PLOS BIOLOGY

RESEARCH ARTICLE

A molecular atlas of plastid and mitochondrial proteins reveals organellar remodeling during plant evolutionary transitions from algae to angiosperms

Parth K. Raval, Alexander I. MacLeod, Sven B. Gould*

Institute for Molecular Evolution, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

* gould@hhu.de

Citation: K. Raval P, MacLeod AI, Gould SB (2024) A molecular atlas of plastid and mitochondrial proteins reveals organellar remodeling during plant evolutionary transitions from algae to angiosperms. PLoS Biol 22(5): e3002608. https://doi.org/10.1371/journal.pbio.3002608

Academic Editor: Andrew J. Tanentzap, University of Cambridge, UNITED KINGDOM

Received: September 18, 2023

Accepted: March 28, 2024

Published: May 7, 2024

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pbio.3002608

Data Availability Statement: Supplementary figures are available in the supplementary information file. Additional supplementary data, in-house scripts and source data for the main and

Abstract

Algae and plants carry 2 organelles of endosymbiotic origin that have been co-evolving in their host cells for more than a billion years. The biology of plastids and mitochondria can differ significantly across major lineages and organelle changes likely accompanied the adaptation to new ecological niches such as the terrestrial habitat. Based on organelle proteome data and the genomes of 168 phototrophic (Archaeplastida) versus a broad range of 518 non-phototrophic eukaryotes, we screened for changes in plastid and mitochondrial biology across 1 billion years of evolution. Taking into account 331,571 protein families (or orthogroups), we identify 31,625 protein families that are unique to primary plastid-bearing eukaryotes. The 1,906 and 825 protein families are predicted to operate in plastids and mitochondria, respectively. Tracing the evolutionary history of these protein families through evolutionary time uncovers the significant remodeling the organelles experienced from algae to land plants. The analyses of gained orthogroups identifies molecular changes of organelle biology that connect to the diversification of major lineages and facilitated major transitions from chlorophytes en route to the global greening and origin of angiosperms.

Introduction

Fewer natural phenomena have been as transformative to planet Earth as the global greening through plants [1,2]. The proliferation of plants on land rests on the emergence and expansion of the Chloroplastida, also referred to as the Viridiplantae or simply the green lineage. The Chloroplastida are made up of 3 phyla: chlorophytes, streptophytes, and the prasinodermophytes that are thought to be the sister lineage to the 2 former [3]. Chloro- and prasinodermophytes are represented by algae only, whereas streptophytes are made up of algae and embryophytes, the latter uniting all land plants [3–5]. The list of key adaptations that fostered land plant expansion in a macroevolutionary context are multiple: roots, a mutualistic symbiosis with fungi, stomata, a cuticle, polyplastidy, and an expansion of many metabolite families such as flavonoids to name a few [1,3–10]. These innovations, evolving gradually in the

supplementary figures are available on Zenedo: https://zenodo.org/records/10855592.

Funding: We thank the Deutsche Forschungsgemeinschaft for grants awarded to SBG (SFB 1208-2672 05415 and SPP2237–440043394) and the Moore and Simons Initiative grant (9743) that was awarded to William Martin, who was so generous to financially support AIM. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. https://www.dfg.de/ https://www.simonsfoundation.org/.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: ASR, ancestral state reconstruction; GOG, green orthogroup; HMM, hidden Markov model; KEGG, Kyoto Encyclopedia of Genes and Genomes; KOID, KEGG orthology identification; MOG, mitochondrial orthogroup; PAP, PEP associated protein; PEP, plastid-encoded RNA polymerase; PGI, phosphoglucose isomerase; POG, plastid orthogroup; RING, really interesting new gene; UV, ultraviolet.

common ancestor of land plants (LCA), provided a decisive fitness advantage over the non-terrestrial chloro-, prasinodermato-, and streptophyte algal relatives [1,11].

The eponymous organelle of plants, the chloroplast, underwent various changes, too. It adapted in multiple ways to the challenges characterizing the habitat the LCA encountered. Improving stress response was necessary to deal for instance with increased levels of ultraviolet (UV) high light stress and to cope with temperature shifts that change rapidly on land in contrast to in water [12–14]. Polyplastidy, a phenomenon that separates plastid from nuclear division, leading to cells that can harbor more than one plastid per cell, was part of being able to develop larger body plans [12,15,16]. To communicate stress and the need for component biosynthesis, an elaborate retrograde signaling evolved on the basis of messenger proteins such as GUN1 and maybe WHIRLY [17,18]. In combination, these adaptations were decisive for the success of streptophytes, which is evident in the number of species they have evolved and the sheer biomass they produce [1,19].

Plastids do not operate autonomously, but are part of an intricate metabolic network and even physically interact with other compartments such as the endoplasmic reticulum and peroxisomes [20,21]. Marked metabolic and physical interactions of plastids also concern the only other compartment of ancient endosymbiotic origin: the mitochondrion. Plant mitochondria are much less in the focus of plant research. Next to their canonical functions, they are known to be involved in immunity, lipid metabolism, and other (eco) physiological processes that are frequently in crosstalk with the photosynthetic organelle [22,23]. Like plastids, mitochondria were critical in the evolution and continued adaptation of important physiological traits, which characterize the green lineage. A notable example of preadaptation includes malate decarboxylation in the C4 photosynthetic pathway [24]—a trait of the green lineage [25] that improves plant photosynthetic efficiency in warm and dry habitats [26]. Similarly, some components of mitochondrial retrograde signaling also evolved in the land plants and likely contributed to its ROS and draught tolerance [27].

In spite of the importance of these 2 organelles of endosymbiotic origin in coordinating their duties, the evolution of components specific to chloroplast and mitochondrial biology has not been explicitly studied in light of streptophyte evolution or plant terrestrialization. Previous work has determined genes specific to certain plant clades and that are catalogued by valuable resources such as the GreenCut [28]. Such analyses, however, did not focus on organelle biology nor clustered protein families. They were also limited by a low number of archaeplastidal genomes and insufficient methods for orthology inference available at that time. Since then, genome assemblies of members from previously unsampled clades has increased manyfold [11,29–37] and more organelle proteomes and better functional annotations are available. Similarly, and concomitantly, the development of novel and accurate algorithms for orthology inference [38–41], along with advances in experimental biology allow to now identify critical evolutionary changes in an eco-evo context of plastid and mitochondrial biology that underpin the success of the Chloroplastida.

Here, we curate a database of protein families unique to the green lineage. We plot their evolution across the major splits in the evolutionary history of streptophytes, focusing on the biology of the 2 organelles of endosymbiotic origin. We report that the number of plastid- and mitochondria-associated protein families changes most significantly at 2 evolutionary bifurcations: firstly, at the green lineage itself and secondly at the split between Zygnematophyceae and embryophytes at the water to land transition. The newly recruited protein families influenced organellar processes such as carbon and lipid metabolism, information processing, and organelle development. We provide an extensive catalogue of the changes the proteomes of plastid and mitochondria experienced throughout streptophyte evolution, which offers

multiple angles from which to explore major evolutionary transitions such as the conquest of land and embryophyte diversification.

Results

Half of the chloroplastida protein families are unique to embryophytes

Out of a total of 12,862,035 proteins, 95% were categorized from 686 eukaryotes and grouped into 331,570 orthogroups (deposited on Zenodo [42]). From these, 31,650 were present only in chloroplastida and classified as green orthogroups (GOGs) (S1 Fig; [42]). An examination of GOG distribution among green species revealed that around half of all GOGs were unique to terrestrial plants (Fig 1A). Approximately 400 GOGs appeared in more than 90% of species, referred from here on to as the "core GOGs" (Fig 1B). For only 5% of all GOGs, a functional annotation could be identified (Fig 1C; [42]). For embryophyte-specific GOGs, the numbers were comparable, yet they maintained a consistent distribution of identified functions, including a substantial fraction of membrane trafficking and ubiquitination-related proteins (Fig 1D; [42]). Notably, for the core GOGs the number is higher. For 30% functional annotations covering photosynthesis, mitochondrial formation, trafficking, and information processing could be identified (Fig 1E; [42]). The functions for a vast majority of the GOGs remain elusive [42], numbers that mirror those of previous studies [28], and they hence provide an excellent ground for experimental exploration.

Mitochondrial and plastid proteomes of the Chloroplastida expanded with the origin and diversification of the green lineage

To investigate changes in the proteomes of plastids and mitochondria, we curated 1,906 plastid and 825 mitochondrial orthogroups (POGs and MOGs, respectively) based on published

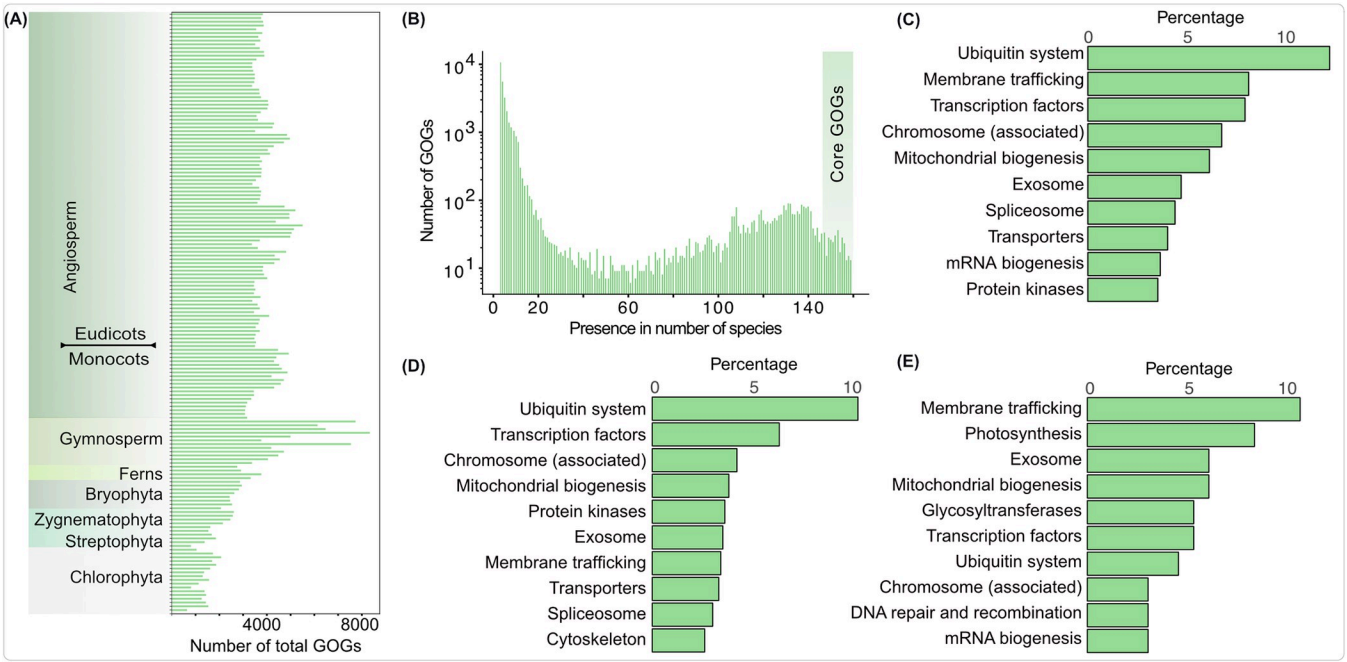

Fig 1. Distribution and functional annotation of GOGs. (A) Total number of GOGs present in each species from major Chloroplastida taxa. (B) Number of GOGs as a function of their presence across 159 Chloroplastida species. Major functional categories of 4.71% of all GOGs (c), 3.96% of the embryophyte GOGs (d), and 27.9% of the core GOGs (e). The underlying data of this figure can be found at https://zenodo.org/records/10855592. GOG, green orthogroup.

https://doi.org/10.1371/journal.pbio.3002608.g001

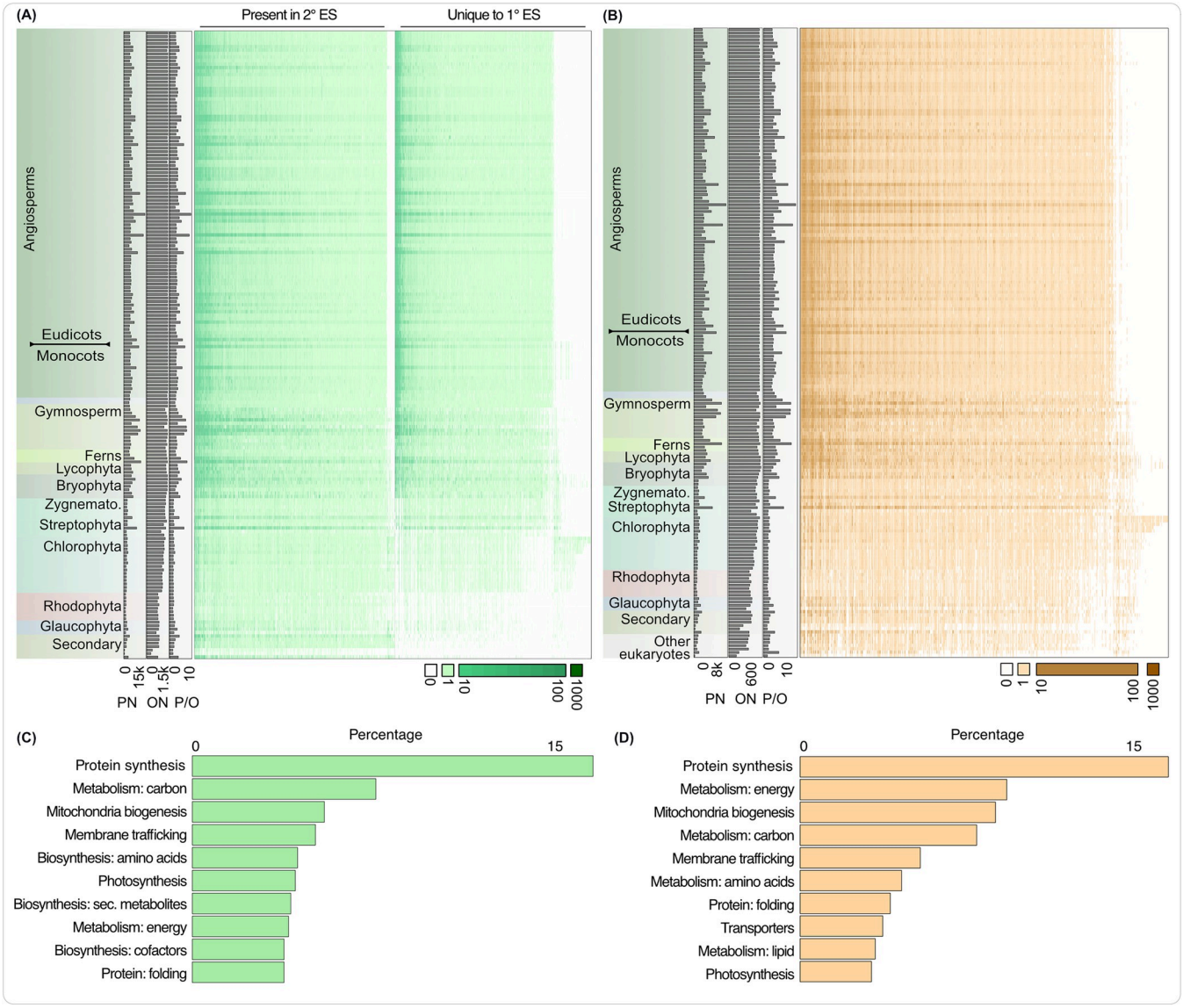

Fig 2. Mitochondrial and plastid orthogroups across archaeplastidal species. Distribution of plastid (POGs; A) and mitochondrial orthogroups (MOGs; B). The distribution of POGs was determined for plastids of primary (1˚ ES) and secondary endosymbiotic origin (2˚ ES). Protein copy numbers within each POG or MOG across species is shown in the heat-map as per the key on the bottom right of the heatmaps. Horizontal bars on the left side of the heatmaps show the total PNs likely localized to organelles, total POG or MOG numbers (ON) and distribution of PN per OG (P/O) for a given species. Major functional categories of POGs and MOGs in (C) and (D), respectively. The underlying data of this figure can be found at https://zenodo.org/records/10855592. MOG, mitochondrial orthogroup; PN, protein number; POG, plastid orthogroup.

