Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte?

Tereza Ševčíková¹,⁎, Aleš Horák²,⁎, Vladimír Klimeš¹, Veronika Zbránková¹, Elif Demir-Hilton⁴, Sebastian Sudek⁴, Jerry Jenkins⁵, Jeremy Schmutz⁶, Pavel Přibyl⁷, Jan Fousek⁸, Čestmír Vlček⁸, B. Franz Lang⁹, Miroslav Oborník²,³, Alexandra Z. Worden⁴,¹⁰ & Marek Eliáš⁵

Algae with secondary plastids of a red algal origin, such as ochrophytes (photosynthetic stramenopiles), are diverse and ecologically important, yet their evolutionary history remains controversial. We sequenced plastid genomes of two ochrophytes, Ochromonas sp. CCMP1393 (Chrysophyceae) and Trachydiscus minutus (Eustigmatophyceae). A shared split of the clpC gene as well as phylogenomic analyses of concatenated protein sequences demonstrated that chrysophytes and eustigmatophytes form a clade, the Limnista, exhibiting an unexpectedly elevated rate of plastid gene evolution. Our analyses also indicate that the root of the ochrophyte phylogeny falls between the recently redefined Khakista and Phaeista assemblages. Taking advantage of the expanded sampling of plastid genome sequences, we revisited the phylogenetic position of the plastid of Vitrella brassicaformis, a member of Alveolata with the least derived plastid genome known for the whole group. The results varied depending on the dataset and phylogenetic method employed, but suggested that the Vitrella plastids emerged from a deep ochrophyte lineage rather than being derived vertically from a hypothetical plastid-bearing common ancestor of alveolates and stramenopiles. Thus, we hypothesize that the plastid in Vitrella, and potentially in other alveolates, may have been acquired by an endosymbiosis of an early ochrophyte.

¹University of Ostrava, Faculty of Science, Department of Biology and Ecology, Life Science Research Centre, Chittussiho 10, 710 00 Ostrava, Czech Republic. ²Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branišovská 31, 370 05 České Budějovice, Czech Republic. ³University of South Bohemia, Faculty of Science, Branišovská 1760, 370 05 České Budějovice, Czech Republic. ⁴Monterey Bay Aquarium Research Institute (MBARI), Moss Landing, CA 95039, USA. ⁵US Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA. ⁶HudsonAlpha Institute for Biotechnology, 601 Genome Way NW, Huntsville, Alabama 35806, USA. ⁷Centre for Algology and Biorefinery Research Centre of Competence, Institute of Botany, Czech Academy of Sciences, Dukelská 135, 379 82 Třeboň, Czech Republic. ⁸Département de Biochimie, Centre Robert-Cedergren, Université de Montréal, 2900 Boulevard Edouard Montpetit, Montréal, Québec, H3C 3J7, Canada. ⁹Integrated Microbial Biodiversity Program, Canadian Institute for Advanced Research, Toronto, M5G 1Z8, Canada. ⁎These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.Z.W. (e-mail: azworden@mbari.org) or M.E. (e-mail: marek.elias@osu.cz)
The evolutionary history of photosynthetic eukaryotes is astonishingly complex. One of the most puzzling aspects is the evolution of plastids derived from red algae (rhodophytes) by eukaryote-to-eukaryote endosymbioses. Phylogenetic analyses and a number of shared features indicate a single origin of these plastids, implying a scenario in which an early red alga became integrated into a heterotrophic eukaryotic host cell and ultimately gave rise to diverse algal lineages as well as non-photosynthetic lineages that subsequently lost their plastids. However, how the distribution of red alga-derived plastids was achieved across these various lineages remains controversial.

The Ochrophyta, a monophyletic phylum within the Stramenopiles (or Heterokonta), are the most diverse algal group with secondary plastids of algal origin in terms of morphology, pigmentation and phylogeny. Diatoms (Bacillariophyceae) and multicellular brown algae (Phaeophyceae) are the best characterized ochrophytes, but at least 15 separate “classes”, plus several isolated smaller lineages of uncertain taxonomic status exist. Several groups within the ochrophytes have important ecological roles, particularly in terms of marine photosynthesis and uptake of CO₂.

Understanding relationships among the ochrophytes is important for both ecological and evolutionary studies. Phylogenetic studies of nuclear SSU rRNA gene sequences proposed the existence of several higher-order clades, specifically: diatoms plus the Bolidophyceae; the Chrysophyceae plus the Symbromophyceae; and the large “PX” clade comprising brown algae, the Xanthophyceae and other less well-known groups. Multi-gene datasets including various combinations of nuclear, plastid and mitochondrial genes have improved resolution of ochrophyte phylogenetic relationships, yet some relationships are still unresolved. Most notably, the position of the root of the ochrophyte phylogeny remains uncertain. Despite these uncertainties and sometimes lack of statistical support, many different putative groupings of ochrophyte classes have been proposed. This includes the Limnista, which contain the classes Eustigmatophyceae and Synchromophyceae along with a few other minor lineages.

Regardless of occasional horizontal movement of plastids between different eukaryotic lineages (see below), it is generally assumed that within most algal and plant lineages plastids are inherited vertically. The growing list of completely sequenced plastid genomes of various algal and plant lineages is thus a resource for inferring relationships among both plastids and host cell lineages. Given the lack of evidence for non-vertical inheritance of ochrophyte plastids, plastid genome sequences could be also very helpful in resolving the ochrophyte phylogeny, but only seven of the approximately 15 known ochrophyte classes are represented by completely sequenced plastid genomes (Table S1). Furthermore, sampling is limited to a single genus in the case of eustigmatophytes, and to a single species in the raphidophytes and xanthophytes. Representatives of other classes, such as the Chrysophyceae, Pingeio phyceae and Dichthyochophyceae, are yet to be sequenced.

