Interactions of photosynthesis with genome size and function

John A. Raven1, John Beardall 2, Anthony W. D. Larkum3 and Patricia Sánchez-Baracaldo4,5

1 School of Plant Biology, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia
2 School of Biological Sciences, Monash University, Clayton, Victoria 3800, Australia
3 Functional Plant Biology and Climate Change Cluster, University of Technology Sydney, New South Wales 2007, Australia
4 School of Biological Sciences, and 5 School of Geographical Sciences, University of Bristol, Woodland Road, Bristol BS8 1SS, UK

Photolithotrophs are divided between those that use water as their electron donor (Cyanobacteria and the photosynthetic eukaryotes) and those that use a different electron donor (the anoxygenic photolithotrophs, all of them Bacteria). Photolithotrophs with the most reduced genomes have more genes than do the corresponding chemoorganotrophs, and the fastest-growing photolithotrophs have significantly lower specific growth rates than the fastest-growing chemoorganotrophs. Slower growth results from diversion of resources into the photosynthetic apparatus, which accounts for about half of the cell protein. There are inherent dangers in (especially oxygenic) photosynthesis, including the formation of reactive oxygen species (ROS) and blue light sensitivity of the water spitting apparatus. The extent to which photolithotrophs incur greater DNA damage and repair, and faster protein turnover with increased rRNA requirement, needs further investigation. A related source of environmental damage is ultraviolet B (UVB) radiation (280–320 nm), whose flux at the Earth’s surface decreased as oxygen (and ozone) increased in the atmosphere. This oxygenation led to the requirements of defence against ROS, and decreasing availability to organisms of combined (non-dinitrogen) nitrogen and ferrous iron, and (indirectly) phosphorus, in the oxygenated biosphere. Differential codon usage in the genome and, especially, the proteome can lead to economies in the use of potentially growth-limiting elements

1. Introduction

The evolution of photosynthesis greatly increased the energy input to the biosphere, supplementing energy from chemolithotrophy and from photochemistry that is not catalysed by organisms, as well as from any less globally significant energy sources [1]. Photosynthetic processes, including energy-transducing rhodopsins as well as (bacterio)chlorophyll-based photochemistry, can bring about some of the proton pumping, and hence adenosine triphosphate (ATP) synthesis, in chemoorganotrophs, which would otherwise involve oxidation of organic compounds. Here, the role of the photosynthetic reactions is to decrease production of carbon dioxide from organic matter, while maintaining, or increasing, the productivity of chemoorganotrophs [2,3]; the global upper limit on these processes in the carbon cycle has been estimated [4]. The predominant biogeochemical role of photosynthetic processes is, however, photolithotrophy; that is, the autotrophic assimilation of carbon dioxide using some inorganic reductant as the electron donor. Unless hydrogen is used as the electron donor, photochemical energy input is needed to produce a reductant (often with additional energy input from photogenerated ATP) capable of reducing carbon dioxide to the redox level of carbohydrate. On Earth today, oxygenic photolithotrophy assimilates carbon dioxide globally at over 100 Pg carbon per year in net primary productivity; the corresponding numbers for anoxygenic photolithotrophy and chemolithotrophy (mainly nitrification) are, respectively, 0.03–0.07 and 0.0001–0.001 Pg carbon per year.
0.3 Pg C per year [5,6], although anoxicogenic photolithotrophy was quantitatively more important in the past [5–8].

However, photosynthetic organisms are not selected in evolution for their contribution to global biogeochemistry: the successful photosynthetic organisms leave more offspring than their less successful competitors. The first question considered in this paper is the extent to which photolithotrophy means that organisms have a larger minimum number of genes, larger minimum genome size, and a larger minimum cell size, than organisms of similar cell organization but living by chemooorganotrophy or chemolithotrophy.

A second area investigated is comparison of the maximum specific growth rate among the three trophic modes in relation to the extent of diversion of resources to catalysts and structures related specifically to chemooorganotrophy, chemolithotrophy or photolithotrophy. Consideration of the involvement in the different trophic modes of highly expressed genes could help explain any mismatch between the small number of genes involved specifically in a given trophic mode and the larger fraction of cell protein involved specifically in that trophic mode. One example of highly expressed genes is those for Rubisco, which occurs in all oxygenic photolithotrophs and also in some anoxicogenic phototrophs and in some chemolithotrophs [9–12]. Another set of highly expressed genes are those encoding the apoproteins, which occur in all (bacterio)chlorophyll-based photosynthetic systems. The third topic addressed is the impact of the requirements for photosynthesis (exposure to solar radiation) on photosynthetic organisms, and the effects of local and global oxygenation on organisms directly as oxygen, and indirectly through the availability of other resources.

