The Origin of Phototrophy Reveals the Importance of Priority Effects for Evolutionary Innovation

Anthony Burnetti[1],William C. Ratcliff[1]

[1]*Georgia Institute of Technology, School of Biological Sciences, Atlanta, GA, USA.*

The history of life on Earth has been shaped by a series of major evolutionary innovations. While some of these innovations occur repeatedly, some of the most important evolutionary innovations (e.g., the origin of life itself, eukaryotes, or the genetic code) are evolutionary singularities, arising just once in the history of life. This historical fact has often been interpreted to mean that singularities are particularly difficult, low-probability evolutionary events, thus making the long-term course of life on Earth highly contingent on their chance appearances. Alternatively, singularities may arise from evolutionary priority effects, where first-movers suppress subsequent independent origins. Here, we disentangle these hypotheses by examining a distinctive innovation: phototrophy. The ability to use light to generate metabolic energy evolved twice, preserving information about the origins of rare, transformative innovations that is lost when examining singular innovations. We show that the two forms of phototrophy occupy opposite ends of several key trade-offs: efficiency of light capture vs. return on investment in protein infrastructure, dependence on limiting nutrients vs. metabolic versatility, and complexity vs. simplicity. Our results suggest that the 'dual singularity' of phototrophy exists due to evolutionary interactions between nascent phototrophs, with phototrophic niche space too large for a first mover to fill all niches and fully suppress fu-

ture innovation but not so large as to support many mature innovations. While often ignored over geological time scales, ecological interactions and evolutionary priority effects may play a fundamental role in the tempo and mode of major evolutionary innovations.

Introduction

Life has been profoundly shaped by a series of evolutionary innovations. From the origin of life via prebiotic chemistry in the Hadean through to the more recent evolution of multicellular organisms, these innovations have extended the upper reaches of organismal complexity and fundamentally changed the state of the biosphere. Some critical innovations have recurred many times across the tree of life, while others have occurred just once in all of history. Given their impact on ecological and evolutionary dynamics, understanding the origin and spread of key biological innovations is fundamental to understanding history of life on Earth.

Some major evolutionary innovations have occurred many times. Multicellularity, for instance, is ubiquitous on today's Earth and has evolved at least 25 times from unicellular ancestors[1]. Perhaps even more surprisingly, complex multicellularity has evolved at least six times among the metazoans, embryophytes, red algae, brown algae, and 8-11 times in fungi[2–4]. Putative multicellular fossils are observed all the way back to 2.1-2.4 billion years ago[5,6], indicating that this innovation has a long history. Other evolutionary transitions in individuality have also evolved repeatedly in diverse lineages, including endosymbiosis and superorganismality[7], as have innovations such as C4 photosynthesis[8] and tetrapod powered flight[9]. Given their repeated evolution, none of these

innovations appear to be evolutionarily ‘difficult’.

Several of the most important innovations and transitions in the history of life, however, are those which have apparently occurred only once (Figure 1). The origin of life from abiotic chemistry is arguably the greatest evolutionary innovation in history, along with the nearly immediate origin and crystallization of the genetic code and a system of stable heredity[10]. The origin of eukaryotes via a symbiosis between an archaea and proteobacterium[11,12] was then perhaps one of the biggest singe innovation since life’s origin. Eukaryogenesis is often considered to have been highly contingent on chance events, more than any other transition. Lane et al. call it a restrictive, singular bottleneck[13] by which an extremely unlikely event (endosymbiosis of mitochondria) is a prerequisite for complex life of any kind[14]. The existence of such unique, impactful innovations has led some to conclude that the history of life on Earth is sensitive to the presence and timing of these rare events, and that most possible biospheres would therefore not possess the complexity and scale that ours does[15].

These evolutionary singularities are notoriously difficult to study[21]. They could of course represent extremely rare chance events or restrictive bottlenecks that we only see due to anthropic selection effects[15]. But they could also represent ‘frozen accidents’ by which a single lineage experienced a winner-take-all effect, deterministic necessities which could only occur one way, or could be the result of evolutionary attrition, in which a large number of original innovators were winnowed down[21]. Unfortunately, there is little information left in the modern day, hundreds of millions or billions of years after the singular event occurred, that would allow us to distinguish

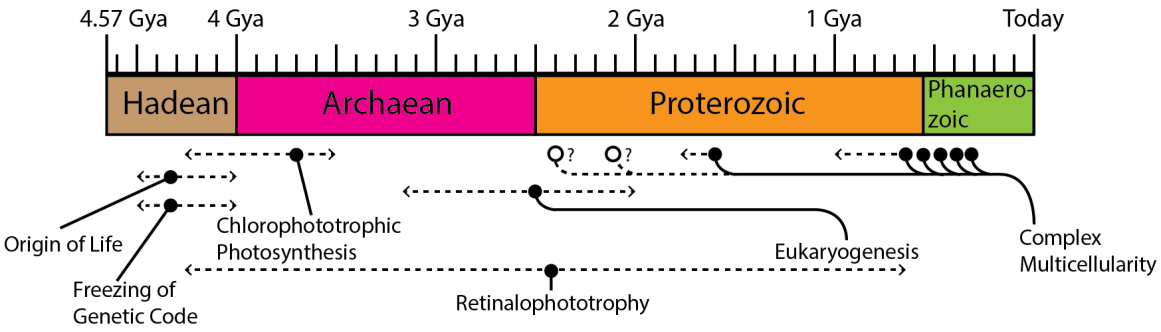

Figure 1: The history and approximate timing of major innovations and evolutionary transitions in Earth's biosphere. The origin of life, the freezing of the genetic code, and eukaryogenesis are evolutionary singularities as they were major innovations which occurred just once. Complex multicellularity has evolved at least six times across at least the last 1.6 billion years[2,3,5]. Phototrophic metabolism has evolved twice, via chlorophototrophy and retinalophototrophy. Chlorophototrophy dates to at least 3.5 billion years ago with the oldest unequivocal photosynthetic microbial mats[16,17], though some argue for older dates[18]. The origin of retinalophototrophy is uncertain due to its lack of preservation in the fossil record, and could date from anywhere between the Hadean to shortly before the rise of animals, but is more likely to be ancient[1920]. This dual singularity provides unique insight into the nature and process of evolutionary innovation.

between these mechanisms. In this paper, we circumvent these limitations by examining evolution of phototrophy (the ability to use light as an energy source), which has independently evolved twice and thus retains information about its origin that has been lost in true singularities.

The evolution of phototrophy is one of the most significant events in the history of life on Earth. It is one of the oldest evolutionary innovations discussed here, occurring at least 3.5 billion years ago[16,17] with some arguing for earlier dates[18]. The capture of light energy into metabolism allowed an enormous increase in the sheer scale of Earth's biosphere. Without the use of radiant light energy to power metabolism in phototrophs and build biomass in photosynthesizers, the only reasonable mechanism for primary production by the early biosphere was chemolithoautotrophy utilizing geologically and atmospherically produced redox couples[22,23]. This puts a low ceiling

on the potential primary production of biomass in a nonphotosynthetic biosphere (Supplemental Figure 1). Photosynthesis is thus the key factor allowing the existence of the large, high-biomass, geochemically significant modern biosphere, transforming the composition of both the atmosphere[24] and the geosphere[25] over geological time.

Unlike other biosphere-transforming innovations, the ability to use light for metabolic energy appears to have evolved independently twice, via mechanisms known as retinalophototrophy and chlorophototrophy[26]. As the only such 'dual singularity', it preserves information on the evolutionary factors underpinning the origin of rare, impactful innovations that have been lost in true singularities. By examining their properties and evolutionary histories, we find that chlorophototrophy and retinalophototrophy have precisely partitioned phototrophic niche space. They occupy opposite ends of critical trade-offs between efficiency per unit resource versus efficiency per unit infrastructure, use of rare limiting nutrients versus metabolic versatility, and complexity versus simplicity. This deep complementarity suggests that phototrophy has evolved twice because phototrophic niche space is too large for an initial first mover to fully suppress future innovation, but too small to support many separate innovations. Together, this work highlights the critical role of evolutionary priority effects in the evolution of biological innovations, and suggests that the origins of evolutionary singularities may be less constrained or contingent than is widely believed.

Chlorophototrophy

Named for the chlorophyll and bacteriochlorophyll pigments that absorb light, chlorophototrophs drive both energy metabolism and redox chemistry via light. Found in cyanobacteria and at least seven other phototrophic clades of bacteria[27,28], it is responsible for the vast majority of primary production of biomass on Earth as well as much of the energy metabolism of organisms which possess it. Approximately 9,000 teramoles of carbon are fixed by chlorophototrophs annually[29], primarily via oxygenic photosynthesis (Supplemental Figure 1).

The functional unit of the chlorophototrophic machinery is the photochemical reaction center, or RC. These large membrane-bound protein complexes are all descended from an ancestral homodimer[30], with some diversifying into heterodimers and some accumulating numerous accessory subunits[30,31]. All chlorophototrophic reaction centers push electrons to more reducing potentials via chlorophyll and bacteriochlorophyll photochemistry, either passing these electrons to electron carriers which can be used to fix biomass or energizing an electron transport chain to produce biologically available energy (Figure 2 A,B). Electrons may be pulled from elsewhere in metabolism via soluble cytochromes, or in the case of cyanobacterial Photosystem II, water itself. Chlorophyll and bacteriochlorophyll pigments are biochemically derived from porphyrins and evolutionarily related to heme, as indicated by the similarity of their biosynthesis[23,32–34]. Three central pairs of chlorophyll molecules in a transmembrane protein core represent the conserved engine of charge separation with one photo-excited chlorophyll donating an electron to another. Additional ‘antenna’ chlorophylls in each reaction center allow absorption of light with a higher cross-section per reaction

center, with energy transferred from chlorophyll to chlorophyll via resonance transfer. The mass of the conserved core reaction center is approximately 150 kilodaltons[35] and when including these integrated antennas it can reach more than 350 kilodaltons[36]. Light-gathering capacity is further enhanced by the presence a remarkably diverse array of independently-evolved pigment-bearing accessory antenna complexes[3738], which further transfer their absorbed energy into the reaction center.

Chlorophototrophy is found only in bacteria and in eukaryotes that have taken up photosynthetic cyanobacteria as plastid organelles, with no known archaeal chlorophototrophs. The distribution of chlorophototrophy within the bacteria is patchy[39], with chlorophototrophic clades scattered across the bacterial tree. Horizontal gene transfer is likely responsible for at least some of the distribution of chlorophototrophy across the tree of life with transfer positively identified into the Gemmatimonadetes, and within clades of the proteobacteria and chloroflexi[40–42]. However this process is rare at best with horizontal transfer requiring over 30 genes to move between species, and the relative importance of horizontal versus vertical transfer outside these examples is ambiguous[43].

