Comparative Analysis and Phylogenetic Implications of Plastomes of Five Genera in Subfamily Amyridoideae (Rutaceae)

Kuo Sun 1, Qiao-Yun Liu 1, Ao Wang 1, Yong-Wei Gao 1, Liang-Cheng Zhao 2,* and Wen-Bin Guan 1,∗

1 School of Ecology and Nature Conservation, Beijing Forestry University, Beijing 100083, China; sunkuo@bjfu.edu.cn (K.S.); lqy8014a@163.com (Q.-Y.L.); bjfuwangao@163.com (A.W.); yongweigao@126.com (Y.-W.G.)
2 Museum of Beijing Forestry University, Beijing Forestry University, Beijing 100083, China
* Correspondence: lczhao@bjfu.edu.cn (L.-C.Z.); swlab@bjfu.edu.cn (W.-B.G.)

Abstract: In the most recent classification of Rutaceae, Amyridoideae is the largest and most diverse subfamily. In Amyridoideae, the genera Phellodendron, Tetradium, Toddalia and Zanthoxylum were proposed as “proto-Rutaceae” due to substantial phytochemical similarities. In this study, we investigated the plastome variations in eight species representing these four genera and Melicope. All plastomes exhibited a typical quadripartite structure with four regions (LSC, SSC, IRa and IRb). The whole chloroplast genome size ranged from 158,383 bp to 159,014 bp and the gene number ranged from 115 to 116. By comparative analyses, we found that there were structural variations at the LSC/IR and SSC/IR borders of the plastomes in the five genera, especially in Melicope. Three most divergent regions (trnH-psbA, trnE-trnT and psaB) were found from the LSC region, which had great potential for developing effective genetic markers. In addition, we conducted a phylogenomic analyses of the “proto-Rutaceae” and related taxa with plastome data from 36 species. Our results showed that (1) Phellodendron, Tetradium, Toddalia and Zanthoxylum were confirmed as close relatives and grouped together as the ‘proto-Rutaceae’, (2) Phellodendron was sister to Tetradium, and Toddalia was deeply nested within Zanthoxylum, and (3) Toddalia asiatica was closely related to the Southwest Pacific and East Asian species of Zanthoxylum, and Melicope pteleifolia was more closely related to Acronychia than it is to Tetradium. This study provided new insights into the plastome structural variations in subfamily Amyridoideae, and demonstrated that the plastomes data were sufficiently robust to explore implications of the phylogeny for the previous phylogenetic hypotheses.

Keywords: Amyridoideae; proto-Rutaceae; chloroplast genome; structural variation; phylogenetic relationship

1. Introduction

Chloroplasts (cp) are one of the essential organelles in plant cells. The chloroplast genomes (plastomes) are usually circular DNA molecules, composed of a large single copy (LSC) region and a small single copy (SSC) region that are separated by a pair of inverted repeats (IRs) regions [1]. However, many plastomes have revealed considerable variation within and between plant species in terms of both sequence and structural variation, which is valuable for understanding of the climatic adaptation of plants [2]. In addition, due to its simple structure, highly conserved sequence, and maternal inheritance characteristics, the plastome has long been a focus of research in plant molecular evolution and systematics [3]. In recent years, the whole plastome genomes have made significant contributions to phylogenetic studies of many plant families, and have been suggested to give higher resolution of evolutionary relationships within phylogenetic clades, compared with small numbers of plastid or nuclear markers [2,4–7].
Rutaceae is a large family of flowering plants, containing 150–162 genera and 1500–2100 species, distributed nearly cosmopolitan, but mainly in tropical and subtropical regions [8,9]. So far, the Rutaceae family has been divided into two [9], three [8], four [10], or seven [11] subfamilies. In the most recent classification [10], Amyridoideae Link is the largest and most diverse subfamily. Among this subfamily, the genera Phellodendron Rupr., Tetradium Lour., Toddalia Juss. and Zanthoxylum L., were found to have the capability to produce the 1-benzyltetrahydroisoquinoline (1-BTIQ) alkaloids [12,13]. Waterman (1983) [14] further proposed to place these four genera in a ‘proto-Rutaceae’. The genus Fagaropsis Mildbr. ex Siebenl. later was added to this group [15]. Among them, Toddalia is a monotypic genus, with the sole species T. asiatica (L.) Lam. widely distributed in Asia and Africa. Phellodendron consists of only two species, P. chinense Schneid. and P. amurense Rupr., mainly distributed in eastern Asia [16]. Tetradium was long confused with the genus Euodia J. R. Forst. & G. Forst, and traditionally placed in the synonymy of Euodia. Euodia sensu lato, formerly comprised about 120 species [17], but in a morphology-based revision, Hartley [18,19] moved the Southeast Asian-Pacific species with pinnately compound leaves from Euodia sensu lato to the reinstated Tetradium, and the majority of the species with unifoliate or trifoliate leaves to the genus Melicope. After his revision, Tetradium consists of nine species and is distributed from the Himalayas east to Japan and south to Java and Sumbawa [18]. Zanthoxylum is the second largest genus of the Rutaceae after Melicope, consists of about 225 species and is distributed widely in tropical and subtropical regions and extended to temperate latitudes in eastern Asia and North America [20]. Fagaropsis is distributed in tropical Africa and Madagascar with four species [8].