https://doi.org/10.1371/journal.pbio.3002608.g002

proteome data and homology-based protein clustering of 204 eukaryotes, including that of secondarily photosynthetic eukaryotes (S1B Fig and S1A–S1D Table). In comparison to rhodophytes and glaucophytes, the green lineage encodes almost twice as many POGs (Fig 2A and S1E Table). Within the green lineage, from the Zygnematophyceae and embryophytes onwards, plastid proteomes further expanded both in terms of the number of proteins within each POG and the number of unique POGs. The former is likely a consequence of genome duplications, while the latter underscores functional divergence that followed gene duplications. The distribution of MOGs appears qualitatively similar to that of POGs (Fig 2B and S1F Table). Approximately 60% of the POGs could be functionally annotated, predominantly

operating in biosynthetic and other metabolic pathways such as photosynthesis (Fig 2C and S1G Table). Around 75% of the MOGs could be annotated, containing proteins for mitochondrial biogenesis, membrane trafficking, and translation (Fig 2D and S1H Table). Protein biosynthesis-related proteins are abundant in both, POGs and MOGs, underscoring their biosynthetic activity. Proteins for mitochondrial biogenesis also appear in both. For example, about 60 POGs are annotated as mitochondrial biogenesis. They encompass numerous PPR and mTERF proteins (crucial for RNA editing and metabolism) and proteins involved in various other information processing activities, probable to function in both organelles. Analysis of the N-terminal 20 amino acids show their charge to range from 0 to 2, indicating they might be dually targeted to plastids and mitochondria [42]. Five of the mTERFs are part of a POG and MOG simultaneously (S3D Fig). Overall, the trends show that in embryophytes the number of protein families associated with an endosymbiotic organelle function increased.

The increased number of POGs and MOGs in the green lineage is explained by a combination of 2 phenomena: (a) new gains in the green ancestor; and (b) secondary losses at the origin of rhodophytes [43]. We used ancestral state reconstruction (ASR) to resolve between these 2 possibilities. The branching order of the archaeplastidal lineages remains challenging [44], as sometimes glaucophytes [45] and sometimes rhodophytes come out as the sister to the other remaining archaeplastidal lineages [4,46]. An inferred eukaryotic tree (with 31 non-Archaeplastida eukaryotes as an outgroup to Archaeplastida) placed the rhodophytes and glaucophytes as sister clades (S2 Fig). This tree and the ASR pipeline were validated using rbcS as a control (S3A Fig), and further undergirded the main results which are consistent with varying thresholds of probability of presence and absence in a given ancestor on this eukaryotic tree (S3B Fig), as well as with manually rooting the Archaeplastida to have glaucophytes or rhodophytes as an outgroup to the Chloroplastida (S4–S7 Figs).

The result suggests that the plastid proteome of the last common ancestor of Archaeplastida united ca. 1,000 POGs (Figs 3A, S3B, and S6, S2A–S2C Table). This inferred proteome witnessed significant gains of protein families at the emergence of the green ancestor (and later speciation). Approximately 50% of these newly gained POGs could be functionally annotated (Fig 3C and S3A Table), showing that at the origin of the green lineage novel photosynthesis- and metabolism-related POGs were recruited, while the transition to land (Z/E and embryophyte ancestors) added metabolism-related, as well as protein synthesis- and ubiquitin-related POGs to the toolkit (S3A Table). Using hidden Markov searches, we verify that more than half of the protein families recruited in embryophyte and Z/E ancestors are absent in non-zygnematophyceae algae (S3C Fig). The mitochondrial proteome followed a qualitatively similar trend of expansion (Figs 3B, S3B, and S7, S2D–S2F Table). About 500 MOGs trace back to the archaeplastidal ancestor, while ca. 700 MOGs were identified at the root of angiosperms (Figs 3C and S3B). Around 50% of the newly gained MOGs could be functionally annotated, showing that the chloroplastidal gains contribute to carbon metabolism, protein synthesis, and mitochondrial biogenesis. Terrestrialization also witnessed a similar gain of MOGs, most of which function in metabolism as well as mitochondrial biogenesis and membrane trafficking (Fig 3C and S3B Table).

In summary, across plant species, plastid and mitochondrial proteomes are predicted to have gained a significant number of protein families reflecting the dynamic nature of organellar proteomes post-endosymbiosis [47,48]. A closer look at the function of the newly gained organelle proteins shows a wide variety, including lipid and carbon metabolism, information processing, development, and division of organelles.

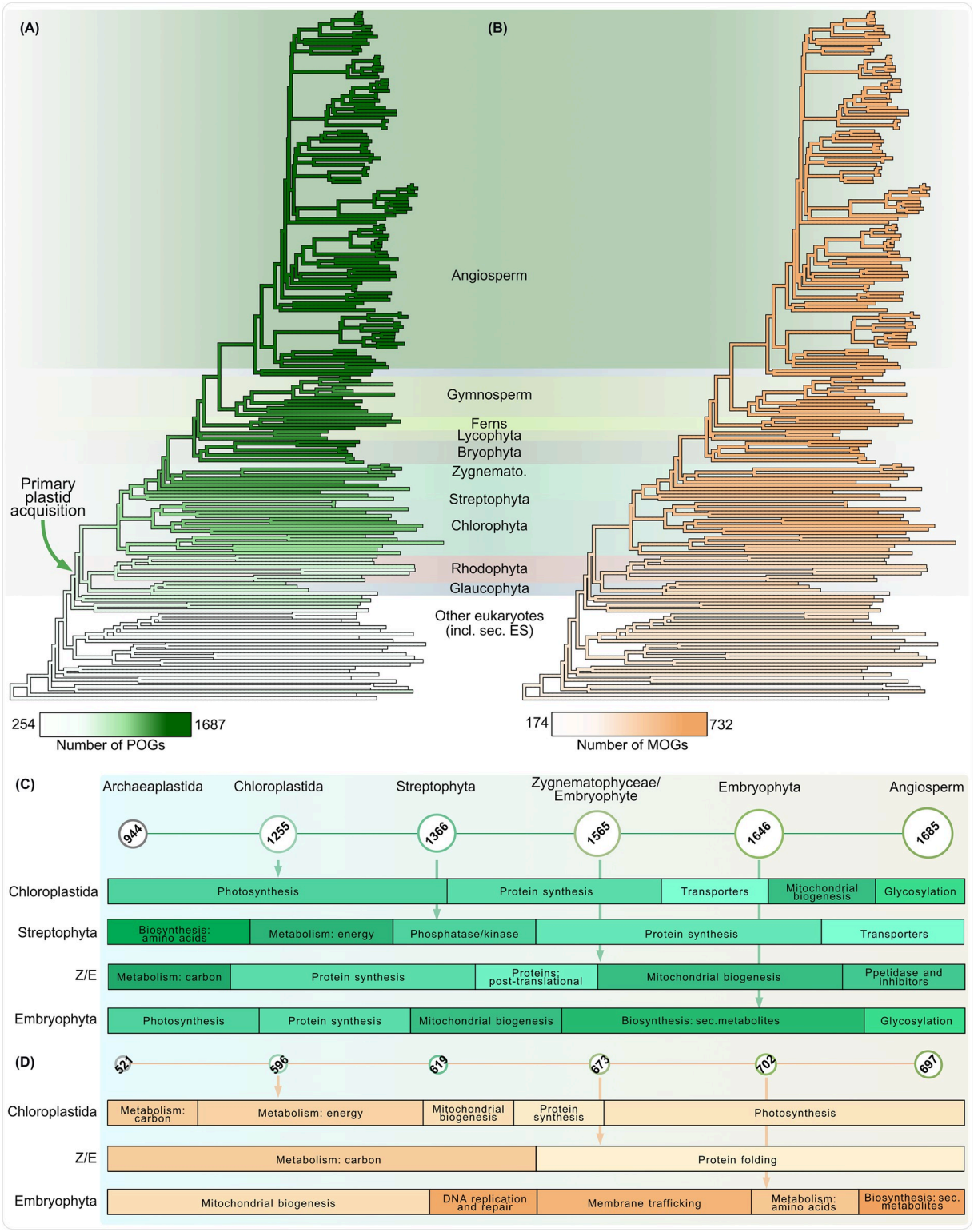

Fig 3. Evolution of organelle proteomes in Archaeplastida. Gains in plastid (POGs; A) and mitochondrial orthogroups (MOGs; B) across all nodes of archaeplastidal evolution and POGs coinciding with primary and secondary plastid acquisitions. Gains across main nodes of interest in (C), where each circle represents an ancestor, with its predicted number of protein families shown in the circle and whose diameter correlates with the number of OGs. Major gains occurred in the chloroplastidal ancestor, the common ancestor of Zygnematophyceae and embryophytes (Z/E), and in embryophytes. In (D) the same as in (C), but for mitochondrial OGs. Their functions are shown in the proportionate bar charts

below the ancestors. The underlying data of this figure can be found at https://zenodo.org/records/10855592. MOG, mitochondrial orthogroup; POG, plastid orthogroup.

https://doi.org/10.1371/journal.pbio.3002608.g003

Increased complexity of RNA metabolism and photosynthetic adaptability

RNA metabolism such as editing intercepts the linear information flow from mRNA to protein and is crucial for organelles to function [49–51]. Two main domains, the PPR and mTERF domain, are associated with RNA editing and metabolism [52,53]. We first screened for organelle orthogroups containing either of these 2 domains in at least 60% of all proteins within each respective orthogroup (S1C Fig). Around 50 POGs and 20 MOGs were found. More than 80% of them were restricted to embryophytes, only few were present in some algae (Fig 4). A closer look revealed that most of the algal homologues lacked PPR and mTERF domains and they are hence unlikely true orthologues. More generally, this shows that any detailed interpretation regarding an inferred orthogroup's function should be supported by screening for functionally relevant domains.

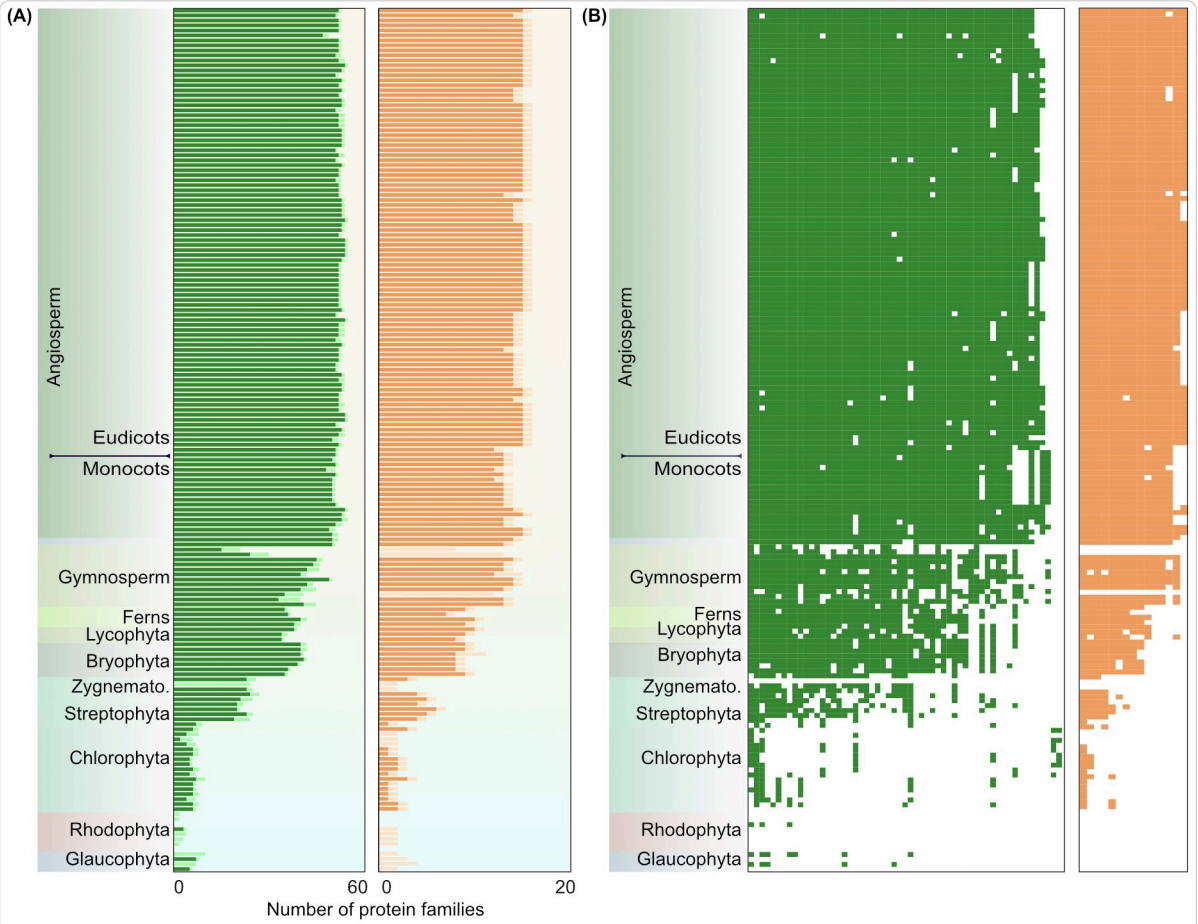

Fig 4. Recruitment of PPR and mTERF domains in organelle proteins. (A) Number of POGs (left) and MOGs (right), where at least 1 protein contains a PPR/mTERF domain, is shown in bars with dark shades of colors. Total number of orthogroups (regardless of presence or absence of PPR/mTERF domain in that particular species) is shown in lighter shade. It shows the presence of the orthogroups in question in algae, but that they only later obtained PPR/mTERF domains in embryophytes. (B) Each cell represents an orthogroup and a colored cell indicates the presence of a PPR or mTERF domain in the protein family (column) of a respective species (rows). The underlying data of this figure can be found at https://zenodo.org/records/10855592. MOG, mitochondrial orthogroup; POG, plastid orthogroup.

