

Communication

# The Complete Chloroplast Genome of the Hare's Ear Root, *Bupleurum falcatum*: Its Molecular Features

Dong-Ho Shin <sup>1</sup>, Jeong-Hoon Lee <sup>2</sup>, Sang-Ho Kang <sup>1</sup>, Byung-Ohg Ahn <sup>3</sup> and Chang-Kug Kim <sup>1,\*</sup>

<sup>1</sup> Genomics Division, National Institute of Agricultural Sciences, Jeonju 54874, Korea; dhshin013@korea.kr (D.-H.S.); hosang93@korea.kr (S.-H.K.)

<sup>2</sup> Department of Herbal Crop Research, National Institute of Horticultural and Herbal Science, Eumseong 55365, Korea; artemisia@korea.kr

<sup>3</sup> Research Policy Planning Division, Rural Development Administration, Jeonju 54875, Korea; boahn@korea.kr

\* Correspondence: chang@korea.kr; Tel.: +82-63-238-4555; Fax: +82-63-238-4554

Academic Editor: Paolo Cinelli

Received: 13 January 2016; Accepted: 9 May 2016; Published: 13 May 2016

**Abstract:** *Bupleurum falcatum*, which belongs to the family *Apiaceae*, has long been applied for curative treatments, especially as a liver tonic, in herbal medicine. The chloroplast (cp) genome has been an ideal model to perform the evolutionary and comparative studies because of its highly conserved features and simple structure. The *Apiaceae* family is taxonomically close to the *Araliaceae* family and there have been numerous complete chloroplast genome sequences reported in the *Araliaceae* family, while little is known about the *Apiaceae* family. In this study, the complete sequence of the *B. falcatum* chloroplast genome was obtained. The full-length of the cp genome is 155,989 nucleotides with a 37.66% overall guanine-cytosine (GC) content and shows a quadripartite structure composed of three nomenclatural regions: a large single-copy (LSC) region, a small single-copy (SSC) region, and a pair of inverted repeat (IR) regions. The genome occupancy is 85,912-bp, 17,517-bp, and 26,280-bp for LSC, SSC, and IR, respectively. *B. falcatum* was shown to contain 111 unique genes (78 for protein-coding, 29 for tRNAs, and four for rRNAs, respectively) on its chloroplast genome. Genic comparison found that *B. falcatum* has no pseudogenes and has two gene losses, *accD* in the LSC and *ycf15* in the IRs. A total of 55 unique tandem repeat sequences were detected in the *B. falcatum* cp genome. This report is the first to describe the complete chloroplast genome sequence in *B. falcatum* and will open up further avenues of research to understand the evolutionary panorama and the chloroplast genome conformation in related plant species.

**Keywords:** *Bupleurum falcatum*; chloroplast genome; quadripartite structure; tandem repeats; phylogenetic analysis

## 1. Introduction

Chloroplasts, also called cp or ct, are distinctly important organelles which have their own genomes and play pivotal roles in generating energy through photosynthesis, fixing carbon, and biosynthesizing starch, fatty acids, pigments, and amino acids in plant cells [1,2]. They have been used as ideal research models, particularly for evolutionary and comparative genomic studies because of their highly conserved gene content, relatively small size, and simple structure.

A complete chloroplast genome sequence was first reported from tobacco, *Nicotiana tabacum*, in 1986 [3] after the existence of the chloroplast DNA was ultrastructurally confirmed from *Chlamydomonas* in 1962 [4]. To date, hundreds of complete chloroplast genome sequences have been identified from plants, algae and bacteria. Plant chloroplasts have been known to be derived from the endosymbiosis of cyanobacteria and their genomes have been reported to inherit uniparentally without recombination, although some plants have been reported to biparentally inherit [5,6]. The chloroplast genome can be

divided into two comprehensive categories including protein-coding genes and non-coding regions; the latter is further divided into introns and intergenic regions [7]. The genome is unified in a conserved quadripartite structure consisting of a large single copy (LSC), a small single copy (SSC), and a set of mutually-inverted repeat sequences (IRa and IRb) [8–10], which are also highly conserved together with gene order and content among species [11]. Therefore, comparative studies of the chloroplast genome have been used not only to solve the jigsaw puzzle of plant evolutionary history but also to identify taxonomical position [12–14], although large-scale genome rearrangement and gene loss have been reported in several angiosperm lineages [15,16]. Information from the chloroplast genome is beneficial to phylogenetics [17], DNA barcoding [18], population biology [19], and transcriptomic [20] studies. It is important to understand how plant species are linked, what features are shared among them, and how they are different from other taxonomic groups. Sequencing the complete chloroplast genome is now inexpensive and efficient using next generation sequencing (NGS) technology.

