diversity

MDPI

Article

Phylotranscriptomic and Evolutionary Analyses of Oedogoniales (Chlorophyceae, Chlorophyta)

Qian Xiong [1,2], Yuxin Hu [3], Xiaoqi Dong [1,4], Yangliang Chen [1,2], Guoxiang Liu [1,*] and Zhengyu Hu [1]

[1] Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China; taxonomyxq@163.com (Q.X.); dongxq@ihb.ac.cn (X.D.); chenyangliang@ihb.ac.cn (Y.C.); huzy@ihb.ac.cn (Z.H.)
[2] Institute of Hydrobiology, University of Chinese Academy of Sciences, Beijing 100039, China
[3] Yangtze River Basin Ecological Environment Monitoring and Scientific Research Center, Yangtze River Basin Ecological Environment Supervision and Administration Bureau, Ministry of Ecological Environment, Wuhan 430072, China; ssshyx@163.com
[4] College of Science, Tibet University, Lhasa 850000, China
* Correspondence: liugx@ihb.ac.cn

Abstract: This study determined the transcriptomes of eight Oedogoniales species, including six species from *Oedogonium* and two species from *Oedocladium* to conduct phylotranscriptomic and evolutionary analyses. 155,952 gene families and 192 single-copy orthogroups were detected. Phylotranscriptomic analyses based on single-copy orthogroups were conducted using supermatrix and coalescent-based approaches. The phylotranscriptomic analysis results revealed that *Oedogonium* is polyphyletic, and *Oedocladium* clustered with *Oedogonium*. Together with the transcriptomes of the OCC clade in the public database, the phylogenetic relationship of the three orders (Oedogoniales, Chaetophorales, Chaetopeltidales) is discussed. The non-synonymous (dN) to synonymous substitution (dS) ratios of single-copy orthogroups of the terrestrial Oedogoniales species using a branch model of phylogenetic analysis by maximum likelihood were estimated, which showed that 92 single-copy orthogroups were putative rapidly evolving genes. Gene Ontology enrichment and Kyoto Encyclopedia of Genes and Genomes pathway analyses results revealed that some of the rapidly evolving genes were associated with photosynthesis, implying that terrestrial Oedogoniales species experienced rapid evolution to adapt to terrestrial habitats. The phylogenetic results combined with evolutionary analyses suggest that the terrestrialization process of Oedogoniales may have occured more than once.

Keywords: Oedogoniales; phylotranscriptomic analysis; rapidly evolving genes; phylogenetics; dN/dS ratios; terrestrialization

Citation: Xiong, Q.; Hu, Y.; Dong, X.; Chen, Y.; Liu, G.; Hu, Z. Phylotranscriptomic and Evolutionary Analyses of Oedogoniales (Chlorophyceae, Chlorophyta). *Diversity* **2022**, *14*, 157. https://doi.org/10.3390/d14030157

Academic Editors: Federico Sebastiani and Christophe Destombe

Received: 9 December 2021
Accepted: 20 February 2022
Published: 23 February 2022

1. Introduction

The order Oedogoniales belonging to the OCC clade (consisting of the Oedogoniales, Chaetophorales, and Chaetopeltidales) of Chlorophyceae, which is within the single family Oedogoniaceae, includes three genera: *Oedogonium* Link ex Hirn, *Oedocladium* Stahl, and *Bulbochaete* Agardh based on the conventional morphological criteria [1–4]. More than 600 species have been described in this order, with most occurring in fresh waters globally, although *Oedocladium* species and a few *Oedogonium* species are terrestrial, predominantly found on soil surfaces [4–14]. The taxonomy of Oedogoniales is mainly based on morphology, and phylogenetic analyses with regard to the group remain limited. Previous studies revealed that Oedogoniales were monophyletic and that *Bulbochaete* may be a sister to the other two genera [15–18]. Using nuclear 18S rDNA of 10 *Oedogonium* species, Alberghina et al. [19] suggested that *Oedogonium* might not be monophyletic; therefore, morphological characteristics may not define phylogenetic groups. Mei et al. [20] analyzed 18S rDNA sequences and observed that *Oedocladium* formed a separate clade within *Oedogonium*, whereas *Bulbochaete* was relatively distant from the other two genera. However,

the researchers suggested that the phylogeny required further investigation based on larger sampling of such taxa, particularly those of *Oedocladium* and *Bulbochaete*. Phylogenetic and evolutionary analyses based on chloroplast genome protein-coding genes have showed that both *Oedocladium* and *Oedogonium* are polyphyletic groups [21,22].

Although phylogenetic analyses based on one or a few genes have been performed extensively in recent decades, they have a few limitations, such as the existence of paralogues, varying evolutionary rates, incomplete lineage sorting, horizontal gene transfer, and gene duplication [23–28], which implies that gene trees may not represent species trees in some cases. Generally, an increase in the information in a dataset and the use of a more appropriate analytical method reduces the influence of such errors [29–32]. With the rapid development of next-generation sequencing technology, transcriptome data is increasingly being applied for phylogenetic analysis, being used widely in plants [33,34], protists [35,36], algae [37–39], and animals [40,41]. However, phylotranscriptomic analysis of Oedogoniales has not been performed yet due to lack of transcriptome data.

Nucleotide substitution rates are often used as a criterion to reflect the selection pressure. According to Yang [42], dN/dS < 1, =1, >1, denote a negative purifying selection, neutral evolution and positive selection, respectively. The dN/dS ratios can be used to determine whether genes are evolving under purifying selection [43,44]. An excess of nonsynonymous substitutions over synonymous ones can be used to detect changes in proteins that provide higher fitness in a given circumstance (e.g., low light) at the molecular level [45]. It has been reported that organisms in energy-rich habitats are often characterized by higher evolutionary rates [46,47], organisms living in low-energy areas, like shaded habitats, have relatively slower rates of molecular evolution [48]. The genus *Oedocladium* (terrestrial) was presumed to have partly originated from *Oedogonium* species, which grow on moist soil surfaces and present underground filaments with slightly unbranched rhizoids [9]. The *Oedocladium* and the terrestrial *Oedogonium* species were detected to have positively selected positions in *psb*A, suggesting that terrestrial Oedogoniales taxa may have undergone adaptive evolution to adjust to the variations in light intensity between aquatic and terrestrial habitats [21,22]. dN/dS analysis in Oedogoniales would further contribute to revealing the evolutionary relationships in this group of algae.