https://doi.org/10.1371/journal.pbio.3002608.g004

True PPR or mTERF domain-containing RNA-editing proteins (and splicing and processing at large) increased significantly in number by recruiting new orthogroups, also through adding the 2 domains to proteins that did not contain these in their algal ancestor. A presence–absence pattern shows that >90% of proteins containing PPR/mTERF domains are exclusive to land plants, except for *Chara braunii* and *Klebsormidium flaccidum* (Fig 4B). These proteins include, but are not limited to, OTP51 and SOT5 (present in embryophytes and *Chara*) as well as SOT1, SVR7, THA8, PDM4 (present only in embryophytes; S9 Fig). Target transcripts of these RNA metabolism factors point to the synthesis and assembly of photosynthesis-related proteins and to proteins of the thylakoid membrane. Likewise, mTERFs, which are crucial for plastid and leaf development, are also uniquely expanded in the terrestrial clade with examples of protein re-targeting across organelles [54]. The dual targeted (plastid and mitochondrion) mTERF6, unique to the land plants (S9 Fig) and the streptophyte alga *Klebsormidium*, takes part in retrograde signaling to the nucleus via ABA and imparts abiotic stress tolerance [55]. Overall, RNA metabolism across plants has undergone major changes and has a significant impact on photosynthesis, improvement of which was key to thriving on land.

Adaptation to the terrestrial habitat and changes in plastid biochemistry

Main terrestrial stresses include draught, high (UV-)light, and swift temperature changes. Cutin and suberin, 2 of the most abundant lipid polymers on Earth [56], evolved as one countermeasure [9]. We find that cutin and suberin evolution was enabled by the recruitment of an organelle-specific GPAT (glycerol-3-phosphate acyltransferases) family in the embryophyte ancestor (Fig 5), which includes GPAT1 (mitochondrial), GPAT 4,6 and 8 of the endoplasmic reticulum [57,58]. Trafficking of these lipids across organelles was made possible by a dual targeted TGD4 [59] that was recruited in the chloroplastida ancestor (Fig 5). Acyl carrier thioesterases, responsible for the export of fatty acids from the plastid, acyl carrier protein desaturases (ACP-desaturase), and acyl-carrier proteins co-factors of fatty acid bio-synthesis were uniquely retained and expanded in the green lineage (S9 Fig). Duplication and divergence of ACP desaturases in embryo- and spermatophytes played an important role in regulating lipid composition shifts in response to temperature and drought, the regulation of seed oil content and development [60]. Likewise, acyl-carrier proteins also increased in copy number (S9 Fig) and adapted towards a light-induced expression and regulation of the seed fatty acid content [61,62]. These changes in organelle lipid synthesis and trafficking underpinned embryophyte adaptations to cope with draught and high temperature stress (wax biosynthesis, deposition on the layer of leaves, and cuticle development), as well as seed development and germination in spermatophytes (Fig 6D).

Changes in starch metabolism mostly pertain to its regulation. ADP-glucose pyrophoshorylase (AGPase), an enzyme responsible for a rate-limiting step in starch metabolism, is uniquely retained in the green lineage and increased in copy number in streptophytes (S9 Fig). AGPases diverged to regulate starch metabolism under osmotic and light stress, as well as the differential regulation of starch synthesis and degradation [63–67]. Another key regulatory enzyme, PGI (phosphoglucose isomerase) evolved a distinct family (PGI1) in Zygnematophyceae (S9 Fig). It likely kickstarted the regulation of starch metabolism at the water-to-land interface and later assumed significant roles in embryophyte fatty acid content regulation and the yield of seeds [68]. PTST3 also emerged around the time of terrestrialization (S9 Fig), which evolved to regulate starch synthesis with significant impact on plastid development [69]. In contrast to the flow of carbon through glycolysis, GSM2 (which originated in streptophytes; S9 Fig), shunts carbon towards the pentose-phosphate pathway and protects plastids from oxidative stress in *Arabidopsis* [70].

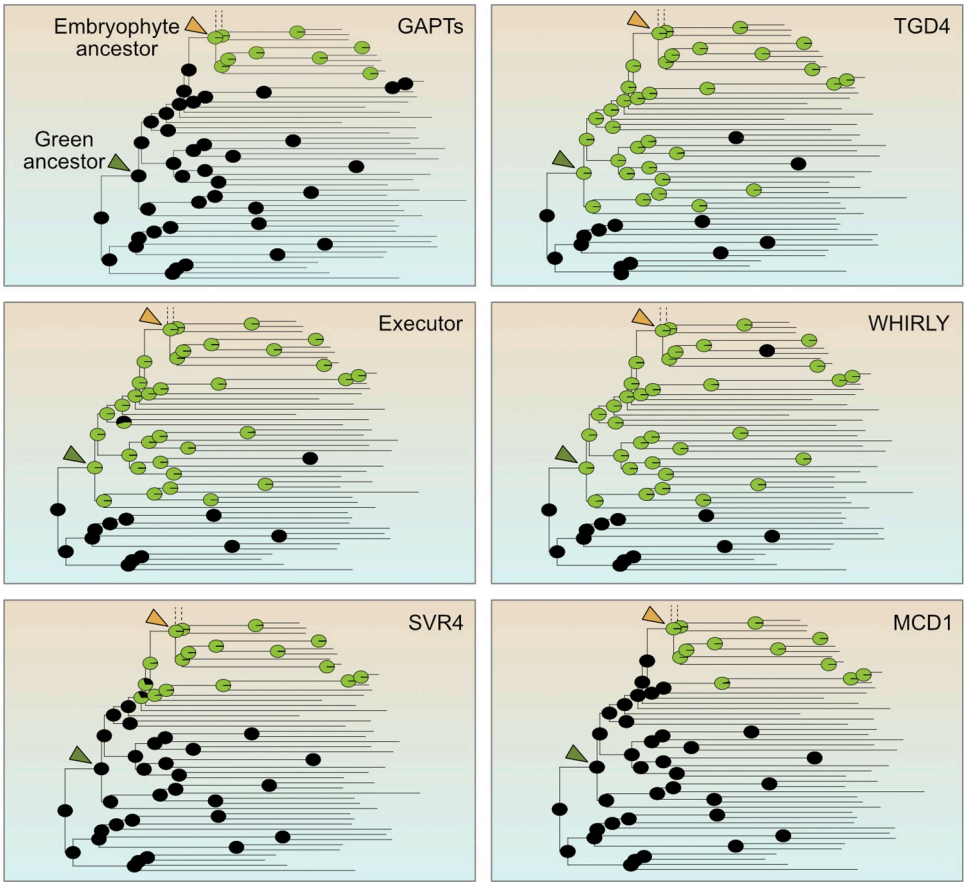

Fig 5. Origins of key proteins involved in metabolism, communication, and development. ASR for selected lipid metabolism (GAPT and TGD4), retrograde signaling (Executor and Whirly), plastid development (SVR4), and division (MCD1)-related proteins. The pie charts at each node represent the probability of presence (green) or absence (black) of a protein family in that node. The underlying data of this figure can be found at https://zenodo.org/records/10855592. ASR, ancestral state reconstruction.

https://doi.org/10.1371/journal.pbio.3002608.g005

Emergence of sophisticated antero- and retrograde communication cascades

Communication across compartments is critical for a concerted response to environmental stimuli. Plastids are key environmental sensors that interconnect cellular metabolism with physiological requirements and stress responses, and terrestrial stressors are key triggers of plastid-to-nucleus retrograde signaling [12,13,22]. We screened for the origin and diversification of EXECUTOR and SVR4, both components of retrograde signaling. We also screened for WHIRLY, a protein family that acts on RNA splicing and ribosome biogenesis, but also relocates between compartments and remains a disputed candidate for retrograde signaling [18,71–75]. EXECUTOR, key to regulating retrograde signaling, oxygen and light stress regulation [76–78], originated in the ancestor of the Chloroplastida and so did WHIRLY (Fig 5); the latter underwent copy number expansion in embryophytes and was likely lost in some bryophytes (S9 Fig). Divergence of these copies led to a localization across multiple organelles and today they are crucial for maintaining functional respiration, photosynthesis, and the response of mitochondria and plastids to biotic and abiotic stresses [79–81]. Additional

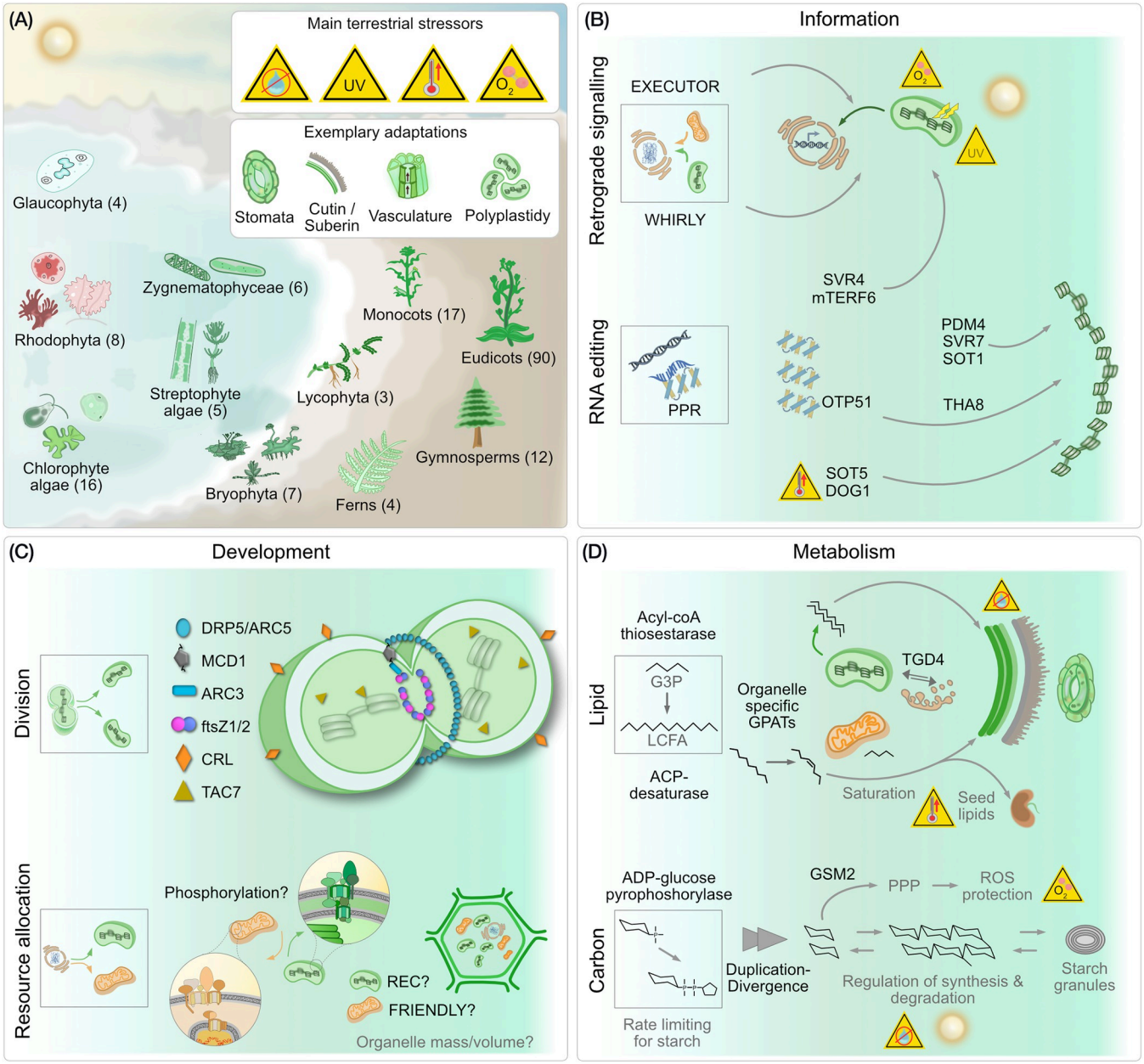

Fig 6. The global greening and endosymbiotic organelles. (A) After the endosymbiotic origin of the plastid, 3 aboriginal lineages emerged that form the Archaeplastida: the glaucophytes, rhodophytes, and chlorophytes. From the latter, streptophyte algae evolved, including the zygnematophyceae, that represent the algal sister clade to land plants (embryophytes). Abiotic stresses encountered during terrestrialization (water scarcity, high UV, swiftly altering temperatures and higher levels of O_2) selected for adaptive features such as stomata and a cutin layer. The numbers in parenthesis indicate the number of genomes from each major group that was screened. Recruitment of new organelle proteins improved 3 key aspects of organelle biology in light of terrestrialization: (B) information processing, (C) development, and (D) metabolism. Details for each tile are discussed in the main text. UV, ultraviolet.

https://doi.org/10.1371/journal.pbio.3002608.g006

paralogs evolved, each with a specific function in the main green lineages, and they likely aided in the colonization of the terrestrial habitat by the ancestor of land plants (Fig 6B).

