pp. 3-20.
33. Ben-Amotz A., Avron M. (1990) The biotechnology of cultivating the halotolerant alga *Dunaliella*. Trends in Biotechnology, 8 (C), pp. 121-126.
34. Ben-Amotz A. (2009) Bioactive compounds: Glycerol production, carotenoid production, fatty acids production. Chapter 8, In: The Alga *Dunaliella*: Biodiversity, Physiology, Genomics, and Biotechnology. Edited by Ben-Amotz A, Polle JEW, Rao DVS. Enfield, New Hampshire, USA, pp. 1-13. ISBN 9781578085453.
35. Borowitzka M.A. (2013) *Dunaliella*: Biology, Production, and Markets. In: Handbook of Microalgal Culture: Applied Phycology and Biotechnology: Second Edition, Eds. Richmond A. & Hu Q., John Wiley and Sons Ltd., West Sussex, UK, pp. 359-368.
36. Borowitzka L. J., M. A. Borowitzka and T. P. Moulton (1984) The mass culture of *Dunaliella salina* for fine chemicals: From laboratory to pilot plant. Hydrobiologia, 116-117(1), 115-121
37. Polle J.E.W., Barry K., Cushman J., Schmutz J., Tran D., (...), Magnuson J. (2017) Draft nuclear genome sequence of the halophilic and β-Carotene-accumulating green alga *Dunaliella salina* strain CCAP19/18. Genome Announcements, 5 (43), art. no. e01105-17.
38. González M.A., Gómez P.I., and J.E.W. Polle 2009 Taxonomy and phylogeny of the genus *Dunaliella*. In: The Alga *Dunaliella*: Biodiversity, Physiology, Genomics, and Biotechnology. Edited by Ben-Amotz A, Polle JEW, Rao DVS. Enfield, New Hampshire, USA pp. 15-44
39. Pick U., L. Karni, M. Avron, (1986) Determination of ion content and ion fluxes in the halotolerant alga *Dunaliella salina*. Plant Physiology 81: 92-96.
40. Sambrook J., and Russell D.W. (2000) Molecular Cloning, 3rd Edition. Cold Spring Harbor Laboratory Press. B017MYTYJO.
41. Goodstein D.M., Shu S., Howson R., Neupane R., Hayes R.D., (…), Rokhsar D.S. (2012) Phytozome: a comparative platform for green plant genomics, Nucleic Acids Research, 40 (D1), D1178-D1186.
42. Lovell J.T., Jenkins J., Lowry D.B., Mamidi S., Sreedasyam A., (...), Juenger T.E. (2018) The genomic landscape of molecular responses to natural drought stress in *Panicum hallii*. Nature Communications, 9 (1), art. no. 5213.
43. Nielsen H., Engelbrecht J., Brunak S., von Heijne G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering 10, 1-6.
44. Melen K., Krogh A., von Heijne G. (2003). Reliability measures for membrane protein topology prediction algorithms. Journal of Molecular Biology 327, 735-744.
45. Quevillon E., Silventoinen V., Pillai S., Harte N., Mulder N., Apweiler R., Lopez R. (2005). InterProScan: protein domains identifier. Nucleic Acids Research 33, W116-120.

46. UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Research 47, D506-D515.
47. Kanehisa M., Sato Y., Furumichi M., Morishima K., Tanabe M. (2018). New approach for understanding genome variations in KEGG. Nucleic Acids Research 47, D590-D595.
48. Koonin E.V., Fedorova N.D., Jackson J.D., Jacobs A.R., Krylov D.M., (...), Natale D.A. (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. Genome Biology, 5 (2), pp. R7.
49. Gene Ontology Consortium (2015). Gene Ontology Consortium: going forward. Nucleic Acids Research, 43, D1049-1056.
50. De Bie T., Cristianini N., Demuth J. P., Hahn M. W. (2006) CAFE: a computational tool for the study of gene family evolution. Bioinformatics 22.
51. Han M. V., Thomas G. W. C., Lugo-Martinez J., Hahn M. W. (2013) Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Molecular Biology and Evolution 30, 8.
52. Merchant S.S., Prochnik S.E., Vallon O., Harris E.H., Karpowicz S.J., (...), Zhou K. (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science, 318 (5848), pp. 245-251.
53. Prochnik S.E., Umen J., Nedelcu A.M., Hallmann A., Miller S.M., (...), Rokhsar D.S. (2010) Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. Science, 329 (5988), pp. 223-226.
54. Hanschen E.R., Marriage T.N., Ferris P.J., Hamaji T., Toyoda A., (...), Olson B.J.S.C. (2016) The *Gonium pectorale* genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. Nature Communications, 7, art. no. 11370.
55. Roth M.S., Cokus S.J., Gallaher S.D., Walter A., Lopez D., (...), Niyogi K.K. (2017) Chromosome-level genome assembly and transcriptome of the green alga *Chromochloris zofingiensis* illuminates astaxanthin production. Proceedings of the National Academy of Sciences of the United States of America, 114 (21), pp. E4296-E4305.
56. Bogen C., Al-Dilaimi A., Albersmeier A., Wichmann J., Grundmann M., (...), Kruse O. (2013) Reconstruction of the lipid metabolism for the microalga *Monoraphidium neglectum* from its genome sequence reveals characteristics suitable for biofuel production. BMC Genomics, 14 (1), art. no. 926.
57. Parra G., Bradnam K., Ning Z., Keane T., Korf I. (2009) Assessing the gene space in draft genomes. Nucleic Acids Research, 37(1): 298-297.
58. Seppey M., Manni M., Zdobnov E.M. (2019) BUSCO: Assessing Genome Assembly and Annotation Completeness. In: Kollmar M. (eds) Gene Prediction. Methods in Molecular Biology, (1962), Humana, New York, NY. 2019 doi.org/10.1007/978-1-4939-9173-0_14. PMID:31020564.
59. Schloer S., Pajonczyk D., Rescher U. (2018) Annexins in translational research: Hidden treasures to be found. International Journal of Molecular Sciences, 19 (6), art. no. 1781.
60. Jami S.K., Clark G.B., Ayele B.T., Ashe P., Kirti P.B. (2012) Genome-wide comparative analysis of annexin superfamily in plants. PLoS ONE, 7 (11), art. no. e47801.
61. Clark G.B., Morgan R.O., Fernandez M.-P., Roux S.J. (2012) Evolutionary adaptation of plant annexins has diversified their molecular structures, interactions and functional roles. New Phytologist, 196 (3), pp. 695-712.

62. Kudla J., Becker D., Grill E., Hedrich R., Hippler M., (...), Schumacher K. (2018) Advances and current challenges in calcium signaling. New Phytologist, 218 (2), pp. 414-431.
63. Zhang X., Tang X., Wang M., Zhang W., Zhou B., Wang Y. (2017) ROS and calcium signaling mediated pathways involved in stress responses of the marine microalgae *Dunaliella salina* to enhanced UV-B radiation. Journal of Photochemistry and Photobiology B: Biology, 173, pp. 360-367.
64. Chen H., Chen S.-L., Jiang J.-G. (2011) Effect of Ca^{2+} channel block on glycerol metabolism in *Dunaliella salina* under hypoosmotic and hyperosmotic stresses. PLoS ONE, 6 (12), art. no. e28613.
65. Yeats T.H., Bacic A., Johnson K.L. (2018) Plant glycosylphosphatidylinositol anchored proteins at the plasma membrane-cell wall nexus. Journal of Integrative Plant Biology, 60 (8), pp. 649-669.
66. Zhou K. (2019) Glycosylphosphatidylinositol-Anchored Proteins in *Arabidopsis* and One of Their Common Roles in Signaling Transduction. Frontiers in Plant Science, 10, art. no. 1022.
67. Schwarz M., Zamir A., Pick U. (2003) Iron-binding properties of TTf, a salt-induced transferrin from the alga *Dunaliella salina*. Journal of Plant Nutrition, 26 (10-11), pp. 2081-2091.
68. Schwarz M., Sal-Man N., Zamir A., Pick U. (2003) A transferrin-like protein that does not bind iron is induced by iron deficiency in the alga *Dunaliella salina*. Biochimica et Biophysica Acta - Proteins and Proteomics, 1649 (2), pp. 190-200.
69. Lambert L.A. (2012) Molecular evolution of the transferrin family and associated receptors. Biochimica et Biophysica Acta - General Subjects, 1820 (3), pp. 244-255.
70. Bai L., Qiao M., Zheng R., Deng C., Mei S., Chen W. (2016) Phylogenomic analysis of transferrin family from animals and plants. Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics, 17, pp. 1-8.
71. Choudhry Z., Rikani A.A., Choudhry A.M., Tariq S., Zakaria F., (...), Mobassarah N.J. (2014) Sonic hedgehog signaling pathway: A complex network. Annals of Neurosciences, 21 (1), pp. 28-31.
72. Sheffer M., Fried A., Gottlieb H.E., Tietz A., Avron M. (1986) Lipid composition of the plasma-membrane of the halotolerant alga, *Dunaliella salina*. BBA - Biomembranes, 857 (2), pp. 165-172.
73. Zelazny A.M., Shaish A., Pick U. (1995) Plasma membrane sterols are essential for sensing osmotic changes in the halotolerant alga *Dunaliella*. Plant Physiology, 109 (4), pp. 1395-1403.
74. Katz A., Pick U. (2001) Plasma membrane electron transport coupled to Na+ extrusion in the halotolerant alga *Dunaliella*. Biochimica et Biophysica Acta - Bioenergetics, 1504 (2-3), pp. 423-431.
75. Wu H., Zhang X., Giraldo J.P., Shabala S. (2018) It is not all about sodium: revealing tissue specificity and signaling roles of potassium in plant responses to salt stress. Plant and Soil, 431 (1-2), pp. 1-17.
76. Hohmann S. (2009) Control of high osmolarity signalling in the yeast *Saccharomyces cerevisiae.* FEBS Letters, 583 (24), pp. 4025-4029.
77. Tang Z., Cao X., Qiao J., Huang G., Lian W., (...), Cao, Y. (2019) MAPKKK gene family in *Dunaliella salina*: identification, interaction network, and expression patterns under abiotic stress. Journal of Applied Phycology, in press.

78. Fachet M., Witte C., Flassig R.J., Rihko-Struckmann L.K., McKie-Krisberg Z., (...), Sundmacher K. (2020) Reconstruction and analysis of a carbon-core metabolic network for *Dunaliella salina*. BMC Bioinformatics 21(1), pp. 1-9.
79. Alkayal F., Albion R.L., Tillett R.L., Hathwaik L.T., Lemos M.S., Cushman J.C. (2010) Expressed sequence tag (EST) profiling in hyper saline shocked *Dunaliella salina* reveals high expression of protein synthetic apparatus components. Plant Science, 179 (5), pp. 437-449.
80. Zhao L.-N., Gong W.-F., Chen X.-W., Chen D.-F. (2013) Characterization of genes and enzymes in *Dunaliella salina* involved in glycerol metabolism in response to salt changes. Phycological Research, 61 (1), pp. 37-45.
81. Craigie J.S, McLachlan J. (1964) Glycerol as a photosynthetic product in *Dunaliella tertiolecta* Butcher. Canadian Journal of Botany 42, pp. 777-778.
82. Avron M. (1986) The osmotic components of halotolerant algae. Trends in Biochemical Sciences, 11 (1), pp. 5-6.
83. Avron M., Ben-Amotz A. (1992) *Dunaliella*: physiology, biochemistry, and biotechnology. CRC Press, Boca Raton, FL, USA. ISBN-10: 084936647X.
84. Ben-Amotz A., Avron M. (1981) Glycerol and ß-carotene metabolism in the halotolerant alga *Dunaliella*: a model system for biosolar energy conversion. Trends in Biochemical Sciences, 6 (C), pp. 297-299.
85. Liska A.J., Shevchenko A., Pick U., Katz A. (2004) Enhanced photosynthesis and redox energy production contribute to salinity tolerance in *Dunaliella* as revealed by homology-based proteomics. Plant Physiology, 136 (1), pp. 2806-2817.
86. Wu Q., Lan Y., Cao X., Yao H., Qiao D., (...), Cao Y. (2019) Characterization and diverse evolution patterns of glycerol-3-phosphate dehydrogenase family genes in *Dunaliella salina*. Gene, 710, pp. 161-169.
87. He Q., Qiao D., Bai L., Zhang Q., Yang W., Li Q., Cao Y. (2007) Cloning and characterization of a plastidic glycerol 3-phosphate dehydrogenase cDNA from *Dunaliella salina*. Journal of Plant Physiology, 164 (2), pp. 214-220.
88. Cai M., He L.-H., Yu T.-Y. (2013) Molecular Clone and Expression of a NAD+-Dependent Glycerol-3-Phosphate Dehydrogenase Isozyme Gene from the Halotolerant alga *Dunaliella salina*. PLoS ONE, 8 (4), art. no. e62287.
89. He Q., Toh J.D., Ero R., Qiao Z., Kumar V., Serra A., Tan J., (...), Gao Y.-G. (2019) The unusual di-domain structure of *Dunaliella salina* glycerol-3-phosphate dehydrogenase enables direct conversion of dihydroxyacetone phosphate to glycerol. Plant Journal, in press.
90. He, Y., Meng, X., Fan, Q., Sun, X., Xu, Z., Song, R., (2009) Cloning and characterization of two novel chloroplastic glycerol-3-phosphate dehydrogenases from *Dunaliella viridis*. Plant Molecular Biology, 71 (1–2), 193–205.
91. Herrera-Valencia, V.A., Macario-González, L.A., Casais-Molina, M.L., Beltran-Aguilar, A.G., Peraza-Echeverría, S., (2012) In silico cloning and characterization of the glycerol-3-phosphate dehydrogenase (GPDH) gene family in the green microalga *Chlamydomonas reinhardtii*. Current Microbiology, 64 (5), 477–485.
92. Morales-Sanchez D., Kim Y., Terng E.L., Peterson L., Cerutti H. (2017) A multidomain enzyme, with glycerol-3-phosphate dehydrogenase and phosphatase activities, is involved in a chloroplastic pathway for glycerol synthesis in *Chlamydomonas reinhardtii.* Plant Journal, 90 (6), pp. 1079-1092.

93. Casais-Molina M.L., Peraza-Echeverria S., Echevarria-Machado I., Herrera-Valencia V.A. (2016) Expression of *Chlamydomonas reinhardtii* CrGPDH2 and CrGPDH3 cDNAs in yeast reveals that they encode functional glycerol-3-phosphate dehydrogenases involved in glycerol production and osmotic stress tolerance. Journal of Applied Phycology, 28 (1), pp. 219-226.
94. Sussman I., Avron M. (1981) Characterization and partial purification of dl-glycerol-1-phosphatase from *Dunaliella salina*. BBA - Enzymology, 661 (2), pp. 199-204.
95. Ruan K., Duan J., Bai F., Lemaire M., Ma X., Bai L. (2009) Function of *Dunaliella salina* (*Dunaliellaceae*) enolase and its expression during stress. European Journal of Phycology, 44 (2), pp. 207-214.
96. Xia B.-B., Wang S.-H., Duan J.-B., Bai L.-H. (2014) The relationship of glycerol and glycolysis metabolism patway under hyperosmotic stress in *Dunaliella salina*. Central European Journal of Biology, 9 (9), pp. 901-908.
97. Polle J.E.W, Neofotis P., Huang A., Chang W., Sury K., and Wiech E.M. (2014) Carbon partitioning in green algae (Chlorophyta) and the Enolase enzyme. Metabolites, 4(3), pp. 612–628.
98. Gruning N.-M., Du D., Keller M.A., Luisi B.F., Ralser M. (2014) Inhibition of triosephosphate isomerase by phosphoenolpyruvate in the feedback-regulation of glycolysis. Open Biology, 4 886 art. no. 130232.
99. Muller W., Wegmann K. (1978) Sucrose biosynthesis in *Dunaliella* - I. Thermic and osmotic regulation. Planta, 141 (2), pp. 155-158.
100. Muller W., Wegmann K. (1978) Sucrose biosynthesis in *Dunaliella* - II. Isolation and properties of sucrose phosphate synthetase. Planta, 141 (2), pp. 159-163.
101. McKie-Krisberg Z.M., Laurens L.M.L., Huang A., Polle J.E.W. (2018) Comparative energetics of carbon storage molecules in green algae. Algal Research, 31, pp. 326-333.
102. Lohr M., Schwender J., Polle J.E.W. (2012) Isoprenoid biosynthesis in eukaryotic phototrophs: A spotlight on algae. Plant Science, 185-186, pp. 9-22.
103. Luo Q., Bian C., Tao M., Huang Y., Zheng Y., Lv Y., Li J., (...), Hu Z. (2019) Genome and transcriptome sequencing of the Astaxanthin-producing green microalga, *Haematococcus pluvialis*. Genome Biology and Evolution, 11 (1), pp. 166-173.
104. Guerin M., Huntley M.E., Olaizola M. (2003) *Haematococcus* astaxanthin: Applications for human health and nutrition. Trends in Biotechnology, 21 (5), pp. 210-216.
105. Jin E., Lee C.-G., Polle J.E.W. (2006) Secondary carotenoid accumulation in *Haematococcus* (Chlorophyceae): Biosynthesis, regulation, and biotechnology. Journal of Microbiology and Biotechnology, 16 (6), pp. 821-831.
106. Rodriguez-Concepcion M. (2010). Supply of precursors for carotenoid biosynthesis in plants. Archives Biochemistry Biophysics, 504 (118–122), 10.1016/j.abb.2010.06.016.