In the present study, we sequenced the transcriptomes of eight Oedogoniales species and subsequently conducted phylotranscriptomic analysis to understand the phylogenetic relationship of this group and the OCC clade (Oedogoniales, Chaetophorales and Chaetopeltidales). In addition, dN/dS ratios of gene families in Oedogoniales were estimated to explore the evolutionary relationship between the aquatic and terrestrial habitats, which could provide valuable information with regard to the phylogenetic and evolutionary relationships of Oedogoniales.

2. Materials and Methods

2.1. Cultures

Eight Oedogoniales species, namely, *Oedogonium crispum* (FACHB-3310), *Oedogonium dentireticulatum* (FACHB-3309), *Oedogonium* sp. (FACHB-3311), *Oedogonium* sp. (FACHB-3313), *Oedogonium* sp. (FACHB-3317), *Oedogonium capilliforme* (FACHB-3312), *Oedocladium prescottii* (FACHB-2452), and *Oedocladium carolinianum* (FACHB-2453) were obtained from culture collections of previous studies and stored in the Culture Collection of Freshwater Algae at the Institute of Hydrobiology, Chinese Academy of Sciences. All the strains were cultured in liquid BG11 medium at 25 °C under a 12–12 h light–dark cycle and light intensity of 15–30 μmol/(m^2·s).

2.2. Library Preparation and Sequencing

Total RNA was extracted using TRIzol reagent (Thermo Fisher Scientific, Waltham, MA, USA), and poly-A+ mRNA was isolated using oligo dT magnetic beads. The mRNA was fragmented using divalent cations under a high temperature in NEBNext First Strand Synthesis Reaction Buffer (5×) and was used as a template for random hexamer-primed

first-strand cDNA synthesis. Afterward, the second strand of cDNA was synthesized. The sequencing library was generated via the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA) and sequenced on the NovaSeq 6000 platform (Illumina, San Diego, CA, USA). Finally, 150 bp paired-end reads were generated.

2.3. Quality Control, De Novo Assembly, and Sequence Annotation

The quality of the raw reads was initially checked using FastQC v0.11.6 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ accessed on 24 June 2021) and the raw reads were subjected to quality control using Trimmomatic v0.39 [49] (LEADING:5, TRAILING:5, SLIDINGWINDOW:4:5, MINLEN:25). Trinity v2.8.5 [50] was used to conduct *de novo* assembly of the clean reads with default parameters. BUSCO v3.0.2 [51] was used to assess the completeness of the final transcripts. Subsequently, TransDecoder v5.5.0 was used to predict the open reading frame (ORF) of each transcript. BLASTP [52] searches of the longest ORFs (the longest coding region in each transcript) were conducted against the Uniref90 database using Diamond v0.8.22.84 [53]. In addition, Pfam searches of the longest ORFs in the Pfam database were conducted using HMMER v3.1b2 [54]. Finally, TransDecoder v5.5.0 was used to integrate the BLASTP and Pfam search results into coding regions. The nucleotide sequences (CDS) and amino acid sequences (PEP sequences) of the regions were used for subsequent analyses. Additionally, CDS and PEP sequences of *Chlamydomonas reinhardtii* were downloaded from NCBI using the assembly accession number GCF_000002595.1. All raw reads were deposited in the NCBI Sequence Read Archive (BioProject PRJNA771938).

2.4. Orthologous Group Identification and Phylotranscriptomic Analysis

Single-copy orthologues were determined using OrthoFinder v2.5.2 [55], and PEP sequences of single-copy orthologues were selected for phylotranscriptomic analysis. The PEP sequences of each single-copy orthologue were aligned using MAFFT v7.394 [56], with the options -maxiterate 1000 and -globalpair. Regions showing poor alignment were trimmed with TrimAl v1.2 [57] using the parameter -automated1. The trimmed alignment of orthologous groups was used for subsequent phylotranscriptomic analysis. Supermatrix and coalescent-based analyses were used to construct the phylogenetic tree. With regard to supermatrix analysis, PhyloSuite [58] was used to concatenate all orthologous groups, the concatenate sequence including 35,684 amino acids, and PartitionFinder 2 [59] was used to determine the evolutionary models and partitioning of the concatenated PEP dataset. Supermatrix analysis was conducted based on Bayesian inference (BI) and maximum likelihood (ML) methods. Bayesian analysis was conducted using MrBayes v3.2.6 [60], and the dataset was partitioned as shown in Table S1 (Supplementary Materials). Markov chain Monte Carlo analyses were run with four Markov chains (three heated, one cold) for 3,000,000 generations, and trees were sampled every 1000 generations. In each round of calculation, a fixed number of samples (burn-in = 1000) was discarded at the beginning of the chain. ML analysis was carried out using the IQ-TREE web server [61] with 1000 ultrafast bootstraps [62] and 1000 SH-aLRT tests [62,63] to test nodal support. For the coalescent-based analyses, RAxML [64] was used to conduct ML analysis of each single-copy orthologue based on the PROTGAMMA GTR model. Furthermore, ASTRAL [65] was used to infer coalescent-based species tree (ST) phylogeny. The 18S rDNA sequences were aligned using MAFFT v7.0 [56], and ambiguous regions were manually edited and adjusted by eye using MEGA7 [66]. The 18S rDNA sequences were determined using jModelTest2 [67] and the best model was GTR + I + G. Similarly, ML and BI methods were used to infer phylogenies. To understand the phylogenetic relationships of the OCC clade, the transcriptomes of the three orders were downloaded from the public database, with the same methods as above, we conducted phylotranscriptomic analyses based on the single-copy orthologues by BI and ML analyses (the concatenate sequence including 8,291 amino acids), and the assembly accession numbers are listed in Table S2.

2.5. Evolutionary Analyses Based on Phylogenetic Analysis by Maximum Likelihood

The CODEML program of PAML v4.9 [42] with the ML model (runmode = −2, CodonFreq = 2) was used to measure the values of dS and dN. The analysis was based on all single-copy orthologues and orthologues with dS values > 5 were excluded from further analyses. The branch model was employed in the calculation of dN/dS ratios for terrestrial Oedogoniales species and aquatic ones with *Oe. Prescottii*, *Oe. Carolinianum, and O.* sp. FACHB-3313, were labeled as foreground branches. A null model (model = 0), where one dN/dS ratio was fixed across all strains, was compared with an alternative model (model = 2), where *Oe. Prescottii*, *Oe. Carolinianum*, and *O.* sp. (FACHB-3313), were allowed to have a different dN/dS ratio. Likelihood ratio tests were performed to examine model fit, a chi-squared test was used to analyze p values, and multiple testing was corrected using false discovery rate (FDR). The genes were considered putative rapidly evolving genes if they had an FDR-adjusted p value < 0.05 and a higher dN/dS ratio in the foreground branch than in the background branches.