SVR4, a dual targeted (plastid and nucleus) protein recruited during the time of terrestrialization (Fig 5), likely communicates required gene expression changes needed for light-induced plastid development, thylakoid stacking, and thermomorphogenesis [82,83]. In combination,

this facilitates light-induced photomorphogenesis, a process key for surviving on land. An increase in the complexity of retrograde signaling was a precursor for terrestrialization [12], for instance via innovations associated with the the 3′-phosphoadenosine-5′-phosphate family, which promoted the emergence of stomatal closing in land plants [84]. The recruitment and diversification of the proteins we highlight were quintessential for responding to 2 major stressors that are more pronounced and more rapidly changing on land than in water: light and temperature (Fig 6B).

Recruitment of new proteins and changes in organelle development

The coordination of tissue and plastid development is linked to ensure an appropriate response to biotic and abiotic factors, especially in morphologically complex plants [85–87]. Polyplastidy is a trait of land plants and some macroscopic algae such as *Bryopsis* or *Chara* [15], and known molecular determinants include MinD, MinE, ARC3, and the FtsZ proteins [16,86]. Our data supports that MULTIPLE CHLOROPLAST DIVISION SITE 1 (MCD1), another known component of the plastid division machinery [88], originated in the ancestral embryophyte (Fig 5). The cotyledon chloroplast biogenesis factor CYO1 and the transcriptionally active chromosome factor 7 (TAC7) are important components of thylakoid biogenesis and the plastid translation machinery, respectively. Both originated in the streptophyte ancestor (S9 Fig) and play key roles in chloroplast, cotyledon, thylakoid and leaf development in *Arabidopsis* [89–91]. Lastly, CRUMPLED LEAF (CRL), a protein residing in the outer plastid membrane, emerged during terrestrialization, too (S9 Fig), likely for regulating plastid division and securing correct plastid inheritance during embryogenesis [92,93].

Crucial for plastid biogenesis, especially in light of an expanding proteome, is the import of proteins. The membrane GTPase TOC159 is essential for chloroplast biogenesis via the selective recognition and import of the photosynthetic proteins [94] and is unique to the green lineage (S9 Fig). The membrane recruitment of this protein requires TOC75, of which a special variant evolved in the green ancestor after the duplication of OEP80 [14,95]. The copy number of TOC159 expanded from the Zygnematophyceae onwards (S9 Fig), hinting at its functional diversification. Unlike in the chlorophyte alga *Chlamydomonas*, land plant TOC159 homologues possess an N-terminal acidic domain that gets phosphorylated to alter substrate specificity [94,96]. Furthermore, TOC159, along with TOC132 and TOC120, play important roles in regulating plastid lipid synthesis and membrane fluidity and in *Arabidopsis* show tissue-specific expression (The Arabidopsis Information Resource) [97–99]. Further on the course of evolution, the J-domain-containing protein TOC12 [100] was likely recruited in the ancestral embryophyte for supporting the import machinery at the intermembrane space (S9 Fig). The terrestrial habitat demands a highly efficient and fluid import of proteins, for example, upon high light and other abiotic stresses [14,101]. The expansion of the TOC/TIC system in the embryophyte ancestor reflects how the organelle dealt with an ever-increasing diversity of substrates that were required to be processed.

Discussion

The settling of land by a streptophyte alga and the subsequent evolution and spreading of plants (Fig 6A) was pivotal in the transformation of the terrestrial habitat and it laid the foundation for the concomitant evolution and diversification of animals [1,2]. Throughout the hundreds of millions of years of plant evolution, both organelles of endosymbiotic origin underwent a multitude of molecular adaptations, hereby evolving into the plastid and mitochondrion of modern plants. We identified 31,650 protein families unique to the green lineage, approximately 50% of which are unique to embryophytes. It demonstrates an expansion and

divergence of protein families at the time of plant terrestrialization and in line with a recent study that identified around 10,000 duplications at the birth of embryophytes [102].

Expansion of proteins families is evident in both organellar proteomes at the origin of the green lineage itself and at the water-to-land transition. The gain of protein families at the origin of the Chloroplastida needs to be treated with caution due to the documented genetic bottleneck that characterizes rhodophyte origin [103–107] and the sparse availability of glaucophyte genome data. Some of the newly recruited protein families at the origin of the green lineage might rather be explained by a loss in rhodophytes and a retention in the chloroplastidal ancestor instead of a gain. Regardless, this has little bearing on the biological significance of a given protein family with respect to the overall increase in complexity of organelle biology—both concerning the variety as well as the number of proteins targeted to plastids and mitochondria—throughout streptophyte evolution. It affected the organelles metabolic, informational and developmental complexity, and facilitated the evolutionary successful transition from water to land more than 500 million years ago (Fig 6).

Changes in organelle lipid biochemistry contributed to one of the key adaptations in land plants that is the cuticle. Land plant GPATs (crucial to lipid synthesis for cutin and suberin) contribute to increased hydrophobicity and water retention in embryophytes [9] and their activity in embryophytes differ from that in algae [108,109]. Our analyses pinpoint the origins of organelle-specific GPATs (GPAT 1,4,6, and 8) to the embryophyte ancestor, and of which deleting GPAT4 and GPAT8 distorts cuticles and increases water loss by several fold [57,58]. In parallel, lipid trafficking was mediated by the recruitment or divergence of proteins such as TGD4 and acyl carrier thioesterases, which contributed to wax biosynthesis and deposition on leaves, cuticle development, thylakoid membrane stacking [59], seed development, and germination [60]. As for starch metabolism, the archaeplastidal ancestor likely stored starch in the cytosol [110], but the red and green lineage experienced different fates from there on. Rhodophytes continued to store starch in the cytosol in the form of Floridean starch [111], while in the green lineage, particularly in complex plants, more localized control of starch synthesis and degradation was facilitated by changes in regulatory proteins (e.g., AGPase). Together, organelle metabolism evolved to serve key roles in the synthesis, regulation, and trafficking of lipids involved in wax coating to prevent water loss in the land plant ancestor, as well as synthesis and storage of starch (Fig 6D).

RNA processing and editing is a crucial component of information processing and overall functionality of plant organelles [49,50]. Changes in RNA metabolism are evident at the origin of the green lineage, where RNase-P (tRNA maturation) was replaced by Protein-only RNase P or PROPs [112,113]. Subsequent expansion of PROPs in embryophytes (S9 Fig) led to organelle-localized copies, of which some are essential for maintaining organelle morphology, function, and plant viability [114]. Components associated with plastid-encoded RNA polymerase (PEP associated proteins, PAPs) also show a gradual recruitment from the green ancestor to embryophyte ancestor (S8 Fig). RNA editing of C to U is not found in algae, however, and editing sites in embryophytes are unlike those of any other eukaryote, suggesting they emerged independently [50]. We trace the emergence of many RNA-metabolism proteins to the time of plant terrestrialization and their known targets are transcripts involved in photosynthesis and stress tolerance-related transcripts, both key to colonizing land (Fig 6B). For example, THA8, PDM4, SVR7, and SOT1 associate with transcripts such as ycf2 and ycf3, and contribute to thylakoid development and biogenesis [115,116], the generation of photosynthetic complex proteins, grana stacking, and embryo and plastid development [115,117,118]. OTP51 and SOT5 splice transcripts related to chlorophyll synthesis, photosynthesis, and thylakoid membranes (ycf3, TRNK, and RPL2) [119–121], whereas DOG1 is important for high temperature response and chloroplast development [122]. This elaborate RNA processing in organelles,

especially plastids, appears to serve photosynthesis (and thylakoid)-related transcripts. It is feasible that by benefitting photosynthesis, organelle RNA editing continued to be positively selected for during terrestrialization and was expanded.

One evolutionary step towards efficient photosynthesis, where RNA editing also plays a key role, are grana stacks [85]. The evolutionary origin of grana remains elusive, along with the underlying developmental pathways involved in regulating its formation and maintenance [85,123,124]. Highly organized grana stacks are observed in embryophytes and some Zygnematophyceae (e.g., the *Cosmarium* genus) [125], but not chlorophytes such as *Chlamydomonas* [126]. We noticed a patchy distribution of grana morphology-associated proteins such as CURT1, RIQ1, and RIQ2 (S9 Fig), with both RIQs being present in all streptophytes and some chlorophytes but excluding *Chlamydomonas*. In light of the many key adaptations in Zygnematophyceae discussed here and elsewhere [11,127], we speculate that a sophisticated stacking of grana originated in streptophytes and was beneficial for thriving on land through photosynthesis optimization, in particular, with respect to photosystem repair and the separation of the photosystems and the ATP synthase [128,129].

This expansion of an organelle proteome necessitates improving the capacity to import proteins. Changes in some import receptors within the green lineage and in targeting sequences at its origins are known, with phosphorylation likely emerging as a key regulator for sorting the newly expanded proteome differentially to plastid and mitochondria (Fig 6C) [14,130]. Despite such adaptations, protein sorting is never perfect and some mistargeting might be positively selected for. A regulated distribution of newly recruited proteins (e.g., WHIRLY, TGD4, mTERF6; Fig 6B) to multiple organelles (with distinct organellar functions) hints at adaptive values of this apparent mis-sorting. How many of the newly recruited proteins get "mis-sorted" owing to biological adaptability versus stochasticity remains to be explored together with obtaining a more comprehensive picture of (regulatory) mechanisms associated with sorting in general.

Embryophyte cells target proteins not to a single plastid, but many simultaneously. The presence of multiple plastids per cell, (polyplastidy), in the green lineage, evolved in an embryophyte ancestor, maybe the common ancestor of embryo–and charophytes, and through changes in plastid fission and a decoupling of the latter from the cell cycle [15,16]. We find that MCD1, a core regulator of the plastid division proteins FtsZ2 and ARC3, emerged in the embryophyte ancestor, which corroborates the idea of a mono- to polyplastidy switch during the transition from water to land [12,15,16,131]. A change in the copy number of plastids also requires a mechanism that maintains a functional organelle to cell volume ratio and resource allocation (Fig 6C). The REDUCED CHLOROPLAST COVERAGE (REC) protein is involved in such a mechanism in *Arabidopsis* [132] and the phylogenetically related protein FRIENDLY regulates the distribution of mitochondria, also in plants and non-photosynthetic organisms [133,134]. REC and FRIENDLY share almost all of their domains. How they exactly function and differentiate between the 2 organelles remains elusive. From what we can tell, FRIENDLY emerged during eukaryogenesis and the origin of mitochondria. REC we can trace back to the streptophyte ancestor (S9 Fig) and after a duplication event of FRIENDLY. We speculate that the origin of REC helped to cement polyplastidy, which itself supports larger body plans and the diversification of different plastid types [15]. Lastly, an increase in organelle copy number also requires an overall increase in the capacity to synthesize proteins. The largest fraction of organelle proteins operate in tRNA, amino acid, and ribosomal biosynthesis and undergird the biosynthetic capacity of organelles, an adaptation strategy akin to their bacterial ancestor [135,136].

The accommodation of the early mitochondrial endosymbiont is associated with the origin of the endomembrane system and necessitated the emergence of eukaryotic traits including

mito- and autophagy [137–139]. Our analyses show that the integration of a subsequent endosymbiont, the plastid, coincided with the emergence of proteins that work for the endomembrane system. Salient are changes in the ubiquitin system during terrestrialization, when polyplastidy in the green lineage also emerged (S1G Table). Ubiquitination is key to proteosome-mediated degradation and is performed chiefly by the E3 ubiquitin ligase family, which are important in land plants also for photomorphogenesis [140]. RING (really interesting new gene) E3 ligases contribute to growth, development, and stress response via also mediating protein–protein interactions [141–144]. We trace a number of RING finger and related proteins to terrestrialization (S9 Fig) that include, but are not limited to, *DAL1* and *DAL2* (for *Drosophila* DIAP1 like 1 and 2), KEG (Keep on going), and NIP1 and NIP2. *DAL1* and *DAL2* play a key role in regulation of programmed cell death [145], peroxisome, and chloroplast biogenesis [146–148]. KEG contributes to stress mitigation [149,150], while NIP1 and NIP2 play a role in plastid development by docking plastid RNA polymerase to the thylakoid membrane [151]. The regulated degradation of plastids and other changes in the endomembrane system are a prerequisite for housing multiple plastids per cell and we find many more recruitments broadly affiliated with the endomembrane system, with poorly characterized functions until now. Exploring the functions of these proteins will add valuable insights into the cell biological changes that endosymbiosis stipulates.

While we focus on the evolution of the chloroplast of the green lineage, the rhodoplast of rhodophytes and the cyanelle of glaucophytes, as well as the plastid of secondary plastids outside of the Archaeplastida are worth considering. Based on the available chloroplastidal and glaucophyte data, our analysis predicts about 900 plastid proteins for rhodophytes (Fig 2A). This number is likely skewed by the sequence divergence of over a billion years, however, and even relaxed homology searches—e.g., on the basis of hidden Markov models—are no replacement for experimental validation (S3D Fig). As for plastids of secondary origin, we find that around half of the plastid proteins, or fewer, are shared between primary and secondary plastids (Fig 2A). Experimental proteomes of secondary plastids are available for *Phaeodactylum* [152] and *Plasmodium falciparum* [153], and the numbers match ours quite well (839 versus 934, and 346 versus 329, respectively; Fig 2A). Still, the experimentally validated proteins of secondary plastids remain entangled with chloroplastidal proteomes [152], as the filtering involves matches to just those. The experimental data, as well as our predictions are likely underestimating the true nature of organelle proteomes to a degree due to sequence divergence and a niche-specific trajectory of primary and secondary plastids. This further underscores the need for additional experimental proteomes from species other than chloroplastida, and which could benefit from the use of subcellular localization mapping using LOPIT [157] and proximity labeling [158].