As a special group, the “proto-Rutaceae” has attracted attention not only for its chemical characteristics, but for their phylogenetic relationships between the genera within it. In addition to morphological and biochemical studies, in recent years, molecular phylogenetic studies provided new insights into the intergeneric relationships of the “proto-Rutaceae” and related taxa, including Melicope J.R.Forst. & G.Forst. and Acronychia J.R.Forst. & G.Forst. An earlier cladistic analysis using two plastid and two nuclear markers [21] showed that Phellodendron, Tetradium, Toddalia, and Zanthoxylum were resolved as a clade, supporting the proposal for the ‘proto-Rutaceae’. This study also suggested that Phellodendron is sister to Tetradium and that Toddalia is sister to Zanthoxylum. Furthermore, it revealed that Melicope is closely related to Acronychia but the position of M. vitiflora (F.Muell.) T.G.Hartley is uncertain. Later study using four nuclear and two plastid markers [22] and recent study using two plastid and two nuclear markers [18] supported the sister-group relationship of Phellodendron and Tetradium, but showed that Toddalia is nested with Zanthoxylum. Based on the analyses of two nuclear and plastid sequences, Ling et al. (2009) [23] confirmed that Fagaropsis was a member of the “proto-Rutaceae” clade. Although the “proto-Rutaceae” and some related genera have been sampled in several phylogenetic studies at a broader scale, only a limited number of chloroplast and nuclear gene markers were used to infer the phylogeny, and the resolution of some important nodes in phylogenetic trees is lower. So far, little is known about the plastomes structural characteristics and variation of these genera and it is unknown whether the ‘proto-Rutaceae’ is a natural group in terms of plastomes data. In this study, firstly, we conducted comparative analyses of three newly assembled plastomes and five plastomes obtained from GenBank, representing five genera (Phellodendron, Tetradium, Toddalia, Zanthoxylum, and Melicope) in Amyridoideae, to reveal their plastome characteristics and variation within them. Secondly, we presented a phylogenomic analysis based on plastome sequences from 36 species of Rutaceae, plus three species of Simaroubaceae, and one species of Meliaceae to explore implications of the robust phylogeny for the previous phylogenetic hypotheses of these genera. Due to the lack of material and published complete chloroplast genomes, the genus Fagaropsis was not included in our analysis.
2. Materials and Methods

2.1. Taxon Sampling, DNA Extraction and Sequencing

Five samples of *Phellodendron chinense*, *Tetradium daniellii* (Benn.) T.G. Hartley, and *Toddalia asiatica* were collected in China from Tianquan of Sichuan, Changping of Beijing, and Gushan of Fujian, respectively. Voucher specimens were deposited at the Museum of Beijing Forestry University (Table S1). Total genomic DNA was extracted from silica-dried leaves of the three species using CTAB method [24].

Average 150 bp pair-end reads were performed on an Illumina HiSeq 4000 platform at Novogene Biotech Co. (http://www.novogene.com (accessed on 1 February 2021), China). The filtered reads were used for *de novo* assembly with Geneious Prime [25]. Then the generated contigs were concatenated into larger contigs based on the published plastomes of *P. amurense* (NC_035551.1), *T. ruticarpum* (MT134114) and *Z. pinnatum* (MN968553). By using the Repeat Finder plugin of Geneious, the IR regions were determined and copied in reverse order to complete the chloroplast sequences. These sequences were annotated by the Unix Program Plann [26]. After checking the annotated circular structures, we used the Organellar Genome DRAW tool [27] to generate the illustrations of circular plasmids.

2.2. Codon Usage and Characterization of Repeat Sequences

Apart from the three newly assembled plastomes, we downloaded the plastomes of other five Amyridoideae species from GenBank (Table S1) for comparative analysis. Then we used CodonW1.4.2 program [28] to calculate relative synonymous codon usage (RSCU) value of these eight plastomes. The R package ggplot2 was used to visualize the RSCU values. RSCU = 1 indicated that codon had no use preference, the codons with RSCU value >1 were defined as high frequency codons whereas the reverse was true for RSCU values less than 1.0 [29].

Simple sequence repeats (SSRs) were produced by MISA [30]. The thresholds of mononucleotides, dinucleotides, trinucleotides, tetranucleotides, pentanucleotides, and hexanucleotides repeats were 8 repeat units, 5 repeat units, 4 repeat units, 3 repeat units, 3 repeat units and 3 repeat units, respectively. Repeat sequences were detected by the online REPuter software tool [31], including forward, reverse, complement and palindromic repeats, with minimum repeat size of 30 bp and 90% or more sequence identity.

2.3. Genome Comparison

We analyzed base compositions of the eight plastomes and calculated the GC content by Geneious. To illustrate the IR expansion and contraction, we specifically compared the IR/SC boundary regions by IRscope [32]. In order to compare the similarity, we used the MAFFT [33] to align chloroplast genome sequences. The mVISTA [34], an online tool, was used to evaluate similarity by visualizing the results with the annotation of *P. chinense* as the reference. We made a further analysis in the LAGAN mode, which can produce true multiple alignments regardless of whether or not inversions were contained [35]. To identify the nucleotide diversity (Pi), a sliding window analysis was conducted using DnaSP [36]. And we used Microsoft Excel 2011 to convert results into charts for observation.