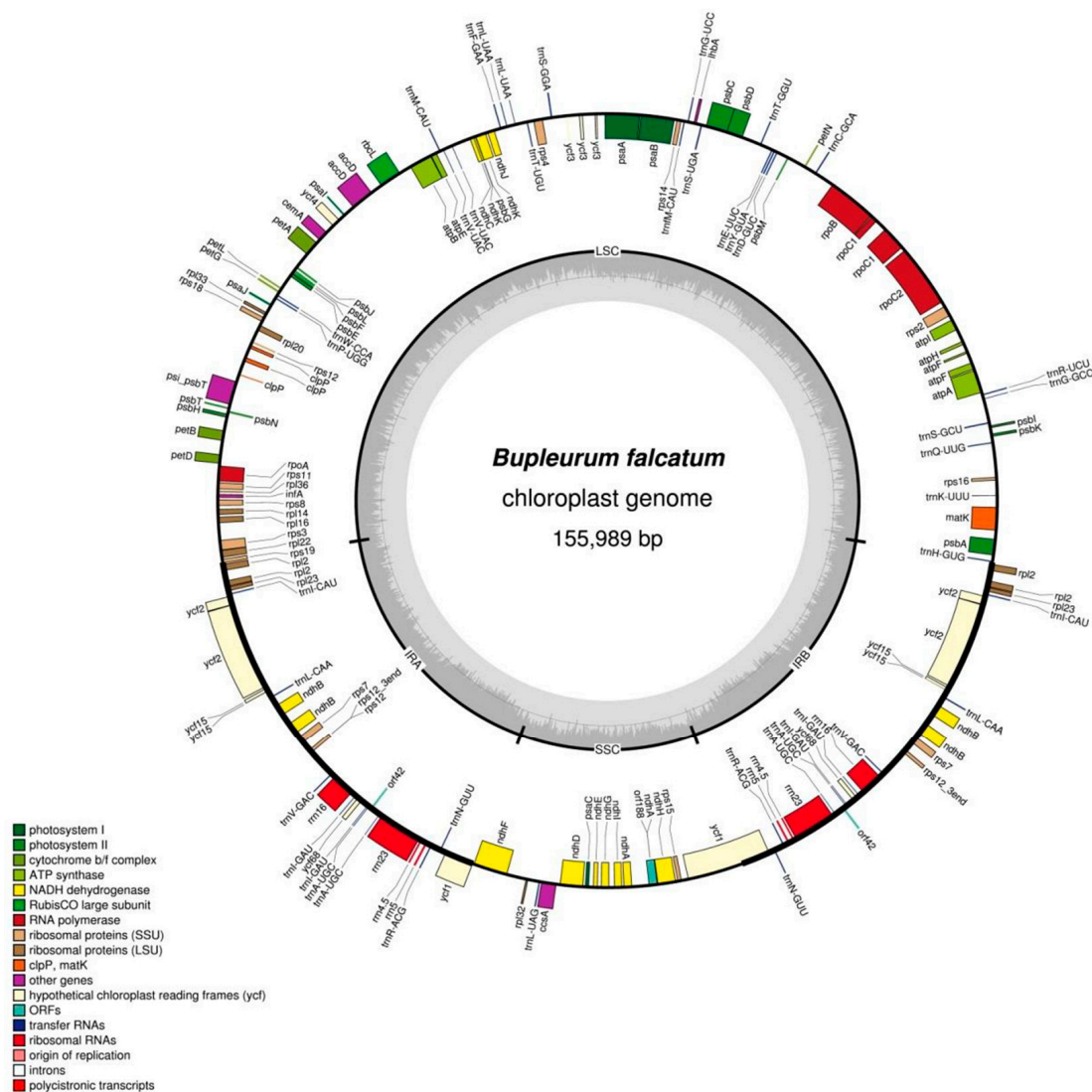
In this study, the complete chloroplast genome of a flowering plant, *Bupleurum falcatum*, which is categorized to the family *Apiaceae*, was investigated. The *Apiaceae* family is taxonomically close to the *Araliaceae* family and there have been numerous complete chloroplast genome sequences reported in the *Araliaceae* family [21,22], while little is known about the *Apiaceae* family. *B. falcatum* (known as Chinese thoroughwax and Sickle hare's ear, called Siho in Korean) is a perennial plant and has long been known to be a medicinal plant because it produces curative compounds that have been reported to possess various beneficial properties [23–26], especially as a liver tonic, in herbal medicine. This report is the first to describe the complete chloroplast genome sequence, sequence analysis, and molecular and phylogenetic comparison of *Bupleurum falcatum* belonging to the *Apiaceae* family. It is believed that the results may contribute to a better understanding of genetic relationships and evolutionary aspects of medicinal plants in the families *Apiaceae* and *Araliaceae*.

