



Article

# Evolutionary Patterns of the Chloroplast Genome in Vanilloid Orchids (Vanilloideae, Orchidaceae)

Young-Kee Kim, Se-Hwan Cheon , Ja-Ram Hong and Ki-Joong Kim \*

Division of Life Sciences, Korea University, Seoul 02841, Republic of Korea

\* Correspondence: kimkj@korea.ac.kr

**Abstract:** The Vanilloideae (vanilloids) is one of five subfamilies of Orchidaceae and is composed of fourteen genera and approximately 245 species. In this study, the six new chloroplast genomes (plastomes) of vanilloids (two *Lecanorchis*, two *Pogonia*, and two *Vanilla* species) were decoded, and then the evolutionary patterns of plastomes were compared to all available vanilloid plastomes. *Pogonia japonica* has the longest plastome, with 158,200 bp in genome size. In contrast, *Lecanorchis japonica* has the shortest plastome with 70,498 bp in genome size. The vanilloid plastomes have regular quadripartite structures, but the small single copy (SSC) region was drastically reduced. Two different tribes of Vanilloideae (Pogonieae and Vanilleae) showed different levels of SSC reductions. In addition, various gene losses were observed among the vanilloid plastomes. The photosynthetic vanilloids (*Pogonia* and *Vanilla*) showed signs of stage 1 degradation and had lost most of their *ndh* genes. The other three species (one *Cyrtosia* and two *Lecanorchis*), however, had stage 3 or stage 4 degradation and had lost almost all the genes in their plastomes, except for some housekeeping genes. The Vanilloideae were located between the Apostasioideae and Cypripedioideae in the maximum likelihood tree. A total of ten rearrangements were found among ten Vanilloideae plastomes when compared to the basal Apostasioideae plastomes. The four sub-regions of the single copy (SC) region shifted into an inverted repeat (IR) region, and the other four sub-regions of the IR region shifted into the SC regions. Both the synonymous (dS) and nonsynonymous (dN) substitution rates of IR in-cooperated SC sub-regions were decelerated, while the substitution rates of SC in-cooperated IR sub-regions were accelerated. A total of 20 protein-coding genes remained in mycoheterotrophic vanilloids. Almost all these protein genes show accelerated base substitution rates compared to the photosynthetic vanilloids. Two of the twenty genes in the mycoheterotrophic species faced strong “relaxed selection” pressure ( $p$ -value < 0.05).

**Keywords:** plastomes; SSC contraction; gene loss; gene relocation; Orchidaceae; Vanilloideae



**Citation:** Kim, Y.-K.; Cheon, S.-H.; Hong, J.-R.; Kim, K.-J. Evolutionary Patterns of the Chloroplast Genome in Vanilloid Orchids (Vanilloideae, Orchidaceae). *Int. J. Mol. Sci.* **2023**, *24*, 3808. <https://doi.org/10.3390/ijms24043808>

Academic Editor: Jen-Tsung Chen

Received: 18 January 2023

Revised: 10 February 2023

Accepted: 11 February 2023

Published: 14 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

One of the largest angiosperm families, Orchidaceae, consists of five subfamilies, 736 genera, and about 28,000 species [1]. The family is distributed worldwide and contains some mycoheterotrophs and many epiphytes. Some 43 genera among 736 genera contain mycoheterotrophs [2]. The mycoheterotrophic orchids are found in three subfamilies; Vanilloideae, Orchidoideae, and Epidendroideae. Vanilloideae comprises two tribes (Pogonieae and Vanilleae) [3]. Pogonieae contains five genera and 76 species, and Vanilleae consists of nine genera and 169 species [3]. Five genera (*Cyrtosia*, *Erythrorchis*, *Galeola*, *Lecanorchis*, and *Pseudovanilla*) of Vanilloideae are mycoheterotrophs [2]. Among the Vanilloideae, two species of *Pogonia*, one species of *Cyrtosia*, and two species of *Lecanorchis* are distributed in South Korea [4].

*Pogonia*, a genus of Pogonieae, consists of five photosynthetic species [3], which are distributed in the temperate regions of East Asia and eastern North America [5]. *P. japonica* and *P. minor* are native to the sunny and wet areas of southern Korea [4]. *Lecanorchis*, a genus of Vanilleae, comprises 20 obligate mycoheterotrophic species [2,3], which are

distributed in Southeast Asia, the Pacific Islands, China, Japan, and South Korea [5,6]. *L. kiusiana* and *L. japonica* are native to South Korea [4]. *Vanilla*, which is the largest genus in Vanilloideae, contains 105 species, which are distributed in Southeast Asia, South America, and Africa [3,7]. *Vanilla* species are diversified into climbing vines, and some species are cultivated in the tropics as fragrant fruit crops [8].

A large number of Orchidaceae plastomes have been reported since the first report of the *Phalaenopsis aphrodite* plastome [9]. However, most of the reported orchid plastomes are confined to some genera containing mycoheterotrophs (*Corallorhiza* and *Neottia*) or horticulturally cultivated species (*Cymbidium*, *Dendrobium*, and *Phalaenopsis*) [10–17]. Orchid plastomes show a wide range of gene contents. For example, *Cattleya* (Epidendreae), *Cephalanthera* (Neottieae), *Cymbidium* (Cymbidieae), *Kuhlhasseltia* (Cranichideae), *Neofinetia* (Vandae), and *Vanilla* (Vanilleae) plastomes do not contain various *ndh* genes [18–22]. The mycoheterotrophic orchid plastomes that are founded in *Corallorhiza* (Epidendreae), *Cyrtosia* (Vanilleae), *Epipogium* (Nervilieae), *Gastrodia* (Gastrodieae), *Neottia* (Neottieae), and *Rhizanthella* (Diurideae) are smaller in size and have fewer genes than photosynthetic orchids [10,11,15,23–27].

Thus far, only four complete plastome sequences of vanilloid orchids have been reported. Three *Vanilla* species (*V. pompona*, *V. planifolia*, and *V. aphylla*) have been sequenced, and their plastomes do not have some of the *ndh* genes [22,28]. Moreover, a small single copy region (SSC) in three *Vanilla* plastomes was largely reduced, and similar reductions have also been reported in *Paphiopedilum* (Cypripedioideae) and *Hetaeria* (Orchidoideae) [16,20,29]. The only sequenced mycoheterotrophic vanilloid, *Cyrtosia septentrionalis*, experienced a plastome size reduction and gene losses similar to other mycoheterotrophic orchids of other subfamilies [25]. The *C. septentrionalis* plastome also shows several gene rearrangements or gene relocations [25]. In order to understand the phylogeny of orchids, complete plastomes were used to construct a phylogenetic tree [20,30,31]. However, a few genes (*psaB*, *rbcL*, or *psbC*) were used to construct a phylogenetic tree of vanilloids [32–35] because of the limitation of whole plastome sequences. A few nuclear regions, such as the internal transcribed spacer (ITS) and *xdh*, were used to search Vanilloideae's phylogenetic relationship [33,36–38]. However, the phylogenetic study of Vanilloideae has not been studied intensively because of the limited taxon sampling.

In this study, first, we documented six new vanilloid plastomes: two photosynthetic *Pogonia* (*P. minor* and *P. japonica*), two mycoheterotrophic *Lecanorchis* (*L. kiusiana* and *L. japonica*), and two photosynthetic *Vanilla* (*V. madagascariensis* and *V. planifolia*). Second, we validated the modifications of vanilloid plastomes and compared them to other orchid plastomes, such as SSC reductions, gene relocations, and gene reductions. Third, we used the elucidated plastome sequences to construct a phylogenetic tree so that the phylogenetic position and the divergence date of major vanilloids could be discussed. Lastly, in order to elucidate the evolutionary modes of vanilloid plastomes, we compared the synonymous (dS) and the nonsynonymous (dN) base substitution rates between the relocated and un-relocated regions both in the photosynthetic species and in the mycoheterotrophic species. In addition, we evaluated differential selection pressures on the different regions and groups.

## 2. Results

### 2.1. General Features of Vanilloideae plastomes

The average coverage depths, voucher specimens, DNA numbers, and lengths of assembled plastomes are given in Table 1, and the NGS results are given in Supplementary Table S1. The coverage depth of newly sequenced species ranged from 76.9× (for *Lecanorchis japonica*) to 4394× (for *Vanilla madagascariensis*). The plastome size of *L. japonica* was the shortest, with 70,498 bp, and the size of *Pogonia japonica* was the longest, with 158,200 bp. Almost all plastomes have a quadripartite structure, which is composed of an large single copy (LSC), an SSC, and two IRs'. The size variations of the plastome regions of ten Vanilloideae and four Apostasioideae species are graphically presented in

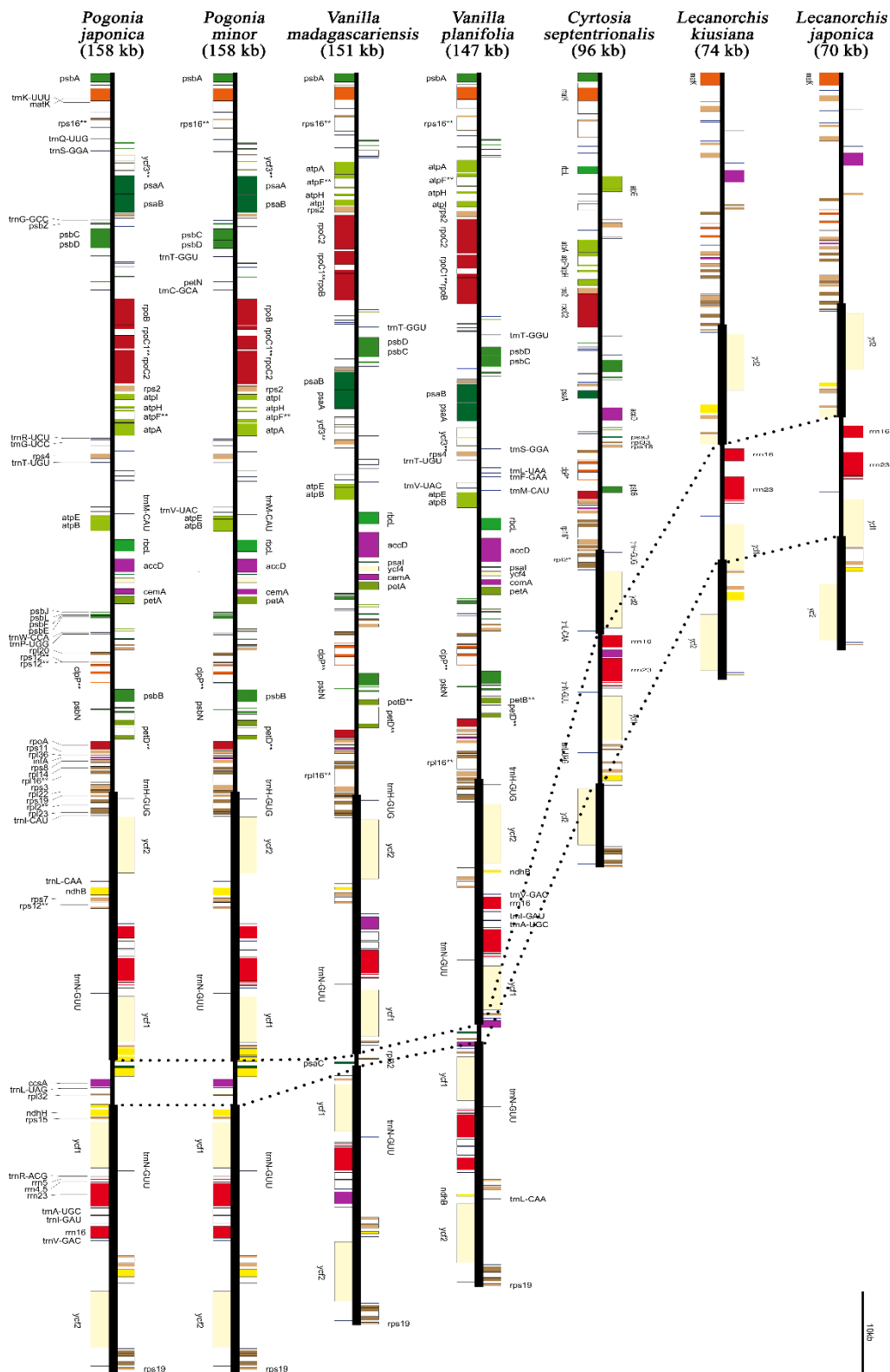
Figure 1. One noticeable size variation is the extensive reduction of the SSC region in the photosynthetic vanilloids. The SSC region ranged from 1254 bp in *V. madagascariensis* to 5885 bp in *P. japonica*. The annotated plastome maps of seven vanilloids belonging to the genera of *Cyrtosia*, *Lecanorchis*, *Pogonia*, and *Vanilla* are shown in Figure 2.