In closing, although experimentally reported plant plastid and mitochondrial proteomes are scarce, we were able to generate a first comprehensive molecular atlas of the changes of plastid and mitochondrial protein families in the evolution of the green lineage. ASR allows to map the organelle transformations that facilitated the major transitions such as terrestrialization and which will improve with every new proteome that is added. By inferring plastid and mitochondrial proteomes for 173 species, we set testable expectations for new proteomes to come and provide a solid database, where origins and across species orthologues of any known (organelle) protein can be searched (S1C and S1D Table). Additional proteomes, once available, will likely solidify the general pattern observed and uncover more lineage-specific curiosities. We identify numerous mitochondrial protein recruitments, whose physiological roles and adaptive values help to better understand plant mitochondrial biology. For plastid proteins, we infer their functions and physiological importance based on the extensively studied *Arabidopsis* system. Utilizing an advanced orthology search technique [40], we postulate that

orthologues of *Arabidopsis* are likely to exhibit similar functions in other species. Our methodologically robust approach maps various changes in evolution associated in particular with terrestrialization that can now be experimentally explored across selected models and with a focus on less-well studied streptophyte algal and bryophyte species [156,157].

Conclusions

Endosymbiotic organelles have a distinct place in the evolutionary tapestry of life. Through the combination of organelle proteome data and phylogeny, we trace the evolution of mitochondria and plastids over a span of a billion years of plant evolution by inferring their proteomes for over a hundred Archaeplastida species. Our comprehensive molecular atlas identifies main changes in their metabolism, communication, information processing, and biogenesis. Key adaptations in plant organelles fostered the emergence of wax and cutin (see organelle lipid synthesis and transport), improved the photosynthetic yield (see organelle RNA metabolism and highly structured grana stacks), and the response to abiotic stressors (see inter-organelle communication), and mediated the transition from mono- to polyplastidy (see division and volume control). By connecting the molecular adaptations of mitochondria and plastids to macroevolutionary trends, we show how important changes in organelles of endosymbiotic origin were for the speciation that gave rise to the Chloroplastida and later the origin of land plants from a charophyte algal ancestor.

Material and methods

Curation of green orthogroups (GOGs)

Input protein sequences from 686 proteomes (from KEGG [158] and Phytozome [29] were clustered using Orthofinder version 2.5.4 [40], after all versus all blasts were conducted (E-value cutoff 10e-10) using diamond blast version2.011 [38]. From orthogroups (OGs) recovered, OGs with at least 3 Chloroplastida species and less than 3 species other than Chloroplastida were annotated as green orthogroup (GOGs). Schematic in S1A Fig. Inhouse python script used for this, the resulting data and other data processing, are available at https://zenodo.org/records/10855592 [42].

Curation of plastid and mitochondria orthogroups (POGs and MOGs)

A total of 5,452,977 proteins from 204 eukaryotes (S1A Table) were clustered using Orthofinder as described above. Orthogroups that contained at least 1 experimentally verified organelle protein from any one of the 4 experimentally verified organelle proteome of *C. reinhardtii* [159], *P. patens* [160], *Z. mays* [161], *A. thaliana* [161], were annotated as organelle (plastid and mitochondria) orthogroups. Schematic in S1B Fig.

Functional annotation of orthogroups

The source of >90% species was Kyoto Encyclopedia of Genes and Genomes (KEGG), which included KEGG orthology identification (KOID) for protein sequences. For all proteins within each GOG, KOIDs were retrieved and the most frequent KOID (i.e., majority rule) was annotated to each GOG (S1C Fig). From the assigned KOIDs, their KO BRITE functional category was assigned to each GOG. KOIDs for POGs and MOGs were retrieved the same way. For each KOID, the pathway names and BRITE categories at various level of resolutions were used for assigning functional categories manually to each OG. Manual assignment was necessary since BRITE names included a large fraction of categories such as "enzymes" and "exosomes." These were either not very informative or were misleading as many of "exosome" annotated

proteins took part in protein synthesis or folding. Lastly, for OGs or proteins discussed with respect to their physiological relevance, the functions were retrieved from the literature (cited in the text).

Inference of ancestral states

A phylogeny of Archaeplastidal species was inferred based on all genes conserved in all species, using "Species tree inference from all genes (STAG)" method [162], as a part of orthofinder analysis. STAG infers a species tree by taking greedy consensus of gene trees from each protein family (including that of multigene families). This phylogeny was rooted using minimal ancestral deviation [163] which places Rhodophyta as the sister to all others. Independently, the same unrooted phylogeny was manually rooted using FigTree (v1.4.4) [164] such that Glaucophyta were at the base. Ancestor state of presence and absence of organelle protein families across nodes were inferred using Phytool [165] package 0.7.80. Based on character state at the tips of the tree, Phytool inferred Bayesian posterior probabilities under a single rate model [166,167] of the character state across nodes of the tree. All Ogs that were present in major ancestors of plant groups with probability higher than 0.75 and absent in the preceding ancestor were considered as newly recruited in that lineage. Ogs or proteins discussed with respect to its physiological role in a given clade, their absence outside the group was verified in our copy number database as well as on homologue database available on TAIR.

Searching for potential RNA metabolism POGs and MOGs

Hidden Markov models (HMM) of PPR and mTERF domains were downloaded from pFAM [168] with the IDs: PF01535, PF12854, PF13041, PF13812, and PF02536. Each of these HMMS was used as a query to search against the full sequences of all proteins within each POG and MOG. If a given OG had more than 60% of individual proteins containing PPR or mTERF, the OG was annotated as RNA metabolism OG. Origin of such Ogs were traced using ASR as described above.

Supporting information

S1 Fig. Sorting and annotating orthogroups. Sorting of **(A)** the green orthogroups (GOGs) and **(B)** plastid orthogroups (POGs). In the first step (top boxes), source protein sequences from available species were clustered into protein families. Based on predetermined criteria (between the 2 boxes), the protein clusters were then separated into the clusters of interest (bottom boxes). Mitochondrial orthogroups (MOGs) were sorted the same as **(B)** and based on their presence in any experimental mitochondrial proteome. **(C)** For the functional annotations (for GOGs, POGs, and MOGs), KEGG KOID were translated from ProteinID of each protein present in each cluster and from across species. For species outside of the KEGG database (or some proteins within the KEGG database), no KOIDs are available, they are indicated by "N/A." The most frequent KOID within a cluster (e.g., KEGG ID 1 in the second circle) was assigned as the KOID for the entire cluster. **(D)** Ogs were sorted into PPR or mTERF containing Ogs by using the hidden Markov profiles of these domains as a query against all proteins in a given OG. If more than 60% of individual proteins inside an OG contained these domains, we labeled it as RNA editing domain.
(TIFF)

S2 Fig. Phylogeny of 204 eukaryotes. Inferred phylogeny of 204 eukaryotes, with major groups and ancestors of Archaeplastida indicated. The underlying data of this figure can be

found at https://zenodo.org/records/10855592.
(TIFF)

S3 Fig. Validation of ancestor state reconstruction (ASR). Validation of ancestor state reconstruction (ASR) approach on a control protein family of rbcS **(A)**. Number of POGs and MOGs gained by major ancestors, as per probability threshold of inclusion 0.65, 0.75, and 0.85 **(B).** Hidden Markov model-based validation of newly gained POGs of Z/E and Embryophyte ancestor **(C).** Charge of the first 20 amino acids across mitochondrial biogenesis related POGs **(D)**, with mTERF containing POGs present also in MOG shown in a darker shade. The underlying data of this figure can be found at https://zenodo.org/records/10855592.
(TIFF)

S4 Fig. Phylogeny of Archaeplastida with rhodophyta at the base. Inferred phylogeny of Archaeplastida (rhodophytes as the sister lineage to all others) with major ancestor nodes indicated with the arrows and major groups highlighted by labels on the right. The underlying data of this figure can be found at https://zenodo.org/records/10855592.
(TIFF)

S5 Fig. Phylogeny of Archaeplastida with laucophyte at the base. Inferred phylogeny of Archaeplastida (glaucophytes as the sister lineage to all others) with major ancestor nodes indicated with the arrows and major groups highlighted and labeled on the right side. The underlying data of this figure can be found at https://zenodo.org/records/10855592.
(TIFF)

S6 Fig. The evolution of POGs. Evolution of plastid orthogroup numbers across the Archaeplastida on a phylogeny with Rhodophyta **(A)** and Glaucophyta **(B)** as a basal branch. The underlying data of this figure can be found at https://zenodo.org/records/10855592.
(TIFF)

S7 Fig. The evolution of MOGs. Evolution of Mitochondrial orthogroup numbers across the Archaeplastida on a phylogeny with Rhodophyta **(A)** and Glaucophyta **(B)** as a basal branch. The underlying data of this figure can be found at https://zenodo.org/records/10855592.
(TIFF)

S8 Fig. ASR for RNA polymerase interacting proteins. Ancestor state reconstruction (ASR) **(A)** and gene copy numbers **(B)** for selected plastid encoded RNA polymerase interacting proteins (PAPs). The pie charts at each node represent the probability of presence (green) or absence (black) of a protein family in that node. The underlying data of this figure can be found at https://zenodo.org/records/10855592.
(TIFF)

S9 Fig. Copy number distribution of key proteins recruited around terrestrialization (white indicates absence of a protein, values above 50 were shown as the darkest shade). The underlying data of this figure can be found at https://zenodo.org/records/10855592.
(TIFF)

S1 Table. Clustering to identify plastid and mitochondria orthogroups (POGs and MOGs).
(XLSX)

S2 Table. Presence-absence of each orthogroup across all ancestors.
(XLSX)

S3 Table. Functions of newly recruited orthogroups. All available at: https://zenodo.org/records/10855592 [42].
(XLSX)

Acknowledgments

We acknowledge support from the high-performance computing cluster (HILBERT) at the Heinrich Heine University Düsseldorf and thank Michael Knopp (HHU Düsseldorf) for his help. We also thank Alice Barkan (University of Oregon) for her useful feedback on our bioRxiv preprint.

Author Contributions

Conceptualization: Parth K. Raval, Sven B. Gould.

Data curation: Parth K. Raval.

Formal analysis: Parth K. Raval, Alexander I. MacLeod.

Funding acquisition: Sven B. Gould.

Investigation: Parth K. Raval, Alexander I. MacLeod.

Methodology: Parth K. Raval, Alexander I. MacLeod.

Project administration: Sven B. Gould.

Resources: Sven B. Gould.

Supervision: Sven B. Gould.

Visualization: Parth K. Raval, Sven B. Gould.

Writing – original draft: Parth K. Raval, Alexander I. MacLeod, Sven B. Gould.

Writing – review & editing: Parth K. Raval, Alexander I. MacLeod, Sven B. Gould.

References

1. Schreiber M, Rensing SA, Gould SB. The greening ashore. Trends Plant Sci. 2022. https://doi.org/10.1016/j.tplants.2022.05.005 PMID: 35739050
2. Bowles AMC, Williamson CJ, Williams TA, Lenton TM, Donoghue PCJ. The origin and early evolution of plants. Trends Plant Sci. 2023; 28:312–329. https://doi.org/10.1016/j.tplants.2022.09.009 PMID: 36328872
3. Li L, Wang S, Wang H, Sahu SK, Marin B, Li H, et al. The genome of Prasinoderma coloniale unveils the existence of a third phylum within green plants. Nat Ecol Evol. 2020; 4:1220–1231. https://doi.org/10.1038/s41559-020-1221-7 PMID: 32572216
4. Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, et al. One thousand plant transcriptomes and the phylogenomics of green plants. Nature. 2019; 574:679–685. https://doi.org/10.1038/s41586-019-1693-2 PMID: 31645766
5. Li X, Hou Z, Xu C, Shi X, Yang L, Lewis LA, et al. Large Phylogenomic Data sets Reveal Deep Relationships and Trait Evolution in Chlorophyte Green Algae. Genome Biol Evol. 2021; 13:evab101–evab101. https://doi.org/10.1093/gbe/evab101 PMID: 33950183
6. Puginier C, Keller J, Delaux P-M. Plant–microbe interactions that have impacted plant terrestrializations. Plant Physiol. 2022; 190:72–84. https://doi.org/10.1093/plphys/kiac258 PMID: 35642902
7. de Vries S, Fürst-Jansen JMR, Irisarri I, Dhabalia Ashok A, Ischebeck T, Feussner K, et al. The evolution of the phenylpropanoid pathway entailed pronounced radiations and divergences of enzyme families. Plant J. 2021; 107:975–1002. https://doi.org/10.1111/tpj.15387 PMID: 34165823
8. Bowman JL. Stomata: Active Portals for Flourishing on Land. Curr Biol. 2011; 21:R540–R541. https://doi.org/10.1016/j.cub.2011.06.021 PMID: 21783030

9. Kong L, Liu Y, Zhi P, Wang X, Xu B, Gong Z, et al. Origins and Evolution of Cuticle Biosynthetic Machinery in Land Plants. Plant Physiol. 2020; 184:1998–2010. https://doi.org/10.1104/pp.20.00913 PMID: 32934149

10. Davies KM, Jibran R, Zhou Y, Albert NW, Brummell DA, Jordan BR, et al. The Evolution of Flavonoid Biosynthesis: A Bryophyte Perspective. Front Plant Sci. 2020; 11. https://doi.org/10.3389/fpls.2020.00007 PMID: 32117358

11. Cheng S, Xian W, Fu Y, Marin B, Keller J, Wu T, et al. Genomes of Subaerial Zygnematophyceae Provide Insights into Land Plant Evolution. Cell. 2019. https://doi.org/10.1016/j.cell.2019.10.019 PMID: 31730849

12. de Vries J, Stanton A, Archibald JM, Gould SB. Streptophyte Terrestrialization in Light of Plastid Evolution. Trends Plant Sci. 2016. https://doi.org/10.1016/j.tplants.2016.01.021 PMID: 26895731

13. de Vries J, Archibald JM. Plant evolution: landmarks on the path to terrestrial life. New Phytologist. 2018. https://doi.org/10.1111/nph.14975 PMID: 29318635

14. Knopp M, Garg SG, Handrich M, Gould SB. Major Changes in Plastid Protein Import and the Origin of the Chloroplastida. iScience. 2020. https://doi.org/10.1016/j.isci.2020.100896 PMID: 32088393

15. de Vries J, Gould SB. The monoplastidic bottleneck in algae and plant evolution. J Cell Sci. 2018. https://doi.org/10.1242/jcs.203414 PMID: 28893840

16. MacLeod AI, Raval PK, Stockhorst S, Knopp MR, Frangedakis E, Gould SB. Loss of Plastid Developmental Genes Coincides With a Reversion to Monoplastidy in Hornworts. Front Plant Sci. 2022; 13. Available from: https://www.frontiersin.org/article/10.3389/fpls.2022.863076. https://doi.org/10.3389/fpls.2022.863076 PMID: 35360315