2.4. Phylogenetic Analysis

A total of 36 complete plastomes (Tables S1 and S2) from 15 genera of Rutaceae (four plastomes newly generated and 32 plastomes obtained from GenBank) were included for phylogenomic analyses. Three species of Simaroubaceae and one species of Meliaceae (plastomes obtained from GenBank) were defined as outgroups. The alignment of these sequences was performed in MAFFT [33], then the nonhomologous sites were deleted for phylogenetic trees. Bayesian inference (BI) was executed with MrBayes v3.2.3 [37] in the GTR + G + I model. Maximum likelihood trees (ML) were conducted by IQtree [38] and 1000 bootstrap replicates were test to evaluate the branch support values. The protein-coding genes of the 36 sequences were also extracted to construct phylogenetic trees using the same methods as the complete chloroplast genome sequences.
3. Results and Discussion

3.1. Structure and Size of Plastomes

Three newly assembled chloroplast genomes (Phellodendron chinense, Tetradium daniellii and Toddalia asiatica) and five published chloroplast genomes (P. amurense, Tetradium ruticarpum (A.Juss.) T.G.Hartley, and Melicope pteleifolia (Champ. ex Benth.) T.G.Hartley, Zanthoxylum calcicola Huang and Z. piperitum (L.) DC.) obtained from GenBank, representing five genera in subfamily Amyridoideae, were used for comparative analyses. All plastomes of these five genera exhibited a similar quadripartite structure (Figure 1), containing a large single copy (LSC) region (85,124 bp–86,387 bp), a small single copy (SSC) region (17,526 bp–18,610 bp), and a pair of inverted repeats (IRs) (26,999 bp–27,644 bp), respectively. The whole genome size of the eight species ranged from 158,383 bp (T. daniellii) to 159,014 bp (T. asiatica) (Table 1), which were similar to or slightly smaller than the reported plastomes of some genera of subfamily Aurantioideae, Rutaceae [39–41].

Figure 1. Gene map of three Rutaceae species plastomes. Genes shown inside the inner circle are transcribed counterclockwise and those outside are transcribed clockwise. Genes belonging to different functional groups are colour-coded. Darker grey in the inner circle corresponds to the GC content of the chloroplast genome. IRs, two inverted repeats; LSC, the large single copy; SSC, the small single copy.
Table 1. General features of the eight chloroplast genomes compared in this study

| Features                        | Phellodendron amurense | Phellodendron chinense | Tetradium daniellii | Tetradium ruticarpum | Melicope ptelefolia | Zanthoxylum calcicola | Zanthoxylum piperitum | Toddalia asiatica |
|---------------------------------|------------------------|------------------------|---------------------|----------------------|--------------------|-----------------------|-----------------------|---------------------|
| Total cp genome size (bp)       | 158,442                | 158,537                | 158,383             | 158,745              | 159,014            | 158,591               | 158,154               | 158,434             |
| Length of LSC (bp)              | 86,144                 | 86,250                 | 86,386              | 86,299               | 85,124             | 86,352                | 85,538                | 86,132              |
| Length of IR (bp)               | 26,999                 | 27,000                 | 27,000              | 27,101               | 27,640             | 27,016                | 27,016                | 27,008              |
| Length of SSR (bp)              | 18,284                 | 18,287                 | 17,997              | 18,244               | 18,610             | 18,207                | 17,526                | 18,288              |
| Coding size (bp)                | 91,058                 | 91,058                 | 90,882              | 92,133               | 91,356             | 91,809                | 92,240                | 92,123              |
| Intron size (bp)                | 19,368                 | 19,372                 | 18,208              | 18,424               | 19,332             | 18,335                | 18,355                | 18,186              |
| Spacer size (bp)                | 48,016                 | 48,107                 | 49,293              | 48,188               | 48,447             | 47,559                | 48,125                | 48,125              |
| Total GC content (%)            | 38.4                   | 38.4                   | 38.4                | 38.3                 | 38.6               | 38.4                  | 38.5                  | 38.5                |
| GC content of LSC (%)           | 36.7                   | 36.6                   | 36.7                | 36.6                 | 36.7               | 36.7                  | 36.8                  | 36.8                |
| GC content of IR (%)            | 42.9                   | 42.9                   | 42.7                | 42.8                 | 42.8               | 42.8                  | 42.5                  | 42.8                |
| GC content of SSC (%)           | 33.2                   | 33.2                   | 33.3                | 33.2                 | 32.8               | 33.4                  | 33.7                  | 33.4                |
| GC content of rRNA (%)          | 55.6                   | 55.6                   | 55.9                | 55.9                 | 55.6               | 55.5                  | 55.5                  | 55.6                |
| GC content of tRNA (%)          | 53.2                   | 53.2                   | 53.1                | 53.1                 | 53.0               | 53.0                  | 53.0                  | 53.1                |
| GC content of Intron (%)        | 38.0                   | 38.0                   | 38.3                | 38.3                 | 38.3               | 38.3                  | 38.3                  | 38.2                |
| Total number of genes           | 116                    | 116                    | 115                 | 115                  | 115               | 116                   | 116                   | 115                 |
| Protein encoding-genes          | 81                     | 81                     | 81                  | 81                   | 80               | 81                    | 81                    | 81                  |
| tRNA genes                      | 30                     | 30                     | 30                  | 30                   | 30               | 30                    | 30                    | 30                  |
| rRNA genes                      | 4                      | 4                      | 4                   | 4                    | 4               | 4                     | 4                     | 4                   |
The DNA GC content is an important indicator of species affinity. The total GC contents of the five genera chloroplast genomes varied from 38.3% to 38.6%, which was obviously lower than AT contents. The average GC contents of SSC region, LSC region and IR regions were 33.3%, 36.7%, and 42.8%, respectively, indicating that the rRNA and tRNA, which occupy mainly in IR regions, prefer to use bases G and C. The total gene content and order were similar across the five genera plastomes. There were 116 genes in *Phellodendron* and *Zanthoxylum*, while 115 genes in *Tetradium*, *Toddalia* and *Melicope*. The number of protein-coding genes in *Melicope* was 80, while other four genera had 81 protein-coding genes. All the genera had 30 tRNA genes and 4 rRNA genes (Table 1).