2.6. Gene Ontology Enrichment and Kyoto Encyclopedia of Genes and Genomes Pathway Analyses of Rapidly Evolving Genes

After acquiring the putative rapidly evolving genes, gene ontology (GO) functional and KEGG pathway enrichment analyses were performed to determine the functions of the genes. For GO enrichment analysis, all the PEP sequences were imported into InterProScan [68] for GO term mapping. In addition, all PEP sequences were subjected to KEGG pathway analyses using EggNOG-mapper [69]. Both analyses were performed using clusterProfiler [70], with a p value cutoff of 0.05, and the FDR method [71] was used for multiple testing.

3. Results

3.1. De Novo Transcriptome Assembly and Ortholog Detection

The assembly and annotation statistics of the eight Oedogoniales species are listed in Table 1. The raw reads of each species generated by Illumina paired-end sequencing technology ranged from 49,070,110 to 69,832,706. After filtering for adapters and low-quality sequences using Trimmomatic v0.39 [49], the number of clean reads ranged from 49,038,062 to 69,639,876. Afterward, de novo assembly was conducted using Trinity v2.8.5. The number of assembled contigs ranged from 69,335 to 143,592 and the average contig length and N50 values for all species were greater than 750 bp and 1000 bp, respectively. After the removal of redundant transcripts using TransDecoder v5.5.0, BUSCO [51] was used for further quantitative assessment of assembly and annotation completeness. A search of BUSCO genes defined for Chlorophyta revealed that the percentages of conserved genes in the transcriptomes of eight Oedogoniales species recovered via BUSCO analysis were >89%, indicating a high completeness level. Finally, TransDecoder v5.5.0 was used to predict coding sequences from the assembled transcripts, and the results revealed that the number of coding sequences ranged from 29,750 to 118,126. The nucleotide coding sequence (CDSs) and amino acid sequence (PEP sequence) of the coding sequences were used for subsequent analysis. The PEP sequences of the eight species were used to detect orthology.

The number of orthogroups was 155,952, where each orthogroup represented a gene family, and all orthogroups were subjected to GO and KEGG enrichment analyses. The number of single-copy orthogroups was 192, and all single-copy orthogroups were subjected to evolutionary analyses. Similarly, the PEP sequences of *C. reinhardtii* were downloaded from the NCBI database as an outgroup, in addition to the eight Oedogoniales species; the PEP sequences of the nine species were used to detect orthology and all single-copy orthogroups were used for phylotranscriptomic analyses. The number of single-copy orthogroups was 99. When analyzing the phylogenetic relationship of the OCC clade, with *C. reinhardtii* as outgroup, and all the PEP sequences of the three orders were used to detect orthology and the number of the single-copy orthogroups was 29.

Table 1. Summary of sequencing and assembly.

Species	Number of Raw Reads	Number of Clean Reads	Sequence Assembled by Tri-ity			Orthogroups	Complete BUSCOs	Number of Coding Sequence Predicted by TransDecoder
			Number of Contigs	Average Contig Length	N50 Length			
Oedogonium sp. (FACHB-3313, terrestrial)	66,116,070	66,020,778	82,567	1084.35	2165	28,539	92.7%	50,244
Oedocladium carolinianum (terrestrial)	49,070,110	49,038,062	112,614	959.71	1897	36,030	92.3%	63,044
Oedogonium capilliforme	67,815,700	67,680,428	118,267	842.10	1439	37,039	93.0%	59,861
Oedocladium prescottii (terrestrial)	64,034,464	64,005,336	139,202	1204.96	2693	66,000	91.9%	118,126
Oedogonium sp. (FACHB-3311)	60,835,586	60,730,652	94,883	849.91	1305	34,131	89.3%	64,417
Oedogonium dentireticulatum	52,801,814	52,695,338	100,195	838.23	1371	30,859	93.7%	48,137
Oedogonium crispum	66,019,080	65,947,940	143,592	753.29	1085	33,704	92.1%	51,654
Oedogonium sp. (FACHB-3317)	69,832,706	69,639,876	69,335	785.33	1129	21,432	90.2%	29,750

3.2. Phylotranscriptomic Analyses

A phylogenetic tree based on 18S rDNA sequences showed that the eight Oedogonium species were in the clade formed by *Oedogonium* and *Oedocladium*, and the *Oedocladium* species were separated by *Oedogonium subplagiostomum* (Figure S1). Phylotranscriptomic analyses were conducted based on 99 single-copy orthogroups using supermatrix and coalescent-based approaches, with *C. reinhardtii* species as the outgroup. The BI tree and ST exhibited similar results, which revealed that the eight Oedogoniales species formed three clades with absolutely high support values (Figure 1); the two *Oedocladium* species and *O. capilliforme* formed the second clade, while *O.* sp. FACHB-3311, formed a separate clade. Based on the ML tree (Figure S2), the position of the *Oedocladium* species was slightly different, with *Oe. carolinianum* forming a separate clade instead of clustering with species *Oe. prescottii*, and the topology was not well supported. Then phylotranscriptomic analyses based on 29 single-copy orthogroups of the OCC clade with *C. reinhardtii* species as the outgroup by BI and ML method showed the same results that Chaetophorales locating at the base of the branch was sister to a clade formed by Oedogoniales and Chaetopeltidales (Figure S3). The phylogenetic tree of the OCC clade also showed the same result as Figure 1, that two *Oedocladium* species formed a clade clustering with *Oedogonium*. The phylogenetic positions other *Oedogonium* species in the present study were also similar with topology based on 29 single-copy orthogroups (Figure 1 and Figure S3).