Table 1. The general features of six newly decoded Vanilloideae species.

Scientific Name	Voucher Specimen and DNA Number	Total Length (bp)	Coverage (x)
<i>Lecanorchis japonica</i>	PDBK2018-0250	70,498	76.9
<i>Lecanorchis kiusiana</i>	PDBK2018-0249	74,084	128.9
<i>Pogonia japonica</i>	PDBK2015-1272	158,200	464.6
<i>Pogonia minor</i>	PDBK2011-1673	158,170	563.2
<i>Vanilla</i> <i>madagascariensis</i>	PDBKTMA2013-1861	151,552	4394
<i>Vanilla planifolia</i>	PDBKTMA2013-1860	147,714	592



Figure 1. Plastome size variation among 14 species of Vanilloideae and Apostasioideae. The red scientific names indicate newly decoded plastomes. Green bars indicate the photosynthetic species, and orange bars indicate the mycoheterotrophic species. Blue bars indicate the four Apostasioideae species.



**Figure 2.** Seven representative plastomes of Vanilloideae. *Pogonia japonica* has the longest plastome, with 158.2 kb. *Pogonia* and *Vanilla* are the photosynthetic species, while *Cyrtosia* and *Lecanorchis* are the nonphotosynthetic mycoheterotrophic species. A few photosynthesis-related genes or pseudogenes remain in *Cyrtosia*, while all photosynthesis-related genes have been lost in *Lecanorchis*. Green gene boxes indicate genes related to the light reaction of photosynthesis. *L. kiusiana* and *L. japonica* hold only 33 and 32 gene regions, respectively. The heavy vertical lines indicate inverted repeat (IR) regions. The dashed lines indicate the single copy (SSC) regions.

Almost all the genes from the nonphotosynthetic vanilloid species were found to be pseudogenized or lost. For example, all photosynthesis-related genes were lost from the plastomes of two mycoheterotrophic *Lecanorchis* species. Only 32 housekeeping genes (*accD*, *clpP*, *infA*, *matK*, *rpl2*, *rpl14*, *rpl16*, *rpl20*, *rpl36*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps11*, *rps12*, *rps14*, *rps18*, *rps19*, 16S rRNA, 23S rRNA, 4.5S rRNA, 5S rRNA, *trnC-GCA*, *trnD-GUC*, *trnF-GAA*, *trnM-CAU*, *trnI-CAU*, *trnN-GUU*, *trnQ-UUG*, *ycf1*, and *ycf2*) are present in *L. japonica*. In addition to the 32 genes, the *trnE-UUC* gene is present in the *L. kiusiana* plastome. The plastome of heterotrophic *Cyrtosia* showed intermediate gene contents between *Vanilla* and *Lecanorchis*. A few photosynthesis-related genes or pseudogenes still remain in the *Cyrtosia* plastome, while all photosynthesis-related genes have been lost in *Lecanorchis*. The detailed gene contents of vanilloids and the other related orchids are given in Supplementary Figure S1.

## 2.2. Comparative and Phylogenetic Analyses

In order to validate the structural changes in the vanilloid plastome, the SC-IR junctions were analyzed among 14 Vanilloideae and Apostasioideae species using an IRScope. In the photosynthetic species, the LSC-IRa and LSC-IRb junctions are located within the *rpl22* gene and the intergenic spacer between *trnH* and *psbA*, respectively. In the nonphotosynthetic species, the LSC-IRa junction is located within *rps19* (*Cyrtosia*) or *rpl2* (*Lecanorchis*), and the LSC-IRb junctions are located near the *matK* gene (Supplementary Figure S2). In the typical orchid plastome, the SSC-IRa and SSC-IRb junctions are located within the *ndhF* gene and the *ycf1* gene, respectively. The SSC-IRa/b junctions of the four Apostasioideae species follow the same typical orchid form. Both the SSC-IRa and the SSC-IRb junctions in *Lecanorchis* are also located within *ycf1*. However, these junctions occurred within *ccsA* (*Vanilla*) and *ndhA* (*Pogonia*) in photosynthetic vanilloids. In the *Cyrtosia* plastome, the SSC-IRa and the SSC-IRb junctions are located near *rrn16* and *ndhB*, respectively.

The gene orders of 14 plastomes were analyzed using progressiveMauve. A total of ten gene rearrangement events were found (Table 2). The longest inversion event occurred between *trnS-GCU* and *ycf3* in the LSC region, with a length of 58.591 kb. The shortest rearrangement occurred between *ycf3* and *trnS-GGA*, and the inverted length was 3.282 kb. In addition to the 10 gene rearrangement events, several gene relocations occurred in mycoheterotrophic vanilloids. For example, the 16S rRNA, 23S rRNA, 4.5S rRNA, and 5S rRNA genes are usually positioned in the IR region. However, those four genes were in the SSC region in *L. kiusiana* and *L. japonica*.

**Table 2.** Rearrangement blocks in the result of progressiveMauve among fourteen species (ten Vanilloideae species and four Apostasioideae species were used).

Rearrangement	Length	Region
A	3.282	<i>ycf3-trnS(GGA)</i>
B	58.591	<i>trnS(GCU)-ycf3</i>
C	24.647	<i>trnS(GCU)-rbcL</i>
D	6.658	<i>accD</i>
E	7.199	<i>ndhB-rps12</i>
F	4.037	<i>trnV(GAC)</i>
G	4.626	<i>rpl32-ndhF</i>
H	8.899	<i>trnL(UAG)-ndhA</i>
I	6.469	<i>ycf1-ndhA</i>
J	7.284	<i>ycf1</i>

The phylogenetic relationship of the vanilloid species was identified using a maximum likelihood (ML) tree and a Bayesian inference tree, which were constructed using 83 genes (Figure 3 and Supplementary Figure S3). Both trees support the relationship of Apostasioideae (Vanilloideae(Cypripedioideae(Orchidoideae,Epidendroideae))). Almost all nodes were supported by a 100% bootstrap value and a 1.0 Bayesian posterior probability. In the subfamily Vanilloideae, two monophyletic groups, corresponding to

the tribe Pogonieae and the tribe Vanilleae, were recognized. In the tribe Vanilleae, two clades corresponding to the photosynthetic group (*Vanilla*) and mycoheterotrophic group (*Cyrtosia* and *Lecanorchis*) were further identified. The *Vanilla* clade was further divided into two subclades corresponding to leafless species (*V. aphylla* and *V. madagascariensis*) and leafy species (*V. pompona* and *V. planitifolia*). The branch lengths leading to the mycoheterotrophic vanilloids are relatively longer than the branch lengths leading to the photosynthetic vanilloids in the ML tree.