3.2. Codon Preference Analysis

Genes in the closely related species generally show a similar codon use pattern. Through the analysis of codon preference, we can better understand the evolution of species [42]. Based on the protein-coding sequences, the frequency of codon usage was estimated for the five genera plastomes. In total, the protein-coding genes were composed of 26,943–27,731 codons, which encoding 21 amino acids (Figure S1). Among them, the most frequent amino acid was Leucine (11.98–15.20%) and the least frequent coded amino acid was Cysteine (1.11–1.84%).

In our analysis, RSCU values of 30 codons were greater than 1, most of them ended with A or U and only five codons ended with G, which indicated that the preferred codons tended to be A/U ending. As shown in Table 2, the effective number of codons (Nc) value ranged from 51.09 to 57.46. When the Nc value >35, there is relatively weak codon usage bias [43]. The content of each base on the third synonymous codon (T<sub>3s</sub>, C<sub>3s</sub>, A<sub>3s</sub>, and G<sub>3s</sub>) should have A = T and C = G when there is no variation or the selection effect on the bias in two complementary strands, and the usage pattern was completely caused by mutation [44]. As can be seen from Table 2, the use frequency between Tr+C and A+G was unbalanced, which indicated that the use pattern of codons in these five genera was also influenced by other factors in addition to mutation. This pattern was similar to that reported for other Rutaceae plastomes such as *Citrus* [45].

| Species                        | T<sub>3s</sub> | C<sub>3s</sub> | A<sub>3s</sub> | G<sub>3s</sub> | Nc       |
|-------------------------------|---------------|---------------|---------------|-------------|----------|
| *Phellodendron amurense*      | 0.3461        | 0.2424        | 0.3882        | 0.312       | 57.46    |
| *Phellodendron chinense*      | 0.4555        | 0.1816        | 0.4206        | 0.1916      | 51.09    |
| *Tetradium daniellii*         | 0.3982        | 0.2174        | 0.4061        | 0.2481      | 55.36    |
| *Tetradium ruticarpum*        | 0.4539        | 0.1831        | 0.4206        | 0.1924      | 51.23    |
| *Zanthoxylum calcicola*       | 0.4536        | 0.1832        | 0.4207        | 0.1924      | 51.2     |
| *Zanthoxylum piperitum*       | 0.4391        | 0.2002        | 0.4159        | 0.2016      | 52.57    |
| *Melicope ptelefolia*         | 0.4599        | 0.182         | 0.4209        | 0.1912      | 51.18    |
| *Toddalia asiatica*           | 0.4382        | 0.2035        | 0.4122        | 0.2062      | 52.97    |

3.3. IR Contraction and Expansion

The contraction and expansion of IR region at the borders can affect the size difference between chloroplast genomes and play important roles in evolution [46]. Although the IR regions of the five genera were relatively conserved, structural variations were found in different degrees, especially in *Melicope* (Figure 2). The IR regions of the eight plastomes ranged from 26,999 bp (*P. amurense*) to 27,644 bp (*Z. piperitum*), among which *rps3, rpl22, rps19, ndhF, ycf1, trnN, rpl2* and *trnH* were located at the junction of the LSC/IR and SSC/IR borders. For the LSC/IRb region, the gene *rpl22* in four genera crossed the junction and extended into the IRb region, with a length from 39 bp to 246 bp. In *Melicope*, however, the gene *rpl22* was completely located in the IRb region, and the gene *rps3* crossed the junction and extended 429 bp into the IRb region. For the IRb/SSC region, the gene *ndhF* was located near the junction or crossed the junction and extended into the IRb regions of *T. daniellii* (6 bp) and *Z. piperitum* (23 bp). All eight plastomes had similar SSC/IRa borders,
and the junction was all crossed by the gene ycf1 with a length from 1069 bp to 1704 bp. For the IRa/LSC region, the junction was within the region between the genes rps19 and trnH in four genera, while in Melicope, the junction was within the region between the genes rpl22 and trnH. In addition, it was found that the gene rpl2 was missing in both the IRa and IRb regions of Melicope.

Figure 2. Comparison of the borders of LSC, SSC, and IR regions among eight plastomes of five genera in Amyridoideae. Different colour boxes indicate specific genes.

### 3.4. Repeat and SSR Analysis

SSRs were an important type of repeated sequence in genomes that are particularly useful molecular markers in genetic diversity research [47]. From Figure 3A, we found that, in all eight species, the palindromic type had the most repeats, and the forward type ranked second. Reverse and complement repeats were much fewer, and even were not found in Z. piperitum and M. pteleifolia. Most SSRs were located in the LSC region (71%), followed by SSC (16%) and IR regions (13%) (Figure 3B). We identified 199, 197, 69, 74, 72, 79, 65 and 62 SSRs in *P. amurense*, *P. chinense*, *T. daniellii*, *T. ruticarpum*, *Z. calcicola*, *Z. piperitum*, *M. pteleifolia* and *T. asiatica*, respectively (Figure 3C). A mononucleotide repeat unit (A/T) was most abundant in all species, accounting for 86.29–97.30%. This was followed by dinucleotide repeat unit (AT/AT) and (AG/CT) (Figure 3C,D). It is noteworthy that *Phellodendron* plastome not only had most SSRs, the trinucleotide repeat units (AAT/ATT and AAG/CTT) and the polynucleotide repeats were only found in this genus.
3.4. Repeat and SSR Analysis