The chlorophototrophic machinery has diversified significantly over time, with different lineages containing machinery that while operating from the same mechanistic basis has been adapted for different purposes. The deepest split in the evolutionary tree of photochemical reaction center proteins is that between type I and type II reaction centers (Figure 2 A and B). Type I reaction centers contain iron-sulfur clusters and are tuned to more reducing redox potentials, pushing electrons from

cytochromes or other soluble electron carriers to ferredoxin using light energy (Figure 2 A). Type II reaction centers are tuned to more oxidizing redox potentials, boosting electrons to membrane-soluble quinones from cytochromes (Figure 2 B) or, in the case of cyanobacterial photosystem II, directly from water. While type I reaction centers produce a highly reduced electron carrier capable of driving either carbon fixation pathways or energy metabolism, the quinone reduced by type II reaction centers cannot drive carbon fixation directly and instead can only directly drive an electron transport chain, typically consisting of a cytochrome bc complex[44,45].

Retinalophototrophy

Retinalophototrophy, the second independent origin of phototrophy, was only discovered in the 1970s via investigation of the phototrophic mechanism of haloarchaea[46]. The retinalophototrophic system is far simpler than chlorophototrophy, consisting of a single 26-28 kilodalton transmembrane protein, known as a microbial or 'type-1' rhodopsin (Figure 2 C). It is covalently bound to a single pigment molecule known as retinal, derived from the oxidative splitting of a carotenoid via a dioxygenase[47]. In a few cases, such as the xanthorhodopsins, a single additional carotenoid molecule is bound to the exterior of the protein and functions as a miniature integral 'antenna'[48].

Microbial rhodopsins directly pump protons across a cell membrane rather than engaging in redox chemistry. Light-driven isomerization of the retinal pigment pumps a single proton per absorbed photon across the membrane through the rhodopsin channel[49], meaning the system is self-contained and does not require additional electron transport chain components to extract energy.

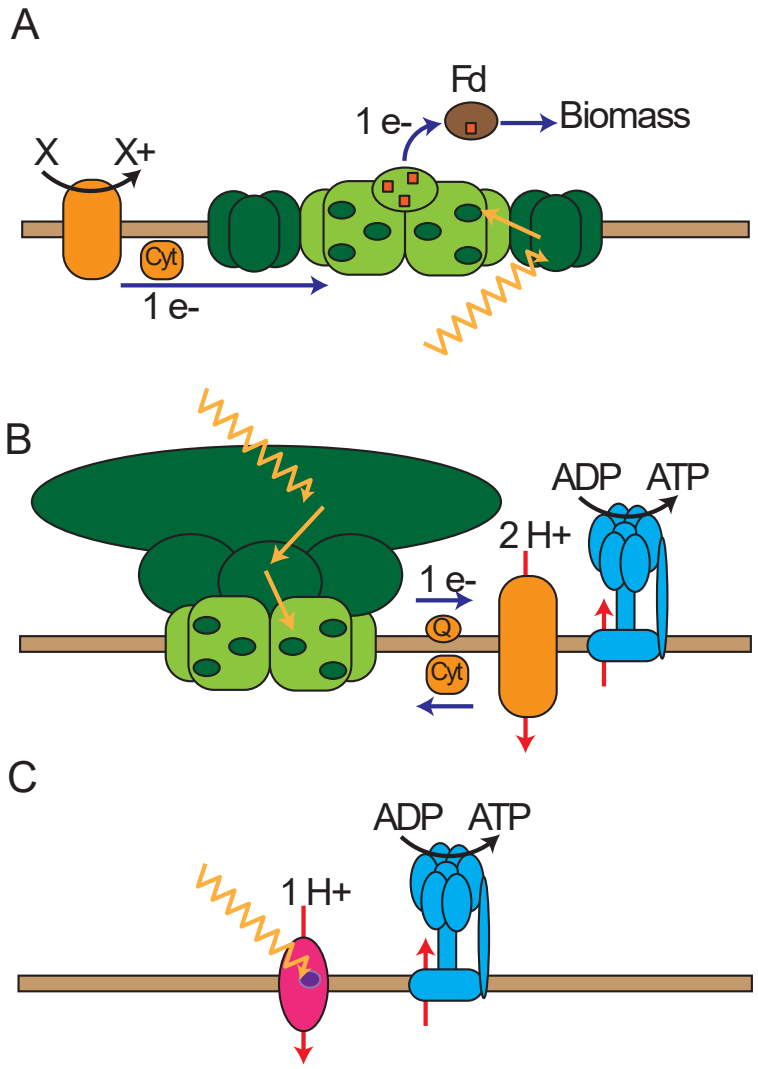

Figure 2: Simplified illustration of the three main types of phototrophic metabolism. **A)** Chlorophototrophy, type I reaction center. A photon is absorbed by one of a diverse array of antenna complexes (dark green) and passed as an exciton via Förster resonance to chlorophyll or bacteriochlorophyll molecules (small dark green spots) within the dimeric photosynthetic reaction center (light green). A type I reaction center is illustrated acquiring an electron via a cytochrome derived from an environmental reducing agent, boosting it via light energy to a low redox potential, and passing it via iron-sulfur clusters (red) to ferredoxin (brown) which can be used to build biomass via carbon and nitrogen fixation. **B)** Chlorophototrophy, type II reaction center illustrated passing electrons to a quinone electron acceptor, allowing for simple cyclic electron transfer via cytochrome bc1 (complex III) (orange) and the pumping of two protons per absorbed photon. **C)** Retinalophototrophy. A single molecule of retinal (purple) is bound to a microbial rhodopsin membrane protein (pink). Absorption of a photon causes one proton to be pumped the exterior of the cell, upon which it can participate in chemiosmotic ATP production via the membrane ATP synthase (blue).

Some rhodopsins, not directly involved in phototrophy, are also capable of pumping ions such as chloride or sodium and others function as light sensors[50]. While there are no known autotrophs able to fix biomass from CO_2 using only the energy derived from microbial rhodopsins, the energy generated by this system appears to be quite important for many photoheterotrophs. This energy can prevent starvation in marine bacteria[51,52], and is extensively used to supplement heterotrophic metabolism: the quantity of light absorbed by retinalophototrophs in the ocean is thought to be at least as large as that absorbed by chlorophototrophs[53].

The phylogenetic ubiquity of microbial rhodopsins, in contrast to the patchy distribution of chlorophototrophy, has only been fully appreciated in the last two decades. Approximately half of marine bacterial cells, from many taxa, bear diverse bacterial rhodopsin genes[54,55]. In addition to haloarchaea, they are present in marine bacteria[56], marine archaea[57], fungi[58], and heterotrophic marine eukaryotes[59–61]. They are known to acidify cellular compartments via pumping protons, and in some taxa are among the most highly expressed proteins[61], contributing significantly to the cell's energy budget. Rhodopsins have even been discovered in metagenomes of Heimdallarchaea, a member of the Asgard archaea considered a likely sister to the archaeal ancestor of eukaryotes[62], and in numerous marine viruses[63,64].

Microbial rhodopsins are exemplars of horizontal gene transfer, explaining its cosmopolitan distribution across the tree of life[65]. If a microbe contains a functional carotenoid synthesis pathway, retinalophototrophy may be transferred into the cell via a simple two-gene cassette consisting of the

rhodopsin itself and an enzyme that oxidatively cleaves a carotenoid into retinal. If no carotenoid synthesis pathway exists, a total of five genes are required, constituting a basic carotenoid synthesis pathway alongside these genes[66,67]. Gene cassettes of these types are widely observed in bacteria and archaea. Due to this ease of horizontal gene transfer, the evolutionary origin of microbial rhodopsins remains unclear.

Mechanistic comparison of phototrophic systems

The differences between chlorophototrophy and retinalophototrophy are manifold. They represent independent origins of phototrophic metabolism, derived from different metabolic cofactors shaped into photoactive pigments and representing different trade-offs and strategies in the space of possible phototrophic metabolisms (Table 1). By abstracting away from their fine details and looking at gross compositions and the products of their metabolisms (Figure 3) the major differences between them may be understood more easily.

One of the greatest differences between retinalophototrophs and chlorophototrophs is the efficiency of conversion of light energy into biologically available energy. Chlorophototrophs have a significantly higher energy yield per captured photon than retinalophototrophs. Retinalophototrophic machinery pumps one proton per photon across the cell membrane, while the chlorophototrophic machinery is capable of pumping multiple protons per photon. Most commonly two protons are pumped per photon, via a cytochome bc proton-pumping complex (related to mitochondrial complex III) or an alternative complex III passing electrons between quinones and cytochromes. Up to

Table 1: Attributes of Chlorophototrophy and Retinalophototrophy

	Chlorophototrophy	Microbial Rhodopsins
Distribution	Bacteria (and plastids)	Bacteria, Archaea, Eukarya
Active unit	$\leq$350 kDa dimeric reaction center	single $\sim$27 kDa protein
Required genes	$\sim$30	2 to 5
Mechanism	Electron transport chain	Direct proton pump
Pigment	Chlorophyll / Bacteriochlorophyll	Retinal
Pigments evolved via	Porphyrins	Carotenoids
Antenna pigments?	Diverse and abundant	One carotenoid in xanthorhodopsins
Products per cycle	One electron (or $\sim$2 protons pumped)	One proton pumped
Used for	Energy, Carbon and Nitrogen fixation	Energy
Speed	Up to 350 electrons s^{-1} RC^{-1}	Up to 50 protons s^{-1}

Comparison of ecologically and evolutionarily relevant differences between chlorophototrophic reaction-center-based and retinalophototrophic microbial-rhodopsin-based phototrophy.

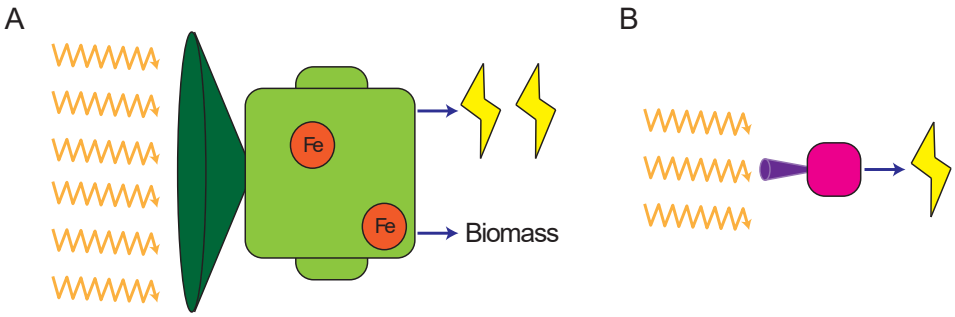

Figure 3: Schematic illustrating relevant functional differences between retinalophototrophy and chlorophototrophy. **A)** Chlorophototrophic functional units have very high absorption cross sections due to light-gathering antenna pigments (dark green), very high protein mass per functional unit (light green), use iron ions in their internal structure (orange) in addition to protein and organic pigments, and either conserve large amounts of energy per photon or are capable of contributing to biomass production. **B)** Functional units in retinalophototrophy have very small absorption cross sections (purple), little protein mass (pink), and conserve low energy per photon.

four protons per photon is possible for some fraction of electrons in oxygenic phototrophs when a type I reaction center is used with electrons passing from ferredoxin through a complex I-like NDH complex and cytochrome b_6f[68–70]. However the difference in available energy may be even greater than this ratio would indicate, as microbial rhodopsins are incapable of pumping against a membrane polarization of 200 mV[71], which is lower than the proton motive force generated by respiratory electron transport chains. The electron transport chains of chlorophototrophs are thus able to conserve more energy per proton than rhodopsins are by reaching a higher membrane voltage. Furthermore, electrons energized by some chlorophototrophic reaction centers may be passed to electron carriers such as ferredoxin and NADPH, or the high proton-motive force they generate can be used to force reverse electron flow through a respiratory electron transport chain into these carriers for carbon and nitrogen fixation. Retinalophototrophs are unable to produce biomass *de novo* using their phototrophic machinery, likely due to their low maximum proton-motive force being insufficient to allow reverse electron flow.