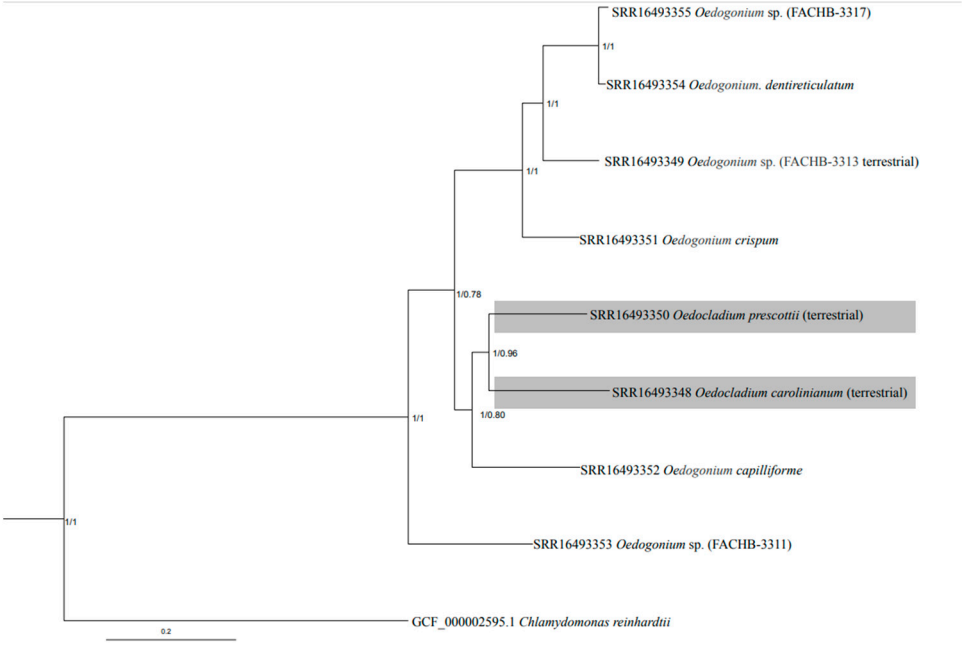

Figure 1. Phylogenetic tree of the Oedogoniales species based on 99 single-copy orthogroups by Mrbayes and ASTRAL. Numbers on branches represent support values of ASTRAL and Bayesian posterior probabilities respectively. Branch lengths are proportional to genetic distances, which are indicated by the scale bar. The grey background indicates the *Oedocladium* species.

3.3. Evolutionary Analyses Based on PAML, GO Enrichment, and KEGG Pathway Analyses of Rapidly Evolving Genes

The ML method was used to calculate the dN and dS substitution rates for 192 single-copy orthogroups of the eight Oedogoniales species. A total of 59 orthogroups were obtained after discarding orthogroups with dS values > 5. Boxplots of species based on the dN and dS rates of the 59 orthogroups are illustrated in Figure 2. We compared dN and dS values between terrestrial and aquatic species across the eight species of Oedogoniales and observed that the genes between the two groups were not significantly different (p = 0.6547 and 0.2967, respectively). Subsequently, a branch model of PAML was used to compare the variations in dN/dS ratios between terrestrial and aquatic species and the results are presented in Table S3. Among the 192 single-copy orthogroups, 93 orthogroups exhibited significant differences and higher dN/dS ratios in the three terrestrial Oedogoniales species, indicating the occurrence of rapid evolution.

To understand the functions of the putative rapidly evolving genes, GO enrichment and KEGG pathway analyses were performed. The results of GO enrichment analysis conducted using InterProScan are illustrated in Figure 3. The putative rapidly evolving genes were enriched in 46 GO terms, three of which being associated with photosystem II oxygen-evolving complex, photosystem II, and photosynthesis (GO: 0009654, GO: 0009523, and GO: 0015979, with adjusted p values = 2.36×10^{-7}, 6.90×10^{-7}, and 1.21×10^{-12}, respectively). The results of the KEGG pathway analysis are illustrated in Figure S4. Notably, five genes were associated with the photosynthetic pathway (adjusted p value = 3.97×10^{-9}).

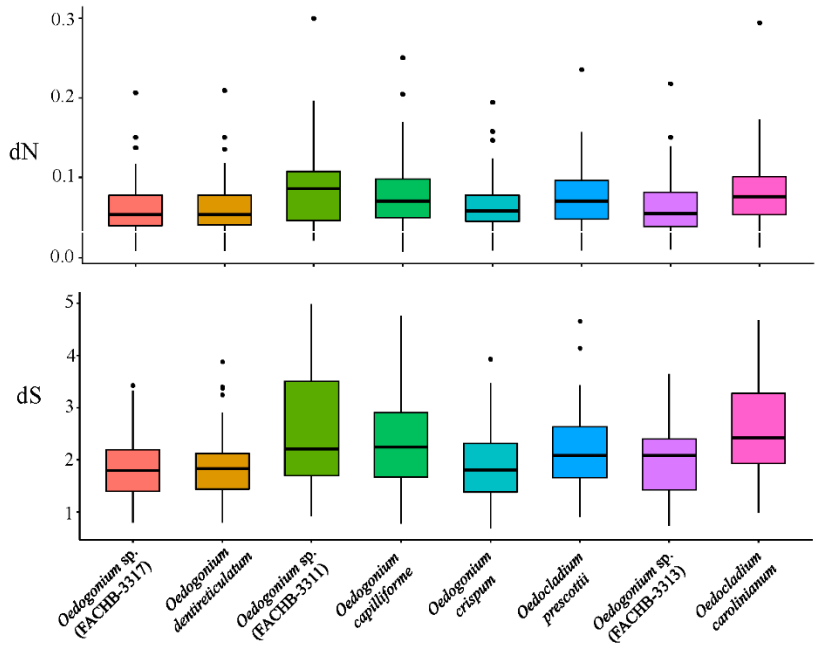

Figure 2. Boxplots of the non-synonymous (dN) and synonymous (dS) substitutions for each species of the eight Oedogoniales species. For each species, the box represents values between quartiles, outliers are shown as black points, and black lines inside the box show median values.