## 2. Results and Discussion

### 2.1. Sequencing, Assembly, and Validation of the *B. falcatum* Chloroplast Genome

A total of eight million paired-end (PE) reads (800 million nucleotides) generated were trimmed and assembled using a CLC genome assembler (CLC Bio, Aarhus, Denmark). A total of 342,316 PE reads were concordantly mapped to the final assembly with an estimation of the cp coverage (194.83×). The mapped cp contigs were selected and merged to construct a complete *B. falcatum* cp genome using a MUMmer (Hamburg University, Hamburg, Germany). The *B. falcatum* cp genome sequence was generated from a combined product of four initial contigs; contigs 8, 27, 76, and 209, respectively with no gaps and no Ns. The assembled *B. falcatum* cp genome showed a high similarity to the reference sequence, *Panax ginseng* Chunpoong (GenBank Accession#: KM088019) and extended its size to 155,989 nucleotides long (Figure 1). The cp genome sequence was registered into the DNA Data Bank of Japan (DDBJ).

European Molecular Biology Laboratory (EMBL), and GenBank with the accession number KC207676. The largest and smallest assembled sequences were derived from contigs 8 and 209 covering the complete regions of the LSC and SSC, respectively. As shown in Figure 1, the *B. falcatum* cp genome is a typical quadripartite structure composed of the LSC, SSC and a pair of IR regions, indicating that the assembled *B. falcatum* cp genome sequence represented full coverage with no abnormalities. *B. falcatum* belongs to the *Apiaceae* family in plant classification. Recently, several chloroplast genome sequences have been reported from the *Apiaceae* and *Araliaceae* families such as *P. ginseng*, *A. undulate*, *D. carota* and *A. cerefolium* [21,22,27]. Both *Apiaceae* and *Araliaceae* families belong to the *Apiales* order and these two families are known to have the most abundant species ranked into the largest and second, respectively, in the order [28]. The overall sequencing results of the *B. falcatum* cp genome compared with genome size and physical shape indicate that there are no significant differences between *B. falcatum* chloroplast genome and those from the two closely-related families.



**Figure 1.** The gene map of the chloroplast genome of *Bupleurum falcatum*. The Organellar Genome Draw (OGDraw) program was applied for drawing the map. LSC, SSC, and IR are abbreviated for large single copy, small single copy, and inverted repeat, respectively. Genes located inside and outside of the outer rim are transcribed clockwise and counterclockwise, respectively. Functionally-annotated genes are seen in colored portions. The dashed gray area in the inner circle shows the proportional GC content of the corresponding genes.

## 2.2. Physical Features of the *B. falcatum* Chloroplast Genome

The complete *B. falcatum* cp genome obtained in this study appears to be a typical circular form and encodes 155,989 nucleotides encompassed in the quadripartite structure which is built in four regions (LSC, SSC, IRa and IRb) described elsewhere [3,29] (Figure 1). The respective four regions of the cp genome occupy 85,912 bp for LSC, 17,517 bp for SSC, and 52,560 bp (26,280 bp each) for a set of the IR regions, accounting for about 55.1% for LSC, 11.2% for SSC and 33.7% for IRs, respectively (Table 1). Table 2 describes a list of the corresponding genes according to their functions, features and participational roles.

**Table 1.** Summary of *de novo* assembly and general features of the chloroplast genome of *B. falcatum*.

Compositional Category	Features in Chloroplast
CP genome length (nt)	155,989
LSC length (nt)	85,912
SSC length (nt)	17,517
IR length (Ira/b) (nt)	52,560
AT/GC contents (%)	62.34/37.66
No. of total/unique genes	129/111
Genic occupancy (nt)	88,065
Intergenic occupancy (nt)	67,924
Protein-coding genes	78
Transfer RNAs (tRNAs)	29
Ribosomal RNAs (rRNAs)	4
No. of genes duplicated in IR regions	16
Total number of genes with intron(s)	17
Gene(s) with single intron	14
Gene(s) with multiple introns	3
tRNAs with intron(s)	5

nt indicates the units of nucleotides.