The material composition of the phototrophic machinery in retinalophototrophs and chlorophototrophs is also quite different. While microbial rhodopsins consist of a single 27 kDa protein molecule attached to one or two photoactive cofactors per functional unit, chlorophototrophic reaction centers consist of 2-4 core protein molecules and a number of accessory proteins per functional unit with a mass of up to 350 kDa[72], with a large number of diverse photopigments bound to each complex. Moreover, nearly every chlorophototrophic reaction center is associated with multiple diverse antenna complexes, which both greatly increases the absorption cross section per

functional unit and can bring the total protein mass per functional unit into the megadaltons or even more[38,73]. This increased absorption cross section leads to a significantly greater efficiency in terms of captured biological energy per unit incident light at low light intensities for chlorophototrophs, at the expense of saturation at relatively low light levels.

Another relevant difference between chlorophototrophs and retinalophototrophs is the requirement for iron in the chlorophototrophic machinery. While bacterial rhodopsins are entirely composed of protein and organic molecules, every known chlorophototrophic reaction center contains iron atoms. All type I reaction centers contain at least 4 in an Fe-S cluster in the core dimer, with 12 known when including accessory subunits in the case of acidobacteria chlorobi and photosystem I of cyanobacteria[36,74,75] and some may be complexed with additional integrated cytochromes bearing heme irons [76].. All type II reaction centers contain one Fe^{2+} ion bound at the interface between subunits with additional heme irons present in Photosystem II of cyanobacteria[72] and an integral cytochrome with additional heme irons present in the reaction centers of many other lineages[28,77,78]. All known electron transport chains that chlorophototrophic reaction centers participate in also utilize iron in their proton-pumping components, with 6 iron atoms present in each subunit of cyanobacterial cytochrome b_6f[79] and the NDH complex used for circular electron flow around Photosystem I containing at least twelve or possibly more[68,69]. Certain picocyanobacteria reduce their electron transport chains to a form which requires nearly only the iron atoms in the photosystems themselves (biased towards photosystem II with fewer iron atoms) and an alternative oxidase, but this comes at the expense of depressing proton yield to only one proton per photon[80].

This constitutive requirement of iron for functional chlorophototrophy but not retinalophototrophy represents a major resource limitation for chlorophototrophs, especially in oligotrophic environments such as the open ocean where iron levels are limiting[81].

Ecological niche partitioning between phototrophic pathways

At first glance the functional differences between retinalophototrophy and chlorophototrophy appear to stem entirely from their disparate evolutionary histories and compositions. However, upon closer inspection these differences appear remarkably coordinated, suggesting that the two of them have precisely partitioned the space of phototrophic ecological niches in two, each filling a different and complementary subset. This has fundamental implications for the early evolution of phototrophy, and evolutionary innovations more broadly.

Most trivially, the light-gathering pigments used by the core machinery of chlorophototrophs and retinalophototrophs are spectrally distinct. Retinal primarily absorbs the green wavelengths of visible light, while chlorophyll primarily absorbs in the red and blue wavelengths. This apparent partitioning of the electromagnetic spectrum is somewhat mitigated by the fact that chlorophototrophs contain many accessory pigments aside from basic chlorophyll which can expand their effective absorption spectrum into the green wavelengths.

Nutrient requirements differ substantially between the two systems. Retinalophototrophy solely uses a small protein and an organic pigment to pump protons while chlorophototrophic reaction

centers contain iron and are dependent upon functionally-coupled electron transport chain components that also require iron. Retainalotrophy is thus favored under low-iron conditions, which are pervasive throughout much of the oceans[82]. Indeed, it appears that up to 50% of individual bacterial cells present in the oligotrophic open ocean express microbial rhodopsins[54], declining in frequency in more nutrient rich environments[53,83]. This iron-dependent niche partitioning is illustrated well in polar diatoms, eukaryotic phototrophs utilizing both oxygenic photosynthesis and proton-pumping rhodopsins: rhosopsin expression is sharply upregulated during iron starvation in a homeostatic response to maintain energy metabolism[84,85]. Iron limitation thus favors retinalophototrophy, even though it is significantly less efficient per unit photon absorbed.

The greater efficiency per unit light intercepted by chlorophototrophs compared to retinalophotrophs (along with their capacity to perform redox reactions and thereby directly fix carbon) would, at first glance, imply they are strictly superior under all circumstances without iron limitation. However, efficiency of energy capture per absorbed photon represents just one element of a complex ecological and biophysical trade-off. All metabolic machinery carries with it an investment cost in protein mass - the infrastructure must be built before it can transduce energy or nutrients, and has a finite lifetime before being either recycled or diluted away by growth and division. As such, every metabolic pathway also has a rate of return on investment. In order to determine the return on investment available from multiple phototrophic pathways, one must take into account the mass per functional unit, the rate of operation of the protein machinery, and the yield per cycle, yielding a specific energy flux per unit protein mass.

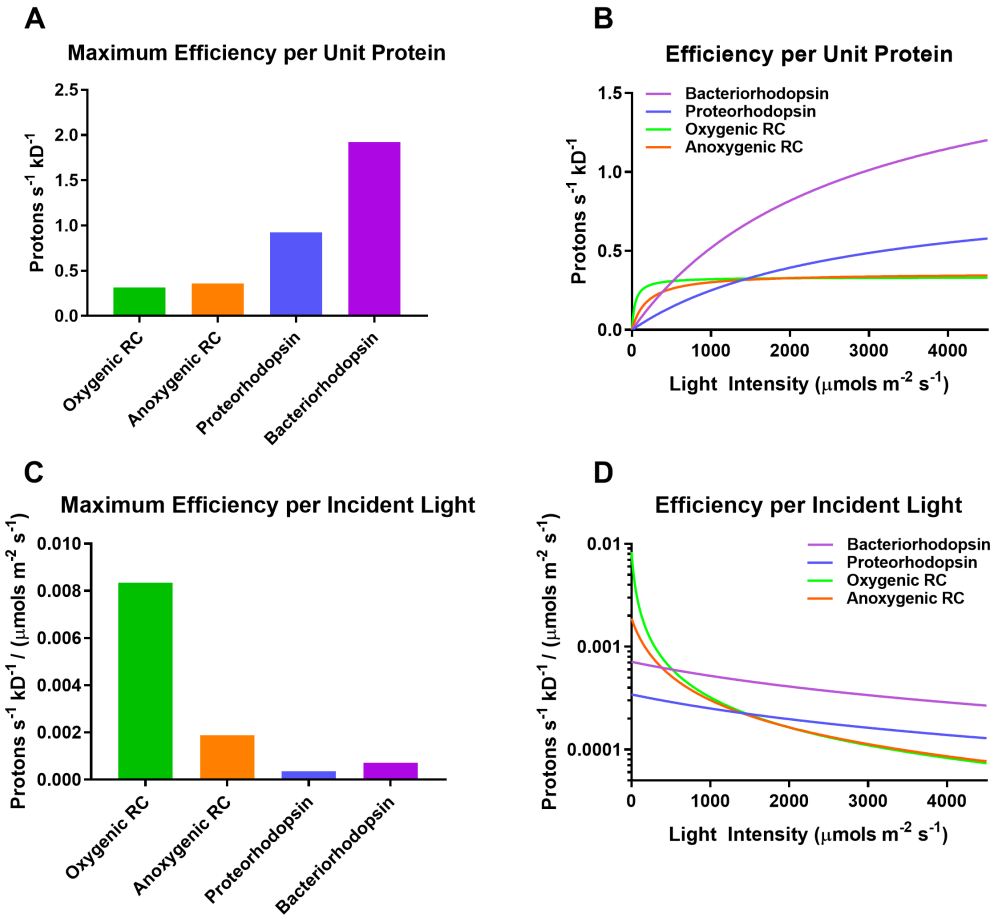

Figure 4: Ecological comparison between chlorophototrophy and retinalophototrophy. **A)** Calculated maximum pumped proton flux available per kDa of protein mass of anoxygenic purple bacterial reaction centers, oxygenic reaction centers, proteorhodopsin, and bacteriorhodopsin at saturating levels of light. Microbial rhodopsins saturate at much higher specific metabolic energy fluxes than reaction centers. **B)** Calculated proton flux available per kilodalton of mass of different phototrophic systems at different light intensities. Chlorophototrophic reaction centers produce more energy flux at low light levels compared to microbial rhodopsins, but saturate quickly, while microbial rhodopsins function best at high light levels with higher specific metabolic energy fluxes. **C)** Calculated maximum pumped proton flux available per kDa per unit incident light in microeinsteins per square meter. Chlorophototrophic reaction centers are capable of extracting much more energy flux per unit incident light. **D)** Calculated energy flux per kilodalton of machinery per microeinstein of incident light. Chlorophototrophic reaction centers are significantly more efficient per unit incident light when light is scarce, but rapidly saturate due to having large absorption cross sections per reaction center, thereby gathering more light than can be converted to energy. Microbial rhodopsins, on the other hand, are significantly less efficient per photon, but use light more efficiently than chlorophototrophy when light levels are high. See Supplement S1 for calculations.

Table 2: Chlorophototrophic and Retinalophototrophic Energy Flux Per Unit Mass

	Anoxygenic RC	Oxygenic RC	Proteorhodopsin	Bacteriorhodopsin
Total Protein mass / RC	∼835 kDa	∼2098 kDa	27 kDa	26 kDa
Electrons s^{-1}	∼150	∼350	0	0
Protons s^{-1}	∼300	∼700	∼25	>50
Protons s^{-1} kDa^{-1}	0.36	0.33	0.92	>1.92
Normalized Protons s^{-1} kDa^{-1}	1.08	1	2.78	>5.76

Proton flux per unit protein mass available to different phototrophic machineries at full light saturation. See supplement S1 for calculations.