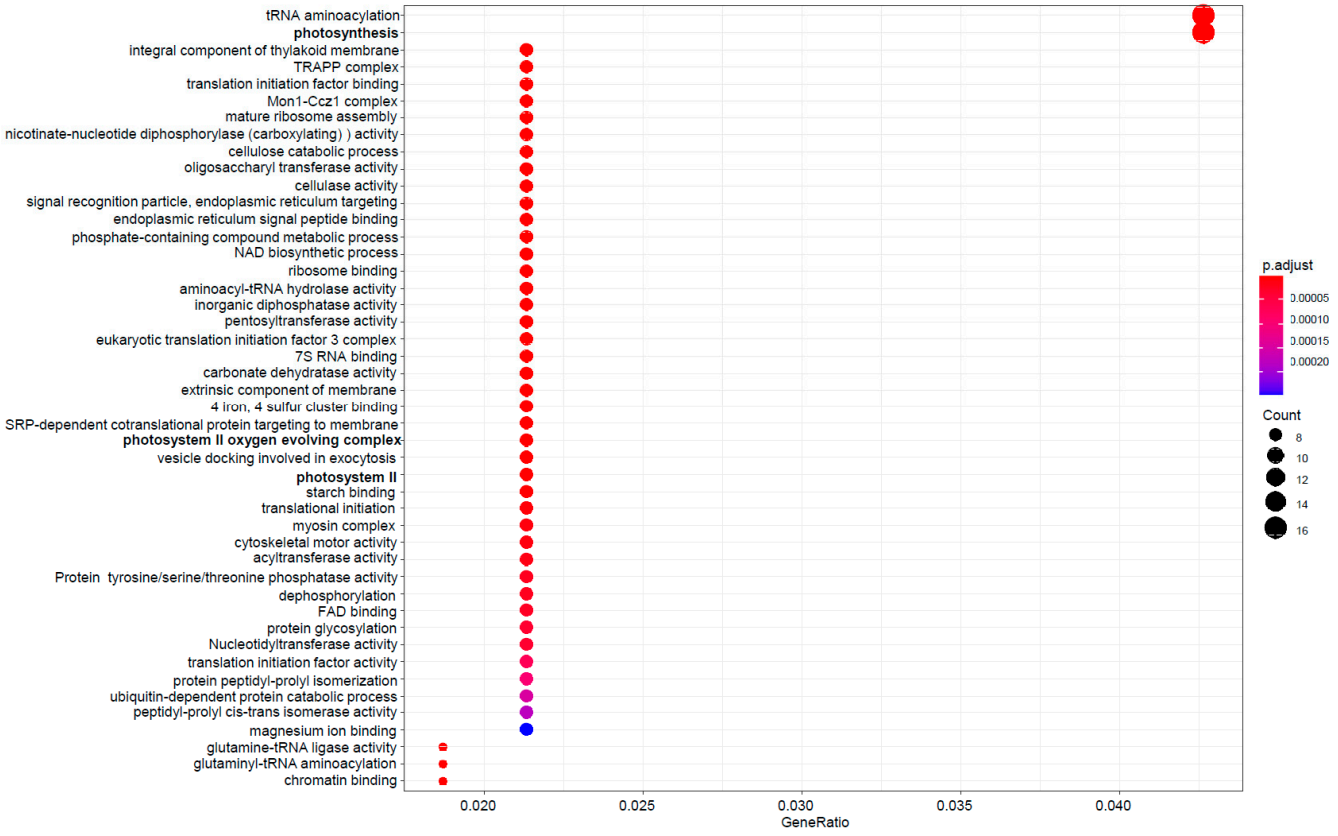

Figure 3. Dot plot showing the enrichment of the 92 putative fast-evolving genes. The dot sizes represent the numbers of genes.

4. Discussion

The phylogenetic results based on the three kinds of dataset (18S rDNA, 99/29 single-copy orthologues based on Oedogoniales and the OCC clade respectively) showed the same results that *Oedogonium* was polyphyletic and *Oedocladium* species formed a separate clade clustering with *Oedogonium*. At the same time, with *C. reinhardtii* as an outgroup, a BI tree based on 99 single-copy orthologues determined by the supermatrix method showed results similar to those of the ST analyzed by ASTRAL, which is consistent with the phylogenetic results based on 18S rDNA sequences, internal transcribed spacers (ITS), and chloroplast protein coding genes [21,22]. The inconsistency in the results observed with regard to the position of *Oe. carolinianum*, we may think the result by BI method more reliable for a relatively high support value, and the topology by the BI method showed high similarity with the ST analyzed by ASTRAL, which enables highly accurate phylogenomic estimation, even in the presence of high levels of gene tree conflict because of incomplete lineage sorting [30] or horizontal gene transfer [31,32]. It has been reported that larger sample sizes could substantially improve phylogenetic results [72]. Previous phylogenetic results based on nuclear genes and chloroplast proteins indicated that in the OCC clade, Oedogoniales was located at the base of the branch, and Chaetophorales and Chaetopeltidales were most closely related [73–77]. While in the present study, phylotranscriptomic results were incongruent in that Chaetophorales, locating at the base of the branch, was sister to a clade formed by Oedogoniales and Chaetopeltidales (Figure S3). However, only one transcriptome of Chaetopeltidales was included in this study and the transcriptome data of the other OCC clade species downloaded from the public database were much smaller than the newly sequenced eight Oedogoniales transcriptomes. More data about Chaetopeltidales and high-quality transcriptomes will contribute to resolve this incongruence.

In the present study, the substitution rates were evaluated using the ML method of PAML, which is the most accurate method currently used for measuring substitution rates [63,78,79]. The statistical analyses results showed no significant differences between the terrestrial and aquatic Oedogoniales species. Furthermore, the dN/dS ratios based on the branch model revealed that 93 orthogroups among the single-copy orthogroups were significantly different, with relatively high dN/dS ratios observed in three terrestrial Oedogoniales species. Previous studies have revealed that low dN/dS ratios ($dN/dS < 1$) denote a strong purifying selection [42], whereas relatively high dN/dS ratios could be interpreted as a weak purifying selection [80]. The relatively high dN/dS ratios observed in the three terrestrial Oedogoniales species suggested the occurrence of rapid evolution. The relatively high dN/dS substitution rates may result in a novel function or adaptive evolution [80–83]. Oedogoniales include over 600 species, with only a few being terrestrial species and others growing in freshwater habitats. The genus *Oedocladium* (terrestrial) was presumed to have originated from *Oedogonium* species, which grow on moist soil surfaces and have underground filaments with slightly unbranched rhizoids [9]. Based on both the GO enrichment and KEGG pathway analyses, the function of the putative rapidly evolving genes is related to photosynthesis, considering the differences between aquatic and terrestrial habitats, we speculated that the terrestrial Oedogoniales species underwent rapid evolution to adapt to the varying conditions, especially for the terrestrialization process. This observation is consistent with the findings of a previous study that was based on chloroplast protein-coding genes [22]. A previous study revealed that the transition from water to land occurred several times and across different taxa, ranging from microorganisms to lichens and green plants, and later, arthropods, mollusks, annelids, and vertebrates [84]. The phylogenetic tree (Figure 1) showed that the three terrestrial Oedogoniales species were distributed across different branches, suggesting that terrestrialization process may occur more than once.