SSRs were an important type of repeated sequence in genomes that are particularly useful molecular markers in genetic diversity research [47]. From Figure 3A, we found that, in all eight species, the palindromic type had the most repeats, and the forward type ranked second. Reverse and complement repeats were much fewer, and even were not found in *Z. piperitum* and *M. pteleifolia*. Most SSRs were located in the LSC region (71%), followed by SSC (16%) and IR regions (13%) (Figure 3B). We identified 199, 197, 69, 74, 72, 79, 65 and 62 SSRs in *P. amurense*, *P. chinense*, *T. daniellii*, *T. ruticarpum*, *Z. calcicola*, *Z. piperitum*, *M. pteleifolia* and *T. asiatica*, respectively (Figure 3C). A mononucleotide repeat unit (A/T) was most abundant in all species, accounting for 86.29%–97.30%. This was followed by dinucleotide repeat unit (AT/AT) and (AG/CT) (Figure 3C,D). It is noteworthy that *Phellodendron* plastome not only had most SSRs, the trinucleotide repeat units (AAT/ATT and AAG/CTT) and the polynucleotide repeats were only found in this genus.

![Figure 3. Simple sequence repeats of the plastomes of the eight species. (A) Number of four repeat types; (B) Proportion of SSRs in different genomic regions; (C) Number of different SSRs types detected in plastome. (D) Number of SSRs in different repeat class types.](image)

3.5. Sequence Divergence Analysis

To characterize genome divergence, multiple sequence alignments were performed between the eight plastomes (Figure 4). The comparison demonstrated that the eight plastomes had the high consistency in genes arrangement, especially in the same genus. Most of sequence variations occurred in the non-coding sequences, which indicated that the coding regions were more conserved than non-coding regions. Particularly, the IR regions were much less divergent than the LSC and SSC regions. We also calculated the nucleotide variability (Pi) to show divergence at the sequence level between the eight plastomes (Figure 5). It was found that the Pi values ranged from 0 to 0.08833 with a mean of 0.02086 in the LSC region, from 0 to 0.07306 with an average of 0.01944 in the SSC regions, and from 0 to 0.04125 with a mean of 0.00539 in the IR regions, indicating that IR regions had much lower variability than the LSC and SSC regions. Three most divergent hotspots, *trnH-psbA*, *trnE-trnT* and *psaB* regions, were observed, which exhibited significantly higher Pi values (>0.08) and were all located in the LSC region. These divergent regions may be undergoing
rapid nucleotide substitution at the species level, indicating great potential as barcode markers for plant identification and further phylogenetic analysis in Amyridoideae.

Figure 4. Sequence alignment of plastomes of the eight species using the LAGAN method. A cut-off of 70% similarity was used for the plot and the Y-scale represents the percent similarity ranging from 50–100%.
3.6. Phylogenetic Analysis

The BI and ML analyses based on both the whole plastomes and protein-coding genes recovered exactly the same tree topology (Figure 6). The phylogenetic trees displayed higher supports at all nodes that provided robust phylogenetic relationships of the selected taxa in Rutaceae. The results showed that the sampled species of *Phellodendron*, *Tetradium*, *Toddalia*, and *Zanthoxylum* were resolved as close relatives and grouped together as a distinct clade, which is consistent with previous phytochemical and phylogenetic studies of the "proto-Rutaceae" [13,14,21]. Moreover, our results showed that these four genera of "proto-Rutaceae" formed two subclades, which were sister to each other.

The first subclade consisted of *Phellodendron* and *Tetradium*. Although *Phellodendron* and *Tetradium* differ radically in fruit structure (the former is dehiscent follicle and the latter is fleshy drupe), Hartley (1981) [18] found that it was impossible to distinguish the two genera using vegetative or staminate material. Previous molecular studies also revealed that these two genera were closely related [18,21,23]. The sister-group relationship of *Phellodendron* and *Tetradium* was also supported by our plastomes data (PP/BS: 1/100). In this subclade, the two sampled species of *Tetradium*, *T. ruticarpum* (A.Juss.) T.G.Hartley, and *T. daniellii* (Benn.) T.G.Hartley, were well supported as sister species (PP/BS: 1/100). *P. chinense* and *P. amurense*, the only two species of *Phellodendron*, formed another strongly supported monophyletic group (PP/BS: 1/100). Unfortunately, as part of 'Proto-Rutaceae', *Fagaropsis* was not included in our analysis, but it most likely belongs to this subclade too. Inclusion of the plastomes data of this genus in future studies will help enhance our understanding of the 'proto-Rutaceae'. The second subclade was composed of *Toddalia* and *Zanthoxylum*. It was interesting that, in 'Proto-Rutaceae', although all four genera produce 1-BTIQ alkaloids, only *Toddalia* and *Zanthoxylum* synthesise anthranilate-derived alkaloids and coumarins, and there was a co-occurrence of furanocoumarins and furoquinolines in *Toddalia* and *Zanthoxylum*, but not in *Phellodendron* [13,15]. Thus, the proximity of these two genera was supported by both the chemical characteristics and the molecular data. In previous phylogenetic studies, *Toddalia* might be sister to *Zanthoxylum*, or nested with it [20,21]. Furthermore, Appelhans et al. (2018) [20] has proposed to change *T. asiatica* as *Z. asiaticum* (L.) Appelhans, Groppo & J.Wen, comb. nov. Our results based on plastomes data clearly showed that *T. asiatica* was deeply nested within *Zanthoxylum*. Therefore, we confirmed that *Toddalia* was part of the clade of *Zanthoxylum*, and suggested that the two genera should be merged. In Appelhans et al.’s (2018) [20] study, *Toddalia* was resolved as sister to the African and Malagasy species of *Zanthoxylum*, however, in another study it was found to be sister to a main Asian clade [48]. In our plastomes analyses, *T. asiatica* was not closely related to the African and Malagasy species of *Z. paniculatum* Balf.f. and *Z. madagascariense* Baker, but was sister to the clade formed by the Southwest Pacific species *Z. pinnatum* (J.R.Forst. & G.Forst.) Oliv. and the East Asian species of *Z. schinifolium* Siebold & Zucc., *Z. oxyphyllum* Edgew., and *Z. calcicola* Huang (PP/BS: 1/100).