We calculated this effective energy flux per unit investment of different phototrophic systems based on a literature review of these values for anoxygenic chlorophototrophic RCs, oxygenic RCs, and two different microbial rhodopsins (proteorhodopsin and bacteriorhodopsin)[31,36,46,56,70,71,73,86–97]. We quantified the effective flux in terms of protons pumped per kilodalton per second at saturating light levels (See Table 2 and Supplement S1). Despite their higher efficiency per photon absorbed and faster photocycle, chlorophototrophic machinery is so much more massive than microbial rhodopsins that their specific energy flux per unit mass is significantly lower. Proteorhodopsin and bacteriorhodopsin are calculated as 2.78-fold and 5.76-fold more efficient per unit investment than oxygenic RCs respectively, with anoxygenic RCs roughly equivalent to oxygenic RCs (Table 2 Figure 4 A, and Supplement S1).

A primary reason for low energy flux per unit investment in chlorophototrophic machinery is the presence of large antenna pigments which feed absorbed light energy into reaction centers. This means that the effective absorption cross section per functional chlorophototrophic unit is much larger than the cross section per retinalophototrophic unit. The relative performance of these

systems thus varies drastically according to ambient light intensity (although the absorption cross section per kilodalton of machinery is very similar between the two - see Supplemental Figure 2). When calculating the effect of these differences between absorption cross sections and saturation of the phototrophic machinery at varied light levels (see Supplement S1), we find that this greater return per unit investment for retinalophototrophs only manifests at high light (Figure 4 B). The small protein mass and presence of only a single retinal pigment in a microbial rhodopsin ensures a small cross section which requires intense ambient light for the machinery to be used effectively. Conversely, the large absorption cross section available to the massive chlorophototrophic system is nearly saturated above low light levels of less than 500 microeinsteins per square meter per second, but at these lower light levels maintains a higher energy flux per unit infrastructure. By dividing the function of the return per unit investment of each phototrophic system by the level of ambient light, we produced functions of the efficiency per unit ambient light in units of protons pumped per kilodalton per second, per microeinstein of light per square meter per second (Figure 4 C, D). While chlorophototrophic reaction centers are more efficient per unit ambient light in the limit of low light, this is reversed at higher light levels.

The differences between chlorophototrophy and retinalophototrophy stem from an intrinsic biophysical trade-off. It is not possible to build a phototrophic system that has both high metabolic efficiency per unit investment (protein infrastructure), and high metabolic efficiency per unit of a rare limiting resource (ambient light). Chlorophototrophy is efficient per unit light at low light levels and requires large amounts of protein investment, while in high light levels retinalophototro-

phy produces higher energy flux at lower levels of investment. This fits the observed physical distribution of phototrophs in the ocean and ecological distribution of these pathways. Retinalophototrophs are observed at their highest levels in surface ocean waters with high light levels, while chlorophototrophs become most common at slightly deeper levels of the ocean at which light has been partially absorbed[53]. Chlorophototrophy requires a significant fraction of the proteome to be invested to result in an effective energy flux and is the only phototrophic pathway observed in obligate phototrophs[39,45]. Retinalophototrophy is observed in fully 50% of bacteria in the open ocean and is frequently present in heterotrophs[54,96,98], which appear to frequently use it as a backstop to prevent starvation and increase biomass yield of heterotrophic metabolism[51–53].

A similar trade-off is observed across the diversity of heterotrophic metabolic machineries. The difference between respiration and fermentation itself is an example - respiration can produce several times the ATP per unit substrate consumed while producing less than half the energy flux per unit protein mass[99]. The two most common glycolytic pathways - the Etner-Doudoroff (ED) and Embden-Meyerhof-Parnas (EMP) pathway - share precisely this relationship as well. The EMP pathway produces twice the ATP per unit carbohydrate consumed as the ED pathway, but requires 5-fold more protein mass, and thus produces approximately 40% the energy flux per kilodalton of protein[100]. Just as chlorophototrophic pathways use more protein then retinalophototrophy to acquire more energy from a small quantity of light and are seen in obligate phototrophs, the EMP pathway is seen more frequently than the ED pathway in obligate anaerobes which cannot switch to aerobic respiration and must obtain more energy from their limited available substrate.

Differences in the cost of metabolic machinery have major implications for growth and ecology. The larger the fraction of a cell's proteome must be put towards the generation and maintenance of energy and resources, the smaller the fraction of the proteome can go towards growth and development[101]. This leads to a series of 'growth laws'[102,103] which dictate that, all else being equal, a larger investment of protein being used to efficiently consume a rare resource leads to a slower growth rate due to less investment in ribosomes and anabolic functions. The optimal allocation of costly metabolic enzymes under situations of differing growth rate and resource availability therefore explains much of the long-observed trade-off between microbial growth rate and biomass yield[99,101,103,104]. Rapid growth and low yield occurs on abundant resources, while slow growth and high yield occurs on scarce resources. Only recently has optimal proteome allocation been analyzed in the context of phototrophy and autotrophy in general[105–107], but the principles are identical when ambient light is treated as a metabolic resource.

Thus, chlorophototrophy and retinalophototrophy have partitioned phototrophic niche space. Chlorophototrophy is a high-investment strategy suitable for environments of low growth rate, low ambient light resources, or for specialists investing heavily in a single pathway in any environment (i.e., obligate photoautotrophs). Retinalophototrophy is a low-investment strategy suitable for situations of higher growth rate, high ambient light resources, or for flexible metabolic generalists capable of using either phototrophy or heterotrophic metabolism. Taken together with other divergent properties (Table 1), including chlorophototrophy's requirement of limiting iron and the ease of horizontal transfer of retinalophototrophic capacity over evolutionary time, the properties of

these two phototrophic pathways are strikingly complementary.

Ecological interference, evolutionary priority effects, and major evolutionary innovations

The complementary nature of Earth's two phototrophic systems suggests that their properties have co-evolved, rather than their properties being independent of each other. In particular, we propose that the evolution of phototrophy has been shaped by the phenomenon of evolutionary priority effects. Much like an ecological priority effect in which the first organisms to colonize a habitat become difficult to displace[108,109], an evolutionary priority effect is a process by which a poorly-adapted newcomer evolves into a new ecological niche, suppressing the evolution of similar newcomers which could fill the same niche[109–111].

Each extreme of the efficiency per unit investment / efficiency per unit light trade-off represents a different emergent phototrophic niche. The set of optimal machineries for a given situation represents a Pareto front on a graph of these two variables against each other[112,113] (Figure 5). Evolution optimizes phototrophic systems towards this front but once it is reached increasing efficiency along one axis requires decreasing it along the other axis, meaning that a mature phototrophic machinery is constrained to evolve along this front. Critically, architectural limitations may have prevented a single phototrophic ancestor from diversifying sufficiently to fill all phototrophic niches along this Pareto front. Microbial rhodopsins are a small, light-driven proton pump driven by isomerization of a single small molecule, which only allows a single proton to be pumped per photocycle[49]. It would be difficult, if not impossible, for it to be reworked into a more efficient form without a complete

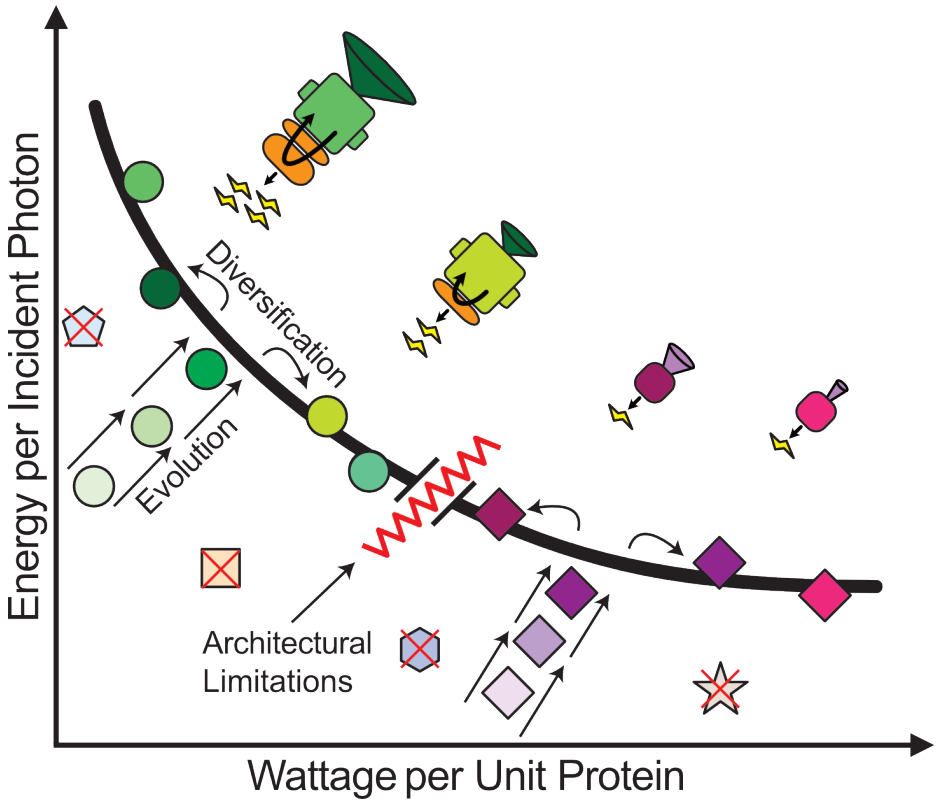

Figure 5: Schematic illustration of hypothesized evolutionary history of phototrophic metabolism on Earth. All modern chlorophototrophs (dark green circles) and retinalophototrophs (dark purple diamonds) lay roughly along a 'Pareto front' representing a trade-off between energy captured per incident photon and wattage available per unit phototrophic infrastructure. Representative differences are illustrated above the curve, with differences in antenna pigments (green and purple funnels) and electron transport chains (orange oval accessory components) contributing to differences in energy flux and energy yield within each class of phototrophs. Early chlorophototrophs (light green circles) and retinalophototrophs (light purple diamonds) lay far away from this Pareto front, and rapidly evolved towards it, subsequently diversifying along the front (arrows). An architectural limitation (red line) prevented whichever evolved first from diversifying to fill all positions on the trade-off curve Pareto front, allowing a second novel phototroph sufficiently different to evolve and fill the rest of the Pareto front. Each phototrophic metabolism suppresses the evolution of novel unrelated phototrophic pathways that are ecologically similar to it but strictly inferior in their initial, unoptimized forms (red Xs).

restructuring. Without any redox-active cofactors in its structure, it cannot be recruited to interact with electron transport chains or redox metabolism. Rhodopsin thus appears to be incapable of evolving to pump more than one proton per photon and efficiently using available light resources, although its small mass means it enjoys a high maximum energy flux per unit mass.

Conversely, the mass of the core machinery of the chlorophototrophic reaction center appears to be constrained, such that it cannot be reduced below a relatively large minimum size. While proteobacterial type II RCs have either lost or never acquired the integrated antenna domains common to other RCs[43], the core catalytic subunit appears to never mass under approximately 150 kilodaltons[90] or contain fewer than a minimum of eight cofactor molecules[90,114]. This minimal unit likely cannot be shrunk further while retaining its function in redox metabolism, limiting its maximum energy flux per unit mass even as it enjoys a high efficiency per unit light captured.