**Table 2.** Gene composition in *B. falcatum* chloroplast genome (total 111 unique genes).

Category of Gene Group	Group of Genes	Name of Genes
Self replication	Ribosomal RNAs	16S ( <i>rrn16</i> )(x2), 23S ( <i>rrn23</i> )(x2) 4.5S ( <i>rrn4.5</i> )(x2), 5S ( <i>rrn5</i> )(x2) <i>trnH-GUG</i> , <i>trnK-UUUU</i> , <i>trnQ-UUG</i> <i>trnS-GCU</i> , <i>trnG-UCC</i> <sup>†</sup> , <i>trnR-UCU</i> <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnY-GUA</i> <i>trnE-UUC</i> , <i>trnT-GGU</i> , <i>trnS-UGA</i> <i>trnM-CAU</i> , <i>trnS-GGA</i> , <i>trnT-UGU</i> <i>trnL-UAA</i> <sup>†</sup> , <i>trnF-GAA</i> , <i>trnV-UAC</i> <sup>†</sup>
	Transfer RNAs	<i>trnM-CAU</i> , <i>trnW-CCA</i> , <i>trnP-UGG</i> , <i>trnI-CAU</i> (x2), <i>trnL-CAA</i> (x2), <i>trnV-GAC</i> (x2), <i>trnI-GAU</i> (x2) <sup>†</sup> , <i>trnA-UGC</i> (x2) <sup>†</sup> , <i>trnR-ACG</i> (x2), <i>trnN-GUU</i> (x2), <i>trnL-UAG</i>
	Small subunit of ribosome	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (x2), <i>rps8</i> , <i>rps11</i> , <i>rps12</i> (x2, <i>part</i> ) <sup>†</sup> , <i>rps14</i> , <i>rps15</i> , <i>rps16</i> <sup>†</sup> , <i>rps18</i> , <i>rps19</i>
	Large subunit of ribosome	<i>rpl2</i> (x2) <sup>†</sup> , <i>rpl14</i> , <i>rpl16</i> <sup>†</sup> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (x2), <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
	RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> <sup>†</sup> , <i>rpoC2</i>
Photosynthesis	NADH-dehydrogenase	<i>ndhA</i> <sup>†</sup> , <i>ndhB</i> (x2) <sup>†</sup> , <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i> , <i>ycf3</i> <sup>#</sup> ,
	Photosystem II	<i>lhbA</i> , <i>psbA</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> <i>psbN</i> , <i>psi-psbT</i> , <i>psbT</i>
	Cytochrome b/f	<i>petA</i> , <i>petB</i> <sup>†</sup> , <i>petD</i> <sup>†</sup> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> <sup>†</sup> , <i>atpH</i> , <i>atpI</i>
	Rubisco	<i>rbcl</i>
Other genes		<i>infA</i> , <i>matK</i> , <i>clpP</i> <sup>#</sup> , <i>cemA</i> , <i>ccsA</i>
Unknown function	ORFs <sup>‡</sup>	<i>ycf1</i> (x2, part), <i>ycf2</i> (x2), <i>ycf4</i>

<sup>†</sup> indicates the existence of single intron in the corresponding genes; <sup>#</sup> indicates the existence of two introns in the corresponding genes; <sup>‡</sup> indicates open reading frame. Genes duplicated in the IR region are represented by the (×2).

The *B. falcatum* cp genome contains a total of 111 unique genes constituting 78 protein-coding genes, 29 transfer RNAs, and four ribosomal RNAs with the overall AT content of 62.34% (Table 1). Seventeen genes are identified to contain at least one intron and two of them (*clpP* and *yfc3*) are shown to contain two introns (Tables 1 and 2). Among the genes containing intron(s), twelve genes appear to belong to the group of the protein-coding genes and five to tRNAs (Table 2). For the nucleotide occupancy, 88,065 bp of the nucleotide length accounts for the genic regions which contain both exonic and intronic portions corresponding to 56.46% and 67,924 bp for intergenic regions corresponding to 43.54%, respectively (Table 1).

The gene composition of the four respective regions in the *B. falcatum* cp genome shows that the LSC region contains 60 genes for the protein-codings and 21 for tRNAs (Figure 1). The cp genome also shows that the SSC region encodes 11 genes for the protein-codings and one for tRNA. Furthermore, it shows that respective IR regions express five genes for the protein-codings, seven for tRNAs, four for rRNAs, and that the border junctions between the SC regions and the IR regions contain two protein-coding genes. As shown in this result, all of the ribosomal RNAs of the *B. falcatum* chloroplast genome are housed only in the IR regions, suggesting that the IRs play a critical role in the existence of the chloroplast in *B. falcatum* as well as in other plant species. It is also speculated that the IR regions have been a very conserved partitive wall to specify ribosomal RNAs for the ribosome build-up through the long-term evolutionary events.

