

## Article

# Comparative Chloroplast Genomics and Phylogenetic Analysis of *Persicaria amphibia* (Polygonaceae)

KyoungSu Choi \*, Yong Hwang and Jeong-Ki Hong

Plant Research Team, Animal and Plant Research Department, Nakdonggang National Institute of Biological Resources, Sangju 37242, Korea

\* Correspondence: choiks010@nnibr.re.kr

**Abstract:** *Persicaria amphibia* (L.) Delarbre, also known as water knotweed, belongs to the Polygonaceae family and has two types: terrestrial and aquatic. We report the chloroplast genome of *P. amphibia* obtained through de novo assembly of Illumina paired-end reads produced by total DNA sequencing. We analyzed the complete chloroplast (cp) genome of *P. amphibia* and found it to be 159,455 bp in length, with a large single-copy region (LSC, 84,281 bp), a small single-copy region (SSC, 13,258 bp), and a pair of inverted repeats (IR, 30,956 bp). It contains 79 protein-coding, 29 tRNA and 4 rRNA genes. Comparative analysis of nine *Persicaria* cp genomes showed a similar genome structure and gene content. However, *ycf3* intron II was lost in three *Persicaria* species (*P. hydropper*, *P. japonica*, and *P. pubescens*) and the SC/IR regions of four species (*P. amphibia*, *P. hydropper*, *P. japonica*, and *P. pubescens*) included the *rps19* gene. Phylogenetic analysis of the nine *Persicaria* species revealed that *P. amphibia* is sister to *P. hydropper*, *P. japonica*, and *P. pubescens*. Moreover, we found sequence divergence regions; the largest were *rps16-trnQ*, *trnQ-psbK*, *trnW-trnP*, *ndhF-rpl32*, and *rpl32-trnL* regions. This study could be useful for phylogenetic tree analysis of *Persicaria* and for the identification of *Persicaria* species.

**Keywords:** *Persicaria amphibia*; chloroplast genome; comparative analysis; phylogenetic analysis; nucleotide diversity



**Citation:** Choi, K.; Hwang, Y.; Hong, J.-K. Comparative Chloroplast Genomics and Phylogenetic Analysis of *Persicaria amphibia* (Polygonaceae). *Diversity* **2022**, *14*, 641. <https://doi.org/10.3390/d14080641>

Academic Editor: Mario A. Pagnotta

Received: 15 July 2022

Accepted: 8 August 2022

Published: 11 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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

*Persicaria* Mill. is a genus in the tribe Persicarieae (family Polygonaceae) that contains approximately 130 species distributed worldwide [1,2]. Based on the macromorphological classification, *Persicaria* is included in the genus *Polygonum* [3]. However, ribulose-bisphosphate carboxylase large subunit (*rbcl*) analysis of Polygonaceae showed that the Polygonoideae subfamily is paraphyletic, and is subdivided into two tribes: Polygoneae and Persicarieae [4].

*Persicaria amphibia*, a species belonging to the genus *Persicaria*, is native to the Northern Hemisphere. It has been described in both aquatic and terrestrial forms, and exhibits complex patterns in leaf morphological variation [5]. *P. amphibia* can grow in aquatic environments as well as terrestrial habitats and has different leaf types. The terrestrial form is erect and has hairy leaves with wavy edges. The aquatic form is hairless with oblong flat leaves. Previous phylogenetic analyses of *Persicaria* used markers of the internal transcribed spacer (ITS) region, and chloroplast DNA (cpDNA) *rbcl* and *trnL-F* [6], which showed that *P. amphibia* formed a clade and was sister to sect. *Echinocaulon* in two cp markers, whereas it was sister to sect. *Tovara* in the ITS marker.

Chloroplasts provide energy to almost all green plants. The chloroplast genome is typically quadripartite in structure, containing a large single copy region (LSC) and a small single copy region (SSC), separated by a pair of inverted repeats (IR) [7] in which organization, gene contents and structures are highly conserved [8]. However, some plants have unique structures owing to gene loss [9–11], rearrangement [12–14], and inversions [15,16].

Recently, cp genome sequences have been used for evolutionary [17,18], taxonomic [19], DNA marker [20,21], and phylogenetic analyses [22–24]. NCBI has uploaded and studied the chloroplast genomes of *Persicaria* species such as *P. chinensis* (NC\_050358) [25], *P. pubescens* (MK234901), *P. hydropiper* (MK234902), *P. japonica* (NC\_056952), *P. filiformis* (NC\_058319), *P. aviculare* (NC\_058892), *P. perfoliata* (NC\_060649), *P. runcinata* (NC\_061176), and *P. maackiana* (NC\_061657).

In this study, we sequenced the complete cp genome of *P. amphibia* and compared it with previously published cp genomes of *Persicaria* species. The aims of this study were to: (1) evaluate the phylogenetic position of *P. amphibia*, (2) evaluate the sequence divergence, and (3) suggest useful markers for future phylogenetic studies in *Persicaria* species.

## 2. Materials and Methods

### 2.1. Sampling, DNA Extraction and Sequencing

Fresh leaves of *P. amphibia* were collected from Gyodongdo, Ganghwa, South Korea, and specimens were deposited in the Herbarium of the Nakdonggang National Institute of Biological Resources (NNIBR). The total DNA content was extracted using the DNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA, USA). Genomic DNA was sequenced using the Illumina Truseq Nano DNA kit (Illumina, San Diego, CA, USA), as per the manufacturer's protocol. Approximately 4.0 Gb of raw data were generated. A total of 31,825,998 reads of the 150 bp paired-end sequence were generated.

### 2.2. Assembly and Annotation

The chloroplast genome was assembled using GetOrganelle [26]. The GeSeq [27] was used to annotate the *P. amphibia* genome. tRNA gene sequences were obtained using tRNAscan-SE [28]. Finally, all published plastid genomes of *Persicaria* were compared using Geneious Prime [29]. OrganellarGenomeDRAW (OGDRAW) was used to draw a circular map of the chloroplast genome of *P. amphibia* [30]. The chloroplast genome of *P. amphibia* was deposited in GenBank (Table 1).