17. Shimizu T, Kacprzak SM, Mochizuki N, Nagatani A, Watanabe S, Shimada T, et al. The retrograde signaling protein GUN1 regulates tetrapyrrole biosynthesis. Proc Natl Acad Sci U S A. 2019; 116:24900–24906. https://doi.org/10.1073/pnas.1911251116 PMID: 31732672

18. Foyer CH, Karpinska B, Krupinska K. The functions of WHIRLY1 and REDOX-RESPONSIVE TRANSCRIPTION FACTOR 1 in cross tolerance responses in plants: a hypothesis. Philos Trans R Soc Lond B Biol Sci. 2014; 369:20130226. https://doi.org/10.1098/rstb.2013.0226 PMID: 24591713

19. Bar-On YM, Phillips R, Milo R. The biomass distribution on Earth. Proc Natl Acad Sci U S A. 2018; 115:6506–6511. https://doi.org/10.1073/pnas.1711842115 PMID: 29784790

20. Pérez-Sancho J, Vanneste S, Lee E, McFarlane HE, Esteban del Valle A, Valpuesta V, et al. The Arabidopsis Synaptotagmin1 Is Enriched in Endoplasmic Reticulum-Plasma Membrane Contact Sites and Confers Cellular Resistance to Mechanical Stresses. Plant Physiol. 2015; 168:132–143. https://doi.org/10.1104/pp.15.00260 PMID: 25792253

21. Gao H, Metz J, Teanby NA, Ward AD, Botchway SW, Coles B, et al. In Vivo Quantification of Peroxisome Tethering to Chloroplasts in Tobacco Epidermal Cells Using Optical Tweezers. Plant Physiol. 2016; 170:263–272. https://doi.org/10.1104/pp.15.01529 PMID: 26518344

22. Wang Y, Selinski J, Mao C, Zhu Y, Berkowitz O, Whelan J. Linking mitochondrial and chloroplast retrograde signalling in plants. Philos Trans R Soc Lond B Biol Sci. 2020; 375:20190410–20190410. https://doi.org/10.1098/rstb.2019.0410 PMID: 32362265

23. Møller IM, Rasmusson AG, Van Aken O. Plant mitochondria–past, present and future. Plant J. 2021; 108:912–959. https://doi.org/10.1111/tpj.15495 PMID: 34528296

24. Kanai R, Edwards GE. The Biochemistry of C4 Photosynthesis. C4 Plant Biology. 1999. https://doi.org/10.1016/b978-012614440-6/50004-5

25. Peers G, Niyogi KK. Pond scum genomics: the genomes of Chlamydomonas and Ostreococcus. Plant Cell. 2008; 20:502–507. https://doi.org/10.1105/tpc.107.056556 PMID: 18359853

26. Christin PA, Osborne CP. The evolutionary ecology of C4 plants. New Phytologist. 2014. https://doi.org/10.1111/nph.13033 PMID: 25263843

27. Khan K, Van Aken O. The colonization of land was a likely driving force for the evolution of mitochondrial retrograde signalling in plants. Jones M, editor. J Exp Bot. 2022; 73:7182–7197. https://doi.org/10.1093/jxb/erac351 PMID: 36055768

28. Heinnickel ML, Grossman AR. The GreenCut: re-evaluation of physiological role of previously studied proteins and potential novel protein functions. Photosynth Res. 2013; 116:427–436. https://doi.org/10.1007/s11120-013-9882-6 PMID: 23873414

29. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 2012; 40:D1178–D1186. https://doi.org/10.1093/nar/gkr944 PMID: 22110026

30. Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, et al. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. Nat Commun. 2014. https://doi.org/10.1038/ncomms4978 PMID: 24865297

31. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016; 44:D733–D745. https://doi.org/10.1093/nar/gkv1189 PMID: 26553804

32. Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, et al. Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. Cell. 2017. https://doi.org/10.1016/j.cell.2017.09.030 PMID: 28985561

33. Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, Haas FB, et al. The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. Plant J. 2018; 93:515–533. https://doi.org/10.1111/tpj.13801 PMID: 29237241

34. Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK, et al. The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. Cell. 2018. https://doi.org/10.1016/j.cell.2018.06.033 PMID: 30007417

35. Li F-W, Nishiyama T, Waller M, Frangedakis E, Keller J, Li Z, et al. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants. 2020; 6:259–272. https://doi.org/10.1038/s41477-020-0618-2 PMID: 32170292

36. Wang S, Li L, Li H, Sahu SK, Wang H, Xu Y, et al. Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. Nat Plants. 2020. https://doi.org/10.1038/s41477-019-0560-3 PMID: 31844283

37. Zhang J, Fu XX, Li RQ, Zhao X, Liu Y, Li MH, et al. The hornwort genome and early land plant evolution. Nat Plants. 2020. https://doi.org/10.1038/s41477-019-0588-4 PMID: 32042158

38. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015; 12:59–60. https://doi.org/10.1038/nmeth.3176 PMID: 25402007

39. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 2015; 16:157–157. https://doi.org/10.1186/s13059-015-0721-2 PMID: 26243257

40. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019; 20:238–238. https://doi.org/10.1186/s13059-019-1832-y PMID: 31727128

41. Emms DM, Kelly S. SHOOT: phylogenetic gene search and ortholog inference. Genome Biol. 2022; 23:85–85. https://doi.org/10.1186/s13059-022-02652-8 PMID: 35346327

42. Raval PK, MacLeod AI, Gould SB. A molecular atlas of plastid and mitochondrial proteins reveals organellar remodeling during plant evolutionary transitions from algae to angiosperms. Zenedo. 2024. https://zenodo.org/records/10855592

43. Janouškovec J, Liu S-L, Martone PT, Carré W, Leblanc C, Collén J, et al. Evolution of Red Algal Plastid Genomes: Ancient Architectures, Introns, Horizontal Gene Transfer, and Taxonomic Utility of Plastid Markers. PLoS ONE. 2013; 8:e59001. https://doi.org/10.1371/journal.pone.0059001 PMID: 23536846

44. Strassert JFH, Irisarri I, Williams TA, Burki F. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. Nat Commun. 2021; 12:1879. https://doi.org/10.1038/s41467-021-22044-z PMID: 33767194

45. Price DC, Goodenough UW, Roth R, Lee J-H, Kariyawasam T, Mutwil M, et al. Analysis of an improved Cyanophora paradoxa genome assembly. DNA Res. 2019; 26:287–299. https://doi.org/10.1093/dnares/dsz009 PMID: 31098614

46. Wang S, Liang H, Xu Y, Li L, Wang H, Sahu DN, et al. Genome-wide analyses across Viridiplantae reveal the origin and diversification of small RNA pathway-related genes. Commun Biol. 2021; 4:412. https://doi.org/10.1038/s42003-021-01933-5 PMID: 33767367

47. Keeling PJ. The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. Annu Rev Plant Biol. 2013. https://doi.org/10.1146/annurev-arplant-050312-120144 PMID: 23451781

48. Ku C, Nelson-Sathi S, Roettger M, Sousa FL, Lockhart PJ, Bryant D, et al. Endosymbiotic origin and differential loss of eukaryotic genes. Nature. 2015; 524:427–432. https://doi.org/10.1038/nature14963 PMID: 26287458

49. Hao W, Liu G, Wang W, Shen W, Zhao Y, Sun J, et al. RNA Editing and Its Roles in Plant Organelles. Front Genet. 2021; 12. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2021.757109. https://doi.org/10.3389/fgene.2021.757109 PMID: 34659369

50. Small ID, Schallenberg-Rüdinger M, Takenaka M, Mireau H, Ostersetzer-Biran O. Plant organellar RNA editing: what 30 years of research has revealed. Plant J. 2020; 101:1040–1056. https://doi.org/10.1111/tpj.14578 PMID: 31630458

51. Barkan A. Expression of Plastid Genes: Organelle-Specific Elaborations on a Prokaryotic Scaffold. Plant Physiol. 2011; 155:1520–1532. https://doi.org/10.1104/pp.110.171231 PMID: 21346173

52. Barkan A, Small I. Pentatricopeptide Repeat Proteins in Plants. Annu Rev Plant Biol. 2014; 65:415–442. https://doi.org/10.1146/annurev-arplant-050213-040159 PMID: 24471833

53. Wobbe L. The Molecular Function of Plant mTERFs as Key Regulators of Organellar Gene Expression. Plant Cell Physiol. 2021; 61:2004–2017. https://doi.org/10.1093/pcp/pcaa132 PMID: 33067620

54. Robles P, Quesada V. Research Progress in the Molecular Functions of Plant mTERF Proteins. Cells. 2021; 10. https://doi.org/10.3390/cells10020205 PMID: 33494215

55. Robles P, Núñez-Delegido E, Ferrández-Ayela A, Sarmiento-Mañús R, Micol JL, Quesada V. Arabidopsis mTERF6 is required for leaf patterning. Plant Sci. 2018; 266:117–129. https://doi.org/10.1016/j.plantsci.2017.11.003 PMID: 29241561

56. Wu L, Zhou Z-Y, Zhang C-G, Chai J, Zhou Q, Wang L, et al. Functional Roles of Three Cutin Biosynthetic Acyltransferases in Cytokinin Responses and Skotomorphogenesis. PLoS ONE. 2015; 10: e0121943. https://doi.org/10.1371/journal.pone.0121943 PMID: 25803274

57. Chen X, Truksa M, Snyder CL, El-Mezawy A, Shah S, Weselake RJ. Three Homologous Genes Encoding sn-Glycerol-3-Phosphate Acyltransferase 4 Exhibit Different Expression Patterns and Functional Divergence in Brassica napus. Plant Physiol. 2011; 155:851–865. https://doi.org/10.1104/pp.110.169482 PMID: 21173024

58. Fernández-Santos R, Izquierdo Y, López A, Muñiz L, Martínez M, Cascón T, et al. Protein Profiles of Lipid Droplets during the Hypersensitive Defense Response of Arabidopsis against Pseudomonas Infection. Plant Cell Physiol. 2020; 61:1144–1157. https://doi.org/10.1093/pcp/pcaa041 PMID: 32219438

59. Xu C, Fan J, Cornish AJ, Benning C. Lipid Trafficking between the Endoplasmic Reticulum and the Plastid in Arabidopsis Requires the Extraplastidic TGD4 Protein. Plant Cell. 2008; 20:2190–2204. https://doi.org/10.1105/tpc.108.061176 PMID: 18689504

60. Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, et al. Acyl-Lipid Metabolism. The Arabidopsis Book. 2013; 2013. https://doi.org/10.1199/tab.0161 PMID: 23505340

61. Huang J, Xue C, Wang H, Wang L, Schmidt W, Shen R, et al. Genes of ACYL CARRIER PROTEIN Family Show Different Expression Profiles and Overexpression of ACYL CARRIER PROTEIN 5 Modulates Fatty Acid Composition and Enhances Salt Stress Tolerance in Arabidopsis. Front Plant Sci. 2017; 8:987. https://doi.org/10.3389/fpls.2017.00987 PMID: 28642782

62. Bonaventure G, Ohlrogge JB. Differential regulation of mRNA levels of acyl carrier protein isoforms in Arabidopsis. Plant Physiol. 2002; 128:223–235. PMID: 11788768

63. Figueroa CM, Asencion Diez MD, Ballicora MA, Iglesias AA. Structure, function, and evolution of plant ADP-glucose pyrophosphorylase. Plant Mol Biol. 2022; 108:307–323. https://doi.org/10.1007/s11103-021-01235-8 PMID: 35006475

64. Liu K, Zou W, Gao X, Wang X, Yu Q, Ge L. Young seedlings adapt to stress by retaining starch and retarding growth through ABA-Dependent and -independent pathways in Arabidopsis. Biochem Biophys Res Commun. 2019; 515:699–705. https://doi.org/10.1016/j.bbrc.2019.06.023 PMID: 31186142

65. Crevillén P, Ventriglia T, Pinto F, Orea A, Mérida Á, Romero JM. Differential Pattern of Expression and Sugar Regulation of Arabidopsis thaliana ADP-glucose Pyrophosphorylase-encoding Genes*. J Biol Chem. 2005; 280:8143–8149. https://doi.org/10.1074/jbc.M411713200 PMID: 15598655

66. Eliyahu E, Rog I, Inbal D, Danon A. ACHT4-driven oxidation of APS1 attenuates starch synthesis under low light intensity in *Arabidopsis* plants. Proc Natl Acad Sci U S A. 2015; 112:12876–12881. https://doi.org/10.1073/pnas.1515513112 PMID: 26424450

67. Pourtau N, Jennings R, Pelzer E, Pallas J, Wingler A. Effect of sugar-induced senescence on gene expression and implications for the regulation of senescence in Arabidopsis. Planta. 2006; 224:556–568. https://doi.org/10.1007/s00425-006-0243-y PMID: 16514542

68. Bahaji A, Almagro G, Ezquer I, Gámez-Arcas S, Sánchez-López ÁM, Muñoz FJ, et al. Plastidial Phosphoglucose Isomerase Is an Important Determinant of Seed Yield through Its Involvement in Gibberellin-Mediated Reproductive Development and Storage Reserve Biosynthesis in Arabidopsis. Plant Cell. 2018; 30:2082–2098. https://doi.org/10.1105/tpc.18.00312 PMID: 30099384

69. Seung D, Boudet J, Monroe J, Schreier TB, David LC, Abt M, et al. Homologs of PROTEIN TARGETING TO STARCH Control Starch Granule Initiation in Arabidopsis Leaves. Plant Cell. 2017; 29:1657–1677. https://doi.org/10.1105/tpc.17.00222 PMID: 28684429

70. Zheng M, Zhu C, Yang T, Qian J, Hsu Y-F. GSM2, a transaldolase, contributes to reactive oxygen species homeostasis in Arabidopsis. Plant Mol Biol. 2020; 104:39–53. https://doi.org/10.1007/s11103-020-01022-x PMID: 32564178

71. Melonek J, Mulisch M, Schmitz-Linneweber C, Grabowski E, Hensel G, Krupinska K. Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids. Planta. 2010; 232:471–481. https://doi.org/10.1007/s00425-010-1183-0 PMID: 20473685

72. Lepage E, Zampini E, Brisson N. Plastid Genome Instability Leads to Reactive Oxygen Species Production and Plastid-to-Nucleus Retrograde Signaling in Arabidopsis. Plant Physiol. 2013; 163:867–881. https://doi.org/10.1104/pp.113.223560 PMID: 23969600