Two different conceptual frameworks are often used to explain the emergence of ochrophytes. One is the “chromalveolate hypothesis”, which posits that all extant groups with red-algal derived plastids, namely the Ochrophyta, Myxozoa (a subgroup of the Alveolata that includes the Apicomplexa, Dinoflagellata, “chromerid” algae and a few additional minor lineages), Haptophyta, and Cryptophyta (a subgroup of the Cryptista along with some non-photosynthetic lineages), inherited their plastids vertically from a common ancestor. Plastid lacking lineages closely related to any of these groups (i.e. plastid lacking stramenopiles, alveolates and cryptists) would have lost the plastid secondarily. However, phylogenetic and phylogenomic analyses have failed to provide evidence for the monophyly of the proposed “chromalveolate” lineages. Whereas stramenopiles and alveolates are specifically related to the Rhizaria and cryptists and haptophytes do not show any robustly supported affiliation and in some analyses are even found nested among eukaryotes with the primary plastid (Archaeplastida).

The second conceptual framework builds on a growing amount of data suggesting that a red algal plastid was acquired by one “chromalveolate” lineage, from which it spread to others through a series of higher-order tertiary or even quaternary endosymbioses (see, e.g., 21–25). This scenario has recently been dubbed the “rhodoplex hypothesis”. However, higher-order endosymbiotic gains of plastids postulated by the rhodoplex hypothesis have only been conclusively demonstrated for some dinoflagellate lineages.

Here, we sequenced and analyzed plastid genomes from two unrepresented ochrophyte groups. Trachydiscus minutus CCALA 838 belongs to a newly recognized eustigmatophyte subgroup (Goniochloridales) and is deeply diverged from the biotechnologically significant genus Nannochloropsis. Ochromonas sp. CCMP1393 is a marine member of the Chrysophyceae (Fig. S1), a group long known from freshwater habitats, but recently shown to be widely distributed in marine environments and important for primary production. We performed comparative and phylogenetic analyses of these genomes in the context of existing plastid sequence data to address questions on relationships among ochrophytes as well as on the plastid evolution in “chromalveolates”.

**Material and methods**

**Sequencing and annotation of the plastid genomes.** Cultivation of the algae and DNA isolation followed standard protocols. Sequencing of total DNA preparations employed the 454 and Illumina platforms. Initial assemblies of the reads were searched to identify scaffolds corresponding to the plastid genome and the final assembly of the plastid genome sequences was achieved by manual gap filling and polishing. MFannot (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl) was used for obtaining an initial automated annotation of the assembled plastid genome sequences, which
was checked and improved manually. Details on the cultivation, DNA isolation, sequencing, assembly, and annotation are provided in Supplementary methods. Annotated plastid genome sequences from *T. minutus* and *Ochromonas* sp. CCMP1393 are deposited in the GenBank database (accession numbers KJ624065 and KJ877675, respectively).

**Phylogenetic analyses**

68 orthologous protein sequences encoded by plastid genomes were aligned and concatenated, leaving 16,948 reliably aligned amino acid (aa) positions in the main ("Full") alignment (hereafter referred to as dataset F). Removal of the fastest-evolving category of sites from the full dataset F yielded the dataset SP ("Slow Positions") comprising 11,208 aa positions. The dataset SG ("Slow Genes" comprising 14,699 reliably aligned aa positions was built by aligning only protein sequences encoded by 34 slowly evolving plastid genes defined in a previous study.

Phylogenetic trees were inferred using two different methods: ML (RAxML 7.4.8a) with the site-homogeneous GTRGAMMA model of amino acid substitution, and Bayesian inference (PhyloBayes 3.3f) with the site-heterogeneous CAT-GTR model. A detailed description of the procedures used to build the alignments and infer trees is available in Supplementary methods.

**Results and Discussion**

The plastid genomes of *Trachydiscus minutus* and *Ochromonas* sp. CCMP1393 reveal split clpC genes and rapid gene evolution shared by eustigmatophytes and chrysophytes. The newly sequenced plastid genomes exhibit size, GC content, and gene content similar to previously sequence plastid genomes (see Tables 1 and S2, Fig. S2). Both genomes also display the typical circular-mapping architecture and the presence of inverted repeat (IR) regions separated by a large single copy and a small single copy region. The IR copies are identical at the nucleotide sequence level except for one-nucleotide deletion in one of the *psaC* copies in *Trachydiscus* introducing a frame-shift into the coding sequence; this copy is thus probably a non-functional pseudogene. The IR region of the *Ochromonas* plastid genome is the second longest among ochrophyte plastid genomes sequenced so far, while the small single copy region is exceptionally short (805 bp) and includes only two protein-coding genes.

The set of genes in the plastid genomes of *Trachydiscus minutus* and *Ochromonas* is similar to other ochrophytes in terms of both their number and identity (Tables 1, S2, and S3). The most significant observations concerning the gene complement are discussed below, while additional details are provided in Supplementary Note 1.

One notable feature is the presence of a predicted group I intron in one of the three tRNA-Leu genes in the *Trachydiscus* genome. A tRNA-Leu gene intron has been postulated to have been present in the cyanobacterial ancestor of plastids and in time was lost, presumably independently, in many plastid-containing lineages. Although overlooked in previous publications, we identified a tRNA-Leu gene in *Nannochloropsis* plastid genomes that also contains a group I intron (Fig. S3). This suggests that the intron may be a common feature in eustigmatophytes. In contrast, no introns were identified in *Ochromonas* plastid tRNA-Leu genes. This result is in accord with the previous indication that the intron was lost in the chrysophyte ancestor.