Moreover, comparative analysis of gene content, interestingly, shows that there are unusual features on the *B. falcatum* cp genome (Supplementary Table S1). There are two gene losses of *accD* and *yfc15* detected in the *B. falcatum* cp genome. As compared in Supplementary Table S1, the *accD* gene is usually located between the *rbcL* and *psaI* genes in the LSC region, and the *yfc15* gene duplicates at the position between *yfc2* and *trnL-CAA* in the IR regions among the compared plant species of *Araliaceae* and *Apiaceae* families. In addition, a gene substitution of *lhbA* is recognized at the position where the *psbZ* gene has generally been shown at the position between the genes *trnS* and *trnG* in related species [22,27]. The *accD* gene encodes a subunit of heteromeric acetyl-CoA carboxylase (ACCase) that is commonly found in plastids of dicots but is not seen in the rice and other *Gramineae* [30]. It will be required to confirm whether the two genes *accD* and *yfc15* have been completely missing from the chloroplast genome of *B. falcatum* or have been transferred to other cellular organelles such as the nucleus, although there have been reports with respect to no *yfc15* gene (GenBank accession number, GU456628) and the *accD* gene's loss [31–33].

Some of the genes in the chloroplast genome have been reported to pseudogenize, particularly genes in the IR regions of plant species. Plant species featuring pseudogenization have recently been reported from the *Apiales* order including *A. cerefolium*, *Panax ginseng*, *A. carota*, *B. hainla*, and *K. septemlobus* [21,22]. However, there were no pseudogenes identified in the *B. falcatum* cp genome in this study (Supplementary Table S1). It is not clear if *B. falcatum* has a specific kinetic mechanism different from the other species. Future investigation will be required to elucidate what mechanism conducts the pseudogenization.

### 2.3. Tandem Repeat Sequences

The repeated sequences come in various sizes in the pattern and are classified according to the length of the core repeat units, the number of contiguous repeat units, and/or the overall length of the repeat region. In order to investigate simple sequence repeats in *B. falcatum*, tandem repeats (TRs) were analyzed. A total of 54 unique consensus sequences (out of 119 copies) of tandem repeats were detected at 55 locations (consensus 16 and 17 are identical but occur at different positions) from the *B. falcatum* chloroplast genome DNA (Supplementary Table S2) using Tandem Repeat Finder, version 4.0 (LBI, Boston University, Boston, MA, USA) [34]. Out of them, 53 repeat locations appear to be arranged by two copies, while one position appears to be arranged by three copies at nucleotides (nt) 49,794–49,811, and the remaining one by ten copies at nt 85,644–85,663. The shortest and longest tandem repeat sequences are identified to be one di- and one pentacontakai-penta-nucleotide repeats,

which are positioned at nt 85,644–85,663 and nt 32,737–32,846, respectively. By length, 13 different sizes of the TR sequences ranging from 2 to 55 nucleotides reside in the *B. falcatum* chloroplast genome. The most abundant repeats lengths are found to be an octa-nucleotide and a nona-nucleotide, which occur at twelve and thirteen locations, respectively. There are seven single locations of tandem repeats found in *B. falcatum* as follows: di-, deca-, tetradeca-, hexadeca-, heptadeca-, nonadeca-, and pentacontakai-penta-nucleotide sequences, respectively. Considering the regional occupancy of the TR sequences, the IR regions have remarkably low frequency of the tandem repeats, showing only four locations corresponding to 7.02% compared to 42 and seven positions corresponding to 73.68% and 12.28% for the LSC and the SSC, respectively. Interestingly, most of the tandem repeats observed in the *B. falcatum* cp genome are identified to be located in the noncoding regions, whereas only a small portion of them is found to be in the genic regions of the following: *psaI*, *rpoA*, *rpl22*, *ycf2*, *ndhE*, *ndhD*, and *ndhI*. Recent reports have described the existence of five and eight protein-coding genes containing the repeat sequences, respectively, in *A. undulate* [22] and *P. ginseng* Damaya [27]. Many reports have mentioned a wide variety of repeat sequences across different plant species [3,21,22], albeit the loci and the genes are quite different.

The results in this study indicate that the LSC region is relatively active in causing sequence variations, while the IR regions have much less kinetic pressure, suggesting that the conservation of the IR regions may be pivotally important to maintaining the functional integrity of plant chloroplasts including *B. falcatum*.

### 3. Experimental Section

#### 3.1. Plant and DNA Sample Preparation

Based on its clinical importance in herbal medicine, *B. falcatum* was employed for the study. The plant was maintained in an indoor facility with an automatically and environmentally controlled system at the National Institute of Horticultural and Herbal Science, Rural Development Administration in Eumseong, Korea. The total plant DNA was extracted from the young leaves of *B. falcatum* using the modified cetyltrimethylammonium bromide (CTAB) method [35]. The entire process of DNA sample preparation for the whole genome shotgun sequencing was performed according to the manufacturer's instructions of the Illumina platform (San Diego, CA, USA).