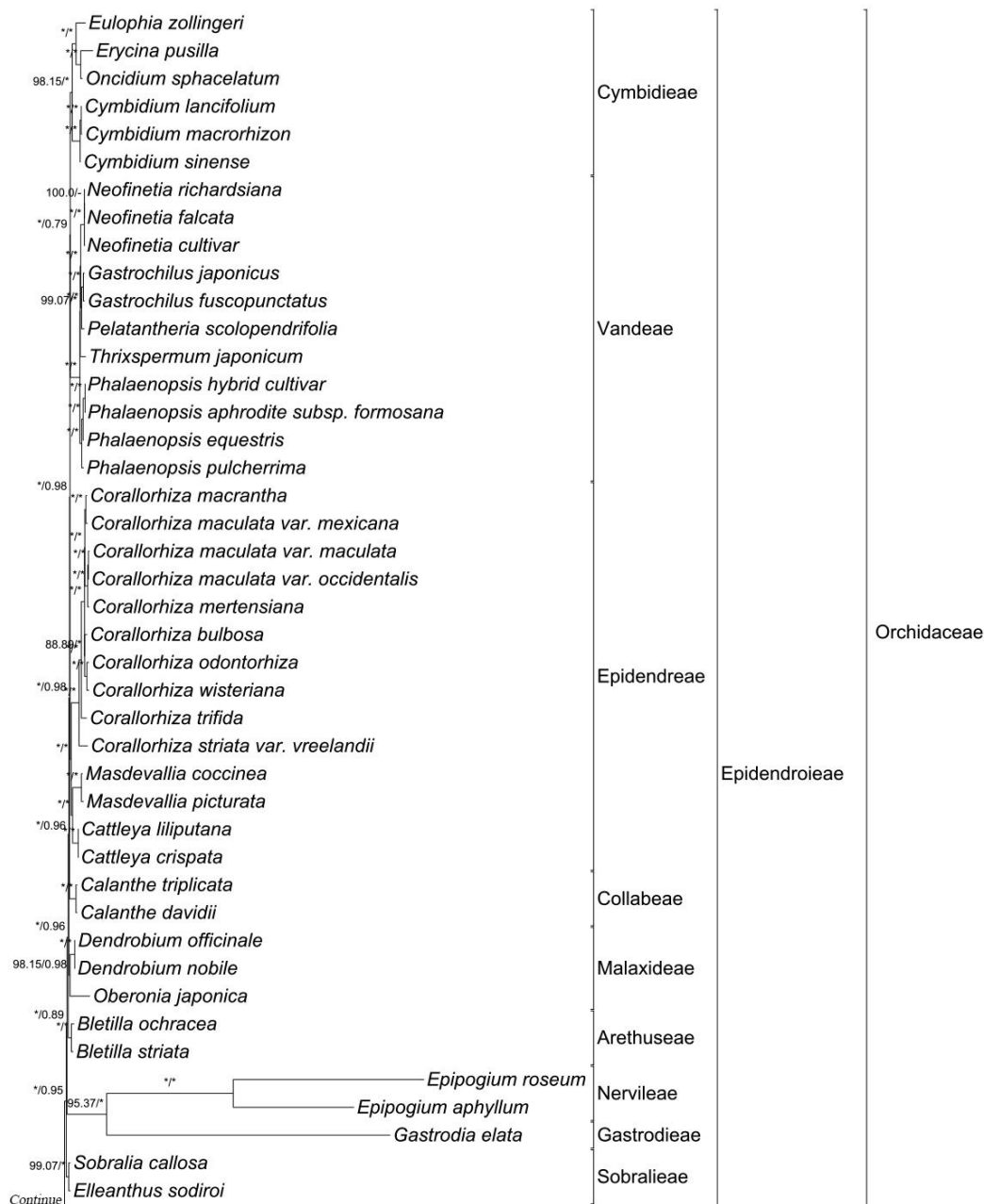
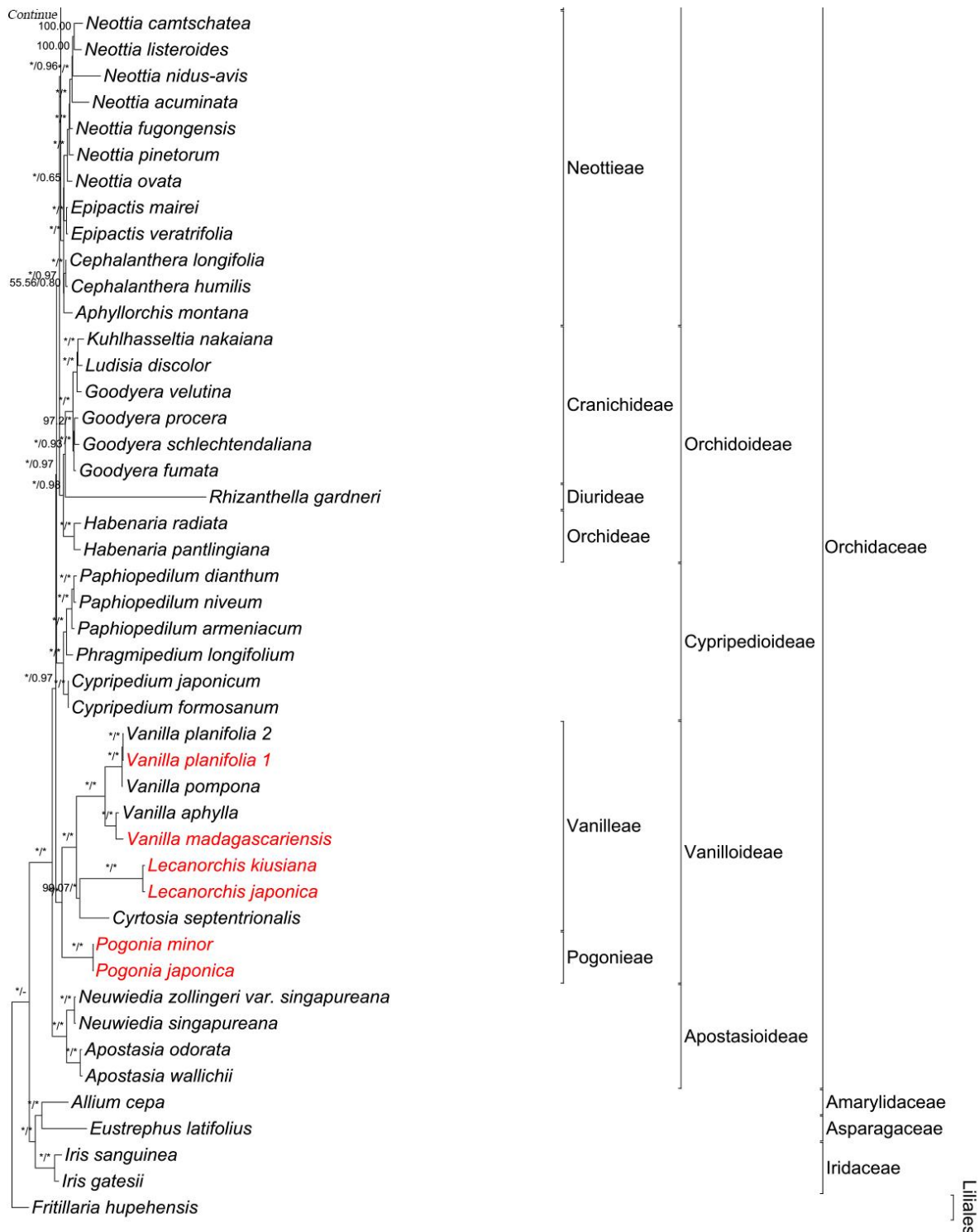


Figure 3. Cont.



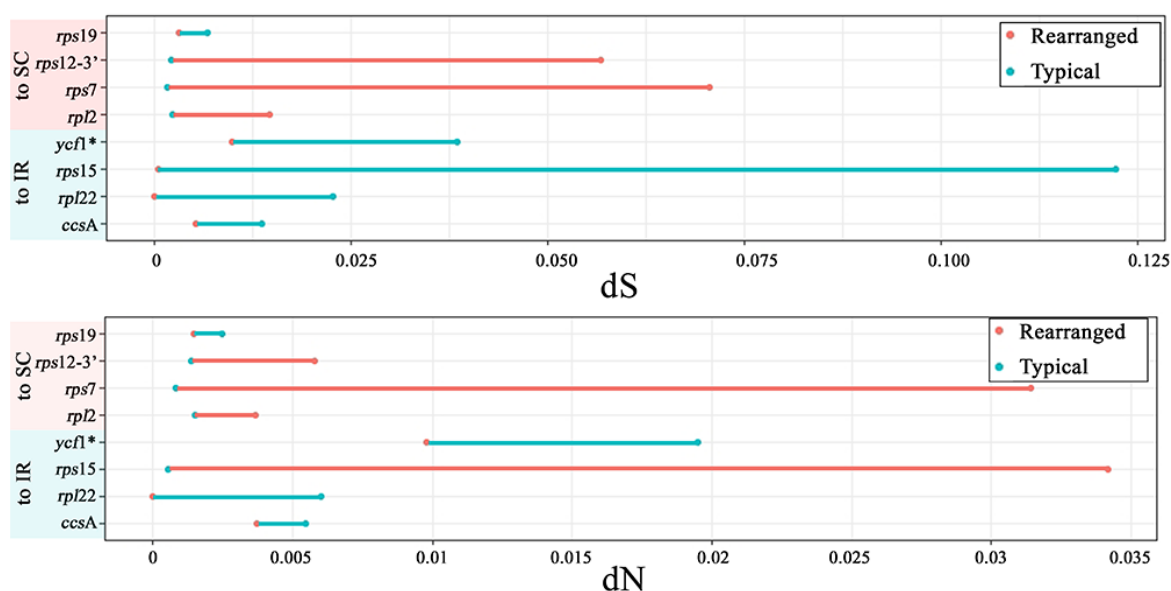
**Figure 3.** A maximum likelihood (ML) tree inferred from 89 species of Orchidaceae. The nucleotide sequences of 79 protein-coding genes and 4 rRNA genes were aligned individually and concatenated to be a single matrix. The aligned matrix was 87,230 bp, and the tree was reconstructed by RaxML with the GTR + G + I model (−563,601.727423 of ML value). The scientific names with red color indicate six newly decoded Vanilloideae species. This is an abbreviated tree, so the taxa of Cyripedioideae, Orchidoideae, and Epidendroideae are only given as a number of used species. The full tree structure is given in Supplementary Figure S3. An asterisk (\*) indicates that 100% of bootstrap value or 1 of Bayesian posterior probability.

### 2.3. Divergence Time of Vanilloideae

Orchidaceae was inferred to have diverged from other Asparagales at 101.1 (85.8–117.2) mya (Supplementary Figure S4). A basal subfamily, Apostasioideae, was estimated to have diverged from other subfamilies at 84.49 (64.5–101.8) mya. Then, the Vanilloideae was inferred to have diverged from other subfamilies at 76.09 (57.6–94.0) mya. The two tribes of Vanilloideae, Vanilleae, and Pogonieae were estimated to have separated from each other at 61.54 (45.1–80.0) mya.

### 2.4. Evolutionary Rate Comparison

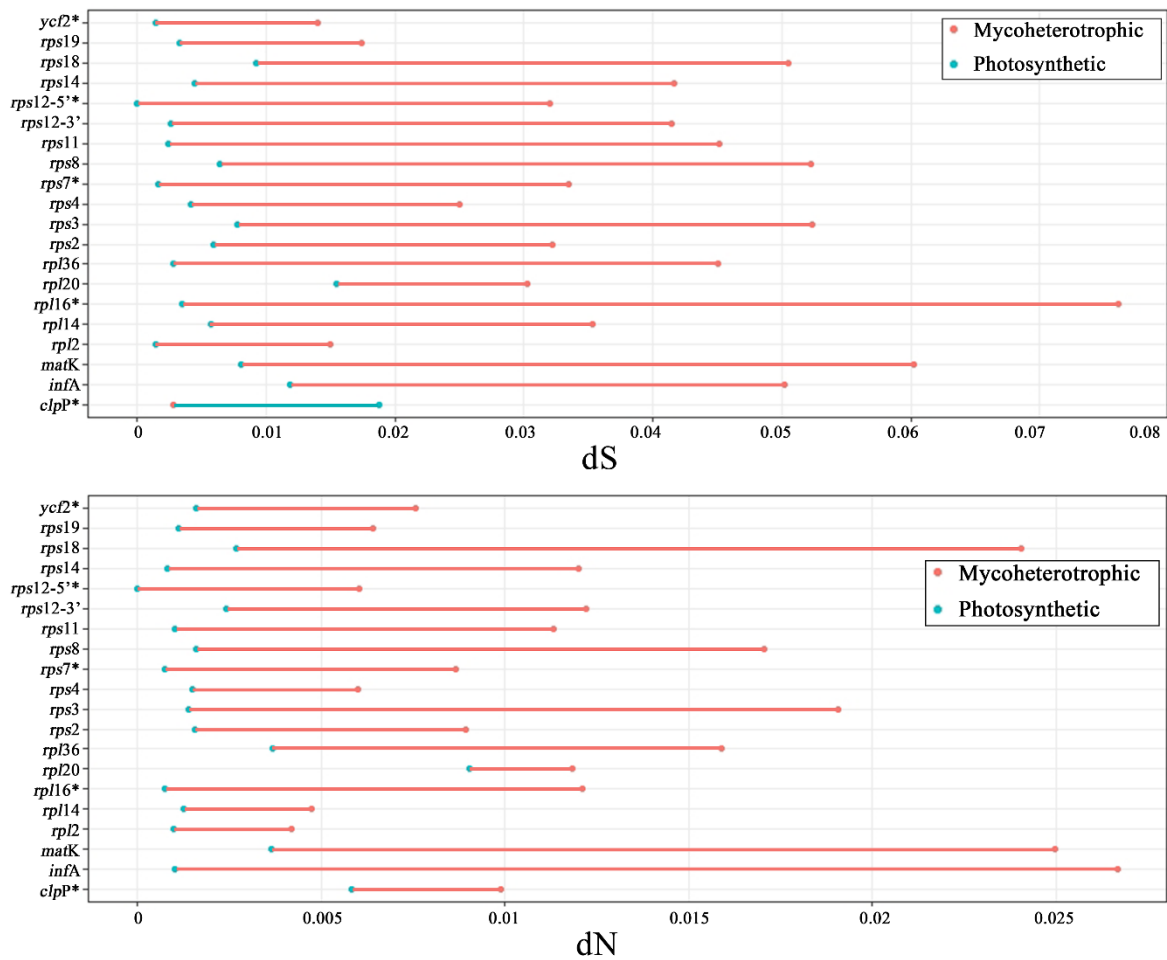
The SSC contraction, mycoheterotrophic life cycle, and rearranged plastome regions are notable features in Vanilloideae plastomes because they are not found in other typical orchid plastomes. The regions that survived in a plastome after a mycoheterotrophic event of *Cyrtosia* and *Lecanorchis* are listed. The rearranged regions are also listed. These listed regions were tested to calculate the nonsynonymous/synonymous substitution rates, and the results are listed (Figures 4 and 5, Supplementary Table S2).



**Figure 4.** Comparison of synonymous (dS) and nonsynonymous (dN) substitution rates between species with rearranged and conserved plastomes in eight regions. A region with a significant  $p$ -value ( $p$ -value < 0.05) was tagged by an asterisk (\*) in its gene name. The red color in the bar graph indicates an acceleration of the substitution rates; the blue color indicates a deceleration of the substitution rates.

The genetic regions that were relocated from the typical Orchidaceae plastome to the Vanilloideae plastome are called rearranged regions in this study. There are eight rearranged regions: *ccsA*, *rpl2*, *rpl22*, *rps7*, *rps12-3'*, *rps15*, *rps19*, and *ycf1*. Among the eight regions, four (*rpl2*, *rps7*, *rps12-3'*, and *rps19*) were moved from the IR region to the SC region. The other four regions (*ccsA*, *rpl22*, *rps15*, and *ycf1*) were moved from the SC region to the IR region. In the case of nonsynonymous substitution rates (dN), the difference between the dN values of the rearranged group and the typical group does not seem related to the region shift. Meanwhile, the synonymous substitution rates (dS) value of the rearranged group differed from the typical group for regions that were shifted to IR. The IR-shifted group showed a much lower dS value than the typical group (Figure 4). A log-likelihood test (LRT) was performed between the null model and the alternative model. The *ycf1* region was the only statistically significant region among the eight shifted regions ( $p$ -value < 0.05).





**Figure 5.** Comparison of synonymous/nonsynonymous substitution rates between mycoheterotrophic species and their photosynthetic relatives in the 20 commonly remaining protein-coding gene or gene regions. A region with a significant  $p$ -value ( $p$ -value < 0.05) is tagged by an asterisk (\*). Red line indicates an acceleration of the substitution rates; blue line indicates a deceleration of the substitution rates.

There are twenty conserved regions in the Vanilloideae mycoheterotrophic plastome: *clpP*, *infA*, *matK*, *rpl2*, *rpl14*, *rpl16*, *rpl20*, *rpl36*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps11*, *rps14*, *rps18*, *rps19*, *rps12-3'*, *rps12-5'*, and *ycf2* (Figure 5). For all those regions, the dN and dS values were calculated using a PAML. As a result, the dN and dS values of nineteen regions, except for the dS value of *clpP*, were higher than regions from photosynthetic plastome. LRT was also performed between the null model and the alternative model so that five regions (*clpP*, *rpl16*, *rps7*, *rps12-5'*, and *ycf2*) were statistically significant regions among twenty conserved regions ( $p$ -value < 0.05).

RELAX was performed using twenty-eight regions to check which evolutionary pressure was applied to the plastome. As a result, a total of six regions (*clpP* and *ycf2* of the mycoheterotrophic conserved regions; *rpl2*, *rps12-5'*, *rps15*, and *ycf2* of the rearranged region) were statistically significant ( $p$ -value < 0.05). Only the *rpl2* region showed an intensification selection, while the other five regions showed a relaxation selection among statistically significant regions (Table 3).

**Table 3.** Estimated selection pressure in the result of RELAX. Regions with significant p-values are listed in the table.

Type	Gene	Selection	k	p-Value	LR	w1	w2	w3
Mycobacteriophages	<i>clpP</i>	Relaxation	0.39	0.022	5.27	0/0 (78.27%)	1/1 (4.27%)	8.02/2.25 (17.47%)
	<i>ycf2</i>	Relaxation	0.45	0	49.96	0.95/0.98 (97.53%)	164.08/9.95 (2.47%)	-
Regional	<i>rpl2</i>	Intensification	4.22	0.015	5.97	0.44/0.03 (20.13%)	0.58/0.10 (58.93%)	1.00/1.46 (20.95%)
	<i>rps15</i>	Relaxation	0	0.003	8.74	0.27 (20.26%)/1.00 (2.64%)	0.30 (77.09%)/1.00 (77.09%)	32.11 (2.64%)/1.00 (77.09%)
	<i>rps12-3'</i>	Relaxation	0	0.048	3.91	0.00/0.00 (80.01%)	0.55/1.00 (15.17%)	292.30/1.00 (4.82%)
	<i>ycf1</i>	Relaxation	0.85	0.022	5.24	0.00/0.01 (24.94%)	1.00/1.00 (72.32%)	192.01/86.05 (2.74%)

### 3. Discussion

#### 3.1. Plastome Structure Evolution

The SSC contraction is an interesting feature of the Vanilloideae plastome (Figures 1 and 2). Though an SSC in a typical Orchidaceae is 10–18 kb in length, it ranges from 1.2–17 kb in length in the Vanilloideae plastome. However, the whole length of the Vanilloideae plastome is not much different from the typical Orchidaceae (Supplementary Table S3). The SSC contraction occurs mostly in the photosynthetic Vanilloideae. *Vanilla madagascariensis*, a photosynthetic member of Vanilleae, possesses the shortest SSC, which is 1.2 kb in length. The longest SSC of the photosynthetic Vanilloideae is 5.3 kb in *Pogonia japonica*. The SSC contraction event in photosynthetic Vanilloideae occurs in two tribes (Pogonieae and Vanilleae). Interestingly, the two tribes experience different levels of SSC contraction. In Pogonieae, *Pogonia's* SSC is 5.3 kb in length. However, the length of *Vanilla's* SSC ranges from 1.2–2.1 kb. *Pogonia* is treated as a basal taxon in the maximum likelihood and Bayesian phylogenetic trees. Because the SSC of *Vanilla* is shorter than *Pogonia's*, SSC contraction may be more extreme in the crown taxa of Vanilloideae.

The SSC contractions are common events in Vanilloideae. However, the SSC contraction in photosynthetic orchids is a rare evolutionary event from three other orchid subfamilies. It has only been reported in three species of *Paphiopedilum* (Cypripedioideae) and one species of *Hetaeria* (Orchidoideae) [16,20,29]. *Paphiopedilum's* SSC ranges from 2.4–5.1 kb, and *Hetaeria's* SSC is 2.3 kb. In photosynthetic orchids, SSC contraction is found in only three genera (*Paphiopedilum*, *Pogonia*, and *Vanilla*). Though very few cases are observed in photosynthetic orchids, most mycoheterotrophic orchids, such as *Neottia* (*N. acuminata*, *N. camtschatea*, *N. listeroides*, and *N. nidus-avis*), *Epipogium* (*E. aphyllum*, and *E. roseum*) have plastomes with shortened SSCs [11,15,26]. However, SSC contraction in mycoheterotrophs is an additional phenomenon that comes after gene reduction. Moreover, mycoheterotrophic Vanilloideae (*Cyrtosia*, *Lecanorchis*) has an SSC of 10–14 kb in length [25]. The SSC contraction event might be an independent event in Vanilloideae.

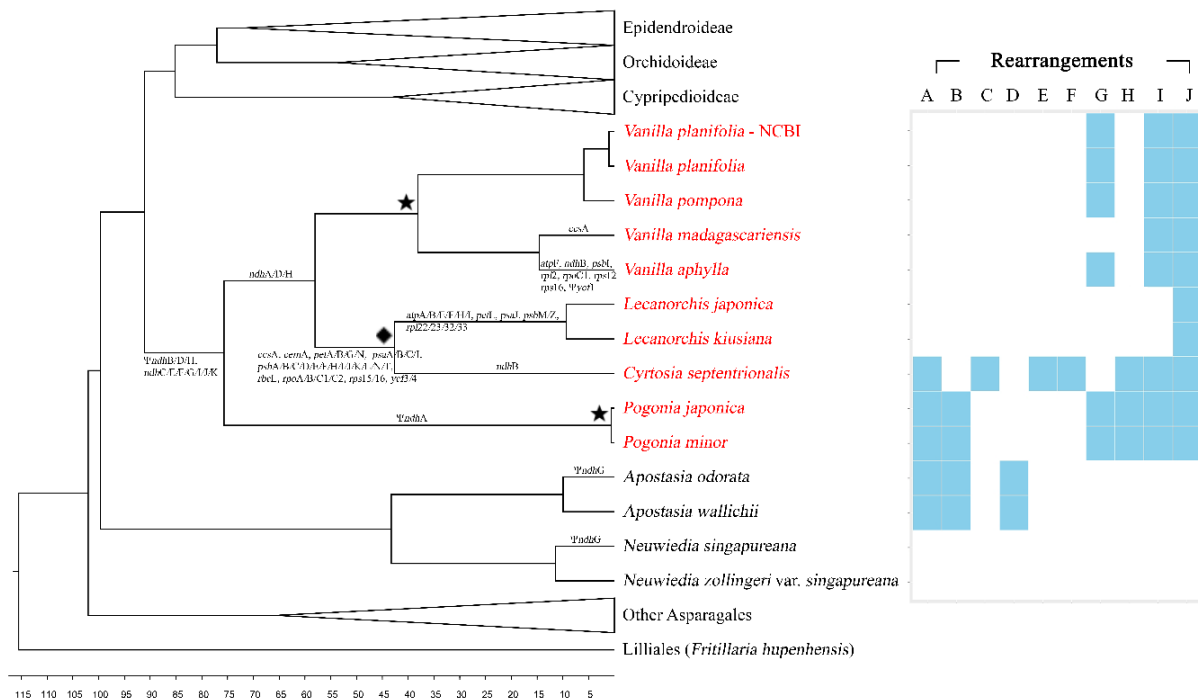
Another important event of the Vanilloideae plastome is gene reduction. In the case of Vanilloideae plastome, *ndh* genes are deleted or pseudogenized like in other Orchidaceae genera, such as *Cymbidium*, *Goodyera*, and *Neofinetia* [17,21,22,29]. Except for the *ccsA* gene of *Vanilla madagascariensis* (which is affected by SSC contraction), gene components of photosynthetic Vanilloideae (*Pogonia* and *Vanilla*) are similar to those of other photosynthetic orchids, such as *Habenaria*, *Phalaenopsis* [9,39].