### 2. Gene number, genome size and cell size

Table 1 gives the number of genes and the number of kbp for a number of genomes from free-living organisms, with an emphasis on those with a small number of genes. While the gene number is very important, the genome size is also worth considering as an indicator of resource requirement to replicate the genome, and, with gene number, the gene density. The organisms with the smallest genomes (as reported in peer-reviewed publications) found in free-living organisms occur in the Archaea and Bacteria. Those photolithotrophic organisms (oxygenic cyanobacteria) with the most reduced genomes have more genes than do the chemooorganotrophs (Proteobacteria) with the most reduced genomes, but fewer genes than the most reduced known genome of a chemolithotroph (autotrophic methanogenic Archaea) (table 1). This suggests that relatively few specific genes are needed for osmochemoorganotrophy (= saprochemoorganotrophy); these genes involve transport of solutes across the membrane and assimilation into core metabolism. Table 1 suggests that a greater number of specific genes are associated with autotrophy (chemolithotrophy and photolithotrophy); one category of such genes involves those coding for autotrophic carbon dioxide assimilation. A further group of genes relate to the energy transformation that converts the energy from reactions of inorganic compounds (in chemolithotrophy) or electromagnetic radiation (in photolithotrophy). It should be emphasized that these very small genomes are evolutionarily derived from larger genomes, e.g. in Cyanobacteria [32].

Table 1 also shows the smallest reported genomes (in peer-reviewed literature) for Eukarya of three trophic modes. The number of protein-coding genes, and the genome size in kbp, are smaller for fungal chemooorganotrophic osmotrophs than for the photolithotrophs; this is the case for the corresponding trophic modes in Bacteria. For Bacteria and Eukarya, the increment in gene numbers for the photolithotrophs relative to chemooorganotrophic osmotrophs is 178 and 613, respectively, while the increment as a percentage is, respectively, 26% and 13%. It will be interesting to see how these values change when more genomes are sequenced. There are no known genetically integrated chemolithotrophic Eukarya, although the known metabolic diversity of protists [33] means that this possibility cannot be ruled out. The other entries for Eukarya in table 1 are for chemoorganotrophic phagotrophs. The genome size and number of genes coding for proteins are much greater for the phagotrophic ciliate Tetrahymena thermophila than for either of the other two trophic groups; the mechanistic basis of this difference is not clear, although the volume of the ciliate is a hundred times that of Schizosaccharomyces pombe and more than a thousand times that of the photolithotrophic eukaryotes in table 1. However, the other phagoochemoorganotroph in table 1, the choanoflagellate (opisthokont) Monosiga brevicollis, has a volume similar to that of the osmochemoorganotroph Schizosaccharomyces pombe, but it has more genes than the fission yeast (or the photolithotrophs) in table 1. To our knowledge, there are no published values of gene number and genome size for rather smaller phagoochemoorganotrophs such as the heterokont flagellate Cafeteria roenbergensis [34]. The total gene number in any genome is likely to be greater than the number of genes in the core genome for that trophic mode (i.e. those genes found in all strains following a given trophic mode); the estimate of the number of genes in the core genome depends on how many organisms are examined and how they are chosen. The core genome concept is central to our understanding of the genomics of the Last Universal Common Ancestor (LUCA) [35,36], with more general discussion in [13–15,22–24,37–40].

It has been suggested that lateral gene transfer is limited to accessory genes, but not core genes. However, this may not be the case in the (photolithotrophic) cyanobacteria Prochlorococcus and the marine strains of Synechococcus, where core photosynthetic genes occur in cyanophages, and in genomic islands [41,42]. Horizontal transfer of genes coding for photochemical energy-transduction processes can occur in plasmids in the mixotrophs (in the second sense in the glossary), giving another means of horizontal gene transfer [2,23]. Also associated with the photochemical reactions of photosynthesis, is the case of the unicellular cyanobacterium Acaryochloris marina, in which chlorophyll d largely replaces chlorophyll a in photochemistry and light-harvesting roles: this organism has a relatively large genome with nine plasmids (but not as large as Nostoc) comprising 8528 protein-coding genes and with a size of 8362 kb [19] (table 1). This has been interpreted in terms of an organism filling a relatively ‘uncompetitive’ niche, enriched in near infrared radiation relative to 400–700 nm, where it is free to diversify its metabolic strategies [19]. Support for this view comes from the demonstration that A. marina has a much higher rate of gene duplication than average cyanobacteria and that certain gene duplications, in e.g. transcription, carbohydrate transport and metabolism, ion transport and metabolism and signal transduction, are positively selected [43].
Having a low gene number and dense gene packing permits cells to be very small (0.5 μm equivalent spherical diameter), without the genome taking up a greater fraction of the cell volume than is the case for larger cells with larger genomes. This relates to ‘scalability’ [44,45]. Genome reduction in Pelagibacter [46] and Prochlorococcus [32,47] may be related to specialization to oligotrophic habitats [16,47,48]. In contrast, larger genomes in some cyanobacteria were found in lineages from more variable habitats, e.g. microbial mats or inter-tidal zones (table 1) [20,32]; these habitats are often associated with more complex morphologies such as filamentous and multicellular forms [49].