The inferred phylogeny based on plastomes data also lent support to the division of *Euodia* sensu lato by Hartley (1981) [18]. *Melicope ptelefolia* (Champ. ex Benth.) T.G.Hartley, a South-East Asian species with trifoliolate (occasionally unifoliolate) leaves, was tradition-
ally regarded as belonging to *Euodia* [49]. In the present study, the results showed that *M. pteleifolia* was not close to *T. ruticarpum* and *T. daniellii*, both of them have odd-pinnately compound leaves and formerly belonged to *Euodia* [50], but was clustered with *A. pedunculata* (L.) Miq., a species that has unifoliolate leaves (PP/BS: 1/100). They formed a strongly supported sister group and appeared basal to the clade of the ‘proto-Rutaceae’.

The genera of the “proto-Rutaceae” are characterized by having 1-BTIQ alkaloids. Waterman (2007) [51] hypothesized that the “proto-Rutaceae” might represent an early branching clade within Rutaceae. However, the molecular studies showed that it was not the basal in the family [11,52,53]. In morphology, instead of fruit types, the phyllotaxis, habit and aculeate protuberances on branches are of taxonomic value in the “proto-Rutaceae” [20]. *Phellodendron*, *Tetradium* and *Fagaropsis* have opposite leaves and do not have any aculeus or spines on branches, while *Zanthoxylum* and *Toddalia* have alternate leaves and are aculeate [8]. Having established good molecular evidence for the clade ‘proto-Rutaceae’ as a natural group, the presence of 1-BTIQ alkaloids can be interpreted as the synapomorphy for these genera of the group. Considering that the “proto-Rutaceae” is not a formal taxon name, we supported Ling et al.’s (2009) [23] suggestion that it may be possible to group these five genera in a re-circumscribed tribe Zanthoxyleae Dumort. in the future.

![Figure 6. Phylogenetic relationships of 36 Rutaceae species constructed by MrBayes (BI) and maximum likelihood (ML) methods. (A) Phylogenetic tree based on the complete chloroplast genomes. (B) Phylogenetic tree based on the protein encoding genes. The numbers at each node show the posterior probabilities (PP) according to BI and bootstrap support (BS) values obtained by ML, respectively.](image)

4. Conclusions

In this study, we investigated plastome variations in eight species representing five genera of the subfamily Amyridoideae, Rutaceae. We found that the plastomes of these five genera are relatively conserved in terms of genome structure and gene content and order.
However, some plastomes, especially in *Melicope*, showed higher structural variations at the LSC/IR and SSC/IR borders. The codon preference, SSRs and sequence divergence of the plastomes were also analyzed, and three most divergent hotspots regions were found from the LSC region. In addition, we conducted a phylogenomic analysis of the "proto-Rutaceae" and related taxa based on plastomes data. Our phylogenetic analysis confirmed the group ‘proto-Rutaceae’, and supported the sister relationship of *Phellodendron* and *Tetradium* and the treatment of merging *Toddalia* into *Zanthoxylum*. This study revealed new insights into the plastome variations in Amyridoideae, and indicated that the plastomes data were sufficiently robust to explore implications of the phylogeny for the previous phylogenetic hypotheses.

**Supplementary Materials:** The following are available online at [https://www.mdpi.com/1999-4907/12/3/277/s1](https://www.mdpi.com/1999-4907/12/3/277/s1), Figure S1: Codon usage analyses of the plastomes. (A) Codon usage of *Phellodendron amurense* and *P. chinense*; (B) Codon usage of *Tetradium daniellii* and *T. ruticarpum*; (C) Codon usage of *Zanthoxylum calcicola* and *Z. piperitum*; (D) Codon usage of *Melicope pteleifolia* and *Toddalia asiatica*. Table S1: Information of three newly sequenced plastomes, Table S2: Plastomes obtained from GenBank in this study. Table S3: Codon usage of the chloroplast genome sequences, Table S4: Simple sequence repeats of eight plastomes.

**Author Contributions:** L.-C.Z. and W.-B.G. conceived and designed the research framework; K.S. and Q.-Y.L. prepared the samples and performed the experiments; K.S. and Q.-Y.L. analyzed the data and wrote the paper. L.-C.Z. made revisions to the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Science and Technology Basic Resources Investigation Program of China “Survey and Germplasm Conservation of Plant Species with Extremely Small Populations in South-west China” (Grant No. 2017FY100100).

**Data Availability Statement:** The data that support the findings of this study are openly available in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov) (accessed on 1 February 2021), reference number NC050949, MW542638, MW542637.