In our model of evolutionary priority effects, whichever pathway evolved first would have been unable to fill all available phototrophic ecological niches. Chlorophototrophy and retinalophototrophy would be architecturally limited to one or another end of the tradeoff between energy flux per unit protein investment and energy flux per unit light. Once either phototrophic pathway had diversified, it would engender an evolutionary priority effect preventing other similar rudimentary phototrophic systems from becoming established (Figure 5). However, it would have been unable to suppress the evolution of a phototrophic pathway sufficiently distinct on one of the key trade-off axes. This new system would then have been able to fill the remaining ecological niches left vacant

by the first system, suppressing the subsequent evolution of phototrophy and resulting in the dual singularity we observe today.

If this model is correct, the fact that a second ecologically-complementary phototrophic pathway evolved suggests that the origin of novel phototrophic systems is not necessarily a low-probability, evolutionarily-difficult innovation. Instead, it suggests that early forays into phototrophy may have occurred many times in the history of life. All but two of these novel, unoptimized pathways would simply have been driven to extinction by competition with the well-adapted first movers.

More generally, our results imply that the evolution of singular innovations may be less difficult than they appear. Easily accessible innovations can be preserved for long periods of time as apparent singularities or near-singularities when evolutionary priority effects strongly inhibit subsequent innovation. The extent to which evolutionary priority effects can constrain subsequent innovation depends on the underlying niche structure. In the absence of either competition or evolutionary priority effects, innovations are not suppressed and are free to evolve repeatedly. Multicellularity, for example, has evolved many times[1], allowing for fundamentally different multicellular life history strategies to evolve in different lineages (as in Figure 6A). The evolution of fungi does not constrain the evolution of plants or animals, for instance. In contrast, singularities are expected when there is a singular niche and no strict architectural limitations, like those which have emergently prevented chlorophototrophs or retinalophototrophs from evolving to dominate all phototrophic niches (Figure 6B). Life itself, an ancient singularity, may in a sense occupy a single, broad niche,

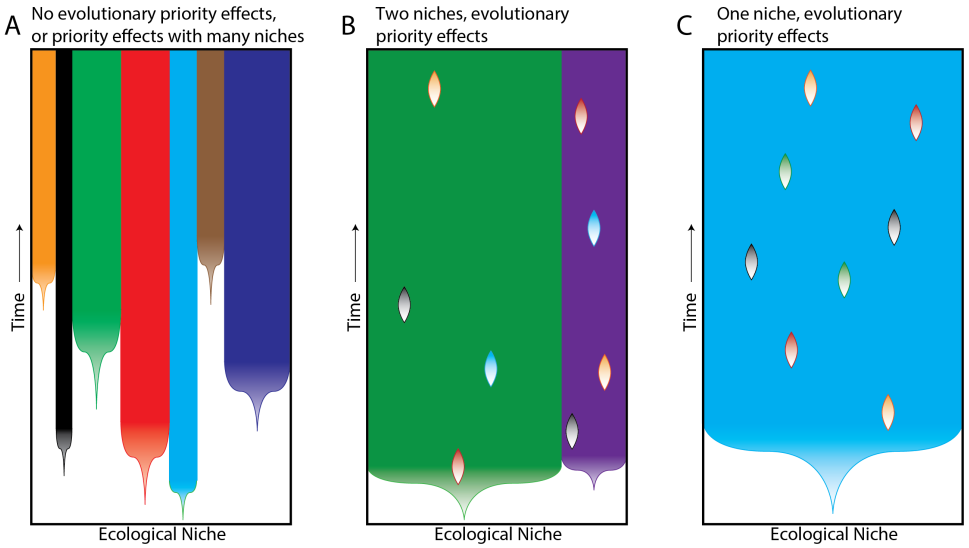

Figure 6: Evolutionary priority effects and their impact on major evolutionary innovation. **A)** The filling of ecological niches in a system with low evolutionary priority effects. A large number of separate innovations fill separate niches. This model fits the multiple origins of complex multicellularity. **B)** The hypothesized impact of evolutionary priority effects on the evolution of phototrophy, perpetuating a 'dual singularity'. Two innovations (chlorophototrophy and retinalophototrophy, green and purple) fill two unbridgeable niches and evolve to stability (darkening shades), while additional innovations which could exploit these niches continually evolve (other colors) but are too evolutionarily immature (light shades) to compete with established players. **C)** Hypothesized circumstance of evolutionary priority effects maintaining the appearance of an 'evolutionary singularity' - a single innovation can fill all available niches of a given type, and continually outcompetes novel evolutionarily young innovations. The singular origin of life and of eukaryogenesis could be represented by this model.

391 in which powerful evolutionary priority effects suppress secondary origins of inefficient and simple

392 novel replicators and protocells (as in Figure 6C). The evolution of eukaryotes, another singularity

393 of profound importance, could represent a similar case of a newcomer inventing a transformative

394 capability - most likely phagocytosis or other capacities for complex and flexible cell morphology[115]

395 - and subsequently suppressing secondary origins.

Evidence for ancient suppressed major innovations

Understanding the origins of evolutionary innovations that occurred billions of years ago poses considerable challenges. Surviving phototrophic pathways do not bear direct evidence of evolutionary priority effects in their structures, but instead only in their relationship to each other. The extinction of prospective newcomers after niches are filled makes it difficult to directly test the hypothesis that additional phototrophic pathways could have evolved but have been suppressed through competition. However, direct evidence of independently originating light-harvesting pathways which have not been refined into niche-defining and biosphere-changing metabolic pathways may survive to the present day in the form of light-driven processes that are not involved directly in energy metabolism. Specifically, they may be preserved if they are co-opted for some ancillary purpose unconnected to any particular phototrophic niche, and thus provide a selective advantage while not competing with entrenched phototrophic metabolisms. Such preserved pathways would likely be comparatively unoptimized, using generic cofactors rather than dedicated pigment molecules, as they have not been subject to the strong selection present when a significant metabolic flux passes through a specialized pathway.

Two modern pathways are of particular interest in their similarity to this template. The most well-studied is DNA repair mediated by photolyases, using light energy to repair pyrimidine dimers created by ultraviolet radiation[116]. A second, more recently discovered class of proteins known as fatty acid photodecarboxylases also use light energy for the production of hydrocarbon oils in algae in a very similar way[117]. Their mechanism of operation resembles that of chlorophototrophic

reaction centers, while being composed of non-homologous components. The active site of both of these proteins contains FAD - a redox cofactor, which happens to absorb and interact with light incidental to its main function as an electron carrier due to its large set of fused aromatic rings. Held nearby in the enzyme is an additional molecule of MTHF (5,10-methenyltetrahydrofolate, a cofactor involved in methyl group metabolism) or 8-HDF (8-hydroxy-5-deazaflavin, a molecule related to other flavins), both of which also happen to be incidentally photoactive. But in photolyases and fatty acid decarboxylases, rather than performing any methyl-group chemistry or redox chemistry, the large absorption cross section for visible light of both of these is instead exploited. They function as an antenna pigment, absorbing photons that would not be absorbed by FAD, and transferring this energy into the FAD and exciting it[116,118]. Once excited, FAD then transfers an electron to the substrate and reduces it, transiently becoming a radical stabilized by the apoprotein and a very strong oxidizing agent[116–118]. This electron triggers a rearrangement of bonds in the substrate, repairing a pyrimidine dimer or decarboxylating a fatty acid, before returning to the oxidized FAD in a form of localized circular electron flow.

Thus, photolyases and fatty acid photodecarboxylases contain repurposed ordinary metabolic and redox cofactors which happen to be photoactive independent of their primary functions. They contain cofactors acting as antenna pigments, transferring excitations into redox-active cofactors at active sites of proteins, analogous to the light-gathering chlorophylls and central redox-active chlorophylls of chlorophototrophic reaction centers. Both drive a form of circular electron flow, much as chlorophototrophic reaction centers drive circular electron trasport chains to capture

biological energy. Their similarities to photosynthetic machinery have been noticed by numerous authors in the astrobiological literature, and used as proof of concept for alternate phototrophic metabolisms that never came to be on Earth[118,119]. It is not difficult to imagine how these light-transducing systems could become more optimized over evolutionary time, with customization of FAD and MTHF into dedicated phototrophic pigments and their electron flow being directed into electron transport chains and redox metabolism for carbon fixation. And yet they never were, and we instead observe this unique bit of photochemistry pushed to the margins of metabolism, directly driving DNA repair and other metabolic reactions that require a small, local circular electron flow rather than a simple reducing or oxidizing agent. These two apparently independent reactions are precisely what would be expected of a remnant of separate origins of prospective phototrophic metabolism if it were to survive by being applied to a purpose independent of phototrophy, and provide evidence for the existence of an alternate evolvable phototrophic metabolism that has been driven to the margins by incumbents.

These ancillary photoactive pathways may recapitulate something of the nature of the earliest phototrophic systems, before they were optimized and became capable of suppressing newcomers. While the nature of the earliest retinalophototrophic machinery is somewhat mysterious, as the small protein's limited homology with anything except eukaryotic G-protein coupled receptors and sensory rhodopsins restricts inferences of their early evolution[20,120], the nature of the earliest precursors of chlorophototrophy are relatively well constrained. Several main structural attributes of the last common ancestor of the reaction center complex can be inferred[30,121–123]. All reaction centers

contain three central pairs of carefully coordinated chlorophyll or bacteriochlorophyll molecules that when excited are able to trigger electron transfer, as well as additional pigment molecules both in the reaction center itself and in associated antenna pigments transferring their energy to the redox-active catalytic center. This combination of a redox-active central complex plus antennas seems to be ancient. It has also been clear for decades that chlorophyll is evolutionarily related to tetrapyrroles such as heme[23,32,33]. Heme is a redox cofactor, binding iron to carry electrons through electron transport chains or perform catalysis in enzymes (its role in binding oxygen in animal globins being a late, derived function). Thus, the chlorophototrophic system is likely ultimately derived from a modified respiratory electron transport chain component[23]. Heme already is mildly photoactive much like flavins and MTHF. It absorbs ultraviolet and short-wavelength visible light, though this usually results in destruction of the molecule[124,125]. Synthesis of chlorophyll involves the modification of the common tetrapyrrole backbone and insertion of a different bound ion which tunes its absorption features further into the visible light range and its available excited states into those which can reversibly transfer electrons. Many scenarios have been proposed for the details the evolution of this process[23]. Ultimately heme or another tetrapyrrole was likely optimized into a dedicated pigment over evolutionary time, in the context of an electron transport chain driven by external redox couples that came to be able to rely on internal generation of redox power.

Marginalized secondary origins of major evolutionary innovations may not be unique to the evolution of phototrophy. Eukaryogenesis has long been considered an exemplar of an evolutionary singularity, with nothing remotely similar to eukaryogenesis having occurred a second time.