#### 3.2. Chloroplast Genome Sequencing, Assembly, and Validation

The *B. falcatum* Illumina PE library was sequenced on a HiSeq 1000 genome analyzer platform (Illumina, San Diego, CA, USA) installed in the Genome Sequencing Center, the Department of Biotechnology, National Academy of Agricultural Science, Rural Development Administration in Suwon, South Korea. Low quality sequences (Phred score < 20) were trimmed and the remaining high quality sequences were assembled in to contigs using a CLC genome assembler beta 4.06 (CLC Inc., Rarhus, Denmark) with parameters of a minimum of 150–500 bp autonomously controlled overlap size at Phyzen Inc. (Seoul, South Korea). The principal contigs representing the chloroplast genome were obtained from the total assembled contigs using MUMmer [36] with the cp genome sequence of *Panax ginseng* cv. ChunPoong (KM088019) [37] as a reference sequence. The obtained chloroplast sequence contigs were ordered and oriented based on the previously reported cp genome sequence, and a complete chloroplast sequence was constructed by connecting overlapping terminal sequences. Validation was performed using the BlastZ program (Miller Lab, Penn State University, University Park, PA, USA) by comparing the structure of the *B. falcatum* cp genome with published cp genomes registered in National Center for Biotechnology Information (NCBI).

#### 3.3. Chloroplast Genome Annotation, Codon Usage and Repeat Sequence Analysis

In order to annotate the chloroplast genome, the program DOGMA (Jansen Lab, UT Austin, TX, USA) [38] equipped with manually-operated corrections for start and stop codons was employed.

The tRNA genes were identified using DOGMA and tRNAscanSE (Lowe Lab, UCSC University, Santa Cruz, CA, USA) [39]. The cp genome map was constructed based on the program OGDRAW (<http://ogdraw.mpimp-golm.mpg.de/>) [40]. Repeat sequences of the *B. falcatum* cp genome were assessed using Tandem repeat finder, version 4.0 (LBI, Boston University, Boston, MA, USA) [34] and REPuter (Faculty of Technology, Bielefeld University, Bielefeld, Germany) [41] with the following parameters: Match 2, Mismatch 7, Delta 7, PM 80, PI 10, Minscore 50, MaxPeriod 500, with similarity 100%, respectively.

#### 4. Conclusions

There has been no report of the complete cp genome sequence of the *B. falcatum* species. We sequenced the complete *B. falcatum* chloroplast genome. The size, structure, gene content, and the compositional organization are not significantly different from most of the other chloroplast genomes reported from closely-related plant species, albeit there are subtle differences in features. It is a noticeable finding that there are no pseudogene(s), particularly in the boundary areas of the IR regions of the *B. falcatum* chloroplast genome. In addition, the *B. falcatum* cp genome showed the losses of *accD* and *ycf15* genes. However, further investigations of pseudogene(s) and the two missing genes remain to be performed. Future comparative analyses will be paid to elucidate how *B. falcatum* has evolved and diverged within the *Araliaceae* family. This report will open up further avenues of research to understand the genomic information and gene contents of the chloroplasts of the genus *Bupleurum* and the related families.

**Supplementary Materials:** The following are available online at: <http://www.mdpi.com/2073-4425/7/5/20/s1>, Table S1: A comparative list of genes present in the chloroplast genomes of *B. falcatum* and others of the *Araliaceae* and *Apiaceae* families, Table S2: The result of tandem repeat sequences detected in the *B. falcatum* chloroplast genome sequence.

**Acknowledgments:** The authors thank the National Institute of Horticultural and Herbal Science (NIHHS) Core Greenhouse for housing, maintaining and growing the plants, and the National Institute of Agricultural Sciences (NAS) Genome Sequencing Core facility for their services. The authors would like to thank Liza Merly of the University of Miami for her help in reviewing the text for accuracy. We also thank Phyzen for assembling the sequencing data. This research was supported by Rural Development Administration (RDA) grant PJ010351

**Author Contributions:** Dong-Ho Shin and Phyzen analyzed the data, Dong-Ho Shin wrote the paper, Jeong-Hoon Lee and Dong-Ho Shin performed experiments, Sang-Ho Kang and Byung-Ohg Ahn contributed the materials, and Chang-Kug Kim conceived and designed the experiments. All authors read and approved the final manuscript.