Recent analyses have shown that the morphologically complex true-branching cyanobacterium Scytonema hofmanii PCC 7110 is the most gene-rich prokaryote currently known, with 12356 protein-coding genes in its 12073 kbp genome [20]. Phylogenetic and genomic studies have shown a clear trend in the reduction of genome size in the evolution of Prochlorococcus [50,51], which dominate in habitats of low nutrient availability. Those Prochlorococcus genotypes growing near the thermocline/pycnocline are exposed to higher nutrient concentrations, but intercept lower fluxes of photosynthetically active radiation [47].

Reduction in genome size has not only been observed within planktonic cyanobacteria (e.g. Prochlorococcus) but also in some symbiotic cyanobacteria [32] in which a reduced genome size might be indicative of their life style. The smallest genome (1400 kbp) found so far is the one from cyanobacterium U-CYNA [52], which is symbiotic with a small-celled prymnesiophycean alga [53]. Cyanobacterium U-CYNA lacks essential components of the photosynthetic machinery (e.g. 127 orthologues present in all other cyanobacterial genomes), including photosystem II, carboxysomes (i.e. the cyanobacterial microcompartments where CO₂ fixation takes place using Rubisco as the terminus of the CO₂ concentrating mechanism) and enzymes specific to the Calvin–Benson cycle, so that it cannot carry out autotrophic CO₂ fixation or O₂ production from water [54]. Less severe genome reductions have been identified in other obligate symbionts such as Nostoc azollae 0708 [32].

Turning to more obvious cases of obligate symbiosis, but one not involving diazotrophy, the chromatophore of the euglyphid (rhizarian) amoeba Paulinella chromatophora is an α-cyanobacterium (similar to Prochlorococcus and open ocean Synechococcus) coding for all photosynthetic components but lacking many enzymes involved in amino acid synthesis, and so depends on the amoeba for these amino acids [55,56]. The chromatophore genome has 867 protein-coding genes and is 1021 kbp in size [55]. There are also apparently genetically integrated diazotrophic cyanobacteria in some freshwater diatoms, such as Epistoma and Rhopalodia [57–59]. The extent of genome reduction in this case is not known, but the estimated genome

### Table 1. Examples of the smallest reported genome size and gene number for free-living Bacteria, Archaea and Eukarya of a range of trophic modes from peer-reviewed literature. For comparison, values are provided for a range of Cyanobacteria.

