Chloroplast Genomic Variation in *Euonymus maackii* Rupr. and Its Differentiation Time in *Euonymus*

Yu-Cheng Wang 1, Hao-Yang Zhou 2 and Xiu-Qun Liu 2,*

1 College of Informatics, Huazhong Agricultural University, Wuhan 430070, China; wangyucheng@webmail.hzau.edu.cn
2 Key Laboratory of Horticultural Plant Biology (Ministry of Education), College of Horticulture and Forestry Science, Huazhong Agricultural University, Wuhan 430070, China; 305zhy@webmail.hzau.edu.cn
* Correspondence: liuxiuqun@mail.hzau.edu.cn

**Abstract:** *Euonymus maackii* Rupr. is a small deciduous tree belonging to family Celastraceae. It is an important ornamental tree and a potential medicinal plant resource. Here, we assembled and annotated the chloroplast (cp) genome of *E. maackii*. By combining this genome with seven available cp genomes from *Euonymus* species, we performed plastome variation analysis of *E. maackii* and *Euonymus*. Furthermore, we reconstructed a phylogenetic tree and estimated the differentiation time of *E. maackii*. The newly assembled cp genome of *E. maackii* was 157,551 bp in size and had a typical quadripartite structure, which consisted of one large single-copy (LSC 86,524 bp) region, one small single-copy (SSC 18,337 bp) region, and a pair of inverted repeat regions (26,345 bp). A total of 652 single nucleotide polymorphisms (SNPs) and 65 insertions/deletions (indels) were detected between the two cp genomes of *E. maackii*, with overall genetic variation of 4.1 SNPs per kb or a π value of 0.00443, reflecting a high level of intraspecific variation. Some coding and noncoding regions with higher variation were identified, including *trnV-UAC*, *petN*, *ycf1-ndhF*, *trnM^CAU-aptE*, *rpl2-rpl23*, *psbZ-trnC^GCC*, *trnY^GLA-trnE-UUC*, *trnW^CCA-trnP^UGG*, *rps16-trnQ^UGU*, and *psbC-trnS^UGA*. The hypervariable coding and noncoding regions in *E. maackii* were not the same as those in *Euonymus*. The phylogenetic tree and divergence time based on the whole cp genomes showed that the seven *Euonymus* species formed a clade, which was sister to that formed with *Catha edulis* and *Maytenus guangxiensis*, and they separated 24.74 million years ago. *E. maackii* and *E. hamiltonianus* were most closely related, having separated from each other only approximately 2.68 million years ago. Our study provides important genetic information for further studies of *E. maackii*, such as studies on its phylogeography, population genetics and molecular ecology, and provides new insights into the evolution of the cp genome in *Euonymus*.

**Keywords:** *Euonymus maackii*; intraspecific variation; chloroplast genome; phylogenetic tree; divergence time

1. Introduction

*Euonymus maackii* Rupr. (synonym: *Euonymus bungeanus*) is a small deciduous tree belonging to the family Celastraceae. It is native to northern China and has the characteristics of cold resistance, drought resistance, salt and alkali resistance, few diseases and insect pests, and strong germination [1,2]; therefore, it is widely distributed in most areas from Heilongjiang Province to the Yangtze River Basin and further extends to a wider area, including the Ussuri Region, southern Siberia, and the Korean Peninsula [3]. *E. maackii* has a beautiful shape due to its long and soft branches with curved drooping. In addition, in autumn, its leaves turn red, and its dehiscent fruit turns orange; therefore, it has high ornamental value and has been cultivated as a horticultural tree for more than 20 years in China [2,4,5].

Furthermore, a series of cultivars of *E. maackii* with colorful leaves have been obtained through artificial cultivation. In addition to the common characteristics of *E. maackii*,
these cultivars also have excellent ornamental characteristics, such as gold leaves, gold branches, and red branches. The planting of these outstanding cultivars has widened in scope [6]. *E. maackii* plants are also a potential medicinal plant resource. The evonine and polysaccharides extracted from the seeds have antitumor activity [7,8], and dihydro-β-agarofuran-type sesquiterpenes purified from the seeds significantly increases the cell viability of Aβ25–35-treated SH-SY5Y cells, with effects similar to those of epigallocatechin gallate [9]. In addition, the seed oil could be used as a nonedible feedstock for biodiesel [3].

Chloroplasts are semiautonomous organelles in plant cells. They have their own genome, which mainly encodes genes related to photosynthesis, transcription and translation, as well as some tRNA and rRNA genes and a few genes with other biological functions [10]. The chloroplast (cp) genome of most angiosperms (including *E. maackii*) is maternally inherited, and thus only genetic information from the female parent can be transmitted to the offspring [11]. It is precisely because of the uniparental inheritance of the cp genome that its DNA generally does not recombine.

In addition, it has relatively high genetic variation; therefore, it is often employed to infer the genetic relationships of species, estimate their differentiation time [12–15], and sometimes identify germplasm [16]. For example, Zeb et al. compared the cp genomes of twenty-four Pinus species and identified a total of thirteen divergent hotspot regions. Moreover, the study revealed phylogenetic relationships and molecular dating information [17]. Kim et al. compared five complete *Vaccinium* cp genomes to identify regions that can be used to distinguish them [18].

Other studies detected cp genome variation within species to investigate the origin of cultivated varieties, estimate genetic diversity, infer population history [19–21], etc. For example, Chen et al., based on cp genome variation, determined the phylogenetic position of Manokami larch within the Japanese larch group, detected intraspecific variation, and determined candidate cpDNA markers [22]. Feng et al. calculated the genetic polymorphism of cp genomes in wild and cultivated rubber trees. Wild rubber trees possessed more variable sites and ~2.8-fold higher levels of nucleotide variation compared with cultivated rubber trees [23].

Different structural regions of cp genomes significantly vary in their genetic diversity. Even within the same coding region or noncoding region, polymorphisms differ among plant groups [24]. In this study, by comparing two cp genomes of *E. maackii*, relatively highly variable regions were detected, which provided genetic information for research on genetic diversity, population history, and variety identification at the species level. In addition, the phylogenetic tree of *Euonymus* was reconstructed based on entire cp genomes, and the divergence times of the species were estimated.

2. Materials and Methods

2.1. Material Collection and Whole-Genome Sequencing

Fresh leaves of *E. maackii* were collected from Henan Agricultural University (34°47′7″ N, 113°39′35″ E), Zhengzhou city, China. The leaf samples were dried with silica gel and stored at −20 °C until DNA extraction. The specimen was deposited in the Herbarium of Henan Agricultural University (EUPEUM202001). The total genomic DNA was extracted from the dried leaves using the modified cetyl trimethyl ammonium bromide (CTAB) method [25]. The purified DNA was then fragmented and used to construct short-insert libraries (~270 bp), which were then sequenced via the Illumina HiSeq double terminal sequencing method.

2.2. De novo Assembly and Annotation

After the raw data were filtered to remove reads with low sequencing quality values, all remaining sequences were assembled into circular genomes by NOVOPlasty-v2.7.2 using the cp genome of *Euonymus szechuanensis* as a reference [26]. GapCloser-1.12 was employed to extend sequences and fill gaps [27]. Annotations and adjustment of the genes were conducted using the program Geneious Prime against the reference genome.
Finally, a circular map of the cp genome was drawn using OGDRAW 1.3.1 (http://ogdraw.mpimp-golm.mpg.de, accessed on 16 March 2021). The newly assembled cp genome was deposited in GenBank under accession number MW771518 (https://www.ncbi.nlm.nih.gov/nuccore/MW771518.1/, accessed on 18 March 2021).

2.3. Comparative Analysis

We downloaded the cp genomes of seven species of *Euonymus* from the NCBI, including the cp genome of another *E. maackii* individual. First, we compared the two sets of cp genomes in Geneious Prime. One set contained two cp genomes of *E. maackii*, one newly assembled and another downloaded from the NCBI. The other set included the newly assembled cp genome and those of six other *Euonymus* species (*E. hamiltonianus, E. phellomanus, E. fortunei, E. japonicus, E. szechuanensis*, and *E. schensianus*) downloaded from the NCBI. Subsequently, DnaSP 5 was used to estimate the genetic diversity (π) of the entire cp genome, four regions (large single-copy (LSC), small single-copy (SSC), inverted repeat a (IRA) and IRb), all coding regions and all noncoding regions. In addition, the indels were counted, and their distribution patterns were analyzed in two sets of cp genomes.

2.4. Phylogenetic Analysis and Differentiation Time Estimation

To detect the origin and differentiation of *E. maackii*, we constructed phylogenetic trees of 11 species based on entire cp genomes, including the newly assembled cp genome and those of 10 other species downloaded from the NCBI. Seven of the eleven species belong to the genus *Euonymus*, and the other four were *Catha edulis, Maytenus guangxiensis, Monimopetalum chinense*, and *Parnassia trinervis* in Celastraceae. The complete cp genomes of the 11 species were aligned by MAFFT in Geneious Prime [28], and then phylogenetic trees were constructed by the maximum likelihood (ML) and Bayesian inference (BI) methods based on their cp genome sequences using the four species outside the genus *Euonymus* as the outgroup.

ML analyses were performed using RAxML-HPC BlackBox v. 8.2.10 with the GTR+G model and 1000 bootstrap replicates on the CIPRES Science Gateway website (https://www.phylo.org/ accessed on 1 February 2022) [29,30]. BI was performed with MrBayes 3.2.6 [31] with the following settings: Markov chain Monte Carlo (MCMC) simulations for 1,000,000 generations with four incrementally heated chains, starting from random trees and sampling one out of every 1000 generations. The first 25% of the trees were regarded as burn-in. The ML tree and BI tree were visualized using FigTree 1.4.3 [32].

In this study, the calibration time was obtained from the TimeTree database as a normal prior to restrain the age of the nodes, such as 79 Ma for the root of the Celastraceae crown, 53 Ma between *Euonymus* and *Parnassia*, and 2.5 Ma between *Catha* and *Couma* (http://www.timetree.org/ accessed on 1 February 2022). Molecular clock dating was performed using BEAST with the GTR+G substitution model, along with a relaxed lognormal molecular clock model. MCMC analyses were run for $1 \times 10^7$ generations with a sampling frequency of 1000 combined after a 10% burn-in.