On the other hand, mycoheterotrophic Vanilloideae (*Cyrtosia* and *Lecanorchis*) have fewer plastome genes than photosynthetic Vanilloideae. Mycoheterotrophs have experienced plastome gene reduction, which can be categorized by stages of degradation [40–43]. Orchids that lack the *ndh* gene can be categorized as degradation stage 1. In degradation stage 2, photosynthesis-related genes (*pet*, *psa*, *psb*, *rpo*) are usually reduced or pseudogenized [40,43]. The mycoheterotrophic Vanilloideae, *Cyrtosia septentrionalis*, was treated as an

intermediate stage between stage 2 and stage 3 reduction [25]. In stage 3, only housekeeping genes, such as *trn*, *rps*, and *rpl*, remain in the plastome. A range of 27–33 genes remains in the orchid plastomes at degradation stage 3. It can be found in *Epipogium*, *Gastrodia*, and *Rhizanthella* [24,26,27]. *Lecanorchis japonica* and *L. kiusiana* have 32 or 33 genes, respectively, indicating that *Lecanorchis* has experienced stage 3 degradation.

Furthermore, rRNA gene relocation from the IR to the SSC region occurred in mycoheterotrophic Vanilloideae (Figure 2). Four rRNAs (4.5s rRNA, 5s rRNA, 16s rRNA, and 23s rRNA) usually stayed in the IR of a typical angiosperm plastome [9,44,45]. However, the relocation of eight genes (4.5s rRNA, 5s rRNA, 16s rRNA, 23s rRNA, *rps7*, *rps12*, *rps15*, and *ycf1*) was observed in *Cyrtosia septentrionalis* [25]. A similar relocation was discovered in the *Lecanorchis kiusiana* and *L. japonica* (Figure 2). In the *Lecanorchis* plastome, *rps7* and *rps12* were not moved to the SSC region, counter to the *Cyrtosia* case. The *rps15* gene was deleted in the *Lecanorchis* plastome. This type of rRNA relocation in Orchidaceae has only been observed in *Rhizanthella gardneri* [24]. Four rRNAs, *rps7*, *rps12*, and *ycf1*, were moved from the IR to the SSC region in *Rhizanthella*. Other mycoheterotrophic orchids, such as *Corallorhiza*, *Epipogium*, *Gastrodia*, or *Neottia*, did not experience this rRNA relocation [10,11,15,23,26,27].

Genome rearrangement is another notable feature of the Vanilloideae plastome. In order to document the structural rearrangement, the Vanilloideae and Apostasioideae plastomes were compared. A total of ten rearrangements were found (Table 2) among species. The ten rearrangements have been plotted in the phylogenetic tree (Figure 6). In the genus *Neuwiedia* of Apostasioideae, the plastome structure is identical to other typical orchid plastomes. No structural rearrangements or gene reduction occurred in this taxon. However, *Apostasia*, the other genus of Apostasioideae, shares two rearrangements with *Pogonia* and *Cyrtosia* of Vanilloideae: rearrangement A (*ycf3-trnS*-GGA, 3.282 kb) and rearrangement B (*trnS*-GCU-*ycf3*, 58.591 kb). One more independent rearrangement (rearrangement D-*accD*, 6.658 kb) is detected in *Apostasia*.



**Figure 6.** Summary of the rearrangements, mycoheterotrophy, SSC contraction, and gene reduction in the time estimation tree. A list of deleted or pseudogenized genes is above or beneath the tree nodes. A diamond mark in the phylogenetic tree indicates mycoheterotrophy of the clade. Star marks in the phylogenetic tree indicate the SSC contraction. Red scientific names indicate a Vanilloideae species.

*Pogonia* and *Vanilla* share three rearrangements: G (*rpl32-ndhF*, 4.626 kb), I (*ycf1-ndhA*, 6.469 kb), and J (*ycf1*, 7.284 kb). *Cyrtosia* shares rearrangement I and J with *Pogonia* and *Vanilla*. *Lecanorchis* shares rearrangement J only with other Vanilloideae. *Pogonia* independently has rearrangement H (*trnL-UAG-ndhA*, 8.899 kb). This different rearrangement sharing may be caused by variations in SSC length. The *Pogonia* plastome has a longer SSC than *Vanilla*. In it, the *ndhA* region is generally located in the SSC region. *Pogonia*'s SSC region is long enough to preserve the *ndhA* region, but *Vanilla*'s SSC region is not. Thus, rearrangement H, which includes the *ndhA* region, occurs in the genus *Pogonia* only. In the case of mycoheterotrophic species, *Cyrtosia* has three independent rearrangements: rearrangements C (*trnS-GCU-rbcL*, 24.647 kb), E (*ndhB-rps12*, 4.626 kb), and F (*trnV-GAC*, 4.037 kb). *Lecanorchis* does not show any other independent rearrangement due to its extremely degraded plastome. Thus, it seems likely that *Lecanorchis* had the common Vanilloideae rearrangements (such as G, H, I, and J) first and later experienced rearrangement losses due to the gene losses (Figure 6). However, it is difficult to argue which events occurred first due to the limited plastome sequence availability.

### 3.2. Evolutionary Rates and Selection Pressure

Eight relocated regions were affected by IR expansion/contraction (the *ccsA*, *rpl22*, *rps15*, and *ycf1* regions moved to IR; the *rpl2*, *rps7*, *rps12-3'*, and *rps19* regions moved to SC). The substitution rates of those eight regions were compared between the reference taxa (typical position of regions) and test groups (relocated). The nonsynonymous substitution rates (dN) are not different between the reference and test groups. However, the synonymous substitution rates (dS) are much lower in the test group than in the reference group in the case of the regions that were moved to the IR region. Similarly, the other four regions of the test group, which were moved to the SC region, show higher dS values than the reference group (Figure 4).

In the general angiosperm plastome, the IR strengthens the stability of the plastome structure. This structural stability is commonly supported by the copy-dependent repair mechanism (flip-flop recombination) stability [46,47]. For that reason, regions in IR show lower substitution rates or genetic diversity. Previous studies of other angiosperms, such as *Sesamum* (Lamiaceae) and *Pelargonium* (Polygonaceae), support the lower substitution rates of IR regions [48,49].

Mycoheterotrophic species usually have decreased plastome size and fewer plastome genes. Previous studies categorized mycoheterotrophic species by the level of plastome degradation [40,43]. According to those criteria, the *Lecanorchis* plastome can be categorized as stage 4 degradation, also known as the "stationary" stage [43]. In twenty retaining gene regions of *Lecanorchis*, the dN and dS substitution rates were compared between the test (mycoheterotrophic taxa; *Cyrtosia* and *Lecanorchis*) and reference groups (other photosynthetic relatives). Nineteen regions, except for *clpP*, have much higher substitution rates. There was a similar result in the genus *Rhizanthella*, which is an obligate mycoheterotroph of Ochidoideae [24]. In the case of the mycoheterotrophic *Corallorhiza*, however, it is reported that the w-ratio is not different from its photosynthetic relatives [50]. Although almost every region (except the only *clpP* region) shows elevated substitution rates, there are just five statistically significant regions in both the dS and dN comparisons (*clpP*, *rpl16*, *rps7*, *rps12-5'*, and *ycf2*). Therefore, this is not enough evidence to claim that the mycoheterotrophic life cycle affects retention regions' elevated substitution rates.

The elevated substitution rates may be affected by selection pressure. In mycoheterotrophs, two regions (*clpP* and *ycf2*) show significant *p*-values (*p*-value < 0.05), which were acquired by RELAX. Those two regions were under relaxation selection pressure ( $K < 1$ ), which means that the surviving regions in mycoheterotrophic species' plastomes tend to be removed. Four regions under the position effect (*rpl2*, *rps15*, *rps12-3'*, and *ycf1*) showed significant *p*-values. Two regions (*rps15* and *ycf1*), which were moved to the IR region from the SC region, experienced relaxation selection ( $K < 1$ ), and the substitution rates were decreased in both. The other two regions (*rpl2* and *rps12-3'*) show accelerated

substitution rates but different selection pressure: *rpl2* experienced intensification selection pressure, and *rps12-3'* experienced relaxation selection pressure.

### 3.3. Phylogenetic Position and Time Estimation

In the maximum likelihood and Bayesian inference phylogenetic trees, *Epipogium*, *Gastrodia*, and *Rhizanthella*, which are mycoheterotrophs, have more elongated branch lengths compared to other photosynthetic taxa (Figure 3). Though the above three genera have extremely degraded plastomes, they have much fewer informative sites of the data matrix than other taxa. This leads to an elongated branch in the phylogenetic tree. However, mixed genera, such as *Corallorhiza* or *Neottia*, which contain both the photosynthetic and mycoheterotrophic species, have similar branch lengths to other typical clades, which are not significantly different. Because Vanilloideae also contains photosynthetic and mycoheterotrophic species, we also expected a similar branch length. However, the actual branch length of Vanilloideae seems to be elongated, especially in the genera with extremely degraded plastomes (*Cyrtosia* and *Lecanorchis*). Moreover, the branch length is longer in Vanilleae than in Pogonieae. This is presumed to be because the mycoheterotrophic genera are included in Vanilleae.

The whole chloroplast genome was used to estimate the derived time of Orchidaceae [20,30,51]. Vanilloideae was derived at 76.09 (57.6–94.0) mya (Supplementary Figure S4) in this study. Though this estimation is much more recent than some from other previous studies, it is within the error range [20,30]. *Cyrtosia* and *Lecanorchis* were derived at 32.75 (18.49–52.45) mya. These two mycoheterotrophic genera have their plastomes at degradation stages 3 to 4. Other genera with plastomes at degradation stage 3 or 4, such as *Gastrodia* (Gastrodieae), *Epipogium* (Nervileae), and *Rhizanthella* (Diurideae), were derived at 39.17 (29.10–50.54) mya, 39.17 (29.10–50.54) mya, and 36.52 (25.45–54.49) mya, relatively. These estimations are very similar to the estimated time of mycoheterotrophic Vanilloideae (*Cyrtosia* and *Lecanorchis*). Old clades tend to have severe plastome degradation since they have had enough time to accumulate mutations [43].

Other mycoheterotrophic taxa, which have only slightly degraded plastomes, were derived much more recently than taxa with stage 3 or 4 degradations. *Corallorhiza* was derived at 21.45 (12.7–30.29) mya, *Aphyllorchis* was derived at 31.62 (14.15–52.01) mya, *Neottia* was derived at 25.66 (15.40–38.76) mya, and *Cymbidium* was derived at 25.36 (15.03–36.61) mya. *Aphyllorchis* (Neottieae), *Cymbidium* (Cymbidieae), *Corallorhiza* (Epidendreae), and *Neottia* (Neottieae) are composed of obligate mycoheterotrophs, mixotrophs, and even green relatives. Except for *Aphyllorchis*, the other three genera were derived slightly recently, before 30 mya. This result supports previous studies that claim that mycoheterotrophic orchids were derived recently, except for those with stage 3 or 4 degradations [20].

## 4. Conclusions

We reported six new chloroplast genomes (plastomes) of vanilloids (two *Lecanorchis*, two *Pogonia*, and two *Vanilla* species). Among these plastomes, *Pogonia japonica* has the longest plastome, with 158,200 bp in genome size. In contrast, *Lecanorchis japonica* has the shortest plastome, with 70,498 bp in genome size. The vanilloid plastomes have a typical quadripartite structure, but the small single copy (SSC) region was drastically reduced. The two different tribes of Vanilloideae (Pogonieae and Vanilleae) showed different levels of SSC reduction. In addition, various gene losses were observed among the vanilloid plastomes. The photosynthetic vanilloids (*Pogonia* and *Vanilla*) showed signs of stage 1 degradation and had lost most of their *ndh* genes. However, the other three species (one *Cyrtosia* and two *Lecanorchis* species) had stage 3 or stage 4 degradation and had lost almost all the genes in their plastomes, except some housekeeping genes. A total of ten genome rearrangements were found among ten *Vanilloideae* plastomes when compared to the basal Apostasioideae plastomes. Both synonymous and nonsynonymous substitution rates of the IR in-cooperated SC sub-regions were decelerated, while the substitution rates of the SC in-cooperated IR sub-regions were accelerated. A total of 20 protein-coding genes

commonly remained in mycoheterotrophic vanilloids. Almost all these protein genes show accelerated base substitution rates compared to photosynthetic vanilloids. Our data clearly supports the different evolutionary modes of plastomes between photosynthetic and nonphotosynthetic lineages.

## 5. Materials and Methods

### 5.1. Plant Material and DNA Extraction

The plant leaf materials used in this study and their voucher information are given in Table 1. Five plant samples were collected in nature, and one plant sample (*Pogonia japonica*) was obtained from artificial cultivation. All voucher specimens and DNAs were deposited into the KUS herbarium and the Plant DNA Bank in Korea. Fresh leaf materials were ground into a powder with liquid nitrogen in a mortar. The ground samples were used to extract genomic DNA, using a G-spin<sup>TM</sup>II for Plant Genomic DNA extraction kit (iNtRON, Seoul, Korea). The quality of genomic DNA was checked with a UV/VIS spectrophotometer (Thermo, Wilmington, DE, USA).

### 5.2. Sequencing, Assembly, and Annotation

Two samples—*Vanilla madagascariensis* and *V. planifolia*—were sequenced by the Illumina HiSeq 2000 (Illumina, San Diego, CA, USA). The other four samples—*Lecanorchis kiusiana*, *L. japonica*, *Pogonia japonica*, and *P. minor*—were sequenced by the Illumina MiSeq (Illumina, San Diego, CA, USA). The raw reads from the Illumina HiSeq were trimmed by Geneious 6.1.8 [52] with an option of 0.05 for the error probability limitation. Trimmed reads were assembled by reference guided assembly using a Geneious assembler; *Vanilla planifolia* (NC026778) was used as a reference sequence. The raw reads from the Illumina MiSeq were trimmed by BBDuk 37.64 and imported in Geneious 11.1.5 (Kmer length of 27). The trimmed reads were normalized by BBNorm 37.64 (with a target coverage level of 30 and a minimum depth of 12). Normalized reads were used to perform the de novo assembly by the Geneious assembler. Both result contigs, which came from HiSeq and MiSeq, were used as reference sequences to filter the trimmed reads. Filtered reads were re-used to conduct the de novo assembly using a Geneious assembler to construct a complete plastome sequence. Obtained plastome sequences were compared with a previous result contig and assembled by reference guided assembly. Two *Lecanorchis* reads were reassembled to check the plastome structure using SPAdes 3.10.0 with error correction and the assembling method.

Complete plastome sequences were annotated with BLASTn, tRNAscan-SE 2.0 [53], ORF finder, and the find annotation function in Geneious 11.1.5 (NC026778—*Vanilla planifolia* was used as a reference). The alternative start codons, such as ACG and TTG, were also included in the ORF finder. In order to judge a protein-coding gene, pseudogene, and its absence, we used the criteria from a previous study [20]. Translated ORFs were extracted to perform psi-blast [54] with orchid plastome CDS. Circular plastome maps were constructed using the OGdraw web server [55].

### 5.3. Phylogenetic Analysis

In addition to our 6 vanilloid plastomes, 83 plastome sequences were downloaded from the NCBI to perform a phylogenetic analysis (Supplementary Table S3, 74 sequences for Orchidaceae, four species for Asparagales, one species for Liliales, and four species for Vanilloideae). The 79 CDS and four rRNA genes were extracted from the 89 plastomes. Extracted gene regions were aligned using MAFFT [56]. Furthermore, 83 independent gene alignments were concatenated to be a single length of 87,230 bp. The concatenated alignment was used to perform jModeltest2 [57] in the CIPRES science gateway [58] to obtain the best model. The best-fit model for concatenated data was the GTR + G + I model. A maximum likelihood (ML) tree was constructed using RaxML-HPC2 on XSEDE in the CIPRES science gateway [59] and 100 bootstrap replicates of ML optimization likelihood (−563,601.727423). The extracted 83 regions were aligned by MUSCLE [60] and used

to perform PartitionFinder2 [61] to estimate the best-fit model by region. Eighty-three alignments were concatenated by the best-fit model (4.5S rRNA, 5S rRNA, *atpH*, *atpI*, and *psbN* for the GTR all-equal model; *clpP*, *petB*, *petG*, *petN*, *psbF*, *psbJ*, *psbL*, and *psbT* for the HKY estimated model; and the other 70 regions for the GTR estimated model). Concatenated alignments (83,665 bp length of the GTR estimated model, 1410 bp of the GTR all-equal model, and 2155 bp of the HKY estimated model) were used to construct a Bayesian inference phylogenetic tree by using MRBayes\_CIPRES api [62,63] with 1,000,000 of chain length.

IRScope [64] was used to describe the SC-IR junction region among the Vanilloideae and four Apostasioideae species. All those fourteen species' general features, such as the length of LSC/IR/SSC, were displayed on a graph by the ggplot2 package in R [65]. The progressiveMauve algorithm [66] was used to check gene relocation and plastome rearrangement among 14 species of Apostasioideae and Vanilloideae. The gene contents of 42 species in Orchidaceae were visualized as a heatmap by R.

Ten species of Vanilloideae and four species of Apostasioideae were used to estimate the synonymous and nonsynonymous substitution rates. Each region was aligned by MUSCLE [60]. The alignments were used to construct the maximum likelihood tree by the RaxML tree builder plugin in Geneious 11.1.5 with the GTR + G + I model. Alignments and phylogenetic trees were used to estimate the synonymous and nonsynonymous substitution rates by pamlX [67]. In order to estimate the position effect, regions were divided into two categories based on their movement from the original position to the IR region and from the original position to the SC region (LSC or SSC). A total of eight regions were counted as rearranged. Four regions—*rpl2*, *rps7*, *rps12-3'*, and *rps19*—moved to the SC region. Meanwhile, another four regions—*ccsA*, *rpl22*, *rps15*, and *ycf1*—moved to the IR region. One ratio omega model along all species was used as a null model. This alternative model adopts a different omega value between moved taxa (to the SC or IR regions) and consistent taxa (runmode = 0, model = 1, codon frequency = 3). The likelihood ratio test was performed to validate the statistical difference between the null model and the alternative model (Supplementary Table S2).

In order to estimate the mycoheterotrophic effect, the remaining 20 regions in all 14 species were used. Each sequence was aligned by MUSCLE, and the alignments were used to construct a maximum likelihood tree using the RaxML tree builder plugin with the GTR + G + I model. Alignments and phylogenetic trees were used to estimate the synonymous and nonsynonymous substitution rates by pamlX. One ratio omega model along all species was used as a null model. The alternative model adopts different omega values between mycoheterotrophic and photosynthetic taxa (runmode = 0, model = 1, codon frequency = 3). The likelihood ratio test was performed to validate the statistical difference between the null model and the alternative model (Supplementary Table S2).

A total of 28 regions (20 regions from the mycoheterotrophic constant region and eight regions for rearrangement) were used to perform RELAX in the Datamonkey server [68–71]. The selection intensity parameter and synonymous/nonsynonymous substitution rates between the reference branch (photosynthetic plants or constant regions) and test branch (mycoheterotrophic plants or rearranged regions) were checked.

#### 5.4. Time Estimation

The alignments used in the Bayesian inference tree construction were used to estimate the divergence time estimation. Eighty-three regions were used (4.5S rRNA, 5S rRNA, *atpH*, *atpI*, and *psbN* for the GTR all-equal model; *clpP*, *petB*, *petG*, *petN*, *psbF*, *psbJ*, *psbL*, and *psbT* for the HKY estimated model; and the other 70 regions for the GTR estimated model). The XML file was prepared by BEAUti 2.5.2 with three fossil data to calibrate the nodes (*Asparagales*, normal distribution, mean 105.3, sigma 8.0; *Dendrobium*, log-normal distribution, sigma 2.0, offset 23.2; *Goodyera*, log-normal distribution, sigma 2.0, offset 15.0) [51,72,73]. A relaxed clock log-normal model [74] and a yule model were chosen to perform Markov chain Monte Carlo (MCMC) 100,000,000 times. Two independent analyses

were performed by BEAST2-XSEDE in the CIPRES science gateway [75]. Trees and logs were collected for every 5000 generations.

The total log file was tested to check the effective sample size (ESS) by Tracer v1.6 [76], and the major parameters (posterior, prior, likelihood) were over 100 of the ESS. Tree files were concatenated by LogCombiner v2.5.2 [77] with an option of a 15% burn-in. A concatenated tree file was treated by TreeAnnotator [78] in the CIPRES science gateway with an option of 0.95 posterior probability. The concatenated maximum clade credibility tree was annotated by FigTree v1.4 [79] and two R packages of ‘phytools’ and ‘ape’ [80,81].

**Supplementary Materials:** Supplementary Materials can be found at <https://www.mdpi.com/article/10.3390/ijms24043808/s1>. Table S1: The NGS results of six newly decoded Vanilloideae species, Table S2. Log-likelihood ratio (LRT) test result among eight regions under position effect and twenty remaining regions in mycoheterotrophic species. Regions with significant *p*-values (*p*-value < 0.05) are highlighted in red color, Table S3. Used sequence information in this study. The red scientific name indicates that the species’ NGS data was generated in this study, Figure S1. Gene contents heatmap of Orchidaceae, including vanilloid species, Figure S2. IR junctions of ten Vanilloideae and four Apostasioideae species, Figure S3. A maximum likelihood (ML) tree inferred from 89 species of Orchidaceae. The nucleotide sequences of 79 protein-coding genes and 4 rRNA genes were aligned individually and concatenated to be a single matrix, Figure S4. Divergence time estimation of Orchidaceae.

**Author Contributions:** Conceptualization, K.-J.K.; Methodology, Y.-K.K.; Software, Y.-K.K.; Validation, Y.-K.K. and K.-J.K.; Formal analysis, Y.-K.K., S.-H.C. and J.-R.H.; Investigation, Y.-K.K.; Resources, K.-J.K.; Data curation, Y.-K.K., S.-H.C. and J.-R.H.; Writing—original draft, Y.-K.K.; Writing—review & editing, K.-J.K.; Supervision, K.-J.K.; Project administration, K.-J.K.; Funding acquisition, K.-J.K.. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Institute of Biological Resources (NIBR) under the genetic evaluation of vascular plants IV-1 (2018, grand no. NIBR201803102) to K.-J.K. This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2020R1A6A3A01100512) to Y.-K.K.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The six newly sequenced Vanilloideae species (*Lecanorchis kiusiana*, *L. japonica*, *Pogonia japonica*, *P. minor*, *Vanilla madagascariensis*, and *V. planifolia*) were submitted to GenBank with accession numbers of MN200375, MN200374, MN200371, MN200372, MN200374, and MN200375, respectively.

**Acknowledgments:** The authors would like to thank to three anonymous reviewers for their helpful comments to improving the manuscript. We also would like to thank to Kyeongwon Kang of Babo orchid Farm for the materials of *Pogonia japonica*. We thank the curator of the Korea University Herbarium (KUS) for voucher specimen preparation. The genomic DNAs are deposited in the Plant DNA Bank in Korea (PDBK).

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

## Abbreviations

LSC	Large Single Copy
SSC	Small Single Copy
IR	Inverted Repeat
ML	Maximum Likelihood
BI	Bayesian Inference



## References

1. Christenhusz, M.J.M.; Byng, J.W. The number of known plants species in the world and its annual increase. *Phytotaxa* **2016**, *261*, 201–217. [[CrossRef](#)]
2. Merckx, V.S.F.T.; Freudenstein, J.V.; Kissling, J.; Christenhusz, M.J.M.; Stotler, R.E.; Crandall-Stotler, B.; Wickett, N.; Rudall, P.J.; Maas-van de Kamer, H.; Maas, P.J.M. *Taxonomy and Classification BT-Mycoheterotrophy: The Biology of Plants Living on Fungi*; Merckx, V., Ed.; Springer: New York, NY, USA, 2013; pp. 19–101. ISBN 978-1-4614-5209-6.
3. Chase, M.W.; Cameron, K.M.; Freudenstein, J.V.; Pridgeon, A.M.; Salazar, G.; van den Berg, C.; Schuiteman, A. An updated classification of Orchidaceae. *Bot. J. Linn. Soc.* **2015**, *177*, 151–174. [[CrossRef](#)]
4. Flora of Korea Editorial Committee; Park, C.-W. *The Genera of Vascular Plants of Korea*; Academy Publishing: Seoul, Republic of Korea, 2007; ISBN 897616380X.
5. Chen, X.Q.; Liu, Z.J.; Zhu, G.H.; Lang, K.Y.; Ji, Z.H.; Luo, Y.B.; Jin, X.H.; Cribb, P.J.; Wood, J.J.; Gale, S.W. Flora of China: Orchidaceae. *Chen XQ, Wood JJ* **2009**, *175*, 477.
6. Lee, N.S. *Illustrated Flora of Korean Orchids*; Ewha Womans University Press: Seoul, Republic of Korea, 2011.
7. Sheehan, T.; Sheehan, M. *An Illustrated Survey of Orchid Genera*; Cambridge University Press: Cambridge, UK, 1994; ISBN 0521480280.
8. Byng, J.W. *The Flowering Plants Handbook: A Practical Guide to Families and Genera of the World*; Plant Gateway Ltd.: Den Haag, The Netherlands, 2014; ISBN 0992999316.
9. Chang, C.C.; Lin, H.C.; Lin, I.P.; Chow, T.Y.; Chen, H.H.; Chen, W.H.; Cheng, C.H.; Lin, C.Y.; Liu, S.M.; Chang, C.C.; et al. The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): Comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Mol. Biol. Evol.* **2006**, *23*, 279–291. [[CrossRef](#)]
10. Barrett, C.F.; Wicke, S.; Sass, C. Dense intraspecific sampling reveals rapid and independent trajectories of plastome degradation in a heterotrophic orchid complex. *New Phytol.* **2018**, *218*, 1192–1204. [[CrossRef](#)]
11. Feng, Y.L.; Wicke, S.; Li, J.W.; Han, Y.; Lin, C.S.; Li, D.Z.; Zhou, T.T.; Huang, W.C.; Huang, L.Q.; Jin, X.H. Lineage-specific reductions of plastid genomes in an orchid tribe with partially and fully mycoheterotrophic species. *Genome Biol. Evol.* **2016**, *8*, 2164–2175. [[CrossRef](#)]
12. Kim, Y.K.; Kwak, M.H.; Chung, M.G.; Kim, H.W.; Jo, S.; Sohn, J.Y.; Cheon, S.H.; Kim, K.J. The complete plastome sequence of the endangered orchid *Cymbidium macrorhizon* (Orchidaceae). *Mitochondrial DNA Part B Resour.* **2017**, *2*, 725–727. [[CrossRef](#)]
13. Kim, H.T.; Shin, C.H.; Sun, H.; Kim, J.H. Sequencing of the plastome in the leafless green mycoheterotroph *Cymbidium macrorhizon* helps us to understand an early stage of fully mycoheterotrophic plastome structure. *Plant Syst. Evol.* **2018**, *304*, 245–258. [[CrossRef](#)]
14. Liu, Z.J. The genome sequence of the orchid *Phalaenopsis equestris*. *Nat. Genet.* **2015**, *47*, 65–72. [[CrossRef](#)]
15. 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](#)]
16. Niu, Z.; Xue, Q.; Zhu, S.; Sun, J.; Liu, W.; Ding, X. The complete plastome sequences of four orchid species: Insights into the evolution of the Orchidaceae and the utility of plastomic mutational hotspots. *Front. Plant Sci.* **2017**, *8*, 715. [[CrossRef](#)] [[PubMed](#)]
17. Yang, J.-B.; Tang, M.; Li, H.-T.; Zhang, Z.-R.; Li, D.-Z. Complete chloroplast genome of the genus *Cymbidium*: Lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evol. Biol.* **2013**, *13*, 84. [[CrossRef](#)] [[PubMed](#)]
18. da Rocha Perini, V.; Leles, B.; Furtado, C.; Prosdocimi, F. Complete chloroplast genome of the orchid *Cattleya crispata* (Orchidaceae:Laeliinae), a neotropical rupicolous species. *Mitochondrial DNA Part A DNA Mapp. Seq. Anal.* **2016**, *27*, 4075–4077. [[CrossRef](#)]
19. Kim, Y.K.; Kwak, M.H.; Hong, J.R.; Kim, H.W.; Jo, S.; Sohn, J.Y.; Cheon, S.H.; Kim, K.J. The complete plastome sequence of the endangered orchid *Kuhlhasseltia nakaiana* (Orchidaceae). *Mitochondrial DNA Part B Resour.* **2017**, *2*, 701–703. [[CrossRef](#)] [[PubMed](#)]
20. Kim, Y.K.; Jo, S.; Cheon, S.H.; Joo, M.J.; Hong, J.R.; Kwak, M.; Kim, K.J. Plastome evolution and phylogeny of Orchidaceae, with 24 new sequences. *Front. Plant Sci.* **2020**, *11*, 22. [[CrossRef](#)]
21. Kim, Y.K.; Jo, S.; Cheon, S.H.; Kwak, M.; Kim, Y.D.; Kim, K.J. Plastome evolution and phylogeny of subtribe Aeridinae (Vandaeae, Orchidaceae). *Mol. Phylogenet. Evol.* **2020**, *144*, 106721. [[CrossRef](#)]
22. Lin, C.S.; Chen, J.J.W.; Huang, Y.T.; Chan, M.T.; Daniell, H.; Chang, W.J.; Hsu, C.T.; Liao, D.C.; Wu, F.H.; Lin, S.Y.; et al. The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. *Sci. Rep.* **2015**, *5*, 9040. [[CrossRef](#)] [[PubMed](#)]
23. Barrett, C.F.; Davis, J.I. The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. *Am. J. Bot.* **2012**, *99*, 1513–1523. [[CrossRef](#)]
24. Delannoy, E.; Fujii, S.; Colas Des Francs-Small, C.; Brundrett, M.; Small, I. Rampant gene loss in the underground orchid *Rhizanthella Gardneri* highlights evolutionary constraints on plastid genomes. *Mol. Biol. Evol.* **2011**, *28*, 2077–2086. [[CrossRef](#)]
25. Kim, Y.K.; Jo, S.; Cheon, S.H.; Joo, M.J.; Hong, J.R.; Kwak, M.H.; Kim, K.J. Extensive losses of photosynthesis genes in the plastome of a mycoheterotrophic orchid, *Cyrtosia septentrionalis* (Vanilloideae: Orchidaceae). *Genome Biol. Evol.* **2019**, *11*, 565–571. [[CrossRef](#)] [[PubMed](#)]
26. Schelkunov, M.I.; Shtratnikova, V.Y.; Nuraliev, M.S.; Selosse, M.A.; Penin, A.A.; Logacheva, M.D. Exploring the limits for reduction of plastid genomes: A case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. *Genome Biol. Evol.* **2015**, *7*, 1179–1191. [[CrossRef](#)] [[PubMed](#)]