<table>
<thead>
<tr>
<th>trophic mode, organism</th>
<th>number of protein-coding genes</th>
<th>genome size (kbp)</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>chemoorganotrophic bacteria</td>
<td>1357 – 1541</td>
<td>1237 – 1457</td>
<td>[13,14]</td>
</tr>
<tr>
<td>SAR11 (e.g. Pelagibacter ubique)</td>
<td>1338</td>
<td>1304</td>
<td>[15]</td>
</tr>
<tr>
<td>β-Proteobacterium HTCC2181</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oxygenic photolithotrophic cyanobacteria</td>
<td>1716 – 3022</td>
<td>1643 – 2683</td>
<td>[16]</td>
</tr>
<tr>
<td>Prochlorococcus marinus</td>
<td>3968</td>
<td>3890</td>
<td>[17]</td>
</tr>
<tr>
<td>Raphidiopsis brookii</td>
<td>3088</td>
<td>3200</td>
<td>[17]</td>
</tr>
<tr>
<td>Cylindrospermopsis raciborskii</td>
<td>6501</td>
<td>8941</td>
<td>[17,18]</td>
</tr>
<tr>
<td>Nostoc punctiforme</td>
<td>8528</td>
<td>8362</td>
<td>[19]</td>
</tr>
<tr>
<td>Acaryochloris marina</td>
<td>12 356</td>
<td>12 073</td>
<td>[20]</td>
</tr>
<tr>
<td>Scytonema hofmannii</td>
<td>2288</td>
<td>2154</td>
<td>[21]</td>
</tr>
<tr>
<td>anoxygenic photolithotrophic bacterium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorobium tepidum TLS</td>
<td>4718</td>
<td>9200</td>
<td>[25]</td>
</tr>
<tr>
<td>chemolithotrophic methanogenic archaeans</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanocorpusculum labreanum</td>
<td>1828</td>
<td>1805</td>
<td>[22,23]</td>
</tr>
<tr>
<td>Methanobacterium thermooautotrophicum</td>
<td>1855</td>
<td>1751</td>
<td>[24]</td>
</tr>
<tr>
<td>chemoorganotrophic osmotrophic eukaryotes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ashbya gosypii</td>
<td>1196</td>
<td>41 600</td>
<td>[27]</td>
</tr>
<tr>
<td>Schizosaccharomyces pombe</td>
<td>27 000</td>
<td>104 000</td>
<td>[28]</td>
</tr>
<tr>
<td>chemoolrganotrophic phagotrophic eukaryotes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monasiga brevicollis</td>
<td>5331</td>
<td>16 520</td>
<td>[29]</td>
</tr>
<tr>
<td>Tetrahymena thermophila</td>
<td>7651</td>
<td>13 200</td>
<td>[30]</td>
</tr>
<tr>
<td>Ostreococcus lucimanus</td>
<td>7892</td>
<td>12 699</td>
<td>[30,31]</td>
</tr>
<tr>
<td>photolithotrophic eukaryotes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyanidioschyzon merolae</td>
<td>1382</td>
<td>16 520</td>
<td>[29]</td>
</tr>
<tr>
<td>Ostreococcus lucimanus</td>
<td>13 200</td>
<td>[30]</td>
<td></td>
</tr>
<tr>
<td>Ostreococcus tauri</td>
<td>12 699</td>
<td>[30,31]</td>
<td></td>
</tr>
</tbody>
</table>
size is 2600 kbp, which is less than that of free-living diazotrophic cyanobacteria but more than that of the smallest genome (table 1) in non-diazotrophic free-living cyanobacteria [57].

Some of the consequences of genome reduction involving significant decreases in metabolic capabilities are indicated by the following three examples, two from oxygenic photolithotrophs and one from osmochemooorganotrophs. (i) Many Prochlorococcus genotypes cannot use NO\textsubscript{3} as a N-source and some cannot use NO\textsubscript{2}; these genotypes need less energy for growth on reduced N, but do not have the option of using the more energy-expensive N source when it is available [60]. (ii) Prochlorococcus has no catalase, and so cannot dispose of photosynthetically produced H\textsubscript{2}O\textsubscript{2} without ‘helper’ chemooorganotrophic bacteria [61,62]. (iii) The α-proteobacterial clade SAR11 (e.g., Pelagibacter ubique) can only use reduced S sources such as methionine and dimethylsulfoniopropionate; other marine saprochemooorganotrophic α-proteobacteria can use SO\textsubscript{4}\textsuperscript{2−} [46]. The use of reduced S sources needs less energy for growth, but SAR-11 does not have the option of using the more energy-expensive SO\textsubscript{4}\textsuperscript{2−}, which is much more abundant in the sea and many other aquatic and terrestrial habitats [46].

It must, however, be pointed out that loss of genes, and hence loss of the associated function, is not restricted to minimal genomes. An example from the eukaryotic algae is the loss of the vitamin B\textsubscript{12}-independent pathway of methionine synthesis involving MetH, leaving the vitamin B\textsubscript{12}-dependent pathway involving MetE [63,64]. Since eukaryotes are unable to synthesize vitamin B\textsubscript{12}, the loss of MetH means that the growth of the cells in the absence (as is usually the case in nature) of external methionine depends on the provision of vitamin B\textsubscript{12} from the organisms which can synthesize it, i.e. Archaea and Bacteria. This provision can either occur through the aqueous environment or by a symbiotic relationship of the producer organism and eukaryotic algae [63]. Of the 306 algal species from a range of clades examined, more than half needed vitamin B\textsubscript{12} [63]. Estimates of the times at which MetH was lost in the chlorophycean order Volvocales shows that this process of gene loss is continuing [64].

Minimization of nuclear genome size does not seem to be a rationale for gene loss in the Volvocales, since there are an order of magnitude more protein-coding genes (14 516–14 520) in the nuclear genome of the Volvocales [65] than in Prochlorococcus (table 1). For comparison, the smallest known nuclear genome in green algae, that of the prasinophycean genome in Palenik et al., according to Develle et al. [31], while Palenik et al. [30] quote 7892 and 7651 protein-coding genes and 12.6 and 13.2 Mbp for O. tauri and O. lucimarinus, respectively. Palenik et al. [30] point out that Ostreococcus lacks MetH and so requires vitamin B\textsubscript{12}. In this instance, with the smallest nuclear gene number known for green algae, it is possible to argue that loss of MetH might, in this case, be related to genome minimization.