3. Results

3.1. Characterization of Plastome Features

The cp genome assembled (*E. maackii* 1) in this study was 157,551 bp in length and had a typical quadripartite structure, which consisted of one LSC region (86,525 bp), one SSC region (18,337 bp), and a pair of IR regions (26,345 bp) (Figure 1). Compared with another previously reported cp genome (*E. maackii* 2), *E. maackii* 1 was 691 bp longer (Table 1). The size variation was largely due to the increase in LSC length, with that of *E. maackii* 1 being 699 bp longer than that of *E. maackii* 2. The GC contents of the two cp genomes were similar, at 37.2% and 37.3%.
Figure 1. Map of the chloroplast genome of *Euonymus maackii* Rupr.

Table 1. Comparison of the chloroplast genome features of *Euonymus* species.

|                  | *Euonymus maackii 1* | *Euonymus maackii 2* | *Euonymus japonicus* | *Euonymus schensianus* | *Euonymus hamiltonianus* | *Euonymus szechuanensis* | *Euonymus phellomanus* | *Euonymus fortunei* |
|------------------|----------------------|----------------------|----------------------|------------------------|--------------------------|--------------------------|-----------------------|---------------------|
| Accession number | MW771518             | MW288091             | NC_026067            | NC_026019              | NC_026058                | NC_047463.1              | MW288092             | MW288090            |
| Total plastome size (bp) | 157,551             | 156,860              | 157,637              | 157,702                | 157,360                  | 157,465                  | 157,543               | 157,611             |
| LSC (bp)         | 86,525               | 85,826               | 85,941               | 86,026                 | 86,399                   | 86,261                   | 86,299                | 85,892              |
| IR (bp)          | 26,345               | 26,328               | 26,678               | 26,484                 | 26,322                   | 26,366                   | 26,354                | 26,700              |
| SSC (bp)         | 18,337               | 18,378               | 18,340               | 18,708                 | 18,317                   | 18,472                   | 18,472                | 18,319              |
| Total number of genes | 131                 | 131                  | 130                  | 130                    | 131                      | 133                      | 133                   | 130                 |
| CDSs             | 86                   | 86                   | 85                   | 85                     | 86                       | 88                       | 88                    | 85                  |
| tRNAs            | 37                   | 37                   | 37                   | 37                     | 37                       | 37                       | 37                    | 37                  |
| rRNAs            | 8                    | 8                    | 8                    | 8                      | 8                        | 8                        | 8                     | 8                   |
| GC content (%)   | 37.2                 | 37.3                 | 37.3                 | 37.3                   | 37.3                     | 37.3                     | 37.3                  | 37.3                |

Similar to *E. maackii* 2, the newly assembled cp genome encoded a total of 131 genes, including 86 protein-coding genes, 37 tRNA genes, and eight rRNA genes (Table 2). Among these genes, seven protein-coding genes, six tRNA genes, and four rRNA genes were
duplicated in the IR regions. In addition, seventeen genes contained introns, fourteen of which had a single intron, and the other three genes (ycf3, rps12, and clpP3) had two introns (Table 2).

### Table 2. Genes in the chloroplast genomes of the two *Euonymus maackii* Rupr. species.

| Category of Genes       | Group of Genes | Names of Genes       |
|-------------------------|----------------|----------------------|
| Photosynthesis-related genes |                |                      |
| photosystem I           |                |                      |
|                         | psaA           | psaB                 |
|                         | ycf3 **        | ycf4                 |
| photosystem II          |                |                      |
|                         | psbA           | psbB                 |
|                         | psbF           | psbH                 |
|                         | psbL           | psbM                 |
|                         | psbC           | psbI                 |
|                         | psbD           | psbJ                 |
|                         | psbE           | psbK                 |
|                         | psbT           | psbZ                 |
| NADPH dehydrogenase     |                |                      |
|                         | ndhA *         | ndhB **              |
|                         | ndhE           | ndhF                 |
|                         | ndhK           | ndhG                 |
| ATP synthase            |                |                      |
|                         | atpA           | atpB                 |
|                         | atpI           | atpE                 |
|                         | atpF *         | atpH                 |
| Rubisco                 |                |                      |
|                         | rbcL           |                      |
| cytochrome b/f complex  |                |                      |
|                         | petA           | petB **              |
|                         | petN           | petD                 |
|                         |               | petG                 |
|                         |               | petL                 |
| cytochrome c synthesis  |                |                      |
|                         |               | ccsA                 |
| Transcription- and translation-related genes |             |                      |
| ribosomal proteins      |                |                      |
|                         | rpl2 *         | rpl14                |
|                         | rpl23          | rpl16 **             |
|                         | rps3           | rpl33                |
|                         | rps12 **       | rpl36                |
|                         | rps19          | rps8                 |
|                         |               | rps11                |
|                         |               | rps18                |
| transcription           |                |                      |
|                         | rpoA           | rpoB                 |
|                         |               | rpoC1 **             |
|                         |               | rpoC2                |
| translation initiation factor |            |                      |
|                         | infA           |                      |
| RNA genes               |                |                      |
| transfer RNA            |                |                      |
|                         | trnA-UGC *     | trnC-GCA             |
|                         | trnF-M-CALI    | trnG-GCC             |
|                         | trnL-GAU *     | trnG-GCC             |
|                         | trnM-CALI      | trnG-GCC             |
|                         | trnR-UCA       | trnG-GCC             |
|                         | trn-T-UCA      | trnV-GAC             |
|                         |               |                      |
| ribosomal RNA           |                |                      |
|                         | trn4.5         | trn5                 |
|                         |               | trn16                |
|                         |               | trn23                |
| Other genes             |                |                      |
| RNA processing          |                |                      |
|                         | matK           |                      |
| fatty acid synthesis    |                |                      |
|                         | accD           |                      |
| carbon metabolism       |                |                      |
|                         | cemA           |                      |
| proteolysis             |                |                      |
|                         | clpP **        |                      |
| Genes of unknown function |              |                      |
|                         | ycf1           | ycf2                 |

* Gene containing a single intron, ** Gene containing two introns.

### 3.2. Sequence Variation

A total of 652 SNPs were detected between the two cp genomes of *E. maackii*, with overall genetic variation of 4.1 SNPs per kb or a π value of 0.00443, which is one-half of the plastome diversity (0.00887) of *Euonymus* (Table 3). The IR region in the *E. maackii* plastomes was very conserved, and its genetic diversity was only 0.00086, which was 0.15 times that of the LSC region and 0.09 times that of the SSC region.
Table 3. The nucleotide diversity (π) of different chloroplast genome regions in Euonymus maackii Rupr. And Euonymus.

| Group            | Genome     | Structural Region | Noncoding Region | Coding Region |
|------------------|------------|-------------------|------------------|--------------|
| Euonymus maackii | 0.00443    | 0.00562           | 0.00086          | 0.00914      |
|                  | 0.00299    |                   |                  |              |
| Euonymus         | 0.00887    | 0.01131           | 0.00379          | 0.01139      |
|                  | 0.01149    |                   |                  | 0.00619      |

The π value of the IR region was 0.00379 in the genus Euonymus, which was 0.34 times that of the LSC region and 0.33 times that of the SSC region. The genetic diversity in non-coding regions was significantly higher than that in coding regions both at the species level and in the genus Euonymus, and the ratios of polymorphisms between the noncoding region and the coding region were similar, with that of the former being 2.03 (0.00607/0.00299) and that of the latter being 1.86 (0.01149/0.00619).

Among the 652 SNPs, 102 were distributed in 57 coding regions (exons) across the cp genome of E. maackii. The genetic diversity (π value) of these coding regions ranged from 0.00036 to 0.02703, with an average of 0.00438 (Figure 2). The coding region with the highest genetic diversity was trnV-UAC, with a π value of 0.02703, followed by petN, with π = 0.01111, and the other eight coding regions in the top ten in terms of genetic polymorphisms were psaI, ycf1, psaJ, rps2, rpoC1, ndhE, and ycf3. Their genetic diversity was greater than or equal to 0.00654 (Table S1, Figure 2).

However, the hypervariable regions in Euonymus were very different from those in E. maackii. Only ycf1 and trnV-UAC were shared by the two groups, while the other eight hypervariable genes were different (Table S1, Figures 2 and 3). In Euonymus, the coding region with the highest genetic diversity was ycf1, which had a π value of 0.02074, and the other nine highly variable coding regions had a genetic diversity that varied from 0.01253 to 0.01727 (Table S1, Figure 3).

Among the 652 SNPs, 450 were located in 96 noncoding regions (introns and intergenic regions) in the E. maackii plastome. The nucleotide diversity (π value) of the noncoding regions ranged from 0.00105 to 0.14286, with an average of 0.01274 (Figure 4). Fifteen noncoding regions with high genetic diversity were identified: ycf1-ndhF was the most diverse, with a π value of 0.14286, followed by trnM^{CAU}-atpE, with a π value of 0.09375.
Figure 3. The nucleotide diversity ($\pi$) of coding regions among the seven Euonymus species.

Figure 4. Nucleotide diversity ($\pi$) of noncoding regions between two Euonymus maackii Rupr. individuals.

Among the top ten genetic polymorphisms were rpl2-rpl23, psbZ-trnG-GCC, trnY-GUA, trnE-UUC, trnW-CCA-trnF-UUG, rps16-trnQ-UUG, psbC-trns-LCA, ccsA-ndhD, and trnD-GUC-trnY-GUA (Table S2). Seven of fifteen highly variable regions in E. maackii were found to be highly variable in Euonymus: trnM-CAI-atpE, psbZ-trnG-GCC, trnY-GUA-trnE-UUC, trnW-CCA, trnB-UUG, ccsA-ndhD, trnH-UUG-psbA, and ndhC-trnV-UAC. The remaining eight highly variable regions in E. maackii were different from those in the genus Euonymus (Table S2). Within Euonymus, the 15 noncoding regions with the highest polymorphism were all intergenic regions, and their genetic diversity ranged from 0.02859 to 0.07804, with an average of 0.02364, which was significantly higher than the average in E. maackii (Table S2, Figure 5).

Figure 5. The nucleotide diversity ($\pi$) of noncoding regions among the seven Euonymus species.

In addition to nucleotide mutations, insertions/deletions (indels) are another common type of genetic variation in cp genomes. Sixty-five indels were detected in the two cp
genomes of *E. maackii*, with a total indel density of 0.41 per kb. Fifty-seven of these indels occurred in noncoding regions, and the remaining eight indels were located in coding regions. The vast majority (58 indels, 89.2%) of the indels were in the LSC region, with a density of 0.67 per kb. Three (4.6%) were in the IR region, and four (6.2%) were in the SSC region, with densities of 0.11 per kb and 0.27 per kb, respectively.

Among these indels, 11 were 1 bp long, two were 2 bp long, three were 3 bp long, and all of the others were 4 or more bp long. The longest indel occurred in the noncoding region *trnS-GCU-trnG-UCC* with a difference of 101 bp, followed by an indel of 68 bp in *rps11-rpl36*. In the coding regions, the longest indels appeared in two genes, *rpoC2* and *YCF1*, both of which had an indel of 21 bp. The bp numbers of the indels were multiples of three, which did not cause a change in the subsequent reading frame.

3.3. Phylogenetic Relationships and Divergence Time Estimation

The ML and BI phylogenetic trees constructed based on the whole cp genomes had the same topological structure. Phylogenetic trees and divergence time estimation showed that seven species of the genus *Euonymus* formed a single clade (*Euonymus* clade), which was sister to that formed with *C. edulis* and *M. guangxiensis*, and they separated 24.74 million years ago (Figures 6 and 7). In the *Euonymus* clade, *E. hamiltonianus*, *E. maackii* and *E. phellomanus* formed one subclade (I); the other four species, *E. fortunei*, *E. japonica*, *E. szechuanensis*, and *E. schensianus*, clustered into another subclade (II); and the two subclades diverged approximately 14.42 million years ago.

---

![Phylogenetic tree](image_url)

**Figure 6.** Phylogenetic tree based on the whole plastome. The numbers to the left of the slashes on the branches show the bootstrap values obtained by maximum likelihood analysis, and those on the right show the posterior probabilities according to Bayesian inference.
Figure 7. Estimation of divergence time based on the whole chloroplast genome.

In subclade I, *E. phellomanus* was the first to diverge, splitting from the other two species 12.22 million years ago. *E. maackii* and *E. hamiltonianus* were the most closely related, having separated from each other only approximately 2.68 million years ago. In subclade II, *E. fortunei* and *E. japonicas* were the latest diverging species in the genus *Euonymus*, splitting approximately 0.24 million years ago. They clustered on a branch sister to that formed with *E. szechuanensis* and *E. schensianus*. The phylogenetic trees reconstructed here exhibited a high level of support, with a 100% bootstrap value and 1 as the posterior probability for each node, indicating that the phylogenetic relationships revealed in this study were reliable.

4. Discussion

The size of the newly assembled cp genome of *E. maackii* is consistent with that of other species in the genus *Euonymus*. In previous studies, *E. schensianus* had the largest cp genome (157,702 bp) in *Euonymus*, while *E. hamiltonianus* had the smallest (157,360 bp), and the newly assembled genome was 157,551 bp, which is approximately the median size of these cp genomes [33]. In addition, the new cp genome is similar to the plastomes of others in terms of gene content, order, and orientation. At the same time, our results revealed that there was a high level of cp genome variation in *E. maackii*. The two main types of variation were nucleotide mutations and indels.

4.1. Intraspecific Nucleotide Mutations and Highly Variable Regions

In most species, nucleotide mutations are a common type of DNA variation in the cp genome [34,35]. There were abundant nucleotide mutations (652 SNPs) in the cp genome of *E. maackii*, and thus it had a high level of genetic polymorphism, with an overall genetic variation of 4.1 SNPs per kb or a π value of 0.00443. This level of intraspecific sequence variation is smaller than that observed in a few plant species, such as *Selaginella tamariscina* (1.213 SNPs, 9.6 SNPs per kb) [21], *Goodyera achlechtendaliana* (740 SNPs, 4.8 SNPs per kb) [22].

[Diagram of phylogenetic tree with node labels for each species and divergence times]

**Figure 7.** Estimation of divergence time based on the whole chloroplast genome.

In subclade I, *E. phellomanus* was the first to diverge, splitting from the other two species 12.22 million years ago. *E. maackii* and *E. hamiltonianus* were the most closely related, having separated from each other only approximately 2.68 million years ago. In subclade II, *E. fortunei* and *E. japonicas* were the latest diverging species in the genus *Euonymus*, splitting approximately 0.24 million years ago. They clustered on a branch sister to that formed with *E. szechuanensis* and *E. schensianus*. The phylogenetic trees reconstructed here exhibited a high level of support, with a 100% bootstrap value and 1 as the posterior probability for each node, indicating that the phylogenetic relationships revealed in this study were reliable.

4. Discussion

The size of the newly assembled cp genome of *E. maackii* is consistent with that of other species in the genus *Euonymus*. In previous studies, *E. schensianus* had the largest cp genome (157,702 bp) in *Euonymus*, while *E. hamiltonianus* had the smallest (157,360 bp), and the newly assembled genome was 157,551 bp, which is approximately the median size of these cp genomes [33]. In addition, the new cp genome is similar to the plastomes of others in terms of gene content, order, and orientation. At the same time, our results revealed that there was a high level of cp genome variation in *E. maackii*. The two main types of variation were nucleotide mutations and indels.

4.1. Intraspecific Nucleotide Mutations and Highly Variable Regions

In most species, nucleotide mutations are a common type of DNA variation in the cp genome [34,35]. There were abundant nucleotide mutations (652 SNPs) in the cp genome of *E. maackii*, and thus it had a high level of genetic polymorphism, with an overall genetic variation of 4.1 SNPs per kb or a π value of 0.00443. This level of intraspecific sequence variation is smaller than that observed in a few plant species, such as *Selaginella tamariscina* (1.213 SNPs, 9.6 SNPs per kb) [21], *Goodyera achlechtendaliana* (740 SNPs, 4.8 SNPs per kb) [22].
per kb) [21] and *Hevea brasiliensis* (725 SNPs, 4.5 SNPs per kb) [23] but larger than those of most other plant species.

For example, *Macadamia integrifolia* is a tree nut crop in Australia, and 407 SNPs were identified across its cp genomes, for an average intraspecific SNP density of 3.8 per kb [19]. A set of 298 SNPs were detected in the cp genome of *Brachypodium distachyon*, a gramineous weed, and its SNP density was only 2.2 per kb [36]. A total of 258 and 139 SNPs were found in *Agrimonia pilosa* and *Tagetes erecta*, with 1.7 and 0.91 SNP densities, respectively [37, 38]. In other species, such as *Chenopodium album*, *Dioscorea polystachya*, and *Ricinus communis*, the level of nucleotide variation in the cp genome was even lower [34, 35, 39]. We hypothesized that the wide distribution range and complex and diverse habitats of *E. maackii* might be the reasons for the high polymorphism rate in its cp genome because these factors promote the accumulation of genetic variation to ensure adaptability to diverse environments [3].

The genetic diversity in the intergenic regions of the cp genome of *E. maackii* was significantly higher than that in the coding regions, and these noncoding regions contained more variation, which could provide more genetic information. In this study, a set of highly variable intergenic regions were screened according to the levels of genetic diversity, and these regions also showed high variation in other species. For example, Su et al. identified four highly variable regions, including *rpl32-trnL-ULAG* and *ccsA-ndhD*, in the cp genomes of seventeen *Aegilops taushii* individuals [24]. Park et al. detected *rpl32-trnL-ULAG*, *ycf1-ndhF*, and *trnH-GUC-trnV-UAC* as highly variable sites in *C. album* plastomes, and Mikhaylova et al. found that *trnH-GUG-psbA*, *ndhC-trnV-ULAC*, and *psbE-pslE* were highly variable sites in the cp genome of *Silene latifolia* [40].

Silva et al. also reported that *trnH-GUC-psbA*, *rps16-trnQ-ULIG*, *trnD-GUC-trnY-GUA* and *ndhF-rpl32* showed markedly higher π values in *Utricularia amethystina* [41]. These highly variable fragments can be selected as powerful molecular markers for studies on phylogeography, population history, and molecular ecology. For example, Wei and Zhang employed three cp hypervariable regions to reveal that the split of the mainland and island lineages of *Lemmaphyllum rostratum* and *L. carnosum* var. *microphyllum* may have resulted from ancestral isolation whereby land bridges acted as a “barrier” rather than as a “corridor” between them [42].

Our research showed that the genetic diversity of the cp genome of *E. maackii* was significantly lower than that of the genus *Euonymus*, and the hypervariable coding and noncoding regions in *E. maackii* were also not completely consistent with those in *Euonymus*. For example, 8 of the 15 coding regions with the highest variation were different from each other in the genus *Euonymus* and *E. maackii*. This result suggests that we should not only rely on comparisons among species within a genus but also carry out a comparative analysis of cp genomes among individuals within the species when selecting highly variable regions suitable for the species.

### 4.2. Intraspecific Indels and Their Distribution Patterns

In this study, 65 indels were detected in the cp genome of *E. maackii*, and its indel density was 0.41 per kb, which was higher than that in most other species, such as *Chenopodium album* [34], *D. polystachya* [39], *Viburnum erosum* [43], *Fagus multinervis* [44], and *Suaeda japonica* [45], but lower than that in *Marchantia polymorpha* and *R. communis* [35, 46]. Among the species reported, *M. polymorpha* had the most abundant indels, and the cp genome of its Polish isolate contained 660 indels (indel density = 5.49 per kb) relative to the reference genome.

In *Quercus acutissima*, 255 indels were scattered across 67 regions (1.6 per kb) of its cp genomes [47]. In *R. communis*, 92 indels were uncovered across the plastomes, with an indel density of 0.56 per kb, and the variation level was also higher than that in *E. maackii*. However, a low level of indel density has been found in many other species. In total, 57 indels were detected in two *Zoysia matrella* cp genomes [48], and 49, 43, and 26 indels were identified in the *V. erosum*, *D. polystachya*, and *C. album* cp genomes, respectively [34, 39, 43]. Only
two indels each were identified in *S. japonica* and *F. multinervis*, indicating a low level of intraspecific variation [45,46].

Although the number of indels in cp genomes varied widely among species, the distribution of these indels showed a similar pattern, with most indels occurring in the LSC region and a smaller number in the SSC and IR regions. Our study revealed that 89.2% of indels were distributed in the LSC region of the *E. maackii* plastome, while only 6.2% and 4.6% of indels were distributed in the SSC and IR regions, respectively. In addition, the genetic diversity of the IR region was only 0.00086, which was 0.15 times that of the LSC region and 0.09 times that of the SSC region. Similar results were also found in *C. album*; most of the indels (61.2%) also appeared in the LSC region, and only a small number of indels appeared in the SSC (19.2%) and IR (19.6%) regions [34]. This distribution pattern of DNA variation suggests that IR regions are more evolutionarily conserved than LSC and SSC regions.

In summary, although only two individuals were used to evaluate cp genome variation in *E. maackii* in our study, these results can represent the overall trend of genetic variation in the species. In future studies, more groups and more samples should be used for comparative analysis to confirm these results.

### 4.3. Phylogenetic Relationships and Molecular Dating

In this study, we constructed a highly reliable phylogenetic tree based on the whole cp genome. Our results are in line with those from some previous studies [49]. For example, according to the characteristics of leaf epidermal cells, Zheng et al. clustered *E. maackii*, *E. hamiltonianus*, and *E. phellomanus* in Sect. *Euonymus*, which was equivalent to subclade I in our study. *E. fortunei* has epidermal cell characteristics similar to those of *E. japonicas*, both of which belong to Sect. *Illiciolina*. Similarly, our results showed that *E. fortunei* and *E. japonicas* were the most recently diverging species in the genus *Euonymus* [49].

However, other studies have obtained results inconsistent with ours. In our study, *E. maackii* was in one subclade, and *E. japonicus* and *E. schensianus* were in another subclade, but a numerical taxonomic analysis indicated that *E. maackii* and *E. schensianus* were merged into one group, and *E. japonicus* belonged to a different group [50]. In addition, our study showed that *E. fortunei* and *E. hamiltonianus* belonged to different clades, but Yao et al. clustered them on one branch in a study based on floral features [51].

### 5. Conclusions

We assembled the complete cp genome of *E. maackii* and analyzed intraspecific and interspecific cp genome variation. A high level of intraspecific variation was detected in *E. maackii*, and hypervariable coding and noncoding regions were identified in *E. maackii* and the genus *Euonymus*, respectively. A phylogeny based on the whole cp genomes revealed that all *Euonymus* species formed one clade, which separated from *C. edulis* and *M. guangxiensis* 24.74 million years ago, and *E. maackii* split from its related species only approximately 2.68 million years ago.

**Supplementary Materials:** The following supporting information can be downloaded at: [https://www.mdpi.com/article/10.3390/f13020265/s1](https://www.mdpi.com/article/10.3390/f13020265/s1), Table S1: Highly variable coding regions in *Euonymus maackii* and *Euonymus*; Table S2: Highly variable noncoding regions in *Euonymus maackii* and *Euonymus*.

**Author Contributions:** Conceptualization, X.-Q.L. and Y.-C.W.; methodology, Y.-C.W.; software, H.-YZ.; validation, Y.-C.W., H.-YZ. and X.-Q.L.; formal analysis, Y.-C.W., H.-YZ.; investigation, Y.-C.W., X.-Q.L.; resources, X.-Q.L.; data curation, Y.-C.W., H.-YZ.; writing—original draft preparation, Y.-C.W.; writing—review and editing, X.-Q.L.; visualization, Y.-C.W., H.-YZ.; supervision, X.-Q.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 31870193.

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

Data Availability Statement: The chloroplast genome sequence of *Euonymus maackii* has been deposited in GenBank under accession number MW771518.

Acknowledgments: We thank Y.-B. Y and C.-T. Li at Huazhong Agricultural University for collecting samples.

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

References

1. Weng, H.; Liu, Y.H.; Zhu, Y.W.; Sun, W.Z.; Liu, K.; Zuo, L.M.; Gu, L.N. A Study on the Drought Resistance and Salt Tolerance of *Euonymus Bungensis* Maxim. *Arid. Zone Res.* 1996, 13, 67–72.

2. Su, W.G.; Meng, Q.T.; Wang, W. Study on Extraction and Stability of Yellow Pigment from *Euonymus Maackii* Rupr. *North. Hortic.* 2011, 2011, 52–55.

3. Liu, J.; Cui, Q.; Kang, Y.; Meng, Y.; Gao, M.; Efferth, T.; Fu, Y. *Euonymus Maackii* Rupr. Seed Oil as a New Potential Non-Edible Feedstock for Biodiesel. *Renew. Energ.* 2019, 133, 261–267. [CrossRef]

4. Li, N.W.; Teng, W.J.; Shu, X.C. Growth Adaptability of *Euonymus Maackii* in Coastal Areas. *J. Anhui Agr. Sci.* 2018, 46, 76–78.

5. Ning, K.; Zhou, T.; Jiang, C.Z.; Wu, H.M.; Jiang, J.L.; Chen, J.; Ei-kassaby, Y.A.; Ma, Y. Rapid and Efficient Leaf Regeneration Propagation System for *Euonymus Bungensis*. *Biol. Plant.* 2021, 65, 118–125. [CrossRef]

6. Yang, H.T.; Wang, R.J.; Wang, S.L.; Li, J.R.; Zhang, D.J.; Wei, H.; Zhang, K.Z.; Li, Y. Horticultural Characteristics and Cultivation Techniques of 3 Species of Colored Leas Silk Cotton Tree in Puyang. *Agr. Sci.* 2018, 7, 87–89.

7. Ding, Y.; Zhang, J.; Liu, C.; Li, Z. Preparation of Polysaccharides from *Euonymus Bungensis* Fruit and Anti-Tumor Activity. *J. Northwest AF Univ.* 2018, 46, 122–129.

8. Pang, C.; Zuo, Y.; Zhang, G.; Zhen, P. Chemical Constituents from the Seeds of *Euonymus Maackii* and their Antitumor Activities. *Chin. Tradit. Pat. Med.* 2020, 42, 1208–1214.