27. Yuan, Y.; Jin, X.; Liu, J.; Zhao, X.; Zhou, J.; Wang, X.; Wang, D.; Lai, C.; Xu, W.; Huang, J.; et al. The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nat. Commun.* **2018**, *9*, 1615. [[CrossRef](#)] [[PubMed](#)]
28. Amiryousefi, A.; Hyvönen, J.; Poczai, P. The plastid genome of Vanillon (*Vanilla pompona*, Orchidaceae). *Mitochondrial DNA Part B* **2017**, *2*, 689–691. [[CrossRef](#)] [[PubMed](#)]
29. Kim, H.T.; Kim, J.S.; Moore, M.J.; Neubig, K.M.; Williams, N.H.; Whitten, W.M.; Kim, J.H. Seven new complete plastome sequences reveal rampant independent loss of the *ndh* gene family across orchids and associated instability of the inverted repeat/small single-copy region boundaries. *PLoS ONE* **2015**, *10*, e0142215. [[CrossRef](#)]
30. Givnish, T.J.; Spalink, D.; Ames, M.; Lyon, S.P.; Hunter, S.J.; Zuluaga, A.; Iles, W.J.D.; Clements, M.A.; Arroyo, M.T.K.; Leebens-Mack, J.; et al. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. R. Soc. B Biol. Sci.* **2015**, *282*, 20151553. [[CrossRef](#)]
31. Li, Y.X.; Li, Z.H.; Schuitman, A.; Chase, M.W.; Li, J.W.; Huang, W.C.; Hidayat, A.; Wu, S.S.; Jin, X.H. Phylogenomics of Orchidaceae based on plastid and mitochondrial genomes. *Mol. Phylogenet. Evol.* **2019**, *139*, 106540. [[CrossRef](#)]
32. Cameron, K.M. Utility of plastid *psaB* gene sequences for investigating intrafamilial relationships within Orchidaceae. *Mol. Phylogenet. Evol.* **2004**, *31*, 1157–1180. [[CrossRef](#)]
33. Cameron, K.M.; Chase, M.W.; Whitten, W.M.; Kores, P.J.; Jarrell, D.C.; Albert, V.A.; Yukawa, T.; Hills, H.G.; Goldman, D.H. A phylogenetic analysis of the Orchidaceae: Evidence from *rbcL* nucleotide sequences. *Am. J. Bot.* **1999**, *86*, 208–224. [[CrossRef](#)]
34. Cameron, K.M.; Carmen Molina, M. Photosystem II gene sequences of *psbB* and *psbC* clarify the phylogenetic position of *Vanilla* (Vanilloideae, Orchidaceae). *Cladistics* **2006**, *22*, 239–248. [[CrossRef](#)]
35. Cameron, K.M.; Chase, M.W. Phylogenetic relationships of Pogoniinae (Vanilloideae, Orchidaceae): An herbaceous example of the eastern north America-eastern Asia phytogeographic disjunction. *J. Plant Res.* **2006**, *112*, 317–329. [[CrossRef](#)]
36. Cameron, K.M. On the value of nuclear and mitochondrial gene sequences for reconstructing the phylogeny of vanilloid orchids (Vanilloideae, Orchidaceae). *Ann. Bot.* **2009**, *104*, 377–385. [[CrossRef](#)] [[PubMed](#)]
37. Górnjak, M.; Paun, O.; Chase, M.W. Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, *Xdh*: Congruence with organellar and nuclear ribosomal DNA results. *Mol. Phylogenet. Evol.* **2010**, *56*, 784–795. [[CrossRef](#)] [[PubMed](#)]
38. Pansarin, E.R.; Salatino, A.; Salatino, M.L.F. Phylogeny of south American Pogonieae (Orchidaceae, Vanilloideae) based on sequences of nuclear ribosomal (ITS) and chloroplast (*psaB*, *rbcL*, *rps16*, and *trnL-F*) DNA, with emphasis on *Cleistes* and discussion of biogeographic implications. *Org. Divers. Evol.* **2008**, *8*, 171–181. [[CrossRef](#)]
39. Kim, Y.K.; Kwak, M.H.; Hong, J.R.; Kim, H.W.; Jo, S.; Sohn, J.Y.; Cheon, S.H.; Kim, K.J. The complete plastome sequence of the endangered orchid *Habenaria radiata* (Orchidaceae). *Mitochondrial DNA Part B Resour.* **2017**, *2*, 704–706. [[CrossRef](#)] [[PubMed](#)]
40. Lam, V.K.Y.; Darby, H.; Merckx, V.S.F.T.; Lim, G.; Yukawa, T.; Neubig, K.M.; Abbott, J.R.; Beatty, G.E.; Provan, J.; Soto Gomez, M.; et al. Phylogenomic inference in extremis: A case study with mycoheterotroph plastomes. *Am. J. Bot.* **2018**, *105*, 480–494. [[CrossRef](#)]
41. Wicke, S.; Müller, K.F.; de Pamphilis, C.W.; Quandt, D.; Wickett, N.J.; Zhang, Y.; Renner, S.S.; Schneeweiss, G.M. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell* **2013**, *25*, 3711–3725. [[CrossRef](#)]
42. Wicke, S.; Müller, K.F.; dePamphilis, C.W.; Quandt, D.; Bellot, S.; Schneeweiss, G.M. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9045–9050. [[CrossRef](#)] [[PubMed](#)]
43. Wicke, S.; Naumann, J. *Molecular Evolution of Plastid Genomes in Parasitic Flowering Plants*, 1st ed.; Elsevier Ltd.: Amsterdam, The Netherlands, 2017; Volume 85, ISBN 9780128134573.
44. Kim, K.-J.; Lee, H. Complete chloroplast genome sequences from Korean ginseng (*Panax ginseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Res.* **2004**, *11*, 247–261. [[CrossRef](#)] [[PubMed](#)]
45. Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.; Matsubayashi, T.; Zaita, N.; Chunwongse, J.; Obokata, J.; Yamaguchi-Shinozaki, K.; et al. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **1986**, *5*, 2043–2049. [[CrossRef](#)] [[PubMed](#)]
46. Kolodner, R.; Tewari, K.K. Inverted repeats in chloroplast DNA from higher plants. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 41–45. [[CrossRef](#)]
47. Palmer, J.D.; Thompson, W.F. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **1982**, *29*, 537–550. [[CrossRef](#)] [[PubMed](#)]
48. Yi, D.K.; Kim, K.J. Complete chloroplast genome sequences of important oilseed crop *Sesamum indicum* L. *PLoS ONE* **2012**, *7*, e35872. [[CrossRef](#)] [[PubMed](#)]
49. Zhu, A.; Guo, W.; Gupta, S.; Fan, W.; Mower, J.P. Evolutionary dynamics of the plastid inverted repeat: The effects of expansion, contraction, and loss on substitution rates. *New Phytol.* **2016**, *209*, 1747–1756. [[CrossRef](#)] [[PubMed](#)]
50. Barrett, C.F.; Freudenstein, J.V.; Li, J.; Mayfield-Jones, D.R.; Perez, L.; Pires, J.C.; Santos, C. Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. *Mol. Biol. Evol.* **2014**, *31*, 3095–3112. [[CrossRef](#)]
51. Gustafsson, A.L.S.; Verola, C.F.; Antonelli, A. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evol. Biol.* **2010**, *10*, 177. [[CrossRef](#)]

52. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]
53. Lowe, T.M.; Chan, P.P. tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* **2016**, *44*, W54–W57. [CrossRef]
54. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef]
55. Lohse, M.; Drechsel, O.; Bock, R. OrganellarGenomeDRAW (OGDRAW): A tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* **2007**, *52*, 267–274. [CrossRef]
56. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [CrossRef]
57. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef] [PubMed]
58. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 14 November 2010; pp. 1–8.
59. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef] [PubMed]
60. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef]
61. Lanfear, R.; Calcott, B.; Ho, S.Y.W.; Guindon, S. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* **2012**, *29*, 1695–1701. [CrossRef] [PubMed]
62. Ronquist, F.; Teslenko, M.; Van Der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, 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*, 539–542. [CrossRef] [PubMed]
63. Miller, M.A.; Schwartz, T.; Pickett, B.E.; He, S.; Klem, E.B.; Scheuermann, R.H.; Passarotti, M.; Kaufman, S.; O’Leary, M.A. A RESTful API for access to phylogenetic tools via the CIPRES science gateway. *Evol. Bioinforma.* **2015**, *11*, EBO-S21501. [CrossRef]
64. Amiryousefi, A.; Hyvo, J.; Poczai, P. Genome analysis IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [CrossRef]
65. Wickham, H.; Chang, W.; Wickham, M.H. Package ‘ggplot2’. *Creat. Elegant Data Vis. Using Gramm. Graph. Version* **2016**, *2*, 1–189.
66. Darling, A.E.; Mau, B.; Perna, N.T. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **2010**, *5*, e11147. [CrossRef]
67. Xu, B.; Yang, Z. PAMLX: A graphical user interface for PAML. *Mol. Biol. Evol.* **2013**, *30*, 2723–2724. [CrossRef]
68. Delpont, W.; Poon, A.F.Y.; Frost, S.D.W.; Kosakovsky Pond, S.L. Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **2010**, *26*, 2455–2457. [CrossRef] [PubMed]
69. Pond, S.L.K.; Frost, S.D.W. Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **2005**, *21*, 2531–2533. [CrossRef] [PubMed]
70. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Kosakovsky Pond, S.L. Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Mol. Biol. Evol.* **2018**, *35*, 773–777. [CrossRef]
71. Wertheim, J.O.; Murrell, B.; Smith, M.D.; Kosakovsky Pond, S.L.; Scheffler, K. RELAX: Detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* **2014**, *32*, 820–832. [CrossRef]
72. Conran, J.G.; Bannister, J.M.; Lee, D.E. Earliest orchid macrofossils: Early Miocene *Dendrobium* and *Earina* (Orchidaceae: Epidendroideae) from New Zealand. *Am. J. Bot.* **2009**, *96*, 466–474. [CrossRef] [PubMed]
73. Iles, W.J.D.; Smith, S.Y.; Gandolfo, M.A.; Graham, S.W. Monocot fossils suitable for molecular dating analyses. *Bot. J. Linn. Soc.* **2015**, *178*, 346–374. [CrossRef]
74. Drummond, A.J.; Ho, S.Y.W.; Phillips, M.J.; Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **2006**, *4*, e88. [CrossRef]
75. Bouckaert, R.; Heled, J.; Kühnert, D.; Vaughan, T.; Wu, C.-H.; Xie, D.; Suchard, M.A.; Rambaut, A.; Drummond, A.J. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **2014**, *10*, e1003537. [CrossRef]
76. Rambaut, A.; Suchard, M.A.; Xie, D.; Drummond, A.J. Tracer 1.6. 2014. Available online: <https://github.com/beast-dev/tracer> (accessed on 12 February 2023).
77. Rambaut, A.; Drummond, A.J. LogCombiner v2.1.3. 2014. Available online: <http://www.beast2.org/> (accessed on 12 February 2023).
78. Rambaut, A.; Drummond, A.J. TreeAnnotator. Program and Documentation Distributed by the Author. Available online: <http://www.beast2.org/> (accessed on 12 February 2023).
79. Rambaut, A. FigTree v1.4. 2012. Available online: <http://tree.bio.ed.ac.uk/software/figtree/> (accessed on 12 February 2023).

80. Revell, L.J. Phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **2012**, *3*, 217–223. [[CrossRef](#)]
81. Paradis, E.; Claude, J.; Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **2004**, *20*, 289–290. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.