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

#### Abbreviations

The following abbreviations are used in this manuscript:

LSC	Large Sequence Copy
IR	Inverted Repeat
SSC	Small Sequence Copy
TR	Tandem Repeats
CP	Chloroplast

#### References

1. Chloroplast. Available online: [http://en.wikipedia.org/wiki/Chloroplast\\_DNA](http://en.wikipedia.org/wiki/Chloroplast_DNA) (accessed on 7 December 2015).
2. Neuhaus, H.E.; Emes, M.J. Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2000**, *51*, 111–140. [[CrossRef](#)] [[PubMed](#)]
3. Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.; Matsubayashi, T.; Zaita, N.; Chunwonse, J.; Obokata, J.; Yamaguchi-Shinozaki, K.; *et al.* The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **1986**, *5*, 2043–2049. [[CrossRef](#)] [[PubMed](#)]
4. Ris, H.; Plaut, W. Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas*. *J. Cell Biol.* **1962**, *13*, 383–391. [[CrossRef](#)] [[PubMed](#)]

5. Birky, C.W., Jr. Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proc. Natl. Acad. Sci. USA* **1995**, *95*, 11331–11338. [[CrossRef](#)]
6. Ilona, M. Non-nuclear genes and their inheritance. *Nat. Educ.* **2008**, *1*, 135.
7. Shaw, J.; Lickey, E.B.; Schilling, E.E.; Small, R.L. Comparison of whole chloroplast genome sequence to choose noncoding regions for phylogenetic studies in Angiosperms: The tortoise and the hare III. *Am. J. Bot.* **2007**, *94*, 275–288. [[CrossRef](#)] [[PubMed](#)]
8. Sugiura, M. The chloroplast genome. *Plant Mol. Biol.* **1992**, *19*, 149–168. [[CrossRef](#)] [[PubMed](#)]
9. Gary, M.W. The evolutionary origins of organelles. *Trends Genet.* **1989**, *5*, 294–299. [[CrossRef](#)]
10. Leseberg, C.H.; Duvall, M.R. The complete chloroplast genome of coix lacryma-jobi and a comparative molecular evolutionary analysis of plastomes in cereals. *J. Mol. Evol.* **2009**, *69*, 311–318. [[CrossRef](#)] [[PubMed](#)]
11. Gary, M.W.; Sankoff, D.; Cedergren, R.J. On the evolutionary descent of organisms and organelles: A global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Res.* **1984**, *12*, 5837–5852. [[CrossRef](#)]
12. Chumley, T.W.; Palmer, J.D.; Mower, J.P.; Fourcade, H.M.; Calie, P.J.; Boore, J.L.; Jansen, R.K. The complete chloroplast genome sequence of *Pelargonium x hortorum*: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol. Biol. Evol.* **2006**, *23*, 2175–2190. [[CrossRef](#)] [[PubMed](#)]
13. Palmer, J.D.; Thompson, W.F. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **1982**, *29*, 537–550. [[CrossRef](#)]
14. Palmer, J.D. Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* **1985**, *19*, 325–354. [[CrossRef](#)] [[PubMed](#)]
15. Wolfe, K.H.; Morden, C.W.; Ems, S.C.; Palmer, J.D. Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: Loss or accelerated sequence evolution of tRNA and ribosomal protein genes. *J. Mol. Evol.* **1992**, *35*, 304–317. [[CrossRef](#)] [[PubMed](#)]
16. Lee, H.L.; Jansen, R.K.; Chumley, T.W.; Kim, K.J. Gene relocations within chloroplast genomes of *Jasminum* and *Menodora* (Oleaceae) are due to multiple, overlapping inversions. *Mol. Biol. Evol.* **2007**, *24*, 1161–1180. [[CrossRef](#)] [[PubMed](#)]
17. Jansen, R.K.; Cai, Z.; Raubeson, L.A.; Daniell, H.; dePamphilis, C.W.; Leebens-Meck, J.; Muller, 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](#)] [[PubMed](#)]
18. Hollingsworth, P.M.; Graham, S.W.; Little, D.P. Choosing and using a plant DNA barcode. *PLoS ONE* **2011**, *6*, e19254. [[CrossRef](#)] [[PubMed](#)]
19. Powell, W.; Morgante, M.; McDevitt, R.; Vendramin, G.G.; Rafalski, J.A. Polymorphic simple sequence repeat regions in chloroplast genomes: Applications to the population genetics of pines. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7759–7763. [[CrossRef](#)] [[PubMed](#)]
20. Bock, R.; Khan, M.S. Taming plastids for a green future. *Trends Biotechnol.* **2004**, *22*, 311–318. [[CrossRef](#)] [[PubMed](#)]
21. Kim, K.J.; Lee, H.L. Complete chloroplast genome sequences from Korean Ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Res.* **2004**, *11*, 247–261. [[CrossRef](#)] [[PubMed](#)]
22. Li, R.; Ma, P.F.; Wen, J.; Yi, T.S. Complete sequencing of five Araliaceae chloroplast genomes and the phylogenetic implications. *PLoS ONE* **2013**, *8*.
23. Lee, B.B.; Shim, I.; Lee, H.J.; Hahm, D.H. Effect of *Bupleurum falcatum* on the stress-induced impairment of spatial working memory in rats. *Biol. Pharm. Bull.* **2009**, *32*, 1392–1398. [[PubMed](#)]
24. Lee, B.B.; Yun, H.Y.; Shim, I.S.; Lee, H.J.; Hahm, D.H. *Bupleurum falcatum* prevents depression and anxiety-like behaviors in rats exposed to repeated restraint stress. *J. Microbiol. Biotechnol.* **2012**, *22*, 422–430. [[CrossRef](#)] [[PubMed](#)]
25. Okamoto, H.; Chino, A.; Hirasaki, Y.; Ueda, K.; Raimura, M.; Namiki, T. A valid approach in refractory glossodynia: single-institution 5-year experience treating with Japanese traditional herbal (kampo) medicine. *Evid. Based Complement. Altern. Med.* **2013**, *2013*. [[CrossRef](#)] [[PubMed](#)]
26. Park, J.S.; Bang, O.S.; Kim, J.H. Screening of Stat3 inhibitory effects of Korean herbal medicines in the A549 human lung cancer cell line. *Integr. Med. Res.* **2014**, *3*, 67–73. [[CrossRef](#)]