**Table 1.** Comparison of features of nine *Persicaria* cp genomes.

	<i>Persicaria amphibia</i>	<i>Persicaria pubescens</i>	<i>Persicaria hydropiper</i>	<i>Persicaria chinensis</i>	<i>Persicaria japonica</i>	<i>Persicaria filiformis</i>	<i>Persicaria perfoliata</i>	<i>Persicaria runcinata</i>	<i>Persicaria maackiana</i>
<b>Length</b>									
Total	159,455 bp	159,502 bp	159,054 bp	159,981 bp	159,747 bp	159,741 bp	160,585 bp	159,220 bp	160,595 bp
LSC	84,281 bp	84,555 bp	83,835 bp	84,347 bp	85,013 bp	84,432 bp	85,439 bp	84,461 bp	85,376 bp
SSC	13,258 bp	13,385 bp	13,357 bp	12,890 bp	13,178 bp	13,073 bp	12,879 bp	12,807 bp	13,055 bp
IR	30,956 bp	30,781 bp	30,931 bp	30,872 bp	30,778 bp	31,118 bp	31,135 bp	30,884 bp	31,082 bp
<b>Genes</b>									
Total	113	113	113	113	113	113	113	113	113
Protein-coding genes	79	79	79	79	79	79	79	79	79
tRNA	29	29	29	29	29	29	29	29	29
rRNA	4	4	4	4	4	4	4	4	4
<b>GC contents</b>	38.2%	38.2%	38.2%	38.0%	38.1%	37.8%	37.5%	37.9%	37.9%
<b>Accession number</b>	This study (ON938209)	MK234901	MK234902	NC_050358	NC_056952	NC_058319	NC_060649	NC_061176	NC_061657

### 2.3. Repeat and Divergence Hotspot Analysis

REPuter was used to visualize forward, palindrome, reverse, and complement sequences with a minimum repeat size of 30 base pairs (bp) and a sequence identity greater than 90% [31]. Simple sequence repeats (SSRs) were detected using MISA [32]. SSRs with minimum numbers of repetitions of 10, 5, 4, 3, 3, 3 for mono-, di-, tri-, tetra-, penta- and hexa-nucleotides were detected. The alignment of nine *Persicaria* complete chloroplast genome sequences was visually compared using the Shuffle-LAGAN model of mVISATA [33]. The nucleotide diversity (Pi) was determined using DnaSP [34]. The step size was set to 200 bp and the window length was set to 600 bp.

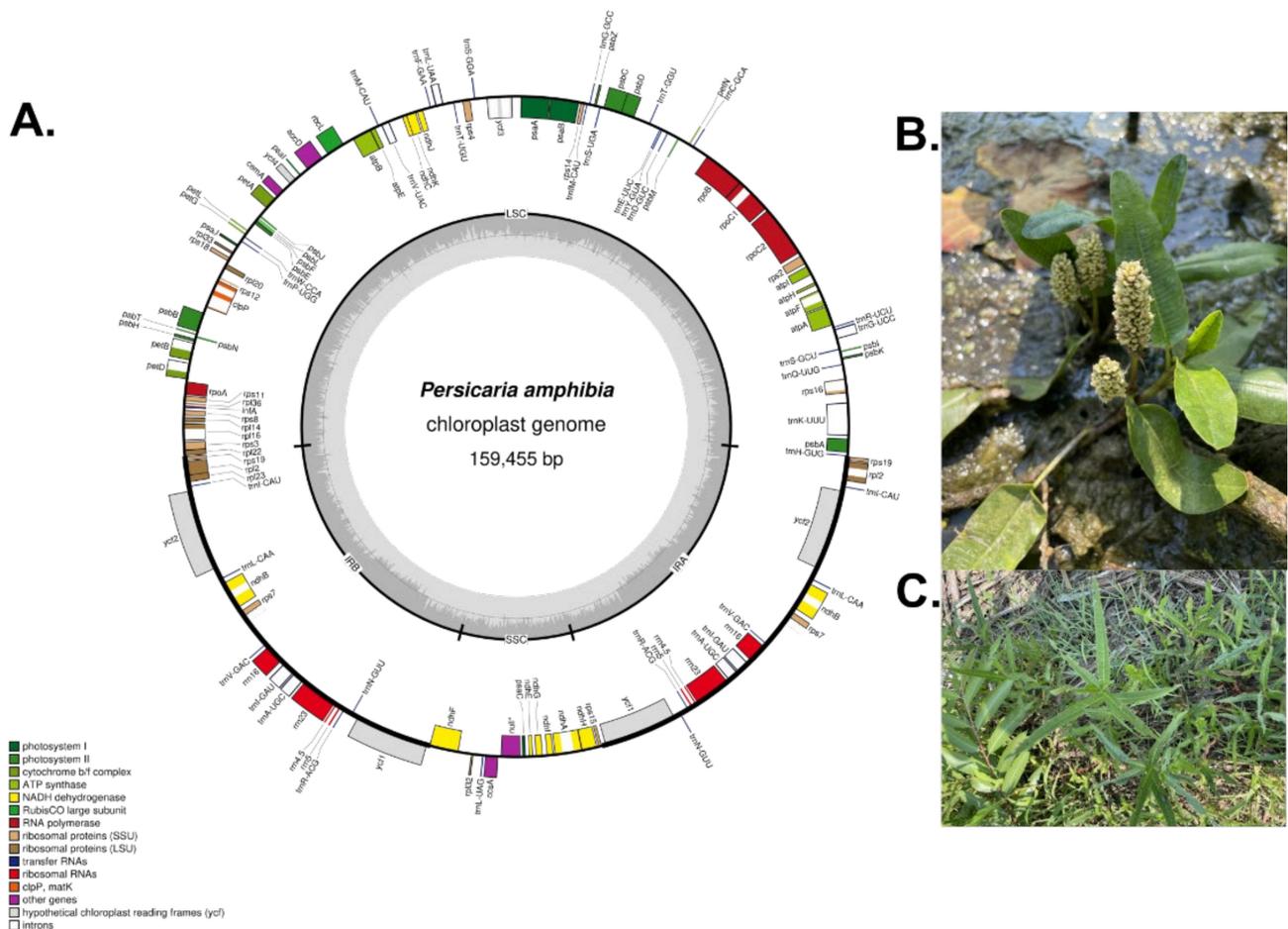
### 2.4. Phylogenetic Analysis

Seventy-seven protein-coding genes of nine *Persicaria* species (Table 1) and one out-group (NC\_58892, *Polygonum aviculare*) were compiled into a single file of 66,731 bp and aligned using MAFFT [35]. Maximum likelihood (ML) analyses were conducted using RAxML v.8 [36] with the GTR+GAMMA I model with 1000 bootstrap replications.

## 3. Results

### 3.1. Characteristic of the *P. amphibia* cp Genome

The complete cp genome of *P. amphibia* (NCBI accession number: ON938209) comprises 159,455 bp with a quadripartite structure and two IRs (30,956 bp) separated by the LSC (84,281 bp) and SSC (13,258 bp) regions (Figure 1, Table 1). The average GC content was 38.2%. The cp genome of *P. amphibia* encoded a total of 113 genes, including 79 protein-coding genes, 29 tRNA genes, and 4 rRNA genes (Table 1). The IR region contained 18 duplicated genes, including seven protein-coding genes (*ycf1*, *rps7*, *ndhB*, *ycf2*, *rpl23*, *rpl2*, and *rps19*), seven tRNA genes (*trnN*-GUU, *trnR*-ACG, *trnA*-UGC, *trnI*-GAU, *trnV*-GAC, *trnL*-CAA, and *trnI*-CAU), and four rRNA genes (*rrn5*, *rrn4.5*, *rrn23*, and *rrn16*).



**Figure 1.** The complete chloroplast genome of *P. amphibia* (A). Genes drawn inside the circle are transcribed clockwise, and those outside are counterclockwise. The darker gray in the inner circle corresponds to GC contents. Aquatic form of *P. amphibia* (B) and terrestrial form of *P. amphibia* (C).

Fifteen genes (*rps16*, *atpF*, *rpoC1*, *petB*, *petD*, *rpl16*, *rpl2*, *ndhB*, *ndhA*, *trnA*-UGC, *trnI*-GAU, *trnV*-UAC, *trnL*-UAA, *trnG*-UCC, and *trnK*-UUU) had one intron, whereas three genes (*ycf3*, *clpP*, and *rps12*) contained two introns.

### 3.2. Comparison of Other *Persicaria* Species

Nine *Persicaria* cp genomes showed a typical quadripartite structure, consisting of a pair of IRs (30,778–31,135 bp), separated by the LSC (84,281–85,439 bp) and SSC (12,879–13,385 bp) regions (Table 1). The cp genome of *P. aviculare* exhibits the longest genome, and *P. hydropiper* is smaller than other *Persicaria* species (Table 1). Nine *Persicaria* species had the same number of genes (79 protein-coding genes, 30 tRNA, and 4 rRNA genes). However, the second intron of *ycf3* was lost in three species (*P. pubescens*, *P. japonica* and *P. hydropiper* Figure 2).



**Figure 2.** The sequences of *ycf3* gene among nine *Persicaria* species. The yellow boxes indicate the exon and red boxes indicate the loss of *ycf3* intron II.

The IR/LSC and IR/SSC junctions of nine *Persicaria* cp genomes were compared. We found two different junction types: IR/LSC and IR/SSC. Type 1 in four species (*P. amphibia*, *P. filiformis*, *P. perfoliate*, and *P. maackiana*) of IR included the intact *rps19* gene. Type 2 in the other *Persicaria* species (except for *P. amphibia*, *P. filiformis*, *P. perfoliate*, and *P. maackiana*) of IR included the partial *rps19* gene. The intact *ycf1* gene was duplicated in IR regions (IRa and IRb). The junctions of IRb and SSC were located in the *ndhF* gene, and the junction of SSC/IRA was located between the *rps15* and *ycf1* genes (Figure 3). However, the total length of IR regions in all of the species was similar (Table 1).

### 3.3. Repeat Sequence Analysis

Simple sequence repeats (SSRs) were detected using MISA [32]. *P. amphibia* had 38 repeats. *P. chinensis* had the highest number of SSRs (59), whereas *P. amphibia* and *P. pubescens* (38) had a smaller number of SSRs than other *Persicaria* species. Most SSRs were A/T mononucleotide repeats (Figure 4A).

Repeats (forward, palindrome, reverse, and complement) were identified using RE-Puter [31]. A total of 25 repeats were present in *P. amphibia* (13 palindrome and 12 forward repeats). The SSRs of other *Persicaria* species were in the range of 19–28. Most of the repeats ranged in size from 30 to 48 bp. *P. japonica* had two long repeats, 154 and 156 (Figure 4B).

### 3.4. Sequences Divergence Hotspots in *Persicaria*

The cp genomes of nine *Persicaria* species were compared using mVISTA [33]. The results show that protein-coding genes were more conserved than the non-coding regions, and SC regions had lower variation than IR regions (Figure 5).

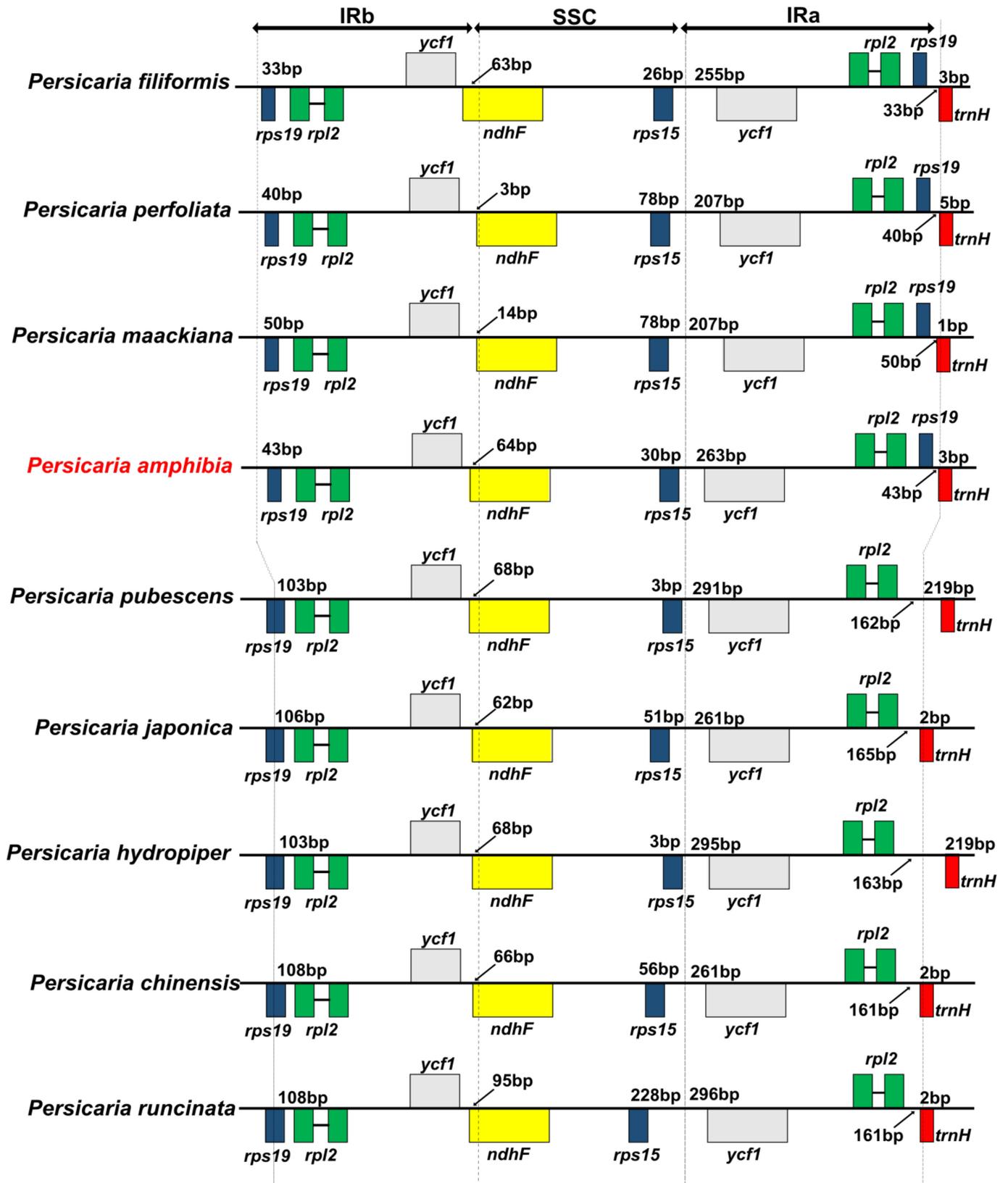


Figure 3. Comparison of the LSC/IR and IR/SSC junction among nine *Persicaria* cp genomes.