**Acknowledgments:** We sincerely thank Wenpan Dong for his help that greatly improved our manuscript.

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

**References**

1. Jansen, R.K. Plastid Genomes of seed plants. In *Advances in Photosynthesis and Respiration*; Bock, R., Knoopp, V., Eds.; Springer: Dordrecht, The Netherlands, 2012; pp. 103–126.
2. Daniell, H.; Lin, C.S.; Yu, M.; Chang, W.J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* 2016, 17, 134. [PubMed]
3. Raubeson, L.; Jansen, R. Chloroplast genomes of plants. In *Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants*, 1st ed.; Robert, J., Ed.; CABI: London, UK, 2005; pp. 53–76.
4. Ma, P.F.; Zhang, Y.X.; Zeng, C.X.; Guo, Z.H.; Li, D.Z. Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable Bamboo tribe Arundinarieae (Poaceae). *Syst. Biol.* 2014, 63, 933–950. [CrossRef] [PubMed]
5. Foster, C.S.P.; Henwood, M.J.; Ho, S.Y.W. Plastome sequences and exploration of tree-space help to resolve the phylogeny of riceflowers (Thymelaeaceae:Pimelea). *Mol. Phylogenet. Evol.* 2018, 127, 156–157. [CrossRef]
6. Song, Y.; Yu, W.B.; Tan, Y.H.; Jin, J.J.; Wang, B.; Yang, J.B.; Liu, B.; Corlett, R.T. Plastid phylogenomics improve phylogenetic resolution in the Lauraceae. *J. Syst. Evol.* 2020, 58, 423–439. [CrossRef]
7. Yang, X.Y.; Wang, Z.F.; Luo, W.C.; Guo, X.Y.; Zhang, C.H.; Liu, J.Q.; Ren, G.P. Plastomes of Betulaceae and phylogenetic implications. *J. Syst. Evol.* 2019, 57. [CrossRef]
8. Kubitzki, K.; Kallunki, J.; Duretto, M.; Wilson, P. Rutaceae. In *The Families and Genera of Vascular Plants*; Kubitzki, K., Ed.; Springer: Berlin, Germany, 2011; Volume 10, pp. 276–356.
9. Groppo, M.; Kallunki, J.A.; Pirani, J.R.; Antonelli, A. Chilean *Pitavia* more closely related to Oceania and Old World Rutaceae than to Neotropical groups: Evidence from two cpDNA non-coding regions, with a new subfamilial classification of the family. *PhytoKeys* 2012, 19, 9–29. [CrossRef]
10. Morton, C.M.; Telmer, C. New subfamily classification for the Rutaceae. *Ann. Mo. Bot. Gard.* 2014, 99, 620–641. [CrossRef]
11. Engler, A. Rutaceae. In *Die natürlichen Pflanzenfamilien—Band 19a*; Engler, A., Harms, H., Eds.; Wilhelm Engelmann: Leipzig, Germany, 1931; pp. 187–359.
12. Waterman, P.G. Alkaloids of the Rutaceae: Their distribution and systematic significance. *Biochem. Syst. Ecol.* 1975, 3, 149–180.
13. Waterman, P.G.; Khaldi, S.A. The biochemical systematics of Fagaropsis angolensis and its significance in the Rutales. *Biochem. Syst. Ecol.* 1981, 9, 45–51. [CrossRef]

14. Waterman, P. Phylogenetic implications of the distribution of secondary metabolites in the Rutales. In *Chemistry and Chemical Taxonomy of the Rutales*; Waterman, P., Grundon, M., Eds.; Academic Press: London, UK, 1983; pp. 377–400.

15. Ng, K.M.; But, P.P.H.; Gray, A.I.; Hartley, T.G.; Kong, Y.C.; Waterman, P.G. The biochemical systematics of *Tetradium, Euodia* and *Melicope* and their significance in the Rutaceae. *Biochem. Syst. Ecol.* 1987, 15, 587–593. [CrossRef]

16. Ma, J.S.; Cao, W.; Liu, Q.R.; Yu, M.; Han, L.J. A revision of *Phellodendron* (Rutaceae). *Edinb. J. Bot.* 2006, 63, 131–151. [CrossRef]

17. Engler, A. *Cneoraceae, Rutaceae*. In *Die natürlichem Pflanzenfamilien*, 2nd ed.; Engler, A., Prantl, K., Eds.; W. Engelmann: Leipzig, Germany, 1931; Volume Band 19a, pp. 184–359.

18. Hartley, T.G. A revision of the genus *Tetradium* (Rutaceae). *Gard. Bull. Singap.* 1981, 34, 91–131.

19. Hartley, T.G. On the taxonomy and biogeography of *Euodia* and *Melicope* (Rutaceae). *Allertonia* 2001, 8, 1–341.

20. Appelhans, M.S.; Reichelt, N.; Groppo, M.; Paetzold, C.; Wen, J. Phylogeny and biogeography of the pantropical genus *Zanthoxylum* and its closest relatives in the proto-Rutaceae group (Rutaceae). *Mol. Phylogenet. Evol.* 2018, 126, 31–44. [PubMed]

21. Poon, W.S.; Shaw, P.S.; Simmons, M.P.; But, P.P.H. Congruence of molecular, morphological, and biochemical profiles in Rutaceae: A cladistic analysis of the subfamilies. *Syst. Bot.* 2007, 32, 837–846. [CrossRef]

22. Appelhans, M.S.; Wen, J.; Wagner, W.L. A molecular phylogeny of *Acronychia, Euodia, Melicope* and relatives (Rutaceae) reveals polyphyletic genera and key innovations for species richness. *Mol. Phylogenet. Evol.* 2014, 79, 54–68. [CrossRef]

23. Ling, K.H.; Wang, Y.; Poon, W.S.; Shaw, P.C.; But, P.P.H. The relationship of *Fagaropsis* and *Luessa* in Rutaceae. *Taiwania* 2009, 54, 335–342.

24. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochem. Bull.* 1987, 19, 11–15.