107. Saladie M., Wright L.P., Garcia-Mas J., Rodriguez-Concepcion M., Phillips M.A. (2014) The 2-C-methylerythritol 4-phosphate pathway in melon is regulated by specialized isoforms for the first and last steps. Journal of Experimental Botany, 65 (17), pp. 5077-5092.
108. Wang Y., Yuan X., Li S., Chen W., Li J. (2019) Gene cloning and functional characterization of Three 1-Deoxy-D-Xylulose 5-Phosphate synthases in Simao Pine. BioResources, 13 (3), pp. 6370-6382.

109. Ro D.-K. (2011) Plant Metabolism and Biotechnology, First Edition. Eds. Hiroshi Ashihara, Alan Crozier, and Atsushi Komamine. John Wiley & Sons, Ltd. ISBN: 978-0-470-74703-2.
110. Tran D., Haven J., Qiu W.-G., Polle J.E.W. (2009) An update on carotenoid biosynthesis in algae: Phylogenetic evidence for the existence of two classes of phytoene synthase. Planta, 229 (3), pp. 723-729.
111. Lado J., Zacarias L., Rodrigo M.J. (2016) Regulation of carotenoid biosynthesis during fruit development. Sub-Cellular Biochemistry, 79, pp. 161-198.
112. Fraser P.D., Truesdale M.R., Bird C.R., Schuch W., Bramley P.M. (1994) Carotenoid biosynthesis during tomato fruit development (evidence for tissue-specific gene expression). Plant Physiology, 105, pp. 405–413.
113. Ben-Amotz A., Lers A., Avron M. (1988) Stereoisomers of β-carotene in the alga *Dunaliella bardawil*. Plant Physiology, 86, pp. 1286–1291.
114. Davidi L., Pick U. (2017) Novel 9-cis/all-trans β-carotene isomerases from plastidic oil bodies in Dunaliella bardawil catalyze the conversion of all-trans to 9-cis β-carotene. Plant Cell Reports, 36 (6), pp. 807-814.
115. Lin B., Cui Y., Yan M., Wang Y., Gao Z., (...), Qin, S. (2019) Construction of astaxanthin metabolic pathway in the green microalga *Dunaliella viridis*. Algal Research, 44, 101697.
116. Perozeni F., Cazzaniga S., Baier T., Zanoni F., Zoccatelli G., (…), Ballottari M. (2020) Turning a green alga red: engineering astaxanthin biosynthesis by intragenic pseudogene revival in *Chlamydomonas reinhardtii*. Plant Biotechnology Journal, pp. 1–15.
117. Goodwin T. (1980) The Biochemistry of the Carotenoids. Volume I Plants. Springer, The Netherlands. ISBN 978-94-009-5860-9.
118. Davidi L., Shimoni E., Khozin-Goldberg I., Zamir A., Pick U. (2014) Origin of β-carotene-rich plastoglobuli in Dunaliella bardawil. Plant Physiology, 164 (4), pp. 2139-2156.
119. Polle J.E.W., Roth R., Ben-Amotz A., Goodenough U. (2019) Ultrastructure of the green alga *Dunaliella salina* (Chlorophyta) as investigated by Quick-Freeze Deep-Etch Electron Microscopy. Algal Research, *submitted*.
120. Pick U., Zarka A., Boussiba S., Davidi L. (2019) A hypothesis about the origin of carotenoid lipid droplets in the green algae *Dunaliella* and *Haematococcus*. Planta, 249 (1), pp. 31-47.
121. Shanmugabalaji V., Besagni C., Piller L.E., Douet V., Ruf S., (…), Kessler F. (2013) Dual targeting of a mature plastoglobulin/fibrillin fusion protein to chloroplast plastoglobules and thylakoids in transplastomic tobacco plants. Plant Molecular Biology, 81 (1-2), pp. 13-25.
122. Davidi L., Levin Y., Ben-Dor S., Pick U. (2015) Proteome analysis of cytoplasmatic and plastidic β-carotene lipid droplets in *Dunaliella bardawil*. Plant Physiology, 167 (1), pp. 60-79.
123. Gong W.-F., Zhao L.-N., Hu B., Chen X.-W., Zhang F., (…), Chen D.-F. (2014) Identifying novel salt-tolerant genes from *Dunaliella salina* using a *Haematococcus pluvialis* expression system. Plant Cell, Tissue and Organ Culture, 117 (1), pp. 113-124.