**Figure 4.** Analyses of repeat sequences in the *Persicaria* chloroplast genomes. **(A)** Frequency of SSRs determined by the MISA program. **(B)** Frequency of repeat sequences determined by REPuter.

The number of polymorphic sites and nucleotide diversity were determined using DnaSP [34]. In the complete cp genome, there were 8470 polymorphic sites, and nucleotide diversity ( $P_i$ ) was found to be 0.02746. The  $P_i$  value was calculated with 200 bp steps. These values ranged from 0 to 0.08875 (Figure 6). The high divergence regions were *rps16-trnQ*, *rpl32-trnL*, *trnW-trnP*, *trnQ-psbK*, and *ndhF-rpl32*, three of which (*rps16-trnQ*, *trnW-trnP*, and *trnQ-psbK*) were located in the LSC region, and two (*rpl32-trnL* and *ndhF-rpl32*) were located in the SSC region. The *rps16-trnQ* region exhibited the highest nucleotide diversity (0.0875, Figure 6).

### 3.5. Phylogenetic Analysis

The ML phylogenetic tree of the nine *Persicaria* species was constructed based on 77 protein-coding genes (66,732 bp, Figure 7). *Persicaria* is monophyletic (BS = 100), and is divided into two clades: (1) *P. runcinata* and *P. chinensis*; and (2) *P. filiformis*, *P. amphibia*, *P. hydropiper*, *P. japonica*, and *P. pubescens*. *P. amphibia* was sister to *P. hydropiper*, *P. japonica*, and *P. pubescens*, with a high bootstrap value (BS = 100).

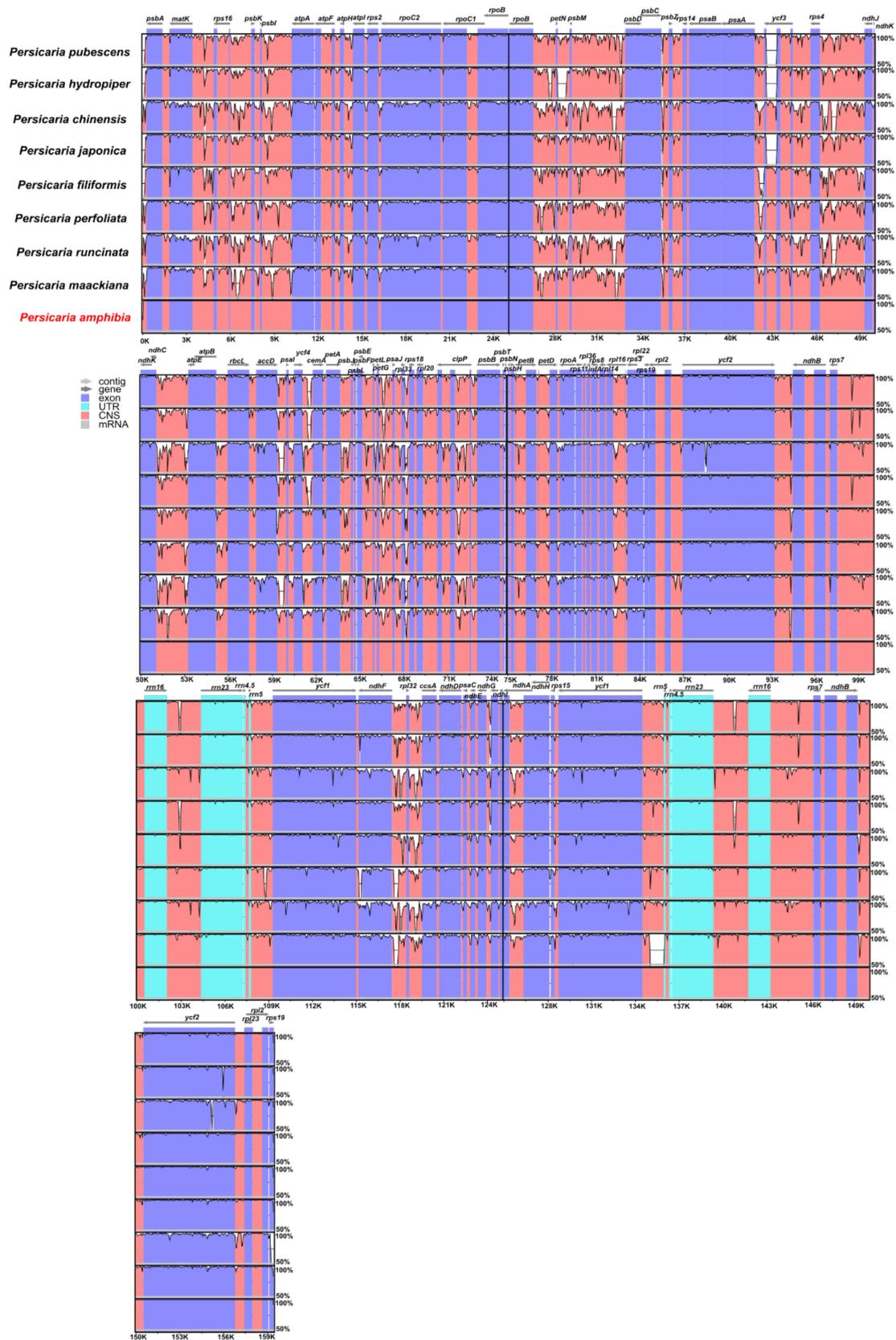


Figure 5. Visualization of the alignment of the 9 *Persicaria* chloroplast genomes. The top grey arrow shows genes in order and the position of each gene. The blue and sky boxes indicate protein-coding genes and rRNA.