25. 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]

26. Huang, D.I.; Cronk, Q.C. Plannt: A command-line application for annotating plastome sequences. *Appl. Plant Sci.* 2015, 3, 150026. [CrossRef]

27. Greiner, S.; Lehrkamp, 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]

28. Shen, Z.N.; Gan, Z.M.; Zhang, F.; Yi, X.Y.; Zhang, J.Z.; Wan, X.H. Analysis of codon usage patterns in citrus based on coding sequence data. *BMC Genom.* 2020, 21, 234. [CrossRef]

29. Sau, K.; Gupta, S.K.; Sau, S.; Mandal, S.C.; Ghosh, T.C. Factors influencing synonymous codon and amino acid usage biases in Mimiviruses. *Biosystems* 2006, 85, 107–113. [CrossRef]

30. Beier, S.; Guth, S.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* 2017, 33, 2583–2585. [CrossRef] [PubMed]

31. Kurtz, S.; Thié, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* 2017, 33, 2583–2585. [CrossRef] [PubMed]

32. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 2018, 34, 3030–3031. [CrossRef] [PubMed]

33. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 2013, 30, 772–780. [CrossRef] [PubMed]

34. Frazer, K.A.; Pachter, L.; Poliaikov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* 2004, 32, W237–W279. [CrossRef] [PubMed]

35. Brudno, M.; Do, C.B.; Cooper, G.M.; Kim, M.F.; Davydov, E.; Green, E.D.; Sidow, A.; Batzoglou, S. LAGAN and Multi-LAGAN: Efficient tools for large-scale multiple alignment of genomic DNA outline of algorithms. *Genome Res.* 2003, 13, 721–731. [CrossRef] [PubMed]

36. 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 datasets. *Mol. Biol. Evol.* 2017, 34, 3299–3302. [CrossRef] [PubMed]

37. Ronquist, F.; Teslenko, M.; Mark, P.V.D.; 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.* 2006, 51, 399–402. [CrossRef] [PubMed]

38. Xu, S.R.; Zhang, Y.Y.; Liu, F.; Tian, N.; Pan, D.M.; Bei, X.J.; Cheng, C.Z. Characterization of the complete plastid genome of the Hongkong kumquat (*Fortunella hisdis Swingle*). *Mitochondrial DNA Part B* 2019, 4, 2612–2613. [CrossRef] [PubMed]

39. Wu, J.W.; Liu, F.; Tian, N.; Liu, J.P.; Shi, X.B.; Bei, X.J.; Cheng, C.Z. Characterisation of the complete plastome genome of *Fortunella crassifolia* and phylogenetic relationships. *Mitochondrial DNA Part B* 2019, 4, 3538–3539. [CrossRef]

40. Yoo, Y.H.; Oh, C.J.; Shin, S.C.; Seo, S.; Kim, H.B. Complete plastid genome sequence of a medicinal landrace citrus *linkyooll* (*Citrus sunki* Hort. ex Tanaka) in Jeju Island, Korea. *Mitochondrial DNA Part B* 2020, 5, 3719–3720. [CrossRef] [PubMed]

41. Angellotti, M.C.; Bhuiyan, S.B.; Chen, G.R.; Wan, X.F. CodonO: Codon usage bias analysis within and across genomes. *Nucleic Acids Res.* 2007, 35, W132–W136. [CrossRef] [PubMed]
43. Wright, F. The effective number of codons' used in a gene. *Gene* **1990**, *87*, 23–29. [CrossRef]

44. Sueoka, N. Near homogeneity of PR2-Bias fingerprints in the human genome and their implications in phylogenetic analyses. *J. Mol. Evol.* **2001**, *53*, 469–476. [CrossRef]

45. Xu, C.; Dong, J.; Tong, C.F.; Gong, X.D.; Wen, Q.; Qiang, Z.G. Analysis of synonymous codon usage patterns in seven different *Citrus* species. *Evol. Bioinform.* **2013**, *9*, 215–228. [CrossRef]

46. Yu, X.Y.; Zuo, L.H.; Lu, D.D.; Lu, B.; Yang, M.S.; Wang, J.M. Comparative analysis of chloroplast genomes of five *Robinia* species: Genome comparative and evolution analysis. *Gene* **2019**, *689*, 141–151. [CrossRef]

47. Cavagnaro, P.F.; Senalik, D.A.; Yang, L.M.; Simon, P.W.; Harkins, T.T.; Kodira, C.D.; Huang, S.W.; Weng, Y.Q. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genom.* **2010**, *11*, 569. [CrossRef]

48. Appelhans, M.S.; Wen, J. Phylogenetic placement of *Ivodea* and biogeographic affinities of Malagasy Rutaceae. *Plant Syst. Evol.* **2020**, *306*, 7. [CrossRef]

49. Huang, C.C. Rutaceae. In *Flora Reipublicae Popularis Sinicae*; Huang, C., Ed.; Science Press: Beijing, China, 1997; Volume 43, pp. 1–250.

50. Zhang, D.; Hartley, T.; Mabberley, D. Rutaceae. In *Flora of China*; Zhang, D., Hartley, T., Mabberley, D., Eds.; Science Press and St. Louis Missouri Botanical Garden Press: Beijing, China, 2008; Volume 11, pp. 51–97.

51. Waterman, P.G. The current status of chemical systematics. *Phytochemistry* **2007**, *68*, 2896–2903. [CrossRef] [PubMed]

52. Chase, M.W.; Morton, C.M.; Kallunki, J.A. Phylogenetic relationships of Rutaceae: A cladistics analysis of the subfamilies using evidence from rbcL and atpB sequence variation. *Am. J. Bot.* **1999**, *86*, 1191–1999. [CrossRef]

53. Groppo, M.; Pirani, J.R.; Salatino, M.L.F.; Blanco, S.R.; Kallunki, J.A. Phylogeny of Rutaceae based on two noncoding regions from cpDNA. *Am. J. Bot.* **2008**, *95*, 985–1005. [CrossRef]