Other interesting features were observed as well. For example, plastid genomes of *Trachydiscus* and *Nannochloropsis* spp. share possession of the gene ycf49. This gene codes for small uncharacterized proteins possessing the DUF2499 domain of an unknown function and is also found in primary plastid genomes from cyanidiophyte red algae and the glaucophyte *Cyanophora paradoxa*, but no other published plastid

| T. minutus CCALA 838 | Ochromonas sp. CCMP1393 |
|----------------------|-------------------------|
| Size (bp)            | 120,090                 | 126,750                  |
| Inverted repeat (bp) | 9,412/ 9,411            | 22,910                   |
| Small single-copy region (bp) | 45,210        | 805                      |
| Large single-copy region (bp) | 56,060         | 80,130                   |
| Total GC content (%)  | 34.0                    | 30.9                     |
| Gene content (total)  | 163                     | 154                      |
| Identified protein-coding genes | 129              | 121                      |
| Unknown or hypothetical ORFs | 3               | 5                        |
| tRNA genes           | 3                       | 3                        |
| tRNA genes           | 28                      | 25                       |

Table 1. Basic characteristics of the newly sequenced plastid genomes. The gene counts ignore the presence of duplicated genes in inverted repeats; the split clpC_A and clpC_B genes in *Trachydiscus minutus* are counted as two separate genes.
genomes (primary or secondary; Table S4). This is the case even when sensitive PSI-BLAST searches are employed. In both cyanidiophytes and eustigmatophytes ycf49 genes are located in the same conserved plastid gene block petL-ycf49-ycf4-trnG-psbE-psbF-psbl-psbL. Thus, the ycf49 genes in the eustigmatophyte plastid genomes were apparently inherited directly from the plastid genome of the red algal ancestor of the "chromalveolate" plastids, while most "chromalveolate" lineages lost the ycf49 gene from their plastid genomes independently on several occasions.

Additionally, results were interesting for the gene clpC, which encodes a member of the Clp/Hsp100 family of AAA + proteins involved in the protein degradation pathway mediated by the ClpP protease. Starkenburg et al. recently demonstrated that in Nannochloropsis spp. clpC is split into three separate genes, annotated as clpN (encoding the N-terminal domain of a typical ClpC protein), clpC1 (encoding the first AAA ATPase domain), and clpC2 (encoding the second AAA ATPase domain). We identified these three genes in the Trachydiscus plastid genome (with clpC2 gene present in two identical copies, since it resides in the IR region). This indicates that the split of clpC into three genes must have occurred before the radiation of known eustigmatophytes. Furthermore, the Ochromonas plastid genome exhibits an intermediate state, as it harbours a separate clpN gene, but the rest of the clpC gene is intact and codes for both AAA ATPase domains. Our results demonstrate that the split of the 5'-end of the clpC gene most likely predates the divergence of eustigmatophytes and chrysophytes, but without a plastid genome from the Pinguioiphyceae, a putative sister group of the Limnista, it cannot yet be ascertained whether this split is exclusive (i.e. synapomorphic) for the Limnista.

While inspecting multiple alignments of protein sequences encoded by the plastid genes we also noted that the sequences from eustigmatophytes and/or Ochromonas tend to harbour unusual indels in regions otherwise well conserved among ochrophytes (see Fig. S4), which may be an indication that the evolutionary rates of plastid genes in eustigmatophytes are elevated compared to the rates exhibited by other ochrophytes. Indeed, phylogenetic analyses using individual genes revealed noticeably longer branches for eustigmatophyte and Ochromonas than for other ochrophytes (Fig. S5), indicating an increased rate of substitutions in the former lineages.

Phylogenomic analyses of plastid genomes support the Limnista clade
The most recent phylogenomic analyses of algal plastid genomes did not include data from chrysophytes and eustigmatophytes. Therefore, we conducted analyses using genes encoded by the new plastid genome sequences along with other recently published sequences. We used three different concatenated alignments (I, SG, and SP) and two methods of phylogenetic inference (ML and Bayesian inference). These employed site-homogeneous and site-heterogeneous substitution models to evaluate the robustness of the results (see Material and Methods and Supplementary methods for technical details).

The six trees obtained from the three different datasets and two methods were generally congruent with the previous analyses and with each other (Fig. 1, S6, and S7, Table 2; informative aspects of our analyses that are not directly related to ochrophytes are provided in Supplementary Note 2). Both ML and Bayesian trees were generally congruent with the previous analyses and with each other (Fig. 1, S6, and S7, Table 2; informative aspects of our analyses that are not directly related to ochrophytes are provided in Supplementary Note 2). Both ML and Bayesian trees were generally congruent with the previous analyses and with each other (Fig. 1, S6, and S7, Table 2; informative aspects of our analyses that are not directly related to ochrophytes are provided in Supplementary Note 2). Both ML and Bayesian trees were generally congruent with the previous analyses and with each other (Fig. 1, S6, and S7, Table 2; informative aspects of our analyses that are not directly related to ochrophytes are provided in Supplementary Note 2). Both ML and Bayesian trees were generally congruent with the previous analyses and with each other (Fig. 1, S6, and S7, Table 2; informative aspects of our analyses that are not directly related to ochrophytes are provided in Supplementary Note 2). Both ML and Bayesian trees were generally congruent with the previous analyses and with each other (Fig. 1, S6, and S7, Table 2; informative aspects of our analyses that are not directly related to ochrophytes are provided in Supplementary Note 2). Both ML and Bayesian trees were generally congruent with the previous analyses and with each other (Fig. 1, S6, and S7, Table 2; informative aspects of our analyses that are not directly related to ochrophytes are provided in Supplementary Note 2).
Figure 1. Plastid phylogeny inferred from protein sequences encoded by 34 slowly-evolving conserved plastid genes (dataset SG, 14,699 aa positions). (A) Maximum-likelihood tree (RAxML, GTRGAMMA model); only the “chromalveolate” subtree is shown for simplicity. Thick branches received 100% bootstrap support, otherwise the bootstrap support values are indicated by numbers when higher than 50%. (B) PhyloBayes tree inferred using the CAT-GTR model. Thick branches were supported by 1.00 posterior probability and 100% bootstrap support values from the ML analyses, otherwise posterior probabilities / bootstrap support values are indicated by numbers when higher than 0.90 / 50%; “d.t.” means that the respective bipartition does not exist in the ML tree.
Table 2. Summary of phylogenomic analyses performed in this study. Support values correspond to bootstrap percentages for ML analyses and posterior probabilities for PhyloBayes analyses. "d.t." means that a different topology not exhibiting the respective grouping was recovered in the analysis. F ("Full") comprised of 69 conserved proteins encoded by a majority of the plastid genomes (16,948 reliably aligned positions). SG ("Slow Genes") included sequences of a subset of 34 slowly-evolving genes (14,699 aa positions). SP ("Slow Positions") was created from the F dataset by removing sites exhibiting the highest substitution rates (11,208 positions). "+V" means that genes from Vitrella brassicaformis were included in the dataset. "-K" means the genes from Karlodinium veneficum were removed from the dataset.