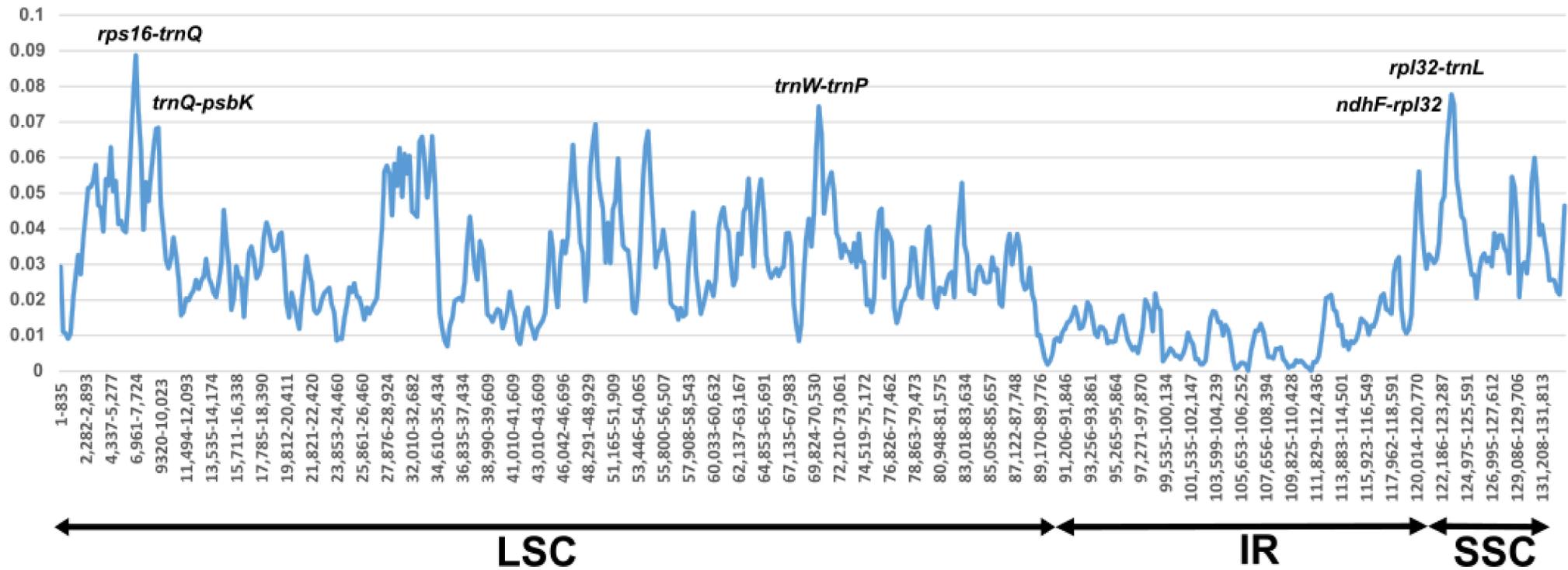
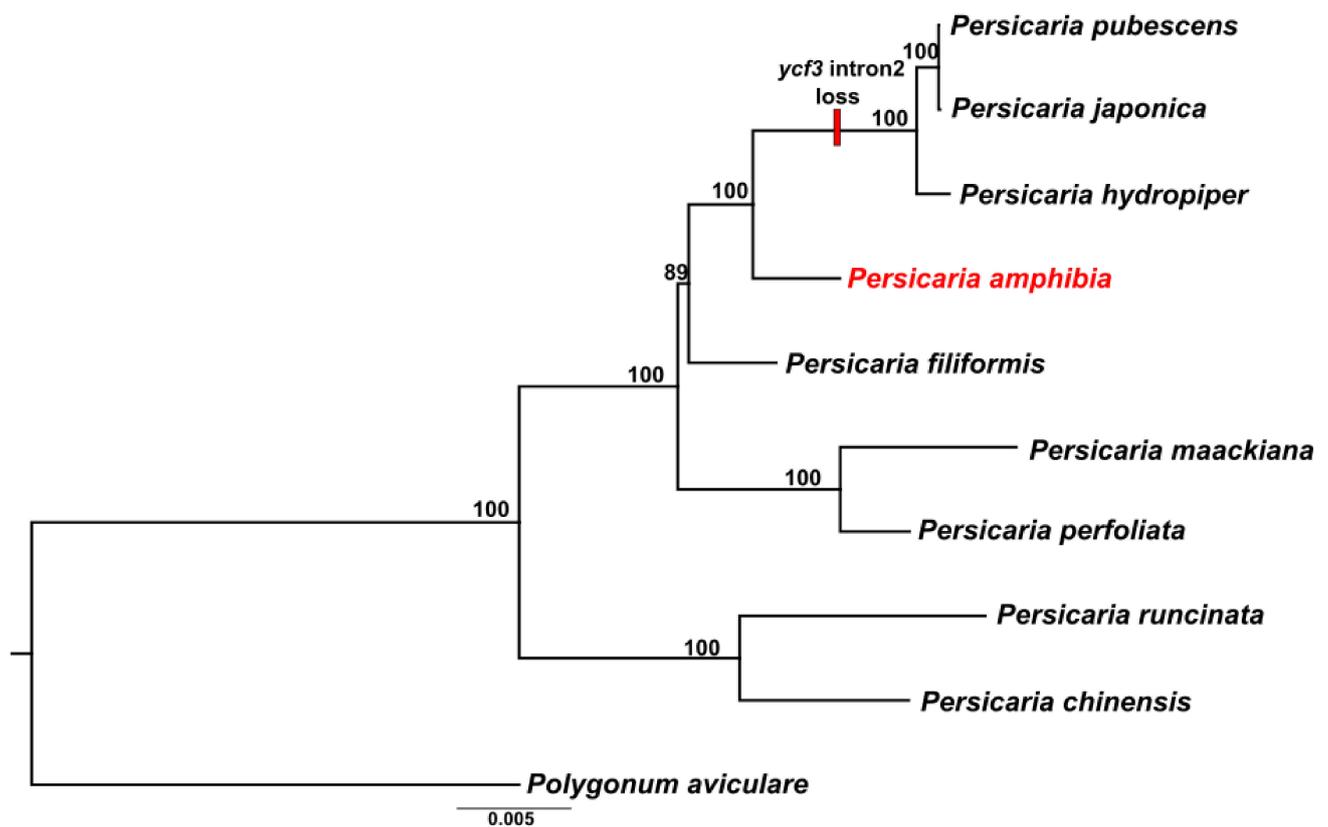


Figure 6. Comparison of the nucleotide variability ( $P_i$ ) values in nine *Persicaria* chloroplast genomes.



**Figure 7.** Phylogenetic tree reconstruction of 9 *Persicaria* species and one outgroup based on the 77 protein-coding genes (66,732 bp). Number above branches are bootstrap value.

#### 4. Discussion

Chloroplast genomes are highly conserved in structures and genes. However, previous studies have reported that gene loss in the chloroplast genome occurs in several lineages, including Geraniaceae [37,38], Orobanchaceae [39–42], Fabaceae [43,44], and Pinaceae [45]. Studies have also shown that introns of *clpP* [44,46], *rps12* [46], and *rpl16* [44] are missing in angiosperms. The gene of *ycf3* normally contains three exons and two introns located in the LSC region of the chloroplast genome. The *ycf3* gene encodes photosystem I, a complex in the thylakoid membrane of plants [47–49]. A previous study reported the loss of *ycf3* intron II in *Grammica* and *Cuscuta nitida* [50]. Our results reveal that among the nine *Persicaria* species, three species (*P. pubescens*, *P. japonica*, and *P. hydropiper*) lost *ycf3* intron II, and *P. amphibia* is sister to these three species. Longevialle et al. [47] suggested that splicing of *ycf3* intron II has an impact on gene *ORGANELLE TRANSCRIPT PROCESSING 51* (OPT51) that specifically promotes the splicing of the only group II intron.

IR/SC junctions are variable sites in the cp genome, such as IR expansion and contraction [51–54]. Comparison of SC/IR junctions in the nine species showed two types: First, an intact *rps19* gene in the IR region of four species (*P. amphibia*, *P. hydropiper*, *P. japonica*, and *P. pubescens*); and second, a partial *rps19* gene in the IR region of the remaining five species. Previous studies have suggested that partial gene duplication or the contraction and expansion of IR regions causes variable types of *rps19* gene [55–57]. However, the IR regions in the nine *Persicaria* are extremely similar in length (Table 1). This suggests that *rps19* in *Persicaria* is more likely to be partially duplicated.