However, with the discovery of the Asgard archaea, a clade bearing numerous proteins previously believed to be specific to eukaryotes[126] including functional cytoskeletal components[127] and even SNARE proteins associated with endomembrane systems[128], the uniqueness of eukaryotes has been cast into doubt. The significance of these archaea, most of which have never been imaged and are entirely inferred from metagenomic samples, is unclear given their apparent close evolutionary affinity with eukaryotes. Recently, an even more striking example of a separate invention of complex cellular architecture has been discovered - a subset of planctomycete bacteria which possesses a phagotrophic lifestyle, consuming other bacteria in a manner previously thought to be unique to eukaryotes[129]. This bacterial group has long been known for large size and an unusually complex if poorly understood cellular architecture[130,131] and has convergently evolved a proteome uniquely rich in gene duplications and large multidomain proteins for a bacterium[132], qualitatively similar to those of simple eukaryotes. Does the evolution of this phagotrophic bacterium represent an independent invention of complex cell architecture and increased genetic complexity analogous to eukaryogenesis, and if so what could have protected it from interference from incumbent eukaryotes in its unique ecological niche? One possibility is metabolic niche partitioning. Planctomycete bacteria are known for having remarkably specialized metabolisms, including the anammox reaction[133]. Eukaryotes are notoriously limited in their metabolic repertoire compared to bacteria, instead relying on size and morphological complexity for adaptation to new niches[115]. One possibility is that a highly specialized metabolic niche could have protected this lineage from competition with eukaryotes, allowing them to evade suppression by evolutionary priority effects and evolve eukaryote-like cellular properties.

Conclusion

Phototrophy is among the most important innovations in the history of life, fundamentally changing the biosphere. It is unique among major biological innovations in that it has evolved not once, and not many times- it arose precisely twice. Here we show that the two origins of phototrophy are mechanistically and ecologically complementary, having partitioned phototrophic niche space along a set of trade-offs that prevent either mechanism from becoming dominant. Under low-light conditions, chlorophorophy captures energy from light more efficiently than retinalophototrophy, but saturates more quickly, becoming less efficient at high irradiance. Chlorophototrophy requires a cell to construct a large iron-containing reaction center, making it less efficient in terms of energy per unit protein, and susceptible to inhibition under oligotrophic conditions.

Architectural limitations inherent to each type of phototrophy appear to prevent either chlorophototrophs or retainalophototrophs from occupying the entire niche space for phototrophs, creating the opportunity for the stable coexistance of both pathways. The fact that phototrophy has evolved just two times over the past 3.5 billion years, particularly given the existence of alternative pathways capable of generating energy from light, suggests that additional independent origins have been suppressed by evolutionary priority effects. We are not the first to argue that priority effects could lead to evolutionary singularities[21], but until now, it has been impossible to disentangle this hypothesis from the possibility that most impactful singularities exist because they are rare and evolutionarily difficult.

The origin of major evolutionary innovations cannot be understood outside of their ecological and evolutionary contexts. Fundamental questions still remain unresolved: how pervasive is competitive suppression due to evolutionary priority effects? What determines an innovation's niche structure? Within the same niche, does competitive exclusion always occur? Three lines of future work stand to be especially informative: First, we should integrate theoretical and empirical approaches to understand the conditions under which evolutionary priority effects constrain innovation. Second, we should make use of Earth's natural experiments, comparing the innovations that have occurred repeatedly (e.g., multicellularity, super-organismality, C4 photosynthesis) to those that have occurred just once or twice (e.g., phototrophy, eukaryogenesis). Finally, we should search for undiscovered vestiges of independent innovations that have survived either by alleviating evolutionary priority effects or by having their function modified to avoid competition with the 'primary' innovation. Together, this work stands to provide significant insight into the nature of evolutionary innovations and the origin of complex life.

1. Herron, M. D., Rashidi, A., Shelton, D. E. & Driscoll, W. W. Cellular differentiation and individuality in the 'minor' multicellular taxa. *Biological Reviews* **88**, 844–861 (2013).

2. Knoll, A. H. The Multiple Origins of Complex Multicellularity. *Annual Review of Earth and Planetary Sciences* **39**, 217–239 (2011). URL `http://www.annualreviews.org/doi/10.1146/annurev.earth.031208.100209`.

3. Nagy, L. G., Kovács, G. M. & Krizsán, K. Complex multicellularity in fungi: evolutionary convergence, single origin, or both? *Biol. Rev* **93**, 1778–1794 (2018).

4. Bengtson, S., Sallstedt, T., Belivanova, V. & Whitehouse, M. Three-dimensional preservation of cellular and subcellular structures suggests 1.6 billion-year-old crown-group red algae. *PLoS Biology* **15**, 1–38 (2017). URL http://dx.doi.org/10.1371/journal.pbio.2000735.

5. Bengtson, S. *et al.* Fungus-like mycelial fossils in 2.4-billion-year-old vesicular basalt. *Nature Ecology and Evolution* **1**, 1–6 (2017). URL http://dx.doi.org/10.1038/s41559-017-0141.

6. El Albani, A. *et al.* The 2.1 Ga old Francevillian biota: Biogenicity, taphonomy and biodiversity. *PLoS ONE* **9** (2014).

7. Smith, J. M. & Szathmary, E. *The major transitions in evolution* (Oxford University Press, 1997).

8. Sage, R. F. The evolution of c4 photosynthesis. *New phytologist* **161**, 341–370 (2004).

9. Alexander, D. E. *On the wing: Insects, pterosaurs, birds, bats and the evolution of animal flight* (Oxford University Press, USA, 2015).

10. Vetsigian, K., Woese, C. & Goldenfeld, N. Collective evolution and the genetic code. *Proc Natl Acad Sci U S A* **103**, 10696–701 %U internal–pdf://vetsigian et al pnas 2 (2006). URL http://www.ncbi.nlm.nih.gov/pubmed/16818880.

11. Esser, C. *et al.* A genome phylogeny for mitochondria among α-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Molecular Biology and Evolution* **21**, 1643–1660 (2004).

12. Martijn, J., Vosseberg, J., Guy, L., Offre, P. & Ettema, T. J. Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature* **557**, 101–105 (2018).

13. Lane, N. Serial endosymbiosis or singular event at the origin of eukaryotes? *Journal of Theoretical Biology* **0**, 1–10 (2017).

14. Lane, N. Bioenergetic constraints on the evolution of complex life. *Cold Spring Harb Perspect Biol* **6**, a015982 (2014).

15. Monod, J. *On Chance and Necessity*, 357–375 (Macmillan Education UK, London, 1974). URL https://doi.org/10.1007/978-1-349-01892-5_20.

16. Walter, M., Buick, R. & Dunlop, J. Stromatolites 3,400–3,500 myr old from the north pole area, western australia. *Nature* **284**, 443–445 (1980).

17. Van Kranendonk, M. J., Philippot, P., Lepot, K., Bodorkos, S. & Pirajno, F. Geological setting of earth's oldest fossils in the ca. 3.5 ga dresser formation, pilbara craton, western australia. *Precambrian Research* **167**, 93–124 (2008).

18. Nutman, A. P., Bennett, V. C., Friend, C. R., Van Kranendonk, M. J. & Chivas, A. R. Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature* **537**, 535–538 (2016).

19. DasSarma, S. & Schwieterman, E. W. Early evolution of purple retinal pigments on earth and implications for exoplanet biosignatures. *International Journal of Astrobiology* 1–10 (2018).

20. Shen, L., Chen, C., Zheng, H. & Jin, L. The evolutionary relationship between microbial rhodopsins and metazoan rhodopsins. *The Scientific World Journal* **2013** (2013).

21. De Duve, C. *Singularities: landmarks on the pathways of life* (Cambridge University Press, 2005).

22. Decker, K., Jungermann, K. & Thauer, R. Energy production in anaerobic organisms. *Angewandte Chemie International Edition in English* **9**, 138–158 (1970).

23. Martin, W. F., Bryant, D. A. & Beatty, J. T. A physiological perspective on the origin and evolution of photosynthesis. *FEMS microbiology reviews* **42**, 205–231 (2018).

24. Lyons, T. W., Reinhard, C. T. & Planavsky, N. J. The rise of oxygen in earth's early ocean and atmosphere. *Nature* **506**, 307–315 (2014).

25. Hazen, R. M. *et al.* Mineral evolution. *American Mineralogist* **93**, 1693–1720 (2008).

26. Bryant, D. A. & Frigaard, N.-U. Prokaryotic photosynthesis and phototrophy illuminated. *Trends in microbiology* **14**, 488–496 (2006).

27. Thiel, V., Tank, M. & Bryant, D. A. Diversity of Chlorophototrophic Bacteria Revealed in the Omics Era. *Annual Review of Plant Biology* **69**, annurev–arplant–042817–040500 (2018). URL `http://www.annualreviews.org/doi/10.1146/annurev-arplant-042817-040500`.

592 28. Ward, L. M., Cardona, T. & Holland-Moritz, H. Evolutionary implications of anoxygenic
593 phototrophy in the bacterial phylum candidatus eremiobacterota (wps-2). *Frontiers in micro-*
594 *biology* **10**, 1658 (2019).

595 29. Field, C. B. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic
596 Components. *Science* **281**, 237–240 (1998). URL http://www.sciencemag.org/
597 cgi/doi/10.1126/science.281.5374.237.1011.1669.

598 30. Cardona, T. A fresh look at the evolution and diversification of photochemical reaction centers.
599 *Photosynthesis Research* **126**, 111–134 (2015). URL https://doi.org/10.1007/
600 s11120-014-0065-x.

601 31. Umena, Y., Kawakami, K., Shen, J. R. & Kamiya, N. Crystal structure of oxygen-evolving
602 photosystem II at a resolution of 1.9Å. *Nature* **473**, 55–60 (2011). URL http://dx.doi.
603 org/10.1038/nature09913.

604 32. GRANICK, S. Evolution of heme and chlorophyll. In Bryson, V. & Vogel, H. J. (eds.)
605 *Evolving Genes and Proteins*, 67 – 88 (Academic Press, 1965). URL http://www.
606 sciencedirect.com/science/article/pii/B9781483227344500140.

607 33. Chew, A. G. M. & Bryant, D. A. Chlorophyll biosynthesis in bacteria: The origins of structural
608 and functional diversity. *Annual Review of Microbiology* **61**, 113–129 (2007). URL https:
609 //doi.org/10.1146/annurev.micro.61.080706.093242. PMID: 17506685.

610 34. Bryant, D. A., Hunter, C. N. & Warren, M. J. Biosynthesis of the modified tetrapyrroles—the
611 pigments of life. *Journal of Biological Chemistry* **295**, 6888–6925 (2020).

612 35. Niwa, S. *et al.* Structure of the LH1-RC complex from Thermochromatium tepidum at 3.0 Å.
613 *Nature* **508**, 228–232 (2014). URL http://dx.doi.org/10.1038/nature13197.