Identifying the root of the ochrophyte phylogeny

In addition to resolving the Limnista, our analyses provide insights into other parts of the ochrophyte phylogeny. All trees provided maximal support for the sisterhood of brown algae and Vaucheria litorea representing the Xanthophyceae. This result agrees with previous single-gene and multi-gene analyses that established the existence of the so-called "PX" clade7,10. Diatoms and pelagophytes consistently formed sister groups within a single clade, generally with strong support (Table 2). The same clade was seen in plastid phylogenies reported by previous analyses lacking representatives of Limnista12,14,34. This clade likely also includes two additional classes with no sequenced plastid genome to date – the Bolidophyceae and Dictyochophyceae7, and was proposed by Riisberg10 to be termed the Khakista (expanding the original meaning of this name coined for a grouping of diatoms and bolidophytes only11). Our results thus add support to an emerging consensus on the existence of this major ochrophyte subclade united by some potential synapomorphies on the ultrastructural and biochemical level, e.g. by the presence of chlorophyll c7,10.

Further relationships within ochrophytes were less clearly resolved and appeared to be more sensitive to the method of inference and dataset employed. Least stable was the relative position of the PX clade, the raphidophyte Heterosigma and the Limnista clade (Fig. 1, S6, and S7, Table 2). All three PhyloBayes analyses provided maximal support for grouping of each of these three lineages, but ML analyses provided only low support or, in the case of the SP dataset, the Limnista clade moved to a position sister to all ochrophytes. The SP dataset yielded another inconsistent result: when analysed with PhyloBayes, it provided only low support or, in the case of the SP dataset, the Limnista clade moved to a position sister to Heterosigma7,10. This conflicts with previous multi-gene analyses that provided strong evidence for the sisterhood of raphidophytes and the PX clade (to the exclusion of limnistan lineages)7,10. It also conflicts with groupings observed in the five remaining trees generated herein (although with only moderate or low support in two of them, Table 2).

The unstable position of the Limnista clade in the different analyses may relate to significantly longer branch lengths for members of this clade compared to most other algal species (Fig. 1, 2, S6-S13), apart from the extremely long branch of the tertiary haptophyte-derived plastid of the dinoflagellate Karlodinium veneficum13. As already mentioned, long branches are known to be difficult to place reliably in phylogenetic trees, but considering that the more complex CAT-GTR model should provide more accurate inferences on the correct phylogenetic position of the rapidly-evolving limnistan branch than
Figure 2. The phylogenetic position of the *Vitrella brassicaformis* plastid. The trees were inferred from a concatenated matrix of 34 slowly-evolving conserved plastid genes (dataset SG, 14,699 aa positions) excluding the rapidly evolving genome of *Karlodinium veneficum*. (A) Maximum-likelihood tree (RAxML, GTRGAMMA model); only the “chromalveolate” subtree is shown for simplicity. (B) PhyloBayes tree inferred using the CAT-GTR model. The convention for indicating branch support values is the same as in Fig. 1.
the site-homogeneous model, we posit that branching of Linnista with the expanded PX clade is the appropriate working hypothesis.

Overall, these results point to an ochrophyte root positioned between the expanded Khakista (diatoms, pelagophytes, most likely also dictyochophytes and bolidophytes) and all other ochrophytes collectively termed the Phaeista. Previous phylogenetic analyses provided conflicting and generally poorly supported results regarding the root position presumably due to the more limited phylogenetic information available at the time. Notably, our inference on the position of the root of the ochrophyte phylogeny is consistent with the recently published phylogenomic analysis based on sequences of 85 nuclear genome-encoded proteins. That study was not focused on ochrophytes and hence its sampling is relatively sparse, but the inferred phylogeny places the root with maximal support between diatoms and a pelagophyte on the one side and a phaeophyte and a eustigmatophyte on the other side. Our analyses suggest that further sampling of ochrophyte plastid genomes will help to completely resolve evolutionary relationships between algae within this immensely important group.

Revisiting the phylogenetic position of plastids in “chromerid” algae

The phylogenetic analyses described above omit data from plastid-bearing alveolates (i.e., myzozoans). Plastid genomes of two prominent myzozoan subgroups – apicomplexans and dinoflagellates – are extremely divergent and have a highly reduced gene content, whereas plastids of some other myzozoans, e.g. of the oyster parasite Perkinsus marinus and various “colpodellids” have lost a plastid genome entirely. The situation is additionally complicated by the fact that several dinoflagellate lineages have more recently acquired algal endosymbionts or plastids from other groups via a higher-order (tertiary or quaternary) endosymbiosis or kleptoplastidy.

However, two recently discovered myzozoan algae, Chromera velia and especially Vitrella brassicaeformis, have plastid genomes that are more similar to conventional plastid genomes. These two species, together hereafter called “chromerid algae” or simply “chromerids” (put in quotation marks, because they are not monophyletic within Myzozoa), have several unusual characteristics that are present in dinoflagellate and apicomplexan plastids as well, supporting the notion that plastids of the different myzozoan lineages (except the dinoflagellates with plastids representing recent replacements, see above) are monophyletic. Furthermore, the lesser divergence of the plastid genome in “chromerids”, particularly Vitrella, allowed phylogenomic analyses of conserved plastid genes, which showed this lineage as a sister branch of ochrophyte plastids as a whole. This echoes the close relationship of stramenopiles and alveolates established on the basis of nuclear genes, and suggests that a plastid was present in a common ancestor of stramenopiles and alveolates, in agreement with the chromalveolate hypothesis.