Previous analysis of cpDNA regions (*rbcL*, *trnL-F*, and *matK*) of *Persicaria* [6] revealed that *P. amphibia* is a sister species to the sect. *Eupersicaria*, such as *P. punctate*, and *P. hydropiper*. However, *P. amphibia* did not show the same results using the nuclear ribosomal DNA (nrDNA) ITS data, indicating that *P. amphibia* is sister to sect. *Tovara* (*P. filiformis* and

*P. virginiana*). Our results show that *P. amphibia* is a sister species to *P. hydropper*, *P. japonica*, and *P. pubescens*.

*Persicaria* plants are highly variable in their morphological characters. For example, *P. amphibia* has variable morphological characteristics, and some authors have reported two varieties, such as *P. amphibia* var. *emersa* and *P. amphibia* var. *stipulacea*. Additionally, the genus *Persicaria* was reported to exhibit self-fertilization and hybridization [1,58]. Partridge [59] suggested that *P. amphibia* has economic qualities as feed for wild animals, herbal medicine and as an ornamental garden plant. Recently, molecular phylogenetic analysis was used to identify plants [60,61]. Previously, cpDNA (*matK*, *rbcL*) and nrDNA ITS regions have been used for species identification and phylogenetic analyses [52–65]. In *Persicaria*, previous studies have used cpDNA (*matK*, *rbcL*), nrDNA ITS, and *LEAFY* regions. However, the relationships between the genus *Persicaria* are not well supported [1,4–6]. Our results reveal that the highest sequence divergence regions were *rps16-trnQ*, *trnQ-psbK*, *trnW-trnP*, *ndhF-rpl32*, and *rpl32-trnL* in *Persicaria* species. These variable DNA regions could be used as molecular markers and will be helpful in the phylogenetic analysis of *Persicaria* species.

## 5. Conclusions

In this study, we analyzed the complete cp genome sequence of *P. amphibia*. It comprises an LSC (84,281 bp), an SSC (13,258 bp), and two IR regions (30,956 bp), and contains 79 protein-coding genes, 29 tRNA genes, and four rRNA genes. Comparative cp genome analysis of the nine *Persicaria* species showed that all have a similar gene structure and content. However, *ycf3* intron II is lost in *P. hydropper*, *P. japonica*, and *P. pubescens*. Additionally, we found that the SC/IR junctions in four species (*P. amphibia*, *P. hydropper*, *P. japonica*, and *P. pubescens*) had an intact *rps19* gene. Phylogenetic relationships based on 77 protein-coding genes showed that *P. amphibia* is sister to *P. hydropper*, *P. japonica*, and *P. pubescens*. We found five sequence divergence regions (*rps16-trnQ*, *trnQ-psbK*, *trnW-trnP*, *ndhF-rpl32*, and *rpl32-trnL*) that will be useful markers for the further phylogenetic analysis of the genus *Persicaria*.

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

**Funding:** This work was supported by the grant [NNIBR202201102] from the Nakdonggang National Institute of Biological Resources, funded by the Ministry of Environment, Korea.

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

**Data Availability Statement:** Not applicable.

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

## References

- Kim, S.-T.; Sultan, S.E.; Donoghue, M.J. Allopolyploid speciation in *Persicaria* (Polygonaceae): Insights from a low-copy nuclear region. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12370–12375. [\[CrossRef\]](#)
- Sanchez, A.; Schuster, T.M.; Kron, K.A. A large-scale phylogeny of Polygonaceae based on molecular data. *Int. J. Plant Sci.* **2009**, *170*, 1044–1055. [\[CrossRef\]](#)
- Decraene, L.-P.R.; Akeroyd, J.R. Generic limits in *Polygonum* and related genera (Polygonaceae) on the basis of floral characters. *Bot. J. Linn. Soc.* **1988**, *98*, 321–371. [\[CrossRef\]](#)
- Galasso, G.; Ban, E.; Grassi, F.; Sgorbati, S.; Labra, M. Molecular Phylogeny of *Polygonum* L. s.l. (Polygonaceae), Focusing on European Taxa: Preliminary Results and Systematic Considerations Based on *rbcL* Plastidial Sequence Data. *Atti Soc. Ital. Nat. Mus. Civ. Stor. Nat. Milano* **2009**, *150*, 113–148.
- Gitsopoulos, T.K.; Vasilakoglou, I.; Tsoktouridis, G. *Persicaria Amphibia*, a serious terrestrial weed in Northern Greece: A combined molecular and morphological approach to identification and taxonomy. *Biotechnol. Biotechnol. Equip.* **2013**, *27*, 4236–4242. [\[CrossRef\]](#)
- Kim, S.-T.; Donoghue, M.J. Molecular phylogeny of *Persicaria* (Persicarieae, Polygonaceae). *Syst. Bot.* **2008**, *33*, 77–86. [\[CrossRef\]](#)

7. Wicke, S.; Schneeweiss, G.M.; dePamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [[CrossRef](#)]
8. Dobrogojski, J.; Adamiec, M.; Luciński, R. The chloroplast genome: A review. *Acta Physiol. Plant.* **2020**, *42*, 98. [[CrossRef](#)]
9. Sun, S.-S.; Fu, P.-C.; Zhou, X.-J.; Cheng, Y.-W.; Zhang, F.-Q.; Chen, S.-L.; Gao, Q.-B. The complete plastome sequences of seven species in *Gentiana* Sect. *Kudoa* (Gentianaceae): Insights into plastid gene loss and molecular evolution. *Front. Plant Sci.* **2018**, *9*, 493. [[CrossRef](#)]
10. Frailey, D.C.; Chaluvadi, S.R.; Vaughn, J.N.; Coatney, C.G.; Bennetzen, J.L. Gene loss and genome rearrangement in the plastids of five hemiparasites in the family Orobanchaceae. *BMC Plant Biol.* **2018**, *18*, 30. [[CrossRef](#)]
11. Cauz-Santos, L.A.; da Costa, Z.P.; Callot, C.; Cauet, S.; Zucchi, M.I.; Bergès, H.; van den Berg, C.; Vieira, M.L.C. A repertory of rearrangements and the loss of an inverted repeat region in *Passiflora* chloroplast genomes. *Genome Biol. Evol.* **2020**, *12*, 1841–1857. [[CrossRef](#)]
12. Lee, H.O.; Joh, H.J.; Kim, K.; Lee, S.-C.; Kim, N.-H.; Park, J.Y.; Park, H.-S.; Park, M.-S.; Kim, S.; Kwak, M.; et al. Dynamic chloroplast genome rearrangement and DNA barcoding for three Apiaceae species known as the medicinal herb “Bang-Poong”. *Int. J. Mol. Sci.* **2019**, *20*, 2196. [[CrossRef](#)] [[PubMed](#)]
13. Han, T.; Li, M.; Li, J.; Lv, H.; Ren, B.; Chen, J.; Li, W. Comparison of chloroplast genomes of *Gynura* species: Sequence variation, genome rearrangement and divergence studies. *BMC Genomics* **2019**, *20*, 791. [[CrossRef](#)] [[PubMed](#)]
14. Oulo, M.A.; Yang, J.-X.; Dong, X.; Wang, V.O.; Mkala, E.M.; Munyao, J.N.; Onjolo, V.O.; Rono, P.C.; Hu, G.-W.; Wang, Q.-F. Complete chloroplast genome of *Rhipsalis baccifera*, the only cactus with natural distribution in the Old World: Genome rearrangement, intron gain and loss, and implications for phylogenetic studies. *Plants* **2020**, *9*, 979. [[CrossRef](#)]
15. Liao, Y.-Y.; Liu, Y.; Liu, X.; Lü, T.-F.; Mbichi, R.W.; Wan, T.; Liu, F. The Complete chloroplast genome of *Myriophyllum Spicatum* reveals a 4-Kb inversion and new insights regarding plastome evolution in Haloragaceae. *Ecol. Evol.* **2020**, *10*, 3090–3102. [[CrossRef](#)] [[PubMed](#)]
16. Choi, K.S.; Ha, Y.-H.; Gil, H.-Y.; Choi, K.; Kim, D.-K.; Oh, S.-H. Two Korean endemic *Clematis* chloroplast genomes: Inversion, reposition, expansion of the inverted repeat region, phylogenetic analysis, and nucleotide substitution rates. *Plants* **2021**, *10*, 397. [[CrossRef](#)]
17. Clegg, M.T.; Gaut, B.S.; Learn, G.H.; Morton, B.R. Rates and patterns of chloroplast DNA evolution. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6795–6801. [[CrossRef](#)] [[PubMed](#)]
18. Wang, R.-J.; Cheng, C.-L.; Chang, C.-C.; Wu, C.-L.; Su, T.-M.; Chaw, S.-M. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of Monocots. *BMC Evol. Biol.* **2008**, *8*, 36. [[CrossRef](#)] [[PubMed](#)]
19. Song, Y.; Zhao, W.; Xu, J.; Li, M.; Zhang, Y. Chloroplast genome evolution and species identification of *Styrax* (Styracaceae). *BioMed Res. Int.* **2022**, *2022*, e5364094. [[CrossRef](#)]
20. Choi, K.S.; Chung, M.G.; Park, S. The complete chloroplast genome sequences of three Veroniceae species (Plantaginaceae): Comparative analysis and highly divergent regions. *Front. Plant Sci.* **2016**, *7*, 355. [[CrossRef](#)] [[PubMed](#)]
21. Raman, G.; Park, K.T.; Kim, J.-H.; Park, S. Characteristics of the completed chloroplast genome sequence of *Xanthium spinosum*: Comparative analyses, identification of mutational hotspots and phylogenetic implications. *BMC Genomics* **2020**, *21*, 855. [[CrossRef](#)] [[PubMed](#)]
22. Jansen, R.K.; Kaittanis, C.; Sasaki, C.; Lee, S.-B.; Tomkins, J.; Alverson, A.J.; Daniell, H. Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: Effects of taxon sampling and phylogenetic methods on resolving relationships among Rosids. *BMC Evol. Biol.* **2006**, *6*, 32. [[CrossRef](#)] [[PubMed](#)]
23. 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](#)]
24. Zhai, W.; Duan, X.; Zhang, R.; Guo, C.; Li, L.; Xu, G.; Shan, H.; Kong, H.; Ren, Y. Chloroplast genomic data provide new and robust insights into the phylogeny and evolution of the Ranunculaceae. *Mol. Phylogenet. Evol.* **2019**, *135*, 12–21. [[CrossRef](#)]
25. Yu, B.; Liu, J.; Liu, X.; Lan, X.; Qu, X. The complete chloroplast genome sequence of *Polygonum chinense* L. *Mitochondrial DNA Part B* **2020**, *5*, 2139–2140. [[CrossRef](#)]
26. Jin, J.-J.; Yu, W.-B.; Yang, J.-B.; Song, Y.; dePamphilis, C.W.; Yi, T.-S.; Li, D.-Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **2020**, *21*, 241. [[CrossRef](#)]
27. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—Versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [[CrossRef](#)]
28. Chan, P.P.; Lin, B.Y.; Mak, A.J.; Lowe, T.M. TRNAscan-SE 2.0: Improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* **2021**, *49*, 9077–9096. [[CrossRef](#)]
29. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)]
30. Greiner, S.; Lehwark, P.; Bock, R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **2019**, *47*, W59–W64. [[CrossRef](#)]
31. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)] [[PubMed](#)]

32. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-Web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
33. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, W273–W279. [[CrossRef](#)] [[PubMed](#)]
34. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)] [[PubMed](#)]
35. Katoh, K. MAFFT: A novel method for rapid multiple sequence alignment based on fast fourier Transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)] [[PubMed](#)]
36. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)]
37. Chris Blazier, J.; Guisinger, M.M.; Jansen, R.K. Recent loss of plastid-encoded *ndh* genes within *Erodium* (Geraniaceae). *Plant Mol. Biol.* **2011**, *76*, 263–272. [[CrossRef](#)]
38. Weng, M.-L.; Blazier, J.C.; Govindu, M.; Jansen, R.K. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. *Mol. Biol. Evol.* **2014**, *31*, 645–659. [[CrossRef](#)]
39. Li, X.; Zhang, T.-C.; Qiao, Q.; Ren, Z.; Zhao, J.; Yonezawa, T.; Hasegawa, M.; Crabbe, M.J.C.; Li, J.; Zhong, Y. Complete chloroplast genome sequence of holoparasite *Cistanche deserticola* (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host *Haloxylon ammodendron* (Chenopodiaceae). *PLoS ONE* **2013**, *8*, e58747. [[CrossRef](#)]
40. Samigullin, T.H.; Logacheva, M.D.; Penin, A.A.; Vallego-Roman, C.M. Complete plastid genome of the recent holoparasite *Lathraea squamaria* reveals earliest stages of plastome reduction in Orobanchaceae. *PLoS ONE*. **2016**, *11*, e0150718. [[CrossRef](#)]
41. Zhou, T.; Ruhsam, M.; Wang, J.; Zhu, H.; Li, W.; Zhang, X.; Xu, Y.; Xu, F.; Wang, X. The complete chloroplast genome of *Euphrasia regelii*, pseudogenization of *ndh* genes and the phylogenetic relationships within Orobanchaceae. *Front. Genet.* **2019**, *10*, 444. [[CrossRef](#)] [[PubMed](#)]
42. Choi, K.-S.; Park, S. Complete plastid and mitochondrial genomes of *Aeginetia indica* reveal intracellular gene transfer (IGT), horizontal gene transfer (HGT), and cytoplasmic male sterility (CMS). *Int. J. Mol. Sci.* **2021**, *22*, 6143. [[CrossRef](#)] [[PubMed](#)]
43. Downie, S.R.; Llanas, E.; Katz-Downie, D.S. Multiple independent losses of the *rpoC1* intron in angiosperm chloroplast DNA's. *Syst. Bot.* **1996**, *21*, 135. [[CrossRef](#)]
44. Li, C.; Zhao, Y.; Xu, Z.; Yang, G.; Peng, J.; Peng, X. Initial characterization of the chloroplast genome of *Vicia sepium*, an important wild resource plant, and related inferences about its evolution. *Front. Genet.* **2020**, *11*, 73. [[CrossRef](#)]
45. Ni, Z.; Ye, Y.; Bai, T.; Xu, M.; Xu, L.-A. Complete chloroplast genome of *Pinus massoniana* (Pinaceae): Gene rearrangements, loss of *ndh* genes, and short inverted repeats contraction, expansion. *Molecules* **2017**, *22*, 1528. [[CrossRef](#)]
46. Jansen, R.K.; Wojciechowski, M.F.; Sanniyasi, E.; Lee, S.-B.; Daniell, H. Complete plastid genome sequence of the Chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps12* and *clpP* intron losses among Legumes (Leguminosae). *Mol. Phylogenet. Evol.* **2008**, *48*, 1204–1217. [[CrossRef](#)]
47. De Longevialle, A.F.; Hendrickson, L.; Taylor, N.L.; Delannoy, E.; Lurin, C.; Badger, M.; Millar, A.H.; Small, I. The Pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid *ycf3* intron 2 in *Arabidopsis thaliana*. *Plant J.* **2008**, *56*, 157–168. [[CrossRef](#)]
48. Albus, C.A.; Ruf, S.; Schöttler, M.A.; Lein, W.; Kehr, J.; Bock, R. Y3IP1, a nucleus-encoded thylakoid protein, cooperates with the plastid-encoded *ycf3* protein in photosystem I assembly of Tobacco and *Arabidopsis*. *Plant Cell* **2010**, *22*, 2838–2855. [[CrossRef](#)]
49. Wang, X.; Yang, Z.; Zhang, Y.; Zhou, W.; Zhang, A.; Lu, C. Pentatricopeptide repeat protein PHOTOSYSTEM I BIOGENESIS FACTOR2 is required for splicing of *ycf3*. *J. Integr. Plant Biol.* **2020**, *62*, 1741–1761. [[CrossRef](#)]
50. McNeal, J.R.; Kuehl, J.V.; Boore, J.L.; Leebens-Mack, J.; dePamphilis, C.W. Parallel loss of plastid introns and their maturase in the genus *Cuscuta*. *PLoS ONE* **2009**, *4*, e5982. [[CrossRef](#)]
51. Sasaki, C.; Lee, S.-B.; Fjellheim, S.; Guda, C.; Jansen, R.K.; Luo, H.; Tomkins, J.; Rognli, O.A.; Daniell, H.; Clarke, J.L. Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. *Theor. Appl. Genet.* **2007**, *115*, 571–590. [[CrossRef](#)] [[PubMed](#)]
52. Choi, K.S.; Park, S. The complete chloroplast genome sequence of *Aster spathulifolius* (Asteraceae); genomic features and relationship with Asteraceae. *Gene* **2015**, *572*, 214–221. [[CrossRef](#)] [[PubMed](#)]
53. Choi, I.-S.; Choi, B.-H. The distinct plastid genome structure of *Maackia fauriei* (Fabaceae: Papilionoideae) and its systematic implications for genistoids and tribe Sophoreae. *PLoS ONE* **2017**, *12*, e0173766. [[CrossRef](#)] [[PubMed](#)]
54. Chen, N.; Sha, L.-N.; Wang, Y.-L.; Yin, L.-J.; Zhang, Y.; Wang, Y.; Wu, D.-D.; Kang, H.-Y.; Zhang, H.-Q.; Zhou, Y.-H.; et al. Variation in plastome sizes accompanied by evolutionary history in monogenomic Triticeae (Poaceae: Triticeae). *Front. Plant Sci.* **2021**, *12*, 741063. [[CrossRef](#)]
55. Lu, R.-S.; Li, P.; Qiu, Y.-X. The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: Comparative genomic and phylogenetic analyses. *Front. Plant Sci.* **2017**, *7*, 2054. [[CrossRef](#)]
56. Wu, C.-X.; Zhai, C.-C.; Fan, S.-J. Characterization of the complete chloroplast genome of *Rumex nepalensis* (Polygonaceae). *Mitochondrial DNA Part B* **2020**, *5*, 2458–2459. [[CrossRef](#)]
57. Yang, Q.; Fu, G.-F.; Wu, Z.-Q.; Li, L.; Zhao, J.-L.; Li, Q.-J. Chloroplast genome evolution in four Montane Zingiberaceae taxa in China. *Front. Plant Sci.* **2022**, *12*, 774482. [[CrossRef](#)]

58. Timson, J. A study of hybridization in *Polygonum* Section *Persicaria*. *J. Linn. Soc. Lond. Bot.* **1965**, *59*, 155–161. [[CrossRef](#)]
59. Patridge, J.W. *Persicaria amphibia* (L.) Gray (*Polygonum amphibium* L.). *J. Ecol.* **2001**, *89*, 487–501. [[CrossRef](#)]
60. Zaman, W.; Ye, J.; Saqib, S.; Liu, Y.; Shan, Z.; Hao, D.; Chen, Z.; Xiao, P. Predicting potential medicinal plants with phylogenetic topology: Inspiration from the research of traditional Chinese medicine. *J. Ethnopharmacol.* **2021**, *281*, 114515. [[CrossRef](#)]
61. Zaman, W.; Ye, J.; Ahmad, M.; Saqib, S.; Shinwari, Z.K.; Chen, Z. Phylogenetic exploration of traditional Chinese medicinal plants: A case study on Lamiaceae. *Pak. J. Bot.* **2022**, *54*, 1033–1040. [[CrossRef](#)]
62. Tamura, M.N.; Yamashita, J.; Fuse, S.; Haraguchi, M. Molecular phylogeny of monocotyledons inferred from combined analysis of plastid *matK* and *rbcl* gene sequences. *J. Plant Res.* **2004**, *117*, 109–120. [[CrossRef](#)] [[PubMed](#)]
63. Topik, H.; Yukawa, T.; Ito, M. Molecular phylogenetics of subtribe *Aeridinae* (Orchidaceae): Insights from plastid *matK* and nuclear ribosomal ITS sequences. *J. Plant Res.* **2005**, *118*, 271–284. [[CrossRef](#)] [[PubMed](#)]
64. Tokuoka, T. Molecular phylogenetic analysis of Euphorbiaceae sensu stricto based on plastid and nuclear DNA sequences and ovule and seed character evolution. *J. Plant Res.* **2007**, *120*, 511–522. [[CrossRef](#)] [[PubMed](#)]
65. Kim, J.S.; Hong, J.-K.; Chase, M.W.; Fay, M.F.; Kim, J.-H. Familial relationships of the monocot order Liliales based on a molecular phylogenetic analysis using four plastid loci: *matK*, *rbcl*, *atpB* and *atpF-H*. *Bot. J. Linn. Soc.* **2013**, *172*, 5–21. [[CrossRef](#)]