Previous studies revealed that the traditional taxonomy of Oedogoniales did not define natural groups [15–20], and the molecular phylogenetic and evolutionary studies of Oedogoniales are still limited due to lack of current molecular data in the public databases. More molecular data will be helpful for further studies. The present study is the first to

perform phylotranscriptomic analysis based on transcriptome data from eight Oedogonilaes species, including six *Oedogonium* and two *Oedocladium* species. Phylogenetic analyses of 99 single-copy orthologues using two phylogenetic analysis approaches revealed that both *Oedogonium* and *Oedocladium* were polyphyletic. According to the dN/dS estimation results, terrestrial Oedogonilaes species had 93 putative rapidly evolving single-copy orthologues, and both GO enrichment and KEGG pathway analyses showed that a few orthologues among the putative rapidly evolving single-copy orthologues were associated with photosynthesis, implying that the terrestrial Oedogoniales species underwent rapid evolution in adapting to terrestrial habitats. Overall, phylogenetic results combined the phylogenetic result suggested that terrestrialization processes of Oegogoniales may have occured several times.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d14030157/s1, Figure S1. Phylogenetic tree of the Oedogoniales species based on 18SrDNA sequences. Figure S2. Phylogenetic tree of the Oedogoniales species based on 99 single-copy orthogroups by ML method. Figure S3. Phylogenetic tree of the OCC clade based on 29 single-copy orthogroups by BI and ML method. Figure S4. Bar chart showing the KEGG pathways of the 92 putative fast-evolving genes. Table S1. Partition scheme of the pep sequences of 99 single-copy orthologues used in this study. Table S2. Accession number of the OCC clade species downloaded from the public database. Table S3. The putative fast-evolving genes performed by branch model of PAML.

Author Contributions: Conceptualization, Q.X., G.L.; methodology, Q.X.; software, Q.X.; validation, G.L.; formal analysis, Q.X.; investigation, Y.H.; resources, G.L.; data curation, Q.X.; writing—original draft preparation, Q.X.; writing—review and editing, Y.H., X.D., Y.C., G.L. and Z.H.; visualization, Q.X.; supervision, Q.X.; project administration, G.L.; funding acquisition, G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Key Research Program of Frontier Sciences, CAS (Grant No. QYZDY-SSWSMC029-6) and the National Natural Science Foundation of China (grant no. 31750001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: https://dataview.ncbi.nlm.nih.gov/?search=SUB10528437, accessed on 24 June 2021.

Acknowledgments: This research was supported by the Wuhan Branch, Supercomputing Center, Chinese Academy of Sciences, China.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ML: Maximum likelihood; mitogenome: mitochondrial genome; O.: *Oedogonium*; Oe.: *Oedocladium*; FACHB: Freshwater Algae Culture Collection at the Institute of Hydrobiology.

References

1. Agardh, C.A. *Synopsis Algarum Scandinaviae: Adjecta Dispositione Universali Algarum*; Kessinger Publishing, LLC: Whitefish, MT, USA, 1817.
2. Bary, A. Ueber die Algengattungen *Oedogonium* und *Bulbochaete*. *Naturf. Ges. Frankf.* **1854**, *1*, 29–105.
3. Stahl, E. *Oedocladium protonema*, eine neue Oedogoniaceen-Gattung. *Wiss Bot.* **1891**, *23*, 339–348.
4. Hirn, K.E. Monographie und Iconographie der Oedogoniaceen. *Acta Soc. Scienti. Fennicae.* **1900**, *27*, 1–395.
5. Tiffany, L.H. *North American Flora: Oedogoniales*; New York Botanical Garden: New York, NY, USA, 1937.
6. Gemeinhardt, K. Oedogoniales. In *Dr. L. Rabenhorst's Kryptogamen-Flora von Deutschland, Oesterreich und der Schweiz*; Rabenhorst, L., Ed.; Akademische Verlagsgesellschaft: Leipzig, Germany, 1939.
7. Islam, A.K.M.N.; Sarma, P. Two new species of terrestrial *Oedogonium* from East Pakistan. *Trans. Am. Micros. Soc.* **1963**, *82*, 74–77. [CrossRef]