9. Fu, Y.; Wang, W.; Gong, Q.; Zhang, H.; Zhao, W. Neuroprotective Dihydro-β-Agarofuran-Type Desquiterpenes from the Seeds of *Euonymus Maackii*. *J. Nat. Prod.* 2019, 82, 3096–3103. [CrossRef][PubMed]

10. 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]

11. Birky, C.W. Uniparental Inheritance of Mitochondrial and Chloroplast Genes-Mechanisms and Evolution. *Proc. Natl. Acad. Sci. USA* 1995, 92, 11331–11338. [CrossRef][PubMed]

12. Zeb, U.; Wang, X.; Fiaz, S.; Azizullah, A.; Shah, A.A.; Ali, S.; Rahim, F.; Ullah, H.; Leghari, U.A.; Wang, W.; et al. Novel Insights into Pinus Species Plastids Genome through Phylogenetic Relationships and Repeat Sequence Analysis. *PLoS ONE* 2022, 17, e0262040. [CrossRef]

13. Jansen, R.K.; Cai, Z.; Raubeson, L.A.; Daniell, H.; Depamphilis, C.W.; Leebens-Mack, J.; Müller, K.F.; Guisinger-Bellian, M.; Haberle, R.C.; Hansen, A.K.; et al. Analysis of 81 Genes from 64 Plastid Genomes Resolves Relationships in Angiosperms and Identifies Genome-Scale Evolutionary Patterns. *Proc. Natl. Acad. Sci. USA* 2007, 104, 19369–19374. [CrossRef]

14. Ding, S.X.; Dong, X.; Yang, J.X.; Guo, C.; Cao, B.B.; Guo, Y.; Hu, G.W. Complete Chloroplast Genome of Clethra Fargesii Fanch., an Original Symptelalous Plant from Central China: Comparative Analysis, Adaptive Evolution, and Phylogenetic Relationships. *Forests* 2021, 12, 441. [CrossRef]

15. Yang, J.; Choi, M.J.; Kim, S.H.; Choi, H.J.; Kim, S.C. Plastome Characterization and Phylogenomic Analysis Yield New Insight into the Evolutionary Relationship among the Species of the Subgenus *Braceoles* (Hosta; Asparagaceae) in East Asia. *Plants* 2021, 10, 180. [CrossRef]

16. Chung, H.Y.; Won, S.Y.; Kim, Y.K.; Kim, J.S. Development of the Chloroplast Genome-Based InDel Markers in Niitaka (*Pyrus Pyrifolia*) and its Application. *Plant Biotechnol. Rep.* 2019, 13, 51–61. [CrossRef]

17. Zeb, U.; Wang, X.; AzizUllah, A.; Fiza, S.; Khan, H.; Ullah, S.; Ali, H.; Shahzad, K. Comparative Genome Sequence and Phylogenetic Analysis of Chloroplast for Evolutionary Relationship among *Pinus* Species. *Saud. J. Bio. Sci.* 2021. [CrossRef]

18. Kim, Y.; Shin, J.; Oh, K.-R.; Kim, A.-K.; Choi, C. Comparative Analysis of Complete Chloroplast Genome Sequences and Insertion-Deletion (Indel) Polymorphisms to Distinguish Five *Vaccinium* Species. *Forests* 2020, 11, 927. [CrossRef]

19. Nock, C.T.; Hardner, C.M.; Motenegro, J.D.; Ahmad, T.; Ainnatul, A.; Hayashi, S.; Playford, J.; Edwards, D.; Batley, J. Wild Origins of Macadamia Domestication Identified through Intraspecific Chloroplast Genome Sequencing. *Front. Plant Sci.* 2019, 10, 334. [CrossRef][PubMed]

20. Xiao, S.; Xu, P.; Deng, Y.; Dai, X.; Zhao, L.; Heider, B.; Zhang, A.; Zhou, Z.; Cao, Q. Comparative Analysis of Chloroplast Genomes of Cultivars and Wild Species of Sweetpotato (*Ipomoea Batatas* [L.] Lam). *BMC Genom.* 2021, 22, 262.

21. Park, J.; Xi, H.; Kim, Y. The Complete Chloroplast Genome of *Arabidopsis Thaliana* Isolated in Korea (Brassicaceae):-An Investigation of Intraspecific Variations of the Chloroplast Genome of Korean *A. Thaliana*. *Int. J. Genom.* 2020, 3236461, 1–18. [CrossRef][PubMed]
48. Lee, B.; Park, J. The Complete Chloroplast Genome of *Zoysia Matrella* (L.) Merr. Isolated in Korea (Poaceae): Investigation of Intraspecific Variations on Chloroplast Genomes. *Mitochondrial DNA Part B* 2021, 6, 572–574. [CrossRef] [PubMed]

49. Zheng, Y.C.; Ming, X.; Tu, Q.; Wu, X.; Wang, R.H. Observation of Leaf Epidermal Cells of 28 Species of *Euonymus* in China by Electron Microscope Scanning. *Hunan Agr. Sci.* 2012, 9, 107–112.

50. Liu, Y.; Meng, S.; Liu, Q. A Numerical Taxonomic Analysis of *Glyptopetalum* and *Euonymus* (Celastraceae). *Guihaia* 2021, 41, 92–102.

51. Yao, C.Y.; Du, C.; Zuo, Y.J.; Liu, Q.R.; Ma, J.S. The Significance of Floral Features within Infrageneric Classification of *Euonymus* (Celastraceae). *Flora* 2018, 242, 53–60. [CrossRef]