Here, we used the expanded sampling of plastid genomes to reanalyze these relationships. We used the same methodology as above to analyze the F, SG and SP datasets with the addition of Vitrella plastid protein sequences. We also analyzed variants of all three datasets from which data from the divergent tertiary plastid of the dinoflagellate K. veneficum were excluded. All ML trees consistently recovered the Vitrella lineage nested within ochrophytes, specifically related to the Eustigmatophyceae (Fig. 2A, S8-S12). Bootstrap support (BS) values for the Vitrella + Eustigmatophyceae clade varied from 97% (F dataset including K. veneficum) to 17% (SG dataset including K. veneficum, Table 2). Most analyses nevertheless provided strong support for a clade comprising the Eustigmatophyceae and Vitrella plus Ochromonas (Table 2). Hence, the ML analyses very consistently showed the Vitrella plastid as branching with the Linnista.

Interestingly, Vitrella and eustigmatophytes have high similarities in their pigment composition, including violaxanthin, as the dominant xanthophyll and the shared absence of chlorophyll ε (missing also from Chromera), which has previously led to speculation that a specific evolutionary connection between these taxa might exist. However, Vitrella plastid gene sequences are still quite divergent, resulting in a long branch in the inferred phylogenetic trees, longer than even that of the tertiary plastid of K. veneficum (Fig. 2). Given the inherent difficulties in placing rapidly evolving lineages in phylogenies the accuracy of these inferences are uncertain, especially since members of the Linnista clade also exhibit relatively divergent plastid genome-encoded proteins.

Indeed, using PhyloBayes and the site-heterogeneous CAT-GTR model, which should be much more effective in coping with long-branch attraction, we recovered the monophyletic Linnista excluding Vitrella with posterior probabilities (PP) of 0.97 to 1.0, depending on the dataset (Fig. 2B, S8-S12, Table 2). In three cases the Vitrella plastid lineage moved to a position sister to all ochrophytes, but the monophyly of ochrophytes to the exclusion of Vitrella was never strongly supported (maximal support was PP of 0.84 in the F dataset including K. veneficum; Table 2, Figs. S8, S10, and S11). However, in the remaining three trees the Vitrella plastid lineage moved from the base of ochrophytes to different positions within them, either sister to diatoms and pelagophytes or sister to the Linnista clade, but support was not attained for specific branching with any ochrophyte lineage (Fig. 2B, S9, and S12, Table 2). Except for the varying position of Vitrella, the topology of the ochrophyte subtree in these analyses was congruent with the results obtained by analyzing the datasets without Vitrella.

We additionally performed an analysis of an expanded SG dataset including not only sequences from Vitrella, but also Chromera (SG + V + C dataset). The resulting trees are portrayed as Fig. S13. As expected, plastids of the two myzozoan algae branch together with maximal support in both the ML and PhyloBayes analyses. However, the branch length of Chromera is more than three times as long as...
the branch length of the already quite divergent Vitrella (measured from the node representing their last common ancestor). In the ML analysis the Vitrella+Chromera clade is sister to Eustigmatophyceae with BS value of 80% (Fig. S13A), consistent with the results of the ML analyses of the SG+V dataset. In the PhyloBayes analysis, the Limmista clade (to the exclusion of Vitrella and Chromera) is recovered with PP of 0.98, but the Vitrella+Chromera clade is sister to Limmista with maximal support (Fig. S13B). Given the extreme divergence of the Chromera plastid protein sequences evident from the analysis of the SG dataset, we did not perform analyses of the other alignment variants (F, SP).

**Did the plastid in alveolates emerge from an ochrophyte endosymbiont?**

Understanding the actual nature of the relationship of plastids in ochrophytes and “chromerids” is relevant to the evolutionary history of plastids in “chromalveolates” in general. Our ML phylogenetic analyses that included Vitrella (or both “chromerid” algae) could suggest a tertiary endosymbiosis of a limnnist (possibly eustigmatophyte-related) alga in an alveolate host cell as the evolutionary origin of the Vitrella plastid. However, the clustering of the divergent “chromerid” plastid genome-encoded proteins with the relatively rapidly evolving sequences of Limmista may be a phylogenetic artifact, as suggested by the results of most analyses with the more realistic site-heterogeneous CAT-GTR model.

Some of these analyses show the Vitrella plastid sister to all ochrophyte plastids. This result seems consistent with the idea that Vitrella and ochrophytes, and by extension all alveolates and stramenopiles, share a plastid-bearing ancestor, as implied by the chromalveolate hypothesis22,16. However, the ochrophyte stem branch in all such trees is very short and poorly supported (Figs. S8, S10, and S11). This is somewhat surprising, since one would expect a relatively long and well-supported stem branch of ochrophyte plastids if the plastids of Vitrella and ochrophytes were vertically inherited from a common ancestor of stramenopiles and alveolates. This is because the phylogenetic distance between the last common ancestor of stramenopiles and alveolates and the last common ancestor of ochrophyte lineages seems to be rather large. Eukaryote SSU rRNA phylogenies with relaxed molecular clock models suggested that the last common ancestor of stramenopiles and alveolates may be older than the last common ancestor of ochrophytes by 250-500 million years48. Dating of the main diversification events in the eukaryote phylogeny based on multigene molecular clocks inferred an even higher difference in the age of these nodes – around 750 million years49.