73. Desveaux D, Subramaniam R, Després C, Mess J-N, Lévesque C, Fobert PR, et al. A "Whirly" Transcription Factor Is Required for Salicylic Acid-Dependent Disease Resistance in Arabidopsis. Dev Cell. 2004; 6:229–240. https://doi.org/10.1016/s1534-5807(04)00028-0 PMID: 14960277

74. Comadira G, Rasool B, Kaprinska B, García BM, Morris J, Verrall SR, et al. WHIRLY1 Functions in the Control of Responses to Nitrogen Deficiency But Not Aphid Infestation in Barley. Plant Physiol. 2015; 168:1140–1151. https://doi.org/10.1104/pp.15.00580 PMID: 25944826

75. Ren Y, Li Y, Jiang Y, Wu B, Miao Y. Phosphorylation of WHIRLY1 by CIPK14 Shifts Its Localization and Dual Functions in Arabidopsis. Mol Plant. 2017; 10:749–763. https://doi.org/10.1016/j.molp.2017.03.011 PMID: 28412544

76. Lee KP, Kim C, Landgraf F, Apel K. EXECUTER1- and EXECUTER2-dependent transfer of stress-related signals from the plastid to the nucleus of Arabidopsis thaliana. Proc Natl Acad Sci U S A. 2007; 104:10270–10275. https://doi.org/10.1073/pnas.0702061104 PMID: 17540731

77. Kim C, Meskauskiene R, Zhang S, Lee KP, Lakshmanan Ashok M, Blajecka K, et al. Chloroplasts of Arabidopsis Are the Source and a Primary Target of a Plant-Specific Programmed Cell Death Signaling Pathway. Plant Cell. 2012; 24:3026–3039. https://doi.org/10.1105/tpc.112.100479 PMID: 22797473

78. Uberegui E, Hall M, Lorenzo Ó, Schröder WP, Balsera M. An Arabidopsis soluble chloroplast proteomic analysis reveals the participation of the Executer pathway in response to increased light conditions. J Exp Bot. 2015; 66:2067–2077. https://doi.org/10.1093/jxb/erv018 PMID: 25740923

79. Maréchal A, Parent J-S, Véronneau-Lafortune F, Joyeux A, Lang BF, Brisson N. Whirly proteins maintain plastid genome stability in Arabidopsis. Proc Natl Acad Sci U S A. 2009; 106:14693–14698. https://doi.org/10.1073/pnas.0901710106 PMID: 19666500

80. Krupinska K, Desel C, Frank S, Hensel G. WHIRLIES Are Multifunctional DNA-Binding Proteins With Impact on Plant Development and Stress Resistance. Front Plant Sci. 2022; 13. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2022.880423. https://doi.org/10.3389/fpls.2022.880423 PMID: 35528945

81. Taylor RE, West CE, Foyer CH. WHIRLY protein functions in plants. Food Energy Secur. 2023; 12: e379. https://doi.org/10.1002/fes3.379 PMID: 38440693

82. Yu F, Park S-S, Liu X, Foudree A, Fu A, Powikrowska M, et al. SUPPRESSOR OF VARIEGATION4, a New var2 Suppressor Locus, Encodes a Pioneer Protein that Is Required for Chloroplast Biogenesis. Mol Plant. 2011; 4:229–240. https://doi.org/10.1093/mp/ssq074 PMID: 21220584

83. Powikrowska M, Khrouchtchova A, Martens HJ, Zygadlo-Nielsen A, Melonek J, Schulz A, et al. SVR4 (suppressor of variegation 4) and SVR4-like: two proteins with a role in proper organization of the chloroplast genetic machinery. Physiol Plant. 2014; 150:477–492. https://doi.org/10.1111/ppl.12108 PMID: 24111559

84. Wang S-W, Li Y, Zhang X-L, Yang H-Q, Han X-F, Liu Z-H, et al. Lacking chloroplasts in guard cells of crumpled leaf attenuates stomatal opening: both guard cell chloroplasts and mesophyll contribute to guard cell ATP levels. Plant Cell Environ. 2014; 37:2201–2210. https://doi.org/10.1111/pce.12297 PMID: 24506786

85. Jarvis P, López-Juez E. Biogenesis and homeostasis of chloroplasts and other plastids. Nat Rev Mol Cell Biol. 2013. https://doi.org/10.1038/nrm3702 PMID: 24263360

86. Osteryoung KW, Pyke KA. Division and dynamic morphology of plastids. Annu Rev Plant Biol. 2014. https://doi.org/10.1146/annurev-arplant-050213-035748 PMID: 24471836

87. Richardson LGL, Schnell DJ. Origins, function, and regulation of the TOC–TIC general protein import machinery of plastids. J Exp Bot. 2020; 71:1226–1238. https://doi.org/10.1093/jxb/erz517 PMID: 31730153

88. Chen L, Sun B, Gao W, Zhang Q, Yuan H, Zhang M. MCD1 Associates with FtsZ Filaments via the Membrane-Tethering Protein ARC6 to Guide Chloroplast Division. Plant Cell. 2018; 30:1807–1823. https://doi.org/10.1105/tpc.18.00189 PMID: 29967285

89. Muranaka A, Watanabe S, Sakamoto A, Shimada H. Arabidopsis cotyledon chloroplast biogenesis factor CYO1 uses glutathione as an electron donor and interacts with PSI (A1 and A2) and PSII (CP43 and CP47) subunits. J Plant Physiol. 2012; 169:1212–1215. https://doi.org/10.1016/j.jplph.2012.04.001 PMID: 22572242

90. Tominaga J, Mizutani H, Horikawa D, Nakahara Y, Takami T, Sakamoto W, et al. Rice CYO1, an ortholog of Arabidopsis thaliana cotyledon chloroplast biogenesis factor AtCYO1, is expressed in leaves and involved in photosynthetic performance. J Plant Physiol. 2016; 207:78–83. https://doi.org/10.1016/j.jplph.2016.10.005 PMID: 27835768

91. Yu Q-B, Lu Y, Ma Q, Zhao T-T, Huang C, Zhao H-F, et al. TAC7, an essential component of the plastid transcriptionally active chromosome complex, interacts with FLN1, TAC10, TAC12 and TAC14 to regulate chloroplast gene expression in Arabidopsis thaliana. Physiol Plant. 2013; 148:408–421. https://doi.org/10.1111/j.1399-3054.2012.01718.x PMID: 23082802

92. Asano T, Yoshioka Y, Kurei S, Sakamoto W, Machida Y. A mutation of the CRUMPLED LEAF gene that encodes a protein localized in the outer envelope membrane of plastids affects the pattern of cell division, cell differentiation, and plastid division in Arabidopsis. Plant J. 2004; 38:448–459. https://doi.org/10.1111/j.1365-313X.2004.02057.x PMID: 15086805

93. Chen Y, Asano T, Fujiwara MT, Yoshida S, Machida Y, Yoshioka Y. Plant Cells Without Detectable Plastids are Generated in the crumpled leaf Mutant of Arabidopsis thaliana. Plant Cell Physiol. 2009; 50:956–969. https://doi.org/10.1093/pcp/pcp047 PMID: 19318374

94. Smith MD, Rounds CM, Wang F, Chen K, Afitlhile M, Schnell DJ. atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. J Cell Biol. 2004; 165:323–334. https://doi.org/10.1083/jcb.200311074 PMID: 15138290

95. Day PM, Potter D, Inoue K. Evolution and targeting of Omp85 homologs in the chloroplast outer envelope membrane. Front Plant Sci. 2014; 5:535. https://doi.org/10.3389/fpls.2014.00535 PMID: 25352854

96. Agne B, Andrès C, Montandon C, Christ B, Ertan A, Jung F, et al. The Acidic A-Domain of Arabidopsis Toc159 Occurs as a Hyperphosphorylated Protein. Plant Physiol. 2010; 153:1016–1030. https://doi.org/10.1104/pp.110.158048 PMID: 20457805

97. Afitlhile M, Fry M, Workman S. The TOC159 mutant of Arabidopsis thaliana accumulates altered levels of saturated and polyunsaturated fatty acids. Plant Physiol Biochem. 2015; 87:61–72. https://doi.org/10.1016/j.plaphy.2014.12.018 PMID: 25557464

98. Afitlhile M, Worthington R, Heda G, Brown L. The TOC159 null mutant of Arabidopsis thaliana is impaired in the accumulation of plastid lipids and phosphatidylcholine. Plant Physiol Biochem. 2021; 159:148–159. https://doi.org/10.1016/j.plaphy.2020.12.011 PMID: 33360238

99. Afitlhile M, Worthington R, Baldric J. The toc132toc120 heterozygote mutant of Arabidopsis thaliana accumulates decreased levels of the major chloroplast lipids. Phytochemistry. 2021; 184:112652. https://doi.org/10.1016/j.phytochem.2020.112652 PMID: 33535085

100. Chiu C-C, Chen L-J, Li H. Pea Chloroplast DnaJ-J8 and Toc12 Are Encoded by the Same Gene and Localized in the Stroma. Plant Physiol. 2010; 154:1172–1182. https://doi.org/10.1104/pp.110.161224 PMID: 20841453

101. Ling Q, Jarvis P. Regulation of Chloroplast Protein Import by the Ubiquitin E3 Ligase SP1 Is Important for Stress Tolerance in Plants. Curr Biol. 2015; 25:2527–2534. https://doi.org/10.1016/j.cub.2015.08.015 PMID: 26387714

102. Harris BJ, Clark JW, Schrempf D, Szöllősi GJ, Donoghue PCJ, Hetherington AM, et al. Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. Nat Ecol Evol. 2022. https://doi.org/10.1038/s41559-022-01885-x PMID: 36175544

103. Brawley SH, Blouin NA, Ficko-Blean E, Wheeler GL, Lohr M, Goodson HV, et al. Insights into the red algae and eukaryotic evolution from the genome of Porphyra umbilicalis (Bangiophyceae, Rhodophyta). Proc Natl Acad Sci U S A. 2017; 114:E6361–E6370. https://doi.org/10.1073/pnas.1703088114 PMID: 28716924

104. Bhattacharya D, Price DC, Chan CX, Qiu H, Rose N, Ball S, et al. Genome of the red alga Porphyridium purpureum. Nat Commun. 2013; 4:1941. https://doi.org/10.1038/ncomms2931 PMID: 23770768

105. Collén J, Porcel B, Carré W, Ball SG, Chaparro C, Tonon T, et al. Genome structure and metabolic features in the red seaweed Chondrus crispus shed light on evolution of the Archaeplastida. Proc Natl Acad Sci U S A. 2013; 110:5247–5252. https://doi.org/10.1073/pnas.1221259110 PMID: 23503846

106. Qiu H, Price DC, Yang EC, Yoon HS, Bhattacharya D. Evidence of ancient genome reduction in red algae (Rhodophyta). J Phycol. 2015; 51:624–636. https://doi.org/10.1111/jpy.12294 PMID: 26986787

107. Lee J, Yang EC, Graf L, Yang JH, Qiu H, Zelzion U, et al. Analysis of the Draft Genome of the Red Seaweed Gracilariopsis chorda Provides Insights into Genome Size Evolution in Rhodophyta. Mol Biol Evol. 2018; 35:1869–1886. https://doi.org/10.1093/molbev/msy081 PMID: 29688518

108. Yang W, Pollard M, Li-Beisson Y, Beisson F, Feig M, Ohlrogge J. A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. Proc Natl Acad Sci U S A. 2010; 107:12040–12045. https://doi.org/10.1073/pnas.0914149107 PMID: 20551224

109. Yang W, Simpson JP, Li-Beisson Y, Beisson F, Pollard M, Ohlrogge JB. A land-plant-specific glycerol-3-phosphate acyltransferase family in arabidopsis: Substrate specificity, sn-2 preference, and evolution. Plant Physiol. 2012; 160:638–652. https://doi.org/10.1104/pp.112.201996 PMID: 22864585

110. Ball S, Colleoni C, Cenci U, Raj JN, Tirtiaux C. The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. J Exp Bot. 2011; 62:1775–1801. https://doi.org/10.1093/jxb/erq411 PMID: 21220783

111. Viola R, Nyvall P, Pedersén M. The unique features of starch metabolism in red algae. Proc Biol Sci. 2001; 268:1417–1422. https://doi.org/10.1098/rspb.2001.1644 PMID: 11429143

112. Lechner M, Rossmanith W, Hartmann RK, Thölken C, Gutmann B, Giegé P, et al. Distribution of ribonucleoprotein and protein-only RNase P in Eukarya. Mol Biol Evol. 2015; 32:3186–3193. https://doi.org/10.1093/molbev/msv187 PMID: 26341299

113. Gutmann B, Gobert A, Giegé P. PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. Genes Dev. 2012; 26:1022–1027. https://doi.org/10.1101/gad.189514.112 PMID: 22549728

114. Gobert A, Gutmann B, Taschner A, Gössringer M, Holzmann J, Hartmann RK, et al. A single Arabidopsis organellar protein has RNase P activity. Nat Struct Mol Biol. 2010; 17:740–744. https://doi.org/10.1038/nsmb.1812 PMID: 20473316

115. Ban T, Ke J, Chen R, Gu X, Tan MHE, Zhou XE, et al. Structure of a PLS-class Pentatricopeptide Repeat Protein Provides Insights into Mechanism of RNA Recognition. J Biol Chem. 2013; 288:31540–31548. https://doi.org/10.1074/jbc.M113.496828 PMID: 24047899

116. Wang X, Zhao L, Man Y, Li X, Wang L, Xiao J. PDM4, a Pentatricopeptide Repeat Protein, Affects Chloroplast Gene Expression and Chloroplast Development in Arabidopsis thaliana. Front Plant Sci. 2020; 11. Available from: https://www.frontiersin.org/article/10.3389/fpls.2020.01198. https://doi.org/10.3389/fpls.2020.01198 PMID: 32849743

117. Zoschke R, Qu Y, Zubo YO, Börner T, Schmitz-Linneweber C. Mutation of the pentatricopeptide repeat-SMR protein SVR7 impairs accumulation and translation of chloroplast ATP synthase subunits in Arabidopsis thaliana. J Plant Res. 2013; 126:403–414. https://doi.org/10.1007/s10265-012-0527-1 PMID: 23076438

118. Wu W, Liu S, Ruwe H, Zhang D, Melonek J, Zhu Y, et al. SOT1, a pentatricopeptide repeat protein with a small MutS-related domain, is required for correct processing of plastid 23S–4.5S rRNA precursors in Arabidopsis thaliana. Plant J. 2016; 85:607–621. https://doi.org/10.1111/tpj.13126 PMID: 26800847

119. De Longevialle AF, Hendrickson L, Taylor NL, Delannoy E, Lurin C, Badger M, et al. The pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid ycf3 intron 2 in Arabidopsis thaliana. Plant J. 2008; 56:157–168. https://doi.org/10.1111/j.1365-313X.2008.03581.x PMID: 18557832

120. Ye J-W, Gong Z-Y, Chen C-G, Mi H-L, Chen G-Y. A Mutation of OSOTP 51 Leads to Impairment of Photosystem I Complex Assembly and Serious Photo-damage in Rice. J Integr Plant Biol. 2012; 54:87–98. https://doi.org/10.1111/j.1744-7909.2012.01094.x PMID: 22353560

121. Huang W, Zhu Y, Wu W, Li X, Zhang D, Yin P, et al. The Pentatricopeptide Repeat Protein SOT5/EMB2279 Is Required for Plastid rpl2 and trnK Intron Splicing. Plant Physiol. 2018; 177:684–697. https://doi.org/10.1104/pp.18.00406 PMID: 29686056

122. Cyrek M, Fedak H, Ciesielski A, Guo Y, Sliwa A, Brzezniak L, et al. Seed Dormancy in Arabidopsis Is Controlled by Alternative Polyadenylation of DOG1. Plant Physiol. 2016; 170:947–955. https://doi.org/10.1104/pp.15.01483 PMID: 26620523

123. Mullineaux CW. Function and evolution of grana. Trends Plant Sci. 2005; 10:521–525. https://doi.org/10.1016/j.tplants.2005.09.001 PMID: 16169274

124. Xu X, Ouyang M, Lu D, Zheng C, Zhang L. Protein Sorting within Chloroplasts. Trends Cell Biol. 2021; 31:9–16. https://doi.org/10.1016/j.tcb.2020.09.011 PMID: 33121860

125. Stamenković M, Woelken E, Hanelt D. Ultrastructure of Cosmarium strains (Zygnematophyceae, Streptophyta) collected from various geographic locations shows species-specific differences both at optimal and stress temperatures. Protoplasma. 2014; 251:1491–1509. https://doi.org/10.1007/s00709-014-0652-x PMID: 24802109

126. Wietrzynski W, Schaffer M, Tegunov D, Albert S, Kanazawa A, Plitzko JM, et al. Charting the native architecture of chlamydomonas thylakoid membranes with single-molecule precision. eLife. 2020. https://doi.org/10.7554/eLife.53740 PMID: 32297859

127. Jiao C, Sørensen I, Sun X, Sun H, Behar H, Alseekh S, et al. The Penium margaritaceum Genome: Hallmarks of the Origins of Land Plants. Cell. 2020. https://doi.org/10.1016/j.cell.2020.04.019 PMID: 32442406

128. Khatoon M, Inagawa K, Pospíšil P, Yamashita A, Yoshioka M, Lundin B, et al. Quality Control of Photosystem II: Thylakoid unstacking is necessary to avoid further damage to the D1 protein and to facilitate D1 degradation under light stress in spinach thylakoids. J Biol Chem. 2009; 284:25343–25352. https://doi.org/10.1074/jbc.M109.007740 PMID: 19617353

129. Kirchhoff H, Hall C, Wood M, Herbstová M, Tsabari O, Nevo R, et al. Dynamic control of protein diffusion within the granal thylakoid lumen. Proc Natl Acad Sci U S A. 2011; 108:20248–20253. https://doi.org/10.1073/pnas.1104141109 PMID: 22128333

130. Garg SG, Gould SB. The Role of Charge in Protein Targeting Evolution. Trends Cell Biol. 2016; 26:894–905. https://doi.org/10.1016/j.tcb.2016.07.001 PMID: 27524662

131. Nakanishi H, Suzuki K, Kabeya Y, Miyagishima S. Plant-Specific Protein MCD1 Determines the Site of Chloroplast Division in Concert with Bacteria-Derived MinD. Curr Biol. 2009; 19:151–156. https://doi.org/10.1016/j.cub.2008.12.018 PMID: 19135368

132. Larkin RM, Stefano G, Ruckle ME, Stavoe AK, Sinkler CA, Brandizzi F, et al. REDUCED CHLOROPLAST COVERAGE genes from Arabidopsis thaliana help to establish the size of the chloroplast compartment. Proc Natl Acad Sci U S A. 2016; 113:E1116–E1125. https://doi.org/10.1073/pnas.1515741113 PMID: 26862170

133. El Zawily AM, Schwarzländer M, Finkemeier I, Johnston IG, Benamar A, Cao Y, et al. FRIENDLY Regulates Mitochondrial Distribution, Fusion, and Quality Control in Arabidopsis. Plant Physiol. 2014; 166:808–828. https://doi.org/10.1104/pp.114.243824 PMID: 25165398

134. Zhu Q, Hulen D, Liu T, Clarke M. The cluA- mutant of Dictyostelium identifies a novel class of proteins required for dispersion of mitochondria. Proc Natl Acad Sci U S A. 1997; 94:7308–7313. https://doi.org/10.1073/pnas.94.14.7308 PMID: 9207087

135. Raval PK, Ngan WY, Gallie J, Agashe D. The layered costs and benefits of translational redundancy. Pilpel Y, Barkai N, Schirman D, editors. eLife. 2023; 12:e81005. https://doi.org/10.7554/eLife.81005 PMID: 36862572

136. Roller BRK, Stoddard SF, Schmidt TM. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. Nat Microbiol. 2016; 1:16160. https://doi.org/10.1038/nmicrobiol.2016.160 PMID: 27617693

137. Raval PK, Garg SG, Gould SB. Endosymbiotic selective pressure at the origin of eukaryotic cell biology. Chacinska A, Perry GH, Chacinska A, Perry GH, editors. eLife. 2022; 11:e81033. https://doi.org/10.7554/eLife.81033 PMID: 36355038

138. Raval PK, Martin WF, Gould SB. Mitochondrial evolution: Gene shuffling, endosymbiosis, and signaling. Sci Adv. 2023; 9:eadj4493. https://doi.org/10.1126/sciadv.adj4493 PMID: 37556561

139. Gould SB, Garg SG, Martin WF. Bacterial Vesicle Secretion and the Evolutionary Origin of the Eukaryotic Endomembrane System. Trends Microbiol. 2016; 24:525–534. https://doi.org/10.1016/j.tim.2016.03.005 PMID: 27040918

140. Hoecker U. The activities of the E3 ubiquitin ligase COP1/SPA, a key repressor in light signaling. Curr Opin Plant Biol. 2017; 37:63–69. https://doi.org/10.1016/j.pbi.2017.03.015 PMID: 28433946

141. Borden KLB. RING domains: master builders of molecular scaffolds? Wright PE, editor. J Mol Biol. 2000; 295:1103–1112. https://doi.org/10.1006/jmbi.1999.3429 PMID: 10653689

142. Stone SL. The role of ubiquitin and the 26S proteasome in plant abiotic stress signaling. Front Plant Sci. 2014; 5. https://doi.org/10.3389/fpls.2014.00135 PMID: 24795732

143. Jiménez-López D, Muñóz-Belman F, González-Prieto JM, Aguilar-Hernández V, Guzmán P. Repertoire of plant RING E3 ubiquitin ligases revisited: New groups counting gene families and single genes. PLoS ONE. 2018; 13:e0203442. https://doi.org/10.1371/journal.pone.0203442 PMID: 30169501

144. Thayale Purayil F, Sudalaimuthuasari N, Li L, Aljneibi R, Al Shamsi AM, David N, et al. Transcriptome Profiling and Functional Validation of RING-Type E3 Ligases in Halophyte Sesuvium verrucosum under Salinity Stress. Int J Mol Sci. 2022; 23. https://doi.org/10.3390/ijms23052821 PMID: 35269961

145. Basnayake BMVS, Li D, Zhang H, Li G, Virk N, Song F. Arabidopsis DAL1 and DAL2, two RING finger proteins homologous to Drosophila DIAP1, are involved in regulation of programmed cell death. Plant Cell Rep. 2011; 30:37–48. https://doi.org/10.1007/s00299-010-0941-6 PMID: 20972793

146. Ling Q, Huang W, Baldwin A, Jarvis P. Chloroplast Biogenesis Is Regulated by Direct Action of the Ubiquitin-Proteasome System. Science. 2012; 338:655–659. https://doi.org/10.1126/science.1225053 PMID: 23118188

147. Pan R, Satkovich J, Hu J. E3 ubiquitin ligase SP1 regulates peroxisome biogenesis in Arabidopsis. Proc Natl Acad Sci U S A. 2016; 113:E7307–E7316. https://doi.org/10.1073/pnas.1613530113 PMID: 27799549

148. Oikawa K, Matsunaga S, Mano S, Kondo M, Yamada K, Hayashi M, et al. Physical interaction between peroxisomes and chloroplasts elucidated by in situ laser analysis. Nat Plants. 2015; 1:15035. https://doi.org/10.1038/nplants.2015.35 PMID: 27247035

149. Liu H, Stone SL. Abscisic Acid Increases Arabidopsis ABI5 Transcription Factor Levels by Promoting KEG E3 Ligase Self-Ubiquitination and Proteasomal Degradation. Plant Cell. 2010; 22:2630–2641. https://doi.org/10.1105/tpc.110.076075 PMID: 20682837

150. Liu H, Stone SL. Cytoplasmic Degradation of the Arabidopsis Transcription Factor ABSCISIC ACID INSENSITIVE 5 Is Mediated by the RING-type E3 Ligase KEEP ON GOING*. J Biol Chem. 2013; 288:20267–20279. https://doi.org/10.1074/jbc.M113.465369 PMID: 23720747

151. Azevedo J, Courtois F, Hakimi M-A, Demarsy E, Lagrange T, Alcaraz J-P, et al. Intraplastidial trafficking of a phage-type RNA polymerase is mediated by a thylakoid RING-H2 protein. Proc Natl Acad Sci U S A. 2008; 105:9123–9128. https://doi.org/10.1073/pnas.0800909105 PMID: 18567673

152. Huang T, Pan Y, Maréchal E, Hu H. Proteomes reveal the lipid metabolic network in the complex plastid of *Phaeodactylum tricornutum*. Plant J. 2024; 117:385–403. https://doi.org/10.1111/tpj.16477 PMID: 37733835

153. Boucher MJ, Ghosh S, Zhang L, Lal A, Jang SW, Ju A, et al. Integrative proteomics and bioinformatic prediction enable a high-confidence apicoplast proteome in malaria parasites. Striepen B, editor. PLoS Biol. 2018; 16:e2005895. https://doi.org/10.1371/journal.pbio.2005895 PMID: 30212465

154. Mulvey CM, Breckels LM, Geladaki A, Britovšek NK, Nightingale DJH, Christoforou A, et al. Using hyperLOPIT to perform high-resolution mapping of the spatial proteome. Nat Protoc. 2017; 12:1110–1135. https://doi.org/10.1038/nprot.2017.026 PMID: 28471460

155. Xu S-L, Shrestha R, Karunadasa SS, Xie P-Q. Proximity Labeling in Plants. Annu Rev Plant Biol. 2023; 74:285–312. https://doi.org/10.1146/annurev-arplant-070522-052132 PMID: 36854476

156. Naramoto S, Hata Y, Fujita T, Kyozuka J. The bryophytes Physcomitrium patens and Marchantia polymorpha as model systems for studying evolutionary cell and developmental biology in plants. Plant Cell. 2021; 34:228–246. https://doi.org/10.1093/plcell/koab218 PMID: 34459922

157. Szövényi P, Frangedakis E, Ricca M, Quandt D, Wicke S, Langdale JA. Establishment of Anthoceros agrestis as a model species for studying the biology of hornworts. BMC Plant Biol. 2015; 15:98. https://doi.org/10.1186/s12870-015-0481-x PMID: 25886741

158. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017; 45:D353–D361. https://doi.org/10.1093/nar/gkw1092 PMID: 27899662

159. Terashima M, Specht M, Hippler M. The chloroplast proteome: a survey from the Chlamydomonas reinhardtii perspective with a focus on distinctive features. Curr Genet. 2011; 57:151–168. https://doi.org/10.1007/s00294-011-0339-1 PMID: 21533645

160. Mueller SJ, Lang D, Hoernstein SNW, Lang EGE, Schuessele C, Schmidt A, et al. Quantitative Analysis of the Mitochondrial and Plastid Proteomes of the Moss Physcomitrella patens Reveals Protein Macrocompartmentation and Microcompartmentation. Plant Physiol. 2014; 164:2081–2095. https://doi.org/10.1104/pp.114.235754 PMID: 24515833

161. Sun Q, Zybailov B, Majeran W, Friso G, Olinares PDB, van Wijk KJ. PPDB, the Plant Proteomics Database at Cornell. Nucleic Acids Res. 2009; 37:D969–D974. https://doi.org/10.1093/nar/gkn654 PMID: 18832363

162. Emms DM, Kelly S. STAG: Species Tree Inference from All Genes. bioRxiv. 2018;267914–267914. https://doi.org/10.1101/267914

163. Tria FDK, Landan G, Dagan T. Phylogenetic rooting using minimal ancestor deviation. Nat Ecol Evol. 2017; 1:0193. https://doi.org/10.1038/s41559-017-0193 PMID: 29388565

164. Rambaut A, Drummond A. FigTree Version 1.4.4. 2018.

165. Revell LJ. phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012. https://doi.org/10.1111/j.2041-210X.2011.00169.x

166. Mooers AØ, Schluter D. Reconstructing Ancestor States with Maximum Likelihood: Support for One- and Two-Rate Models. Syst Biol. 1999; 48:623–633. https://doi.org/10.1080/106351599260193

167. Schluter D, Price T, Mooers AØ, Ludwig D. Likelihood of ancestor states in adaptive radiation. Evolution. 1997; 51:1699–1711. https://doi.org/10.1111/j.1558-5646.1997.tb05095.x PMID: 28565128

168. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. Nucleic Acids Res. 2021; 49:D412–D419. https://doi.org/10.1093/nar/gkaa913 PMID: 33125078