8. Gauthier-Lièvre, L. *Oedogoniacées Africaines*; Cramer: Stuttgart, Gertmany, 1964; pp. 1–104.
9. Jao, C.C. *Monographia Oedogoniales Sinicae*; Beijing Science Press: Beijing, China, 1979.
10. Mrozińska, T. Oedogoniophyceae: Oedogoniales. In *Süßwasserflora von Mitteleuropa 14; Chlorophyta VI*; Gustav Fischer Verlag: Stuttgart, Germany, 1985.
11. van Den Hoek, C.; Mann, D.G.; Jahns, U.M. *Algae: An Introduction to Phycology*; Cambridge University Press: Cambridge, UK, 1995.
12. Graham, L.E.; Wilcox, L.W. *Algae*; Prentice Hall: Upper Saddle River, NJ, USA, 2000.
13. Liu, G.X.; Hu, Z.Y. Predominant occurrence of apical cell divisions in *Oedogoniam pakistanense* and its phylogenetic significance. *Phycologia* **2004**, *43*, 669–671. [CrossRef]
14. Wittrock, V.B. Prodromus Monographiae Oedogoniearum. *Nova Acta Regiae Soc. Sci. Ups.* **1874**, *3*, 1–64.
15. Booton, G.C.; Floyd, G.L.; Fuerst, P.A. Origins and affinities of the filamentous green algal orders Chaetophorales and Oedogoniales based on 18S rRNA gene sequences. *J. Phycol.* **1998**, *34*, 312–318. [CrossRef]
16. Buchheim, M.A.; Michalopulos, E.A.; Buchheim, J.A. Phylogeny of the Chlorophyceae with special reference to the Sphaeropleales: A study of 18S and 26S rDNA data. *J. Phycol.* **2001**, *37*, 819–835. [CrossRef]
17. Krienitz, L.; Hegewald, E.; Hepperle, D.; Wolf, M. The systematics of coccoid green algae. 18S rRNA gene sequence data versus morphology. *Biologia* **2003**, *58*, 437–446.
18. Shoup, S.; Lewis, L.A. Polyphyletic origin of parallel basal bodies in swimming cells of Chlorophycean green algae (Chlorophyta). *J. Phycol.* **2003**, *39*, 789–796. [CrossRef]
19. Alberghina, J.S.; Vigna, M.S.; Confalonieri, V.A. Phylogenetic position of the Oedogoniales within the green algae (Chlorophyta) and the evolution of the absolute orientation of the flagellar apparatus. *Plant. Syst. Evol.* **2006**, *261*, 151–163. [CrossRef]
20. Mei, H.; Luo, W.; Liu, G.X.; Hu, Z.Y. Phylogeny of Oedogoniales (Chlorophyceae, Chlorophyta) inferred from 18S rDNA sequences with emphasis on the relationships in the genus *Oedogonium* based on ITS-2 sequences. *Plant. Syst. Evol.* **2007**, *265*, 179–191. [CrossRef]
21. Xiong, Q.; Hu, Y.; Liu, B.; Zhu, H.; Liu, G.; Hu, Z. Chloroplast genomes and phylogenetic analysis of two species of *Oedocladium* (Oedogoniales, Chlorophyta). *Eur. J. Phycol.* **2021**, *56*, 403–415. [CrossRef]
22. Xiong, Q.; Hu, Y.; Lv, W.; Wang, Q.; Liu, G.; Hu, Z. Chloroplast genomes of five *Oedogonium* species: Genome structure, phylogenetic analysis and adaptive evolution. *BMC Genom.* **2021**, *22*, 707. [CrossRef]
23. Maddison, W.P. Gene trees in species trees. *Syst. Biol.* **1997**, *46*, 523–536. [CrossRef]
24. Huynen, M.A.; Bork, P. Measuring genome evolution. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5849–5856. [CrossRef]
25. Doolittle, W.F. Phylogenetic classification and the universal tree. *Science* **1999**, *284*, 2124–2128. [CrossRef]
26. Degnan, J.H.; Rosenberg, N.A. Discordance of species trees with their most likely gene trees. *PLoS Genet.* **2006**, *2*, e68. [CrossRef]
27. Zou, J.; Yue, W.; Li, L.; Wang, X.; Lu, J.; Duan, B.; Liu, J. DNA barcoding of recently diversified tree species: A case study on spruces based on 20 DNA fragments from three different genomes. *Trees* **2016**, *30*, 959–969. [CrossRef]
28. Gernandt, D.S.; Dugua, X.A.; Vázquez-Lobo, A.; Willyard, A.; Letelier, A.M.; Pérez de la Rosa, J.A.; Pinero, D.; Liston, A. Multi-locus phylogenetics, lineage sorting, and reticulation in Pinus subsection Australes. *Am. J. Bot.* **2018**, *105*, 711–725. [CrossRef]
29. Wolf, Y.I.; Rogozin, I.B.; Grishin, N.V.; Koonin, E.V. Genome trees and the tree of life. *Trends Genet.* **2002**, *18*, 472–479. [CrossRef]
30. Mirarab, S.; Reaz, R.; Bayzid, M.S.; Zimmermann, T.; Swenson, M.S.; Warnow, T. ASTRAL: Genome-scale coalescent-based species tree estimation. *Bioinformatics* **2014**, *30*, i541–i548. [CrossRef] [PubMed]
31. Davidson, R.; Vachaspati, P.; Mirarab, S.; Warnow, T. Phylogenomic species tree estimation in the presence of incomplete lineage sorting and horizontal gene transfer. *BMC Genom.* **2015**, *16*, S1. [CrossRef] [PubMed]
32. Mirarab, S. Species tree estimation using ASTRAL: Practical considerations. *arXiv* **2019**, arXiv:1904.03826.
33. Deng, H.; Zhang, G.Q.; Lin, M.; Wang, Y.; Liu, Z.J. Mining from transcriptomes: 315 single-copy orthologous genes concatenated for the phylogenetic analyses of Orchidaceae. *Ecol. Evol.* **2015**, *5*, 3800–3807. [CrossRef]
34. Wen, J.; Xiong, Z.; Nie, Z.L.; Mao, L.; Zhu, Y.; Kan, X.Z.; Fang, X.D. Transcriptome sequences resolve deep relationships of the grape family. *PLoS ONE* **2013**, *8*, e74394.
35. Cavalier-Smith, T.; Chao, E.E.; Lewis, R. 187-gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosea) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. *Mol. Phylogenet. Evol.* **2016**, *99*, 275–296. [CrossRef]
36. Cavalier-Smith, T.; Chao, E.E.; Snell, E.A.; Berney, C.; Fiore-Donno, A.M.; Lewis, R. Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. *Mol. Phylogenet. Evol.* **2014**, *81*, 71–85. [CrossRef]
37. Saunders, G.W.; Jackson, C.; Salomaki, E.D. Phylogenetic analyses of transcriptome data resolve familial assignments for genera of the red-algal Acrochaetiales-Palmariales complex (Nemaliophycidae). *Mol. Phylogenet. Evol.* **2018**, *119*, 151–159. [CrossRef]
38. Yin, Y.; Johns, M.A.; Cao, H.; Rupani, M. A survey of plant and algal genomes and transcriptomes reveals new insights into the evolution and function of the cellulose synthase superfamily. *BMC Genom.* **2014**, *15*, 260. [CrossRef]
39. Hu, Y.; Xing, W.; Song, H.; Hu, Z.; Liu, G. Comparison of colonial volvocine algae based on phylotranscriptomic analysis of gene family evolution and natural selection. *Eur. J. Phycol.* **2019**, *54*, 1–13. [CrossRef]
40. Dunn, C.W.; Giribet, G.; Edgecombe, G.D.; Hejnol, A. Animal phylogeny and its evolutionary implications. *Annu. Rev. Ecol. Evol. Syst.* **2014**, *45*, 371–395. [CrossRef]