Although these numbers must be interpreted with caution, it is established that the ochrophyte radiation happened hundreds of millions of years after the stramenopile stem lineage separated from the stem lineage of alveolates. Thus, under this scenario, it is unclear why monophyly of ochrophyte plastids to the exclusion of Vitrella is not recovered consistently. Our logic is in principle similar to that of Baurain et al.21, who pointed to the discrepancy between the strong phylogenetic signal in “chromalveolate” plastid genomes that was suggestive of common ancestry and the weak, or perhaps non-existent, phylogenetic signal in nuclear genomes for the monophyly of the “chromalveolates” as such. All these observations argue against a simple vertical inheritance of a plastid from a common ancestor of stramenopiles and alveolates, or from a hypothetical common ancestor of “chromalveolates”.

Therefore, the results of our phylogenetic analyses appear more compatible with the idea that the “chromerid” lineage gained its plastid from an early ochrophyte. The ML analyses and some PhyloBayes analyses (see Fig. 2 and S13) suggest that the donor might have been a Limmista-related alga, although these results may well result from an artefactual attraction of the very divergent protein sequences encoded by “chromerid” plastid genomes to the relatively divergent sequences of the limnnist algae. Regardless of the actual donor ochrophyte lineage, we posit that the “chromerid” plastid represents a tertiary acquisition. The derived nature of the “chromerid” plastid genomes, especially that of Chromera, may then be explained as a consequence of this origin by a higher-order endosymbiosis, in analogy to the highly divergent plastid genome of K. veneficum (note that the plastid genomes of both Chromera and K. veneficum are linear rather than circular50) established by a (presumably) tertiary endosymbiosis of a haptophyte in a dinoflagellate host43.

The actual nature of the “chromerid” plastid depends on whether the plastid in the putative ochrophyte donor is derived from a secondary endosymbiosis itself, which has been challenged21–23,51. Most recently, Stiller and co-workers used novel statistical analyses of the gene content in “chromalveolate” genomes to conclude that the results are incompatible with the presence of a red algal endosymbiont in an ancestor of all “chromalveolates”25. They instead suggested an explicit scenario, in which a secondary red-algal plastid originated in the cryptophyte lineage, from which it moved by serial endosymbioses first into an ancestor of ochrophytes and subsequently from an early ochrophyte into an ancestor of haptophytes (alveolates were not analyzed in their study). If the haptophyte plastid really emerged from a tertiary endosymbiosis, we would modify our interpretation above such that the “chromerid” plastid is not tertiary, but quaternary, as previously speculated by Bodyl and co-workers22. This putative endosymbiosis must have been a different event than the quaternary origin of the haptophyte plastid suggested by Stiller et al.25, since haptophytes and alveolates never branch together in phylogenomic analyses of both plastid and nuclear genes (our results and, e.g.,18). Furthermore, the putative ochrophyte donor for the haptophyte plastid must have represented a very deep ochrophyte lineage preceding the radiation of extant ochrophytes, as our phylogenomic analyses leave no doubt on the monophyly of ochrophyte plastids to the exclusion of haptophyte plastids.
Previously, several lines of evidence suggested that plastids of “chromerid” algae are monophyletic with plastids of other myzozoans, i.e. apicomplexans and peridinin-containing dinoflagellates. Hence, our inference of an ochorphyte origin of the “chromerid” plastid would by extension apply to the plastids in the Myzozoa as a whole. Anecdotal evidence for such a notion was actually available before, for example a phylogenetic analysis of five concatenated plastid genes carried out by Yoon et al. recovered peridinin-containing plastids of dinoflagellates nested within ochorphytes. Our own multi-gene phylogenetic analyses that included sequences from plastid genomes of apicomplexans and peridinin-containing dinoflagellates indeed showed their grouping with “chromerids” (data not shown), but it cannot be excluded that this results is an artifact stemming from the extremely divergent nature of all these plastid genomes. Furthermore, Petersen et al. recently cast some doubt on the common origin of myzozoan plastids based on phylogenies of several nucleus-encoded plastid-targeted proteins. While the significance of their observations remains unclear, we cautiously refrain from making definitive statements concerning all myzozoan plastids on the basis of results obtained by analyzing Vitrella and Chromera only.

Conclusions
Our newly sequenced ochorphyte plastid genomes enabled us to improve our understanding of the ochorphyte phylogeny by providing convincing support for the existence of the Linnista clade (comprised of the classes Chrysophyceae and Eustigmatophyceae). Our studies also support positioning of the root of the ochorphyte phylogeny between the groups Khakista (diatom, bolidophytes, pelagophytes, dictyochophytes) and Phaeista (i.e. raphidophytes, phaeophytes, xanthophytes, eustigmatophytes, chrysophytes, and pinguiophytes along with a few other, smaller lineages) as redefined by Riisberg et al. Sampling of plastid genomes from ochorphyte classes that are as yet unsequenced and additional taxa for those like chrysophytes, for which the Ochromonas plastid genome sequenced herein is the sole representative, will allow testing of hypotheses on other relationships among the ochorphyte classes, such as the sisterhood of chrysophytes and synchromophytes and the probable sister relationship of the Linnista and Pinguioophyceae.

The introduction of chrysophyte and additional eustigmatophyte plastid genome sequences provides a new impetus for revisiting the current ideas about the evolutionary origin of plastids in the Myzozoa. Our analyses suggest that a valid alternative hypothesis is horizontal transfer of a plastid, via endosymbiosis or kleptoplastidy, from an early ochorphyte lineage to an ancestor of “chromerid” algae. Whether this ancestor was a progenitor of the whole Myzozoa groups remains contentious due to the extremely divergent nature of all myzozoan plastid genomes characterized so far. This question may eventually be answered by additional improvements in the methodology of phylogenetic inference, new data from the nuclear genomes of Vitrella and Chromera, and/or characterization of the presently enigmatic additional plastid-bearing myzozoan lineages known by their plastid 16S rRNA genes sequences in environmental surveys.

Finally, the gain of a plastid by one of the major “chromalveolate” lineages (Myzozoa) from another such lineages (Ochrophyta) suggested here does not support the chromalveolate hypothesis, but is consistent with an alternative scenario, “the rhodoplex hypothesis”. This scenario explains, for example, why plastids are not observed in non-myzozoan alveolates (ciliates, and Pinguiophyceae).