Despite the small (1.66 Mbp) genome size, over half of cellular P in P-limited Prochlorococcus MED4 is in DNA [66]. It has been suggested [66] that there is a role of P (and N) limitation in selecting for small genomes in organisms in P-limited environments and that this leaves more P available for RNA, allowing more rapid growth; however, this may not be universally applicable [67–69]. Also, there is often a negative correlation between maximum specific growth rate and genome size, and a positive correlation between specific growth rate and RNA content [67,69].
numbers per cell per second.

Oxygenic photosynthesis is especially susceptible to this danger of producing reactive oxygen species (ROS), with their damaging effects on cell constituents [94,95]. All electron transfer reactions in the presence of molecular oxygen carry the danger of producing reactive oxygen species (ROS), with their damaging effects on cell constituents [94,95].

Table 3. Maximum specific growth rates of microscopic Eukarya of a range of trophic modes; all aerobic. From [80–82]. Rates normalized to 20 °C assuming a Q10 of 2. Units of specific growth rate are increment of cell numbers per cell per second.

<table>
<thead>
<tr>
<th>organism and trophic mode</th>
<th>specific growth rate ($\times 10^6$ s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osmochemoorganotroph Achlya bisexualis (oomycete)</td>
<td>170</td>
</tr>
<tr>
<td>Phagoochemoorganotroph Tetrahymena geleii (ciliate)</td>
<td>88</td>
</tr>
<tr>
<td>Chemophagotroph Parphysomonas perforata (Chrysophyceae)</td>
<td>69</td>
</tr>
<tr>
<td>Photophagotroph Ochromonas sp. (Chrysophyceae)</td>
<td>32</td>
</tr>
<tr>
<td>Photolithotroph Dinobryon divergens (Chrysophyceae)</td>
<td>12</td>
</tr>
<tr>
<td>Photolithotroph Chlorella regularis (Trebouxiophyceae)</td>
<td>40</td>
</tr>
<tr>
<td>Photolithotroph Chlorella regularis (Trebouxiophyceae)</td>
<td>27</td>
</tr>
<tr>
<td>Photolithotroph Chaetoceros salsugineum (Bacillariophyceae)</td>
<td>75</td>
</tr>
</tbody>
</table>

The fastest-growing chemoorganotrophs have significantly higher specific growth rates than the fastest-growing photolithotrophs in both Bacteria and Eukarya. This slower growth of photosynthetic organisms apparently results from diversion of resources into the photosynthetic apparatus that can account for up to half of the cell protein (table 4). A further diversion of protein, and of RNA, in that fraction of the cellular ribosomal complement (a fraction of the cytosolic ribosomes, and all of the ribosomes in the plastids) are related to the synthesis of the photosynthesis-specific proteins. In eukaryotic algae, the plastids occupy up to half of the ‘metabolic’ cell volume, i.e. excluding vacuoles, storage granules and cell walls and other extracellular structures (table 5). The situation regarding protein allocation to chemolithotrophy-specific components of the methanogenic Archaea in table 2 is less clear. Less effort seems to have been made to determine the allocation of proteins (and RNA) to synthesizing and maintaining the chemolithotrophy-specific components of the cell machinery than is the case for photolithotrophy-specific components.

4. Photosynthesis-related damage to nucleic acids and proteins

All electron transfer reactions in the presence of molecular oxygen carry the danger of producing reactive oxygen species (ROS), with their damaging effects on cell constituents [94,95]. Oxidative photosynthesis is especially susceptible to this problem because the redox level needed to drain electrons from water (greater than +0.8 V) by the reaction centre of photosystem II leads to the generation of much singlet oxygen; the fact that the major protein involved, D1, is the fastest-turning-over protein known is a testament to the unavoidable damage incurred by the water splitting mechanism [96]. We will discuss the extent to which other protective measures fail and involve greater DNA damage and repair, and faster protein turnover than in chemooorganotrophs, with implications for energetics and rRNA expression. We will also comment on possible implications for the location of genes within the eukaryotic cell. A related source of environmental damage is UVB radiation, the flux of which decreased at the Earth’s surface as oxygen (and ozone) increased in the atmosphere; possibly leading UV avoidance and repair mechanisms from the times of higher UV fluxes will be discussed.

Local, and then global, oxygenation in the Great Oxygenation Event (GOE) some 2.32 Ga [97] led to the much-discussed requirements of defence against ROS and the opportunity for aerobic respiration for organisms in the oxygenated habitats. However, there are also effects on the nitrogen, phosphorus and iron cycles that decrease the availability to organisms of combined (non-dinitrogen) nitrogen, phosphorus and ferrous iron in the oxygenated biosphere. The effects on combined nitrogen and ferrous iron are relatively direct and occurred at, or soon after, the GOE [98]. The decreased availability of phosphorus was delayed, and related to the decrease in silicic acid concentration in the surface ocean, which decreases the competition between silicic acid and phosphate in the scavenging of phosphate by iron [99]. The removal of silicic acid depended on the evolution of silicified eukaryotes about 550 Ma (radiolarians, sponges) and, especially, diatoms approximately 120 Ma. There are implications of the decreased availabilities of these nutrient elements for the evolution of novel pathways related to enhancing the availability of these elements and/or decreasing the requirement for them. There is also the possibility of an influence of photosynthesis on variations in codon usage, which could economize in the use of potentially growth-limiting elements in the genome and, especially, the proteome, but this is not discussed in detail here [12,100,101].

The GOE involved $O_2$ production by oxidative photosynthesis, which evolved in Cyanobacteria [102]. Evolution of the water splitting apparatus, involving a manganese–calcium complex, exposed all organisms that use this apparatus to two vulnerabilities: (i) production of singlet oxygen, noted above, and (ii) blue light and UVB inhibition [103–105]. Recent work, combining phylogenic and character evolution analyses, has shown that Cyanobacteria originated in freshwater environments [20,49,102]. The small area of freshwaters means that the global biogeochemical impact of oxidative photosynthesis was minimal until Cyanobacteria started colonizing the marine environment (approx. 2.32 Ga), which currently covers more than two-thirds of the Earth’s surface [102]. Furthermore, marine cyanobacteria do not form a natural group or monophyletic clade, but they are clustered with freshwater species within the cyanobacterial radiation, providing further evidence for independent colonization events into marine environments [49,102]. As well as the possibility of respiration with $O_2$ as the terminal electron acceptor, the GOE also increased the possibility of damage to cells by ROS. Decreased ROS production from the interaction of UV radiation with (even anoxic) water [106], as a result of decreased UV flux at the Earth’s surface as atmospheric oxygenation led to increasing stratospheric $O_3$, is outweighed by ROS production from intracellular $O_2$ [94]. A specific
Table 4. Fraction of cell protein content of Cyanobacteria and microscopic eukaryotic algae occupied by the highly expressed photosynthetic proteins Rubisco and apoproteins of reaction centre and light-harvesting complexes, and for comparison the ribosomal proteins.

<table>
<thead>
<tr>
<th>fraction of cell protein</th>
<th>references</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubisco</td>
<td>0.03 – 0.16</td>
<td>[86] most data from algae expressing CMs grown at high light. Some values assume a conversion factor for chlorophyll to cell protein</td>
</tr>
<tr>
<td></td>
<td>0.02 – 0.06</td>
<td>[87] values all based on direct estimates of Rubisco protein and total protein</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[9,11] a saving of about a third of the protein requirement for a given rate of carbon dioxide fixation could be achieved by replacing the Rubisco Benson–Calvin cycle with one of the alternative autotrophic carbon dioxide fixation pathways</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>apoproteins of pigment–protein complexes</th>
<th>references</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04 – 0.4</td>
<td>[88 – 90]</td>
<td>highest values for algae growing at low light</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ribosomal proteins</th>
<th>references</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.09 – 0.21</td>
<td>[90]</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Fraction of ‘metabolic volume’ (cell volume minus walls and vacuoles) occupied by chloroplasts in microalgae (mainly freshwater) grown photolithotrophically [89,91 – 93]. n = number of species examined.

<table>
<thead>
<tr>
<th>Chlorophyta</th>
<th>0.33 – 0.61 (n = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euglenophyta</td>
<td>0.15 – 0.5 (n = 2)</td>
</tr>
<tr>
<td>Cryptophyta</td>
<td>0.69 (n = 1)</td>
</tr>
<tr>
<td>Haptophyta</td>
<td>0.34 – 0.40 (n = 2)</td>
</tr>
<tr>
<td>Bacillariophyceae</td>
<td>0.39 – 0.56 (n = 6)</td>
</tr>
<tr>
<td>Dinophyta</td>
<td>0.07 – 0.32 (n = 2)</td>
</tr>
</tbody>
</table>

The damaging effect of O$_2$ is the irreversible damage to nitrogenase [101,107,108], and through singlet oxygen formation, O$_2$ also contributes to photodamage to the D1 protein of photosystem II, resulting in photoinhibition [96,107,109]. In addition to the energy cost [110] of synthesis of replacement nitrogenase and photosystem II components (mainly D1 protein, the psbA gene product), Raven [101] argues that there is also a significant diversion of rRNA from other significant protein syntheses, at least under P limitation and with optimal allocation of resources, and also the occurrence of additional genes related to protection and repair [101,109]. Competition of O$_2$ with CO$_2$ at the active site of Rubisco, which evolved in anoxic conditions, does not involve damage by O$_2$ [12,107,111].

Despite these energetic and other constraints, some photosynthetic organisms can grow rapidly at up to four times air-equilibrium O$_2$ concentration, e.g. in high intertidal rock pools and stromatolites, and in bundle sheath cells of terrestrial NADme and PEPck C$_4$ flowering plants [112], and non-photosynthetic organisms occur in the rock pools and as endobionts in stromatolites. It would be predicted that there would be more oxygen-related protein damage, hence more protein resynthesis, in oxygenic photosynthetic cells than in non-photosynthetic aerobic cells, as a result of the production of singlet oxygen in photosynthesis, noted above. Moreover, at least in the light, the O$_2$ concentration in oxygenic photosynthetic cells is higher that it is in the medium, while for non-photosynthetic cells the intracellular O$_2$ concentration in the respiring cells is less than that in the medium. It would also be predicted that there would be more protein damage, and hence more protein re-synthesis, in aerobic cells than in anaerobic cells. However, there seem to be no appropriate good comparative data that test these predictions [110].

The green sulfur bacterium Chlorobium tepidum can only grow photolithotrophically in the absence of oxygen and grows at, or just below, the oxycline in water bodies [21]. Although this organism is normally exposed to low fluxes of the wavelengths that it can use in photosynthesis (400–840 nm) and correspondingly low fluxes of UVB, and is generally not exposed to O$_2$, the complete genome sequence shows that the organism has a large suite of genes coding for proteins involved in DNA repair, as well as genes coding for enzymes that scavenge ROS and so might help them to protect various oxygen-damaged enzymes in photosynthesis and nitrogenase [21]. The presence of superoxide dismutase is typical of aerotolerant anaerobes and aerobes, but not generally of obligate anaerobes [113,114]. It is likely that at least some of these DNA repair and free radical scavenging enzymes are expressed in the absence of an increased threat of damage as a precautionary step, and that at least the free radical scavenging enzymes were absent before the GOE.

Turning to damage by UVB, organisms that intercept photosynthetically active radiation (PAR) (including, to a varying extent, ultraviolet A (UVA) radiation) also intercept some damaging ultraviolet radiation (UVR) (UVB + UVA). UVR is attenuated more than PAR by water and the substances dissolved in it [111,115]. UVA (320–400 nm) is damaging to some oxygenic photosynthetic organisms [116] but may be used by others in metabolic and developmental regulation. Regulation related to absorption of UVA involves the long wavelength tail of UVB pigments such as UVBR [117] and the short wavelength tail of the blue light-UVA pigments cryptochrome and phototropin [118], as well as via a specific but poorly characterized photoreceptor [119]. Furthermore, UVA can be a significant source of energy for some anaerobic photosynthetic bacteria (on the basis of bacteriochlorophyll absorption spectra) [120] and algae [121–125]. UVB (280–320 nm) screening and damage repair are required by both anaerogenic [21] and oxygenic phototrophs, with the need for additional genes (see above). Organisms at other trophic levels in habitats in which they intercept solar radiation, e.g. in grazing on, or parasitizing, photolithotrophs, also need UVB protection and repair mechanisms. Small cells have smaller package effect, so (other things being equal) are less readily protected by ultraviolet...
(UV) sunscreens intercepting UVB [126,127]. However, evidence that cell size is the dominant factor in UVB screening is equivocal [44,111], although the evolution of the smallest cyanobacterial cells well past the GOE [102], when an effective ozone screen for UVB was in place, is consistent with the predictions from photophysics [126,127]. The production of thymine dimers is a significant component of UVB damage to DNA, a potentially mutagenic process that could be decreased by a bias toward GC-rich codons rather than AT-rich codons, i.e. an increase in the GC content of the genome. However, there is evidence that there is a bias against AT codons in organisms growing in high-irradiance environments, e.g. strains of Prochlorococcus from different habitats [59,99]. A changed GC:AT ratio could conflict with other evolutionarily favoured outcomes of varied nucleotide use [128]. The conclusion is that GC:AT in DNA is not generally related to high UVB fluxes in the natural environment. The production of stratospheric O3 after the GOE, and the influx of O2 into the atmosphere, resulted in a lower flux of picoplanktonic cellular screens are absent from the late-evolving [102] shallow-water and terrestrial ecosystems [116,129]. Such extra-Scytonemin production in association with sheaths may have preceded the appearance of the ozone shield [130]. UVC is thought to have been much more intense in the Precambrian, before the appearance of the ozone shield [130]. Scytonemin production in association with sheaths may have facilitated the ecological expansion of Cyanobacteria in shallow-water and terrestrial ecosystems [116,129]. Such extra-cellular screens are absent from the late-evolving [102] picoplanktonic Prochlorococcus and marine Synechococcus. Extra-cellular screens sufficient to significantly decrease the UVB flux incident on cells would be so thick as to compromise features of these picophytoanion cells such as a minimal diffusion boundary layer that facilitates nutrient acquisition [44].