614 36. Fromme, P., Jordan, P. & Krauß, N. Structure of photosystem i. *Biochimica et Biophysica*
615 *Acta (BBA)-Bioenergetics* **1507**, 5–31 (2001).

616 37. Kouyianou, K. *et al.* The chlorosome of chlorobaculum tepidum: size, mass and protein
617 composition revealed by electron microscopy, dynamic light scattering and mass spectrometry-
618 driven proteomics. *Proteomics* **11 14**, 2867–80 (2011).

619 38. Bryant, D. A. & Canniffe, D. P. How nature designs light-harvesting antenna systems: design
620 principles and functional realization in chlorophototrophic prokaryotes. *Journal of Physics B:*
621 *Atomic, Molecular and Optical Physics* **51**, 033001 (2018).

622 39. Thiel, V., Tank, M. & Bryant, D. A. Diversity of Chlorophototrophic Bacteria Re-
623 vealed in the Omics Era. *Annual Review of Plant Biology* **69**, annurev–arplant–
624 042817–040500 (2018). URL http://www.annualreviews.org/doi/10.1146/
625 annurev-arplant-042817-040500.

626 40. Zeng, Y., Feng, F., Medová, H., Dean, J. & Koblížek, M. Functional type 2 photosynthetic
627 reaction centers found in the rare bacterial phylum gemmatimonadetes. *Proceedings of the*
628 *National Academy of Sciences* **111**, 7795–7800 (2014).

629 41. Brinkmann, H., Göker, M., Koblížek, M., Wagner-Döbler, I. & Petersen, J. Horizontal operon
630 transfer, plasmids, and the evolution of photosynthesis in rhodobacteraceae. *The ISME journal*
631 **12**, 1994–2010 (2018).

632 42. Ward, L. M., Hemp, J., Shih, P. M., McGlynn, S. E. & Fischer, W. W. Evolution of phototrophy
633 in the chloroflexi phylum driven by horizontal gene transfer. *Frontiers in microbiology* **9**, 260
634 (2018).

635 43. Cardona, T. Thinking twice about the evolution of photosynthesis. *Open biology* **9**, 180246
636 (2019).

637 44. Schütz, M. *et al.* Early evolution of cytochrome bc complexes. *Journal of molecular biology*
638 **300**, 663–675 (2000).

639 45. Hohmann-Marriott, M. F. & Blankenship, R. E. Evolution of photosynthesis. *Annual review*
640 *of plant biology* **62**, 515–548 (2011).

641 46. Oesterhelt, D. & Stoeckenius, W. Rhodopsin-like protein from the purple membrane of
642 halobacterium halobium. *Nature new biology* **233**, 149–152 (1971).

643 47. Sabehi, G. *et al.* New insights into metabolic properties of marine bacteria encoding prote-
644 orhodopsins. *PLOS Biology* **3** (2005). URL `https://doi.org/10.1371/journal.`
645 `pbio.0030273`.

646 48. Balashov, S. P. *et al.* Xanthorhodopsin: the Retinal Protein Proton Pump of Salinibacter ruber
647 with a Light-harvesting Carotenoid Antenna. *Science (New York, N.Y.)* **309**, 2061–2064 (2005).
648 URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3065861/`.

649 49. Ernst, O. P. *et al.* Microbial and animal rhodopsins: structures, functions, and molecular
650 mechanisms. *Chemical reviews* **114**, 126–163 (2014).

50. Govorunova, E. G., Sineshchekov, O. A., Li, H. & Spudich, J. L. Microbial rhodopsins: diversity, mechanisms, and optogenetic applications. *Annual review of biochemistry* **86**, 845–872 (2017).

51. Gómez-Consarnau, L. *et al.* Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation. *PLOS Biology* **8**, 1–10 (2010). URL https://doi.org/10.1371/journal.pbio.1000358.

52. Gómez-Consarnau, L. *et al.* Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation. *PLOS Biology* **8**, 1–10 (2010). URL https://doi.org/10.1371/journal.pbio.1000358.

53. Gómez-Consarnau, L. *et al.* Microbial rhodopsins are major contributors to the solar energy captured in the sea. *Science advances* **5**, eaaw8855 (2019).

54. Finkel, O. M., Béjà, O. & Belkin, S. Global abundance of microbial rhodopsins. *ISME Journal* **7**, 448–451 (2013).

55. Nowicka, B. & Kruk, J. Powered by light: Phototrophy and photosynthesis in prokaryotes and its evolution. *Microbiological Research* **186-187**, 99–118 (2016).

56. Béjà, O. *et al.* Bacterial Rhodopsin : Evidence for a New Type of Phototrophy in the Sea Bacterial Rhodopsin : Evidence for a New Type of Phototrophy in the Sea. *Science* **289**, 1902–1906 (2000).

669 57. Rinke, C. *et al.* A phylogenomic and ecological analysis of the globally abundant Marine
670 Group II archaea (Ca. Poseidoniales ord. nov.). *The ISME Journal* 1 (2018). URL `http:`
671 `//www.nature.com/articles/s41396-018-0282-y`.

672 58. Gleason, F. H., Larkum, A. W., Raven, J. A., Manohar, C. S. & Lilje, O. Ecological implica-
673 tions of recently discovered and poorly studied sources of energy for the growth of true fungi
674 especially in extreme environments. *Fungal ecology* **39**, 380–387 (2019).

675 59. Slamovits, C. H., Okamoto, N., Burri, L., James, E. R. & Keeling, P. J. A bacterial prote-
676 orhodopsin proton pump in marine eukaryotes. *Nature Communications* **2**, 183–186 (2011).
677 URL `http://dx.doi.org/10.1038/ncomms1188`.

678 60. Vader, A., Laughinghouse, H. D., Griffiths, C., Jakobsen, K. S. & Gabrielsen, T. M. Proton-
679 pumping rhodopsins are abundantly expressed by microbial eukaryotes in a high-Arctic fjord.
680 *Environmental Microbiology* **20**, 890–902 (2018).

681 61. Labarre, A., Obiol, A., Wilken, S., Forn, I. & Massana, R. Expression of genes involved
682 in phagocytosis in uncultured heterotrophic flagellates. *Limnology and Oceanography* **65**,
683 S149–S160 (2020).

684 62. Bulzu, P.-A. *et al.* The sunlit microoxic niche of the archaeal eukaryotic ancestor comes to
685 light. *preprint* (2018).

686 63. Yutin, N. & Koonin, E. V. Proteorhodopsin genes in giant viruses. *Biology direct* **7**, 34 (2012).

687 64. Needham, D. M. *et al.* A distinct lineage of giant viruses brings a rhodopsin photosystem
688 to unicellular marine predators. *Proceedings of the National Academy of Sciences* **116**,
689 20574–20583 (2019).

690 65. Sharma, A. K., Spudich, J. L. & Doolittle, W. F. Microbial rhodopsins: functional versatility
691 and genetic mobility. *Trends in Microbiology* **14**, 463–469 (2006).

692 66. Sabehi, G. *et al.* New insights into metabolic properties of marine bacteria encoding prote-
693 orhodopsins. *PLoS Biology* **3** (2005).

694 67. Pinhassi, J., DeLong, E. F., Béjà, O., González, J. M. & Pedrós-Alió, C. Marine Bacterial and
695 Archaeal Ion-Pumping Rhodopsins: Genetic Diversity, Physiology, and Ecology. *Microbiology*
696 *and Molecular Biology Reviews* **80**, 929–954 (2016). URL `http://mmbr.asm.org/`
697 `lookup/doi/10.1128/MMBR.00003-16`.

698 68. Shikanai, T. Chloroplast NDH: A different enzyme with a structure similar to that of respiratory
699 NADH dehydrogenase. *Biochimica et Biophysica Acta - Bioenergetics* **1857**, 1015–1022
700 (2016). URL `http://dx.doi.org/10.1016/j.bbabio.2015.10.013`.

701 69. Zhu, J., Vinothkumar, K. R. & Hirst, J. Structure of mammalian respiratory complex i. *Nature*
702 **536**, 354–358 (2016). `arXiv:1011.1669v3`.

703 70. Nawrocki, W. *et al.* The mechanism of cyclic electron flow. *Biochimica et Biophysica Acta*
704 *(BBA)-Bioenergetics* (2019).

705 71. Walter, J. M., Greenfield, D., Bustamante, C. & Liphardt, J. Light-powering Escherichia coli
706 with proteorhodopsin. *Proceedings of the National Academy of Sciences* **104**, 2408–2412
707 (2007). URL http://www.pnas.org/cgi/doi/10.1073/pnas.0611035104.

708 72. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving
709 photosystem ii at a resolution of 1.9 å. *Nature* **473**, 55–60 (2011).

710 73. Zhang, J. *et al.* Article Structure of phycobilisome from the red alga Griffithsia pacifica.
711 *Nature Publishing Group* **551**, 57–63 (2017). URL http://dx.doi.org/10.1038/
712 nature24278.

713 74. He, G., Zhang, H., King, J. D. & Blankenship, R. E. Structural analysis of the homodimeric
714 reaction center complex from the photosynthetic green sulfur bacterium chlorobaculum
715 tepidum. *Biochemistry* **53**, 4924–4930 (2014).

716 75. Tsukatani, Y., Romberger, S. P., Golbeck, J. H. & Bryant, D. A. Isolation and characterization
717 of homodimeric type-i reaction center complex from candidatus chloracidobacterium ther-
718 mophilum, an aerobic chlorophototroph. *Journal of Biological Chemistry* **287**, 5720–5732
719 (2012).

720 76. He, Z. *et al.* Reaction centers of the thermophilic microaerophile, chloracidobacterium
721 thermophilum (acidobacteria) i: biochemical and biophysical characterization. *Photosynthesis*
722 *research* **142**, 87–103 (2019).

723 77. Nagashima, S. & Nagashima, K. V. Comparison of photosynthesis gene clusters retrieved
724 from total genome sequences of purple bacteria. In *Advances in Botanical Research*, vol. 66,
725 151–178 (Elsevier, 2013).

726 78. Xin, Y. *et al.* Cryo-em structure of the rc-lh core complex from an early branching photosyn-
727 thetic prokaryote. *Nature communications* **9**, 1–10 (2018).

728 79. Kurisu, G., Zhang, H., Smith, J. L. & Cramer, W. A. Structure of the cytochrome b6f complex
729 of oxygenic photosynthesis: Tuning the cavity. *Science* **302**, 1009–1014 (2003). URL `http:`
730 `//science.sciencemag.org/content/302/5647/1009. http://science.`
731 `sciencemag.org/content/302/5647/1009.full.pdf`.

732 80. Braakman, R., Follows, M. J. & Chisholm, S. W. Metabolic evolution and the
733 self-organization of ecosystems. *Proceedings of the National Academy of Sciences*
734 **114**, E3091–E3100 (2017). URL `http://www.pnas.org/lookup/doi/10.1073/`
735 `pnas.1619573114. 1703.08755`.