27. Zhao, Y.; Yin, J.; Guo, H.; Zhang, Y.; Xiao, W.; Sun, C.; Wu, J.; Qu, X.; Yu, J.; Wang, X.; *et al.* The complete chloroplast genome provides insight into the evolution and polymorphism of *Panax ginseng*. *Front. Plant Sci.* **2015**, *5*. [[CrossRef](#)] [[PubMed](#)]
28. Angiosperm Phylogeny Group III. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG. *Bot. J. Linn. Soc.* **2009**, *161*, 105–121.
29. Sugita, M.; Kato, A.; Shimada, H.; Sugiura, M. Sequence analysis of the junctions between a large inverted repeat and single-copy regions in tobacco chloroplast DNA. *Mol. Gen. Genet.* **1984**, *194*, 200–205. [[CrossRef](#)]
30. Konishi, T.; Shinohara, K.; Yamada, K.; Sasaki, Y. Acetyl-CoA carboxylase in higher plants: Most plants other than *Gramineae* have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant Cell Physiol.* **1996**, *37*, 117–122. [[CrossRef](#)] [[PubMed](#)]
31. Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.; Sakamoto, M.; Mori, M.; Kondo, C.; Honji, Y.; Sun, C.R.; Meng, B.Y.; *et al.* The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **1989**, *217*, 185–194. [[CrossRef](#)] [[PubMed](#)]
32. Maier, R.M.; Neckermann, K.; Igloi, G.L.; Kossel, H. Complete sequence of the maize chloroplast genome: Gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J. Mol. Biol.* **1995**, *251*, 614–628. [[CrossRef](#)] [[PubMed](#)]
33. Ogihara, Y.; Isono, K.; Kojima, T.; Endo, A.; Hanaoka, M.; Shiina, T.; Terachi, T.; Utsugi, S.; Murata, M.; Mori, N.; *et al.* Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. *Mol. Genet. Genom.* **2002**, *266*, 740–746.
34. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [[CrossRef](#)] [[PubMed](#)]
35. Allen, G.C.; Flores-Vergara, M.A.; Krasynanski, S.; Kumar, S.; Thompson, W.F. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethyl-ammonium bromide. *Nat. Protoc.* **2006**, *1*, 2320–2325. [[CrossRef](#)] [[PubMed](#)]
36. Kurtz, S.; Phillippy, A.; Delcher, A.L.; Smoot, M.; Shumway, M.; Antonescu, C.; Salzberg, S.L. Versatile and open software for comparing large genomes. *Genome Biol.* **2004**. [[CrossRef](#)] [[PubMed](#)]
37. Kim, K.H.; Lee, S.C.; Lee, J.K.; Lee, H.O.; Joh, H.J.; Kim, N.H.; Park, H.S.; Yang, T.J. Comprehensive survey of genetic diversity in chloroplast genome and 45S nrDNAs within *Panax ginseng* Species. *PLoS ONE* **2015**, *10*. [[CrossRef](#)] [[PubMed](#)]
38. Wyman, S.K.; Jensen, R.K.; Boore, J.L. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **2004**, *20*, 3252–3255. [[CrossRef](#)] [[PubMed](#)]
39. Schattner, P.; Brooks, A.N.; Lowe, T.M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acid Res.* **2005**, *33*, W686–W689. [[CrossRef](#)] [[PubMed](#)]
40. Lohse, M.; Drechsel, O.; Bock, R. OrganellarGenomeDRAW (OGDRAW): A tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* **2007**, *52*, 267–274. [[CrossRef](#)] [[PubMed](#)]
41. 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](#)]