There is also no evidence that oxidative damage to DNA varies among nuclei, mitochondria and plastids [94]. Allen & Raven [131] suggested that one of the reasons in natural selection for the transfer of genes from the genome of endosymbionts, which yielded mitochondria and plastids to the host cell nucleus with subsequent loss of that gene from the organelle genome, is the greater production of free radicals in the energy-transducing organelles than in the nucleus. However, as Allen & Raven [131] point out, the rate of accumulation of nucleotide substitutions in organelle genomes can either be greater, or less, than that in the nuclear genome of the same organism, so other factors are clearly involved in gene transfer.

5. Conclusions

More genes may be needed to support the core processes of autotrophy (chemolithotrophy, photolithotrophy) than for osmochemoorganotrophy, if not for phagoorganotrophy. This agrees with the finding that, on the basis of peer-reviewed data, autotrophs have larger minimal genome sizes than chemoorganotrophs.

Small genome and cell size do not always mean faster growth rate and more rapid energy conversion; the smallest archaeal and bacterial cells with a given trophic mode grow less rapidly than cells with volumes 1–2 orders of magnitude greater. The diversion of a large proportion of resources into the photosynthetic apparatus means that phototrophic organisms tend to exhibit slower growth rates than chemooorganothrophs.

Changes in the global environment with implications for cellular energetics, such as the GOE, might be related to changes in codon usage, leading in turn to economies in C and S usage but not in N. Variations in GC:AT are not generally related to high UVB fluxes in the natural environment. UVB fluxes incident on organisms would have been greater before the GOE, and the GOE:AT effect may have been greater before the evolution of oxygenic photosynthesis and the subsequent GOE.

J.A.R. thanks John Allen, Jason Bragg, Paul Falkowski, Zoe Finkel, Kevin Flynn, Mario Giordano, Andrew Knoll, Hans Lammers, Enid MacRobbie and Antonietta Quigg for their contributions. The University of Dundee is a registered Scottish charity, No. SC015096.

References

10. Fuchs G. 2011 Alternative pathways of carbon dioxide fixation: insights into the early evolution of


42. Smith DR et al. 1997 Complete genome sequence of Methanobacterium thermoautotrophicum ΔH.
Glossary

autotroph an organism using exergonic inorganic reactions, or photos, as the energy source for growth and inorganic chemicals taken up on a molecule-by-molecule basis across the plasmalemma to supply nutrient elements.

chemolithotroph an organism using the catalysis of exergonic inorganic chemical reactions as the energy source for growth and inorganic chemicals taken up on a molecule-by-molecule basis across the plasmalemma to supply nutrient elements.

chemooorganotroph an organism using the catabolism of organic compounds as the energy source for growth, and organic compounds as the source of carbon and, in many cases, the source of nitrogen.

Genetically integrated an organism, such as a genetically integrated cyanobacterium, that uses both exogenous and endogenous energy sources to support its growth and metabolism.

heterotroph a chemoorganotrophic organism taking up organic and inorganic nutrients on a molecule-by-molecule basis across the plasmalemma as the means of acquiring carbon and other elements.

Vertically transmitted. genetic integrated
mixotroph (i) an organism combining autotrophy and chemorganotrophy; usually applied to phototrophic autotrophs. (ii) An osmochemooorganotroph that has no mechanism of autotrophic carbon dioxide fixation, but which has a photochemical apparatus that generates ion gradients and ATP that can supplement and partially replace the generation of ion gradients and ATP by catabolism of organic compounds.

phagochemoorganotroph a chemoorganotrophic organism taking up particles of live or dead organic matter as their means of acquiring carbon and other elements.

photolithotroph an organism using photons as the energy source for growth and inorganic chemicals taken up on a molecule-by-molecule basis across the plasmalemma to supply nutrient elements.

saprochemoorganotroph osmochemoorganotroph.