736 81. Sieradzki, E. T., Fuhrman, J. A., Rivero-Calle, S. & Gómez-Consarnau, L. Proteorhodopsins
737 dominate the expression of phototrophic mechanisms in seasonal and dynamic marine pi-
738 coplankton communities. *PeerJ* **6**, e5798 (2018).

739 82. Behrenfeld, M. J. & Kolber, Z. S. Widespread iron limitation of phytoplankton in the south
740 pacific ocean. *Science* **283**, 840–843 (1999).

741 83. Maresca, J. A., Miller, K. J., Keffer, J. L., Sabanayagam, C. R. & Campbell, B. J. Distribu-
742 tion and diversity of rhodopsin-producing microbes in the chesapeake bay. *Appl. Environ.*
743 *Microbiol.* **84**, e00137–18 (2018).

744 84. Parker, M. S. *et al.* Comparative metatranscriptomics identifies molecular bases for the
745 physiological responses of phytoplankton to varying iron availability. *Proceedings of the*
746 *National Academy of Sciences* **109**, E317–E325 (2012).

747 85. Marchetti, A., Catlett, D., Hopkinson, B. M., Ellis, K. & Cassar, N. Marine diatom prote-
748 orhodopsins and their potential role in coping with low iron availability. *ISME Journal* **9**,
749 2745–2748 (2015). URL `http://dx.doi.org/10.1038/ismej.2015.74`.

750 86. Cunningham, F. X., Dennenberg, R. J., Mustardy, L., Jursinic, P. a. & Gantt, E. Stoichiometry
751 of Photosystem I, Photosystem II, and Phycobilisomes in the Red Alga Porphyridium cruentum
752 as a Function of Growth Irradiance. *Plant physiology* **91**, 1179–1187 (1989).

753 87. Singharoy, A. *et al.* Atoms to phenotypes: Molecular design principles of cellular energy
754 metabolism. *Cell* **179**, 1098–1111 (2019).

755 88. Muench, S. P., Trinick, J. & Harrison, M. A. Structural divergence of the rotary atpases.
756 *Quarterly reviews of biophysics* **44**, 311–356 (2011).

757 89. Scheuring, S. & Sturgis, J. N. Atomic force microscopy of the bacterial photosynthetic
758 apparatus: Plain pictures of an elaborate machinery. *Photosynthesis Research* **102**, 197–211
759 (2009).

90. Niwa, S. *et al.* Structure of the lh1–rc complex from thermochromatium tepidum at 3.0 å. *Nature* **508**, 228–232 (2014).

91. Cherezov, V., Clogston, J., Papiz, M. Z. & Caffrey, M. Room to move: Crystallizing membrane proteins in swollen lipidic mesophases. *Journal of Molecular Biology* **357**, 1605–1618 (2006).

92. Herzen, V. *et al.* Bacterial photosynthesis in surface waters of the open ocean. *Nature* **407**, 177–179 (2000).

93. Lubner, C. E. *et al.* Solar hydrogen-producing bionanodevice outperforms natural photosynthesis. *Proceedings of the National Academy of Sciences* **108**, 20988–20991 (2011).

94. Friedrich, T. *et al.* Proteorhodopsin is a light-driven proton pump with variable vectoriality. *Journal of Molecular Biology* **321**, 821 – 838 (2002). URL `http://www.sciencedirect.com/science/article/pii/S0022283602006964`.

95. Lanyi, J. K. Proton transfers in the bacteriorhodopsin photocycle. *Biochimica et Biophysica Acta - Bioenergetics* **1757**, 1012–1018 (2006).

96. Kirchman, D. L. & Hanson, T. E. Bioenergetics of photoheterotrophic bacteria in the oceans. *Environmental Microbiology Reports* **5**, 188–199 (2013).

97. Sener, M., Strumpfer, J., Singharoy, A., Hunter, C. N. & Schulten, K. Overall energy conversion efficiency of a photosynthetic vesicle. *eLife* **5**, 1–30 (2016).

98. Zubkov, M. V. Photoheterotrophy in marine prokaryotes. *Journal of Plankton Research* **31**, 933–938 (2009).

779 99. Basan, M. *et al.* Overflow metabolism in escherichia coli results from efficient proteome
780 allocation. *Nature* **528**, 99–104 (2015).

781 100. Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W. & Milo, R. Glycolytic strategy as
782 a tradeoff between energy yield and protein cost. *Proceedings of the National Academy of*
783 *Sciences of the United States of America* **110**, 10039–10044 (2013).

784 101. Hui, S. *et al.* Quantitative proteomic analysis reveals a simple strategy of global resource
785 allocation in bacteria. *Molecular systems biology* **11** (2015).

786 102. Scott, M., Klumpp, S., Mateescu, E. M. & Hwa, T. Emergence of robust growth laws from
787 optimal regulation of ribosome synthesis. *Molecular systems biology* **10** (2014).

788 103. Erickson, D. W. *et al.* A global resource allocation strategy governs growth transition kinetics
789 of escherichia coli. *Nature* **551**, 119–123 (2017).

790 104. Wortel, M. T., Noor, E., Ferris, M., Bruggeman, F. J. & Liebermeister, W. Metabolic enzyme
791 cost explains variable trade-offs between microbial growth rate and yield. *PLoS computational*
792 *biology* **14**, e1006010 (2018).

793 105. Faizi, M., Zavřel, T., Loureiro, C., Červeny̌, J. & Steuer, R. A model of optimal protein
794 allocation during phototrophic growth. *Biosystems* **166**, 26–36 (2018).

795 106. Zavřel, T. *et al.* Quantitative insights into the cyanobacterial cell economy. *Elife* **8**, e42508
796 (2019).

797 107. Faizi, M. & Steuer, R. Optimal proteome allocation strategies for phototrophic growth in a
798 light-limited chemostat. *Microbial cell factories* **18**, 165 (2019).

799 108. Connell, J. H. & Slatyer, R. O. Mechanisms of succession in natural communities and their
800 role in community stability and organization. *The American Naturalist* **111**, 1119–1144
801 (1977).

802 109. Fukami, T. Historical contingency in community assembly: integrating niches, species pools,
803 and priority effects. *Annual Review of Ecology, Evolution, and Systematics* **46**, 1–23 (2015).

804 110. De Meester, L., Vanoverbeke, J., Kilsdonk, L. J. & Urban, M. C. Evolving perspectives on
805 monopolization and priority effects. *Trends in Ecology & Evolution* **31**, 136–146 (2016).

806 111. Cavender-Bares, J., Ackerly, D. D., Hobbie, S. E. & Townsend, P. A. Evolutionary legacy
807 effects on ecosystems: Biogeographic origins, plant traits, and implications for management in
808 the era of global change. *Annual Review of Ecology, Evolution, and Systematics* **47**, 433–462
809 (2016).

810 112. Mori, M., Marinari, E. & De Martino, A. A yield-cost tradeoff governs escherichia coli's
811 decision between fermentation and respiration in carbon-limited growth. *NPJ systems biology*
812 *and applications* **5**, 1–9 (2019).

813 113. Li, Y., Petrov, D. A. & Sherlock, G. Single nucleotide mapping of trait space reveals pareto
814 fronts that constrain adaptation. *Nature ecology & evolution* 1–13 (2019).

815 114. Noy, D., Moser, C. C. & Dutton, P. L. Design and engineering of photosynthetic light-
816 harvesting and electron transfer using length, time, and energy scales. *Biochimica et Biophys-*
817 *ica Acta (BBA)-Bioenergetics* **1757**, 90–105 (2006).

818 115. Carlile, M. Prokaryotes and eukaryotes: strategies and successes. *Trends in Biochemical*
819 *Sciences* **7**, 128–130 (1982).

820 116. Sancar, A. Structure and function of photolyase and in vivo enzymology: 50th anniversary.
821 *Journal of Biological Chemistry* **283**, 32153–32157 (2008).

822 117. Sorigué, D. *et al.* An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **357**,
823 903–907 (2017).

824 118. Kritsky, M. S., Telegina, T. A., Vechtomova, Y. L. & Buglak, A. A. Why flavins are not com-
825 petitors of chlorophyll in the evolution of biological converters of solar energy. *International*
826 *journal of molecular sciences* **14**, 575–593 (2013).

827 119. Rothschild, L. J. The evolution of photosynthesis... again? *Philosophical Transactions of the*
828 *Royal Society B: Biological Sciences* **363**, 2787–2801 (2008).

829 120. Shalaeva, D. N., Galperin, M. Y. & Mulkidjanian, A. Y. Eukaryotic g protein-coupled receptors
830 as descendants of prokaryotic sodium-translocating rhodopsins. *Biology direct* **10**, 63 (2015).

831 121. Sánchez-Baracaldo, P. & Cardona, T. On the origin of oxygenic photosynthesis and cyanobac-
832 teria. *New Phytologist* **225**, 1440–1446 (2020).

122. Cardona, T. & Rutherford, A. W. Evolution of photochemical reaction centres: more twists? *Trends in plant science* (2019).

123. Gisriel, C. *et al.* Structure of a symmetric photosynthetic reaction center–photosystem. *Science* **357**, 1021–1025 (2017).

124. Epel, B. & Butler, W. L. Cytochrome a3: Destruction by light. *Science* **166**, 621–622 (1969).

125. Robertson, J. B., Davis, C. R. & Johnson, C. H. Visible light alters yeast metabolic rhythms by inhibiting respiration. *Proceedings of the National Academy of Sciences* **110**, 21130–21135 (2013).

126. Zaremba-Niedzwiedzka, K. *et al.* Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353–358 (2017).

127. Stairs, C. W. & Ettema, T. J. The archaeal roots of the eukaryotic dynamic actin cytoskeleton. *Current Biology* **30**, R521–R526 (2020).

128. Neveu, E., Khalifeh, D., Salamin, N. & Fasshauer, D. Prototypic snare proteins are encoded in the genomes of heimdallarchaeota, potentially bridging the gap between the prokaryotes and eukaryotes. *Current Biology* (2020).

129. Shiratori, T., Suzuki, S., Kakizawa, Y. & Ishida, K.-i. Phagocytosis-like cell engulfment by a planctomycete bacterium. *Nature Communications* **10**, 5529 (2019). URL `http://dx.doi.org/10.1038/s41467-019-13499-2http://www.nature.com/articles/s41467-019-13499-2`.

130. Sagulenko, E. *et al.* Structural studies of planctomycete gemmata obscuriglobus support cell compartmentalisation in a bacterium. *PLoS One* **9** (2014).

131. Boedeker, C. *et al.* Determining the bacterial cell biology of planctomycetes. *Nature communications* **8**, 1–14 (2017).

132. Mahajan, M. *et al.* Paralogization and new protein architectures in planctomycetes bacteria with complex cell structures. *Molecular biology and evolution* **37**, 1020–1040 (2020).

133. Strous, M. *et al.* Missing lithotroph identified as new planctomycete. *Nature* **400**, 446–449 (1999).