While the significance of their observations remains unclear, we cautiously refrain from making definitive statements concerning all myzozoan plastids on the basis of results obtained by analyzing Vitrella and Chromera only.

References
1. Green, B. R. Chloroplast genomes of photosynthetic eukaryotes. Plant J. 66, 34–44 (2011).
2. Keeling, P. J. The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. Annu. Rev. Plant Biol. 64, 583–607 (2013).
3. Andersen, R. A. Biology and systematics of heterokont and haptophyte algae. Annu. Rev. Plant Biol. 59, 1508–1522 (2004).
4. Yoon, H. S., Andersen, R. A., Boo, S. M. & Bhattacharya, D. Stramenopiles. In Eukaryotic Microbes (ed. Schaechter, M.) 373–384 (Academic Press, 2012).
5. Kai, A., Yoshii, Y., Nakayama, T. & Inouye, I. Aurearenophyceae classis nova, a new class of Heterokontophyta based on a new marine unicellular alga Aurearena cruciata gen. et sp. nov. inhabiting sandy beaches. Protist 159, 435–457 (2008).
6. Brown, J. W. & Sorhannus, U. A molecular genetic timescale for the diversification of autotrophic Stramenopiles (Ochrophyta): substantive underestimation of putative fossil ages. PLoS ONE 5, e12759 (2010).
7. Yang, E. C. et al. Supermatrix data highlight the phylogenetic relationships of photosynthetic stramenopiles. Protist 163, 217–231 (2012).
8. Cavalier-Smith, T. & Scoble, J. M. Phylogeny of Heterokonta: Incisomonas marina, a uniciliate gliding opaloozoan related to Solenicola (Nanomonadea), and evidence that Actinophryida evolved from raphidophytes. Eur. J. Protistol. 49, 328–333 (2013).
9. Not, F. et al. Diversity and ecology of eukaryotic marine phytoplankton. In Genomic Insights Gained into the Diversity, Biology and Evolution of Microbial Photosynthetic Eukaryotes. (ed. Piganeau, S.) 1–53 (Elsevier, 2012).
10. Riisberg, I. et al. Seven gene phylogeny of heterokonts. Protist 160, 191–204 (2009).
11. Cavalier-Smith, T. & Chao, E. E. Y. Phylogeny and megasytems of phagotrophic Heterokonts (Kingdom Chromista). J. Mol. Evol. 62, 388–420 (2006).
12. Janouškovec, J., Horák, A., Oborník, M., Lukeš, J. & Keeling, P. J. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc. Natl. Acad. Sci. U.S.A. 107, 10949–10954 (2010).
13. Ruhfel, B. R., Gitzenanner, M. A., Solits, P. S., Solits, D. E. & Burleigh, J. G. From algae to angiosperms—inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. BMC Evol. Biol. 14, 23 (2014).
14. Worden, A. Z. et al. Global distribution of a wild alga revealed by targeted metagenomics. Curr. Biol. 22, R675–677 (2012).
15. Cavalier-Smith, T. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J. Euk. Microbiol. 46, 347–366 (1999).
16. Cavalier-Smith, T. Deep phylogeny, ancestral groups and the four ages of life. Philos. Trans. R. Soc. London B 365, 111–132 (2010).
17. Burki, F. et al. Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, Telenemia and Centroheliozoa, are related to photosynthetic chromalveolates. Genome Biol. Evol. 1, 231–238 (2009).
18. Burki, F., Okamoto, N., Pombert, J.-F. & Keeling, P. J. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proc. R. Soc. B 279, 2246–2254 (2012).
19. Hampl, V. et al. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic “supergroups”. Proc. Natl. Acad. Sci. USA 106, 3859–3864 (2009).
20. Yabuki, A. et al. Paleopitomas bilix represents a basil cryptist lineage: insight into the character evolution in Cryptista. Sci. Rep. 4, 4641 (2014).
21. Sanchez-Puerta, M. V., Bachvaroff, T. R. & Delwiche, C. F. Sorting wheat from chaff in multi-genome analyses of chlorophyll c-containing plasts. Mol. Phylogenet. Evol. 44, 885–897 (2007).
22. Bodily, A., Stiller, J. W. & Mackiewicz, P. Chromalveolate plastids: direct descent or multiple endosymbioses? Trends Ecol. Evol. 24, 119–121 (2009).
23. Baurain, D. et al. Phylogenomic evidence for separate acquisition of plasts in Cryptophytes, Haptophytes, and Stramenopiles. Mol. Biol. Evol. 27, 1698–1709 (2010).
24. Stiller, J. W. Toward an empirical framework for interpreting plastid evolution. J. Phycol. 50, 462–471 (2014).
25. Stiller, J. W. et al. The evolution of photosynthesis in chlorist algae through serial endosymbioses. Nat. Commun. 5, 5764 (2014).
26. Petersen, J. et al. Chromera velia, endosymbioses and the rhodophyta hypothesis — plastid evolution in Cryptophytes, Alveolates, Stramenopiles, and Haptophytes (CASH lineages). Genome Biol. Evol. 6, 666–684 (2014).
27. Gagat, P., Bodily, A., Mackiewicz, P. & Stiller, J. W. Tertiary plastid endosymbiosis in dinoflagellates. In Endosymbiosis (ed. Löffelhardt, W.) 233–290 (Springer-Verlag, 2014).
28. Přibyl, P., Eliáš, M., Cepák, V., Lukavský, J. & Šaštín, P. Zeolites as a tool for the study of evolutionary processes. J. Euk. Microbiol. 57, 666–684 (2014).
29. Starkenburg, S. R., Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288 (2009).
30.Lang, B. F. & Nedelcu, A. M. Plastid genomes of algae. In Genomics of Chloroplasts and Mitochondria. Advances in Photosynthesis and Respiration, Vol. 35 (eds Bock, R. & Knoop, V.) 59–87 (Springer, 2012).
31. Kück, P., Mayer, C., Wägele, J.-W. & Misof, B. Long branch effects distort maximum likelihood phylogenies in simulations despite site-heterogeneous model. BMC Evol. Biol. 14, 231–242 (2012).
32. Febvre, K. P., Eliáš, M. & Fawley, M. W. The diversity and phylogeny of the commercially important algal class Eustigmatophyceae, including the new clade Goniochloridales. J. Appl. Phycol. 26, 1773–1782 (2014).
33. Lin, Y. C. et al. Distribution patterns and phylogeny of marine stramenopiles in the north Pacific ocean. Appl. Environ. Microbiol. 78, 3387–3399 (2012).
34. Kirkham, A. R. et al. A global perspective on marine photosynthetic picoeukaryote community structure. ISME J. 7, 922–936 (2013).
35. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690 (2006).
36. Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288 (2009).
37. Lang, B. F. & Nedelcu, A. M. Plastid genomes of algae. In Genomics of Chloroplasts and Mitochondria. Advances in Photosynthesis and Respiration, Vol. 35 (eds Bock, R. & Knoop, V.) 59–87 (Springer, 2012).
38. Simon, D., Fewer, D., Friedl, T. & Bhattacharya, D. Phylogeny and self-splicing ability of the plastid tRNA-Leu group 1 Intron. J. Mol. Evol. 57, 710–720 (2003).
39. Fawley, M. W. et al. A pangenomic analysis of the Nanochloropsis organellar genomes reveals novel genetic variations in key metabolic genes. BMC Genomics 15, 212 (2014).
40. Tajima, N. et al. Analysis of the complete plastid genome of the unicellular red alga Porphyridium purpureum. J. Plant Res. 127, 389–397 (2014).
41. Philippe, H., Zhou, Y., Brinkmann, H., Rodrigue, N. & Deluc, F. Heterotaxy and long-branch attraction in phylogenetics. BMC Evol. Biol. 5, 50 (2005).
42. Lartillot, N., Brinkmann, H. & Philippe, H. Suppression of long-branch attraction artefacts in the animal using a site-heterogeneous model. BMC Evol. Biol. 7, 1–14 (2007).
43. Kück, P., Mayer, C., Wägele, J.-W. & Misof, B. Long branch effects distort maximum likelihood phylogenies in simulations despite selection of the correct model. PLoS ONE 7, e63932 (2012).
44. Patil, V., Bratré, I., Shalchian-Tabrizi, K. & Jakobsen, K. S. Revisiting the phylogenetic position of Synchoma grande. J. Euk. Microbiol. 56, 394–396 (2009).
45. Grant, J., Teike, Y. I., Anderson, O. R., Patterson, D. J. & Katz, L. A. Multigene evidence for the placement of a heterotrophic amoeboid lineage Leukarachnion sp. among photosynthetic Stramenopiles. Protist 160, 376–385 (2009).
46. Gabrielsen, T. M. et al. Genome evolution of a tertiary dinoflagellate plastid. PLoS ONE 6, e19132 (2011).
47. Janouškovec, J. et al. Molecular clock dating of the plastid endosymbiosis. Mol. Biol. Evol. 17, 1173–1182 (2000).
48. Janouškovec, J. et al. Factors mediating plastid dependency and the origins of parasitism in apicoplasts and their close relatives. Proc. Natl. Acad. Sci. U.S.A. 102, 10373–10378 (2005).
49. Schmitt, P., Pombert, J.-F., Delwiche, C. F. & Goodwillie, C. Are algal genes in nonphotosynthetic prokaryotes evidence of historical plastid endosymbioses? BMC Genomics 10, 484 (2009).
50. Yoon, H. S. et al. Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. Mol. Biol. Evol. 22, 1299–1308 (2005).
51. Janouškovec, J., Horák, A., Barott, K. L., Rohwer, F. L. & Keeling, P. J. Global analysis of plastid diversity reveals apicoplast–endosymbiotic–related lineages in coral reefs. Curr. Biol. 22, R518–R519 (2012).
52. Janouškovec, J. et al. Colponemids represent multiple ancient alveolate lineages. Curr. Biol. 23, 2546–2552 (2013).
53. Tkhtonenkov, D. V. et al. Description of Colponema vietnamica sp. nov. and Acavomonas parvula n. gen. n. sp., two new alveolate phyla (Colponemidia nom. nov. and Acavomonadina nom. nov.) and their contributions to reconstructing the ancestral state of Alveolates and Eukaryotes. PLoS ONE 9, e8095467 (2014).
Acknowledgements
We thank Jan Janouškovec for sharing data on *Acavomonas* prior to publication. This work was supported by grants from the Gordon and Betty Moore Foundation (GBMF3788) and DOE-DE-SC0004765 (to A.Z.W), and from the Czech Science Foundation (13-33093S to M.O. and P506-12-P931 to A.H.). The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Funds also came from the Lucile and David Packard Foundation (to A.Z.W.). T.Š. and M.E. acknowledge support of the Project CZ.1.05/2.1.00/03.0100 (IET) financed by the Structural Funds of the EU and of the Project LO1208 of the National Feasibility Programme I of the Czech Republic.

Author Contributions
M.E. and A.Z.W. designed the experiments and drafted the manuscript. T.Š. and V.Z. annotated plastid genomes. A.H. performed phylogenetic analyses. T.Š. and A.H. prepared figures and tables. V.K., Č.V., J.J., E.D.H., S.S., J.F. and J.S. sequenced and assembled plastid genomes. P.P. provided biological material. F.B.L. and A.Z.W. provided the initial annotation of the *T. minutus* and *Ochromonas* plastid genomes, respectively, and F.B.L. performed an initial phylogenetic analysis. M.O. contributed to the design of the study and to the manuscript. All authors reviewed the manuscript.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Ševčíková, T. et al. Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? Sci. Rep. 5, 10134; doi: 10.1038/srep10134 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/