41. Giribet, G. New animal phylogeny: Future challenges for animal phylogeny in the age of phylogenomics. *Org. Divers. Evol.* **2015**, *16*, 419–426. [CrossRef]
42. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [CrossRef] [PubMed]
43. Logacheva, M.D.; Schelkunov, M.I.; Penin, A.A. Sequencing and analysis of plastid genome in mycoheterotrophic orchid Neottia nidus-avis. *Genome Biol. Evol.* **2011**, *3*, 1296–1303. [CrossRef]
44. McNeal, J.R.; Kuehl, J.V.; Boore, J.L.; De Pamphilis, C.W. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biol.* **2007**, *7*, 57. [CrossRef]
45. Yang, Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol. Biol. Evol.* **1998**, *15*, 568–573. [CrossRef]
46. Davies, T.J.; Savolainen, V.; Chase, M.W.; Moat, J.; Barraclough, T.G. Environmental energy and evolutionary rates in flowering plants. *Proc. R. Soc. B* **2004**, *271*, 2195–2200. [CrossRef]
47. Clarke, A.; Gaston, K.J. Climate, energy and diversity. *Proc. Biol. Sci.* **2006**, *273*, 2257–2266. [CrossRef]
48. Marcelino, V.R.; Ma, C.; Jackson, C.J.; Larkum, A.; Verbruggen, H. Evolutionary dynamics of chloroplast genomes in low light: A case study of the endolithic green alga *Ostreobium quekettii*. *Genome Biol. Evol.* **2016**, *8*, 2939–2951. [CrossRef]
49. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]
50. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644. [CrossRef] [PubMed]
51. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [CrossRef]
52. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [CrossRef]
53. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59. [CrossRef]
54. Mistry, J.; Finn, R.D.; Eddy, S.R.; Bateman, A.; Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **2013**, *41*, e121. [CrossRef]
55. Emms, D.M.; Kelly, S. OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **2015**, *16*, 157. [CrossRef] [PubMed]
56. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]
57. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 972–1973. [CrossRef]
58. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [CrossRef]
59. Lanfear, R.; Frandsen, P.B.; Wright, A.M.; Senfeld, T.; Calcott, B. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **2016**, *34*, 772–773. [CrossRef] [PubMed]
60. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Hohna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 39–42. [CrossRef] [PubMed]
61. Trifinopoulos, J.; Lam-Tung, N.; von Haeseler, A.; Minh, B.Q. W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **2016**, *44*, 232–235. [CrossRef]
62. Minh, B.Q.; Nguyen, M.A.T.; von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **2013**, *30*, 1188–1196. [CrossRef]
63. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef]
64. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]
65. Zhang, C.; Rabiee, M.; Sayyari, E.; Mirarab, S. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* **2018**, *19*, 15–30. [CrossRef]
66. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger sets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
67. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef]
68. Philip, J.; Binns, D.; Chang, H.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240.
69. Huerta-Cepas, J.; Forslund, K.; Coelho, L.P.; Szklarczyk, D.; Jensen, L.J.; von Mering, C.; Bork, P. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* **2016**, *34*, 2115–2122. [CrossRef] [PubMed]

70. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *Omics J. Integr. Biol.* **2012**, *16*, 284–287. [CrossRef] [PubMed]
71. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [CrossRef]
72. Pollock, D.D.; Zwickl, D.J.; McGuire, J.A.; Hillis, D.M. Increased tax on sampling is advantageous for phylogenetic inference. *Syst. Biol.* **2002**, *51*, 664–671. [CrossRef]
73. Buchheim, M.A.; Sutherland, D.M.; Schleicher, T.; Förster, F.; Wolf, M. Phylogeny of Oedogoniales, Chaetophorales and Chaetopeltidales (Chlorophyceae): Inferences from sequence-structure analysis of ITS2. *Ann. Bot.* **2012**, *109*, 109–116. [CrossRef]
74. Caisová, L.; Marin, B.; Melkonian, M. A consensus secondary structure of ITS2 in the Chlorophyta identified by phylogenetic reconstruction. *Protist* **2013**, *164*, 482–496. [CrossRef]
75. Liu, B.; Hu, Y.; Hu, Z.; Liu, G.; Zhu, H. Taxonomic scheme of the order Chaetophorales (Chlorophyceae, Chlorophyta) based on chloroplast genomes. *BMC Genom.* **2020**, *21*, 442. [CrossRef]
76. Brouard, J.S.; Turmel, M.; Otis, C.; Lemieux, C. Proliferation of group II introns in the chloroplast genome of the green alga *Oedocladium carolinianum* (Chlorophyceae). *PeerJ* **2016**, *4*, e2627. [CrossRef]
77. Liu, B.; Zhu, H.; Dong, X.; Yan, Q.; Liu, G.; Hu, Z. Reassessment of Suitable Markers for Taxonomy of Chaetophorales (Chlorophyceae, Chlorophyta) based on Chloroplast Genomes. *J. Eukaryot. Microbiol.* **2021**, *68*, e12858. [CrossRef]
78. Smith, D.R. Mutation rates in plastid genomes. They are lower than you might think. *Genome Biol. Evol.* **2015**, *7*, 1227–1234. [CrossRef]
79. Guisinger, M.M.; Kuehl, J.V.; Boore, J.L.; Jansen, R.K. Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 18424–18429. [CrossRef] [PubMed]
80. Kato-Minoura, T.; Karino, K.; Akimoto, N.; Yoshiga, N.; Ehara, M.; Aoki, S. Phylogenetic analysis of NAP, an unconventional actin of the Volvocales. *Plant Syst. Evol.* **2015**, *301*, 1725–1733. [CrossRef]
81. Vigeland, M.D.; Spannagl, M.; Asp, T.; Paina, C.; Rudi, H.; Rognli, O.-A.; Fjellheim, S.; Sandve, S.R. Evidence for adaptive evolution of low-temperature stress response genes in a Pooideae grass ancestor. *New Phytol.* **2013**, *199*, 1060–1068. [CrossRef] [PubMed]
82. McGowen, M.R.; Grossman, L.I.; Wildman, D.E. Dolphin genome provides evidence for adaptive evolution of nervous system genes and a molecular rate slowdown. *Proc. R. Soc. B* **2012**, *279*, 3643–3651. [CrossRef] [PubMed]
83. Yu, J.; Wang, C.; Gong, X. Degeneration of photosynthetic capacity in mixotrophic plants, *Chimaphila japonica* and *Pyrola decorata* (Ericaceae). *Plant Divers.* **2017**, *39*, 80–88. [CrossRef]
84. Laurin, M. *How Vertebrates Left Water*; University of California Press: Berkeley, CA, USA, 2010.