Peculiarities of the inverted repeats in the complete chloroplast genome of Strobilanthes bantonensis Lindau

Wuwei Wu¹, Jingling Li², Yu Liu³, Mei Jiang³, Mingsheng Lan⁴ and Chang Liu⁵

¹Guangxi Botanical Garden of Medicinal Plants, Nanning, China; ²College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China; ³Key Laboratory of Bioactive Substances and Resource Utilization of Chinese Herbal Medicine from the Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China

ABSTRACT
Strobilanthes bantonensis Lindau belongs to the family Acanthaceae. It is an antiviral herb that can be used to prevent Influenza virus infections in the border areas between China and Vietnam. Local people call it ‘Purple Ban-lan-gen’ because its root is very similar to that of Strobilanthes cusia (Nees) Kuntze, which is called ‘Southern Ban-lan-gen’ and is listed in Chinese Pharmacopoeia. The two species have been used interchangeably locally. However, their pharmacological equivalence has caused concern for years. We have sequenced the chloroplast genome of S. cusia previously. In this study, we sequenced the complete chloroplast genome sequence of S. bantonensis to perform in-depth comparative genetic analysis of the two Strobilanthes species. The chloroplast genome of S. bantonensis is a circular DNA molecule with a total length of 144,591 bp and encodes 84 protein-coding, 8 ribosomes, and 37 transfer RNA genes. The chloroplast genome has a conservative quadripartite structure, including a large single-copy (LSC) region, a small single-copy (SSC) region, and a pair of inverted repeat (IR) regions, with lengths of 92,068 bp, 17,767 bp, and 17,378 bp, respectively. Phylogenetic analysis confirmed that S. bantonensis is closely related to the S. cusia. Compared with other species from Acanthaceae, S. bantonensis has a significantly shortened IR region, suggesting the occurrence of IR contraction events. This study will help future taxonomic, evolutionary, phylogenetic, and bioprospecting studies of the sizeable Strobilanthes genus, which contains over 400 species.

1. Introduction
Strobilanthes bantonensis Lindau is a member of the Acanthaceae family. Acanthaceae consists of about 220 genera and over 4000 species, and the genus Strobilanthes contains approximately 400 species (Deng et al. 2011).

The Strobilanthes species are essential herb sources. Strobilanthes cusia (Nees) Kuntze is the source plants of ‘Southern Ban-lan-gen’ in the Chinese Pharmacopoeia (Pharmacopoeia 2015) with multiple pharmacological functions (Ko et al. 2006; Yu-Chi et al. 2020), particular for treating virus-induced diseases, such as SARS (Gu et al. 2015; Chia-Lin et al. 2019) and most recently, COVID-19. In contrast, Strobilanthes crispus has potent anticancer activities (Chin et al. 2017; Akowuah et al. 2020). Extracts of Strobilanthes species were also used to prevent animal and crop diseases (Manaf et al. 2016). For example, S. bantonensis has been used as a tea or animal feed additive to prevent fever (Lingfei 2016).

Strobilanthes bantonensis is mainly distributed in South China and Vietnam (Han et al. 2011). It has been used as a local herb in the Baise area (Shilin et al. 2011). It is also called ‘Purple Ban-lan-gen’ since the back of its leaf is purple, and has been traded as the herb ‘Southern Ban-lan-gen’ because their roots are very similar. However, there have been concerns regarding the equivalence in their pharmacological benefits. Consequently, distinguishing those two species is very important both in scientific research and the assurance of the effectiveness and safety of the herbal products.

The hypervariable sites from the chloroplast genome can be useful markers in discriminating closely related species (Lei et al. 2016). To distinguish S. cusia from other species in genus Strobilanthes, we had sequenced its chloroplast genome in a previous study (Chen et al. 2018). Here we report the complete chloroplast genome of S. bantonensis and hope this will help discriminate S. bantonensis from other plants of genus Strobilanthes and understand the evolutionary history of these plants.

2. Materials and methods
2.1. Sampling, DNA extraction, and sequencing
The fresh leaves of S. bantonensis were collected from Nianjing, Napo, Baise, Guangxi, China (Geospatial coordinate:...
E105.880213, N23.939937). The voucher samples were deposited in the Herbarium of the Guangxi botanical garden of medicinal plants (#451026101000LM51). Total DNA was extracted using the plant genomic DNA kit and sequenced using the Hiseq2500 platform (Illumina, Inc., San Diego, CA).

2.2. Genome assembly and annotation

The chloroplast genome was assembled using NOVOPlasty (v. 2.7.2) (Dierckxsens et al. 2017) with a k-mer length of 39 bp and a conserved gene (rbcL gene) from Arabidopsis thaliana were used as the seed sequence. A total of 326,982 reads were used in the final plastid assembly and a circle genome was obtained. The average sequence coverage was 588.3. The clean reads were mapped to the assembled genome by using Bowtie2 (v.2.0.1) (Langmead 2012) for validating the correctness of the assembled chloroplast genome. The chloroplast genome was annotated using CPGAVAS2 (Shi et al. 2019) with the second option (2544-chloroplast genome was annotated using REPuter/) was used to calculate palindrome repeats, forward repeats, reverse repeats, and complement repeats with the settings: Hamming Distance was three, and Minimal Repeat Size was 30 bp (Kurtz et al. 2001). Tandem repeats were detected with the Tandem Repeats Finder program (v. 4.07b, http://tandem.bu.edu/trf/trf.html) with default setting.

2.3. Repeats and SSR analyses

The GC content was calculated using the cusp program from EMBOSs (v. 6.3.1) (Rice et al. 2000). The simple sequence repeats (SSRs) were identified using MISA (https://webblast.ipk-gatersleben.de/misa/), including mono-, di-, tri-, tetra-, penta-, and hexanucleotides with the minimum numbers were 10, 6, 5, 5, 5, and 5, respectively (Beier et al. 2017). Additionally, REPiter (https://bibiserv.cebitec.uni-bielefeld.de/reputer/) was used to calculate palindromic repeats, forward repeats, reverse repeats, and complement repeats with the settings: Hamming Distance was three, and Minimal Repeat Size was 30 bp (Kurtz et al. 2001). Tandem repeats were detected with the Tandem Repeats Finder program (v. 4.07b, http://tandem.bu.edu/trf/trf.html) with default setting.

2.3. Genome comparison

We used MEGA (v. 6.0) (Tamura et al. 2013) to calculate the Codon usage of protein-coding sequences in S. bantonensis. RSCU (Relative Synonymous Codon Usage) value was calculated to analyze codon preference. The chloroplast genome sequences of S. cusia, were downloaded from GenBank (NC_037485.1). The two sequences were aligned by using MAFFT (v. 7.450) (Rozewicki et al. 2019), and the command was ‘mafft –thread 8 –thread dont 5 –thread 0 –reorder –auto input > output.’ We conducted a sliding window analysis by using DnaSP (DNA Sequences Polymorphism, v. 6.0) to calculate the nucleotide polymorphism (Pi) among the species (window length: 600 bp, step size: 200 bp). Lastly, Irscope (https://irscope.shinyapps.io/irapp/) was used for visualizing the IR boundaries (Amiryousefi et al. 2018).

2.3. Phylogenetic analysis

The chloroplast genome sequences of 10 species belonging to the Acanthaceae family were downloaded from GenBank (Table S1). Two species, Nicotiana tabacum and Arabidopsis thaliana, were used as outgroups. A total of 68 orthologous genes among the 13 species were identified and extracted using Phylosuite (Zhang et al. 2020). They are ACCD, ATPA, ATPB, ATPE, ATPF, ATPH, ATPI, CCSA, CEMA, CLPP, MATK, NDHA, NDHB, NDHC, NDHD, NDHE, NDHF, NDHG, NDHH, NDHI, NDHJ, NDHK, PETA, PETB, PETD, PETG, PETL, PETN, PSAA, PSAB, PSAC, PSAJ, PSBA, PSBB, PSBC, PSBD, PSBE, PSBF, PSBH, PSBI, PSBJ, PSBK, PSBM, PSBT, PSBZ, RBCL, RPL14, RPL16, RPL20, RPL22, RPL23, RPL33, RPL36, RPOA, RPOB, RPOC1, RPS11, RPS14, RPS15, RPS16, RPS18, RPS2, RPS3, RPS4, RPS7, RPS8, YCF3, YCF4. The corresponding protein sequences were aligned using MAFFT (v. 7.450) (Rozewicki et al. 2019) implemented in Phylosuite. These aligned sequences were used to construct the phylogenetic trees by using the maximum-likelihood (ML) method implemented in RaxML (v8.2.4) (Stamatakis 2014). The parameters were ‘-taxmlHPC-PTHREADS-SS13 -f a -N 1000 -m PROTGAMMALGX/ GTRGAMMA -x 551314260 -p 551314260.’ The bootstrap analysis was performed with 1000 replicates.

3. Results

3.1. Structure of S. bantonensis chloroplast genome

A total of 6 G of raw sequencing data were generated. A chloroplast genome sequence was assembled successfully by using NOVOPlasty software. The correctness of the assembly was validated by mapping the raw sequence reads to the assembled genome and obtained a uniform coverage. The chloroplast genome of S. bantonensis is a circular DNA molecule with a total length of 144,591 bp. It has a conservative quadripartite structure, including a large-single copy (LSC) region, a small-single copy (SSC) region, and a pair of inverted repeat region (IR) regions, with the length of 92,068 bp, 17,767 bp, and 17,378 bp, respectively (Table 1). These results are similar to those of S. cusia, which is 91,666 bp, 17,811 bp, and 17,328 bp, respectively. We also compared the GC content of the two species. The total GC

| Table 1. Basic features of the 2 chloroplast genomes. |
|----------------|----------------|----------------|
| Species        | S. bantonensis | S. cusia       |
| Accession number | MN385242       | NC_037485.1    |
| Length (bp)    | 144,591        | 144,133        |
| LSC            | 92,068         | 91,666         |
| SSC            | 17,767         | 17,811         |
| IR             | 17,378         | 17,328         |
| GC content (%) | 38.18          | 38.20          |
| LSC            | 36.48          | 36.52          |
| SSC            | 32.44          | 32.37          |
| IR             | 45.62          | 45.66          |
| Gene numbers   | 129            | 129            |
| trnA gene      | 37             | 37             |
| rRNA gene      | 8              | 8              |
| protein-coding gene | 84          | 84             |
content is 38.18% in *S. bantonensis*, and 38.20% in *S. cusia*. The GC content in IR regions is much higher than that in the LSC or SSC region, consistent with the results observed in most angiosperms (Asaf et al. 2017).

### 3.2. Gene annotation

The chloroplast genome of *S. bantonensis* comprises 129 genes, includes 84 protein-coding, 37 tRNA, and eight rRNA genes (Table 2 and Figure 1). Ten protein-coding genes (rps16, atpF, rpoC1, petB, petD, rpl16, rpl2, ndhA and ndhB (×2)) contain one intron, and two genes (ycf3, clpP) contain two introns. Eight tRNA genes (trnK-UUU, trnG-CAU, trnL-UAA, trnC-ACA, trnL-GAU (×2), trnA-UGC (×2)) contain one intron (Table 3). The length of the protein-coding, tRNA and rRNA genes in *S. bantonensis* chloroplast genome are 71,352 bp, 2813 bp, and 9078 bp respectively, accounting for 49.35%, 6.28%, and 1.95% of the total chloroplast genome length. Remarkably, a total of 16 genes replicated in the IR regions, of which, only 5 are protein-coding genes: psbA, ndhB, rps7, rps12, and ycf1 (only a fragment). This observation was less than that of other higher plants, suggesting that the IR regions may have undergone abnormal contraction or expansion.

### 3.3. Codon usage in protein-coding regions of the chloroplast genome

In terms of codon usage, a total of 23,784 codons were annotated in the chloroplast genome of *S. bantonensis*. The most common codon, AUU, codes for the amino acid Isoleucine (abbreviated I), was recorded 998 times (Table S2). Other common codes include AAA (936) and UUU (915), encoding Lysine (abbreviated K), and Phenylalanine (abbreviated F), respectively. The relative synonymous codon use (RSCU) values are often used to assess the preference for codon usage in protein-coding. When RSCU value more than 1, it means that this codon is used preferentially. We observed that most amino acids have a codon preference except for the methionine (Met) and tryptophan (Trp) codons (Figure S1), which is similar to most higher plants (Sablok et al. 2011). In codons with high RSCU value, they usually end with A/T. For example, Leucine (Leu, RSCU = 1.87), Arginine (Arg, RSCU = 1.79), andAlanine (Ala, RSCU = 1.77) have a high preference for ‘UUA,’ ‘AGA,’ and ‘GCU,’ respectively. The high frequency of A/T usage in protein-coding regions may be the reason that the GC content is much lower than in tRNA and rRNA sequences.

### 3.4. Repeats analysis

We detected a total of 47 SSRs in the two analyzed species, and the results showed that most SSRs were mononucleotide repeats (Figure 2(a) and Table S3). In the number of mononucleotide repeats, *S. bantonensis* is much higher than that of *S. cusia*. However, in the number of polynucleotide (di-, tri-) repeats, the results were reversed, which indicates that SSRs are highly polymorphism in both species. Detail analysis revealed that 23 SSR loci show polymorphism in these two species (Table S4), which could be potential cpSSR markers for species identification.

In the chloroplast genomes of the two *Strobilanthes* species, four types of interspersed repeats were detected. There are 9 forward repeats, 14 palindrome repeats, 3 reverse repeats, and only 1 complementary repeat in *S. bantonensis*. The numbers of interspersed repeats in *S. cusia* are 11, 11, 1, respectively (Figure 2(b) and Table S5). The numbers of these repeats exhibited significant interspecific differences. Besides, the number of tandem repeats is 27 in *S. cusia* and 22 in *S. bantonensis*, satisfying the conditions of length over 30 bp and similarity over 80% (Table S6). These sequences of tandem repeats and interspersed repeats are ubiquitous components of both prokaryotic and eukaryotic genomes (Heslop-Harrison 2000). It is also thought to be essential for promoting chloroplast genome rearrangements (McCann et al. 2018; Lee et al. 2019).

### 3.5. IR contraction event

The contraction and expansion of IR regions are essential for the length diversity in chloroplast genomes (Goulding et al. 1996). In the chloroplast genomes of *S. bantonensis* and
Figure 1. Graphic representation of features identified in the chloroplast genome of *S. bantonensis* by using CPGAVAS2. The map contains four rings. From the center going outward, the first circle shows the forward and reverse repeats connected with red and green arcs. The next circle shows the tandem repeats marked with short bars. The third circle shows the microsatellite sequences identified using MISA. The fourth circle is drawn using drawgenemap and shows the gene structure on the chloroplast genomes. The genes were colored based on their functional categories, which are shown at the left corner.

Table 3. The lengths of introns and exons for the splitting genes.

| Gene  | Strand | Start | End  | ExonI | IntronI | ExonII | IntronII | ExonIII | IntronsI | IntronsII | IntronsIII |
|-------|--------|-------|------|-------|---------|--------|---------|---------|----------|-----------|------------|
| trnK-UUU | –      | 527   | 3056 | 37    | 2457    | 36     |          |          | 153      |           |            |
| rps16  | –      | 3472  | 4617 | 40    | 879     | 227    |          |          |          |           |            |
| trnS-CGA | +      | 7990  | 8767 | 32    | 685     | 61     |          |          |          |           |            |
| atpF   | –      | 10,701| 12,012| 141   | 700     | 471    |          |          |          |           |            |
| rpoC1  | –      | 19,923| 22,853| 432   | 879     | 1620   |          |          |          |           |            |
| ycf3   | –      | 41,905| 43,836| 126   | 722     | 228    |          |          | 703      | 153       |            |
| trnL-UAA | +     | 46,606| 47,166| 35    | 476     | 50     |          |          |          |           |            |
| trnC-ACA | –     | 51,002| 51,669| 38    | 574     | 56     |          |          |          |           |            |
| clpP   | –      | 69,162| 71,108| 71    | 745     | 300    | 605      | 226     |          |           |            |
| petB   | +      | 73,990| 75,373| 6     | 730     | 648    |          |          |          |           |            |
| petD   | +      | 75,562| 76,778| 8     | 683     | 526    |          |          |          |           |            |
| rpl16  | –      | 80,157| 81,455| 9     | 888     | 402    |          |          |          |           |            |
| rpl2   | –      | 83,099| 84,603| 391   | 644     | 470    |          |          |          |           |            |
| ndhB   | –      | 94,435| 96,647| 886   | 584     | 743    |          |          |          |           |            |
| trnL-GAU | +     | 101,908| 102,917| 36   | 917     | 57     |          |          |          |           |            |
| trnA-UGC | +     | 102,982| 103,863| 37   | 809     | 36     |          |          |          |           |            |
| ndhA   | –      | 118,366| 120,548| 553   | 1091    | 539    |          |          |          |           |            |
| trnA-UAG | –     | 132,798| 133,678| 37   | 808     | 36     |          |          |          |           |            |
| trnL-GAU | –     | 133,743| 134,752| 36   | 917     | 57     |          |          |          |           |            |
| ndhB   | +      | 140,013| 142,225| 886  | 584     | 743    |          |          |          |           |            |
S. cusia, we observed a significant contraction of the IR regions (Figure 3). The IR length of these two Strobilanthes species (~17 kb) are both significantly shorter than those of the other four species from Acanthaceae family (~25 kb), which suggests the occurrence of IR contraction events. The shrunk IR regions might cause the overall length of chloroplast genome in Strobilanthes being much shorter than those of other genera. Moreover, we observed a rare distribution of genes in the IR regions of S. bantonensis, which is very similar to that of S. cusia (Chen et al. 2018). In most angiosperms, the trnH and psbA genes are located in the LSC region, and the ycf2 gene is located in the IR regions with two copies. But in S. bantonensis and S. cusia, the gene, trnH, and psbA are found inside of the IR regions unexpectedly, while the ycf2 gene is found outside of the IR regions. So far, this particular distribution of genes in the IR regions has only been observed in these two Strobilanthes species and might be unique to this genus.

3.6. Divergence of chloroplast genome sequences

To evaluate the sequence divergence in the two Strobilanthes species, we used DnaSP software to quantify the levels of DNA polymorphism. There are three regions showed significant differences, which are psbK-psbl (0.03167), trnE-trnT (0.02333), and ycf1 (0.02333) (Figure 4). The Pi values are listed in the parentheses, and they are all exceeding 0.02. These hypervariable regions could be used as potential barcodes to identify species.

Furthermore, there are 605 variable single nucleotide sites (data not shown) and 116 indels (Table S7) in these two chloroplast genome sequences. Most indels involved only one base. We detected a deletion of 223 bp in the intergenic region rps16-trnQ-UUG in S. cusia. Besides, there are a total of 6 indels with a length of more than 30 bp.
We analyzed the variable sites in all protein-coding genes, and 32 of the 78 unique genes had no variable sites and were completely conserved (Table S8). The gene, *ycf1*, had the most variable sites (42), then followed by *ycf2* (39), *rpoC* (17), and *ccsA* (14). Furthermore, the mutation rate for the three genes was more than 1%; they are *ccsA* (1.44%), *ycf3* (1.19%), and *rps15* (1.11%). Two protein-coding genes, *matK* (7, 0.7%) and *rbcL* (5, 0.35%), which are commonly used for DNA barcoding, are not able to distinguish the two species. In comparison, the gene *ycf1* and *ccsA* look promising.

3.7. Phylogenetic analysis

To examine the phylogenetic position of *S. bantonensis*, we constructed the phylogenetic tree between *S. bantonensis* and 12 other species based on the datasets of chloroplast DNA sequences. As shown in Figure 5, phylogenetic relationships based on the chloroplast genomes of 11 species of 6 genera of Acanthaceae resulted in the placement of *S. bantonensis* in the middle and upper part of the phylogenetic tree, with 100% bootstrap support. *Aphelandra* and *Andrographis* are at the base of the phylogenetic trees in Acanthaceae. Then followed by *Justicia* and *Clinacanthus*. As the main clade, we concern about, *S. bantonensis*, and *S. cusia*, are get together as an unambiguous cluster in the phylogenetic tree and form a monophyletic group. *Echinacanthus* was the closest taxa to *Strobilanthes*.

However, due to the maternal inheritance of the plastid genome, these results are limited. Accurate phylogenetic relationships still require a comprehensive analysis of nuclear and organellar genes. The phylogenetic relationships of *E. attenuates* and *Strobilanthes* species require further study. Furthermore, more genome sequencing are needed in the future to determine the relationships among *Strobilanthes* and other species from the family Acanthaceae.
4. Discussion
In summary, we assembled and analyzed a complete chloroplast genome with data generated by next-generation DNA sequencing for Strobilanthes bantonensis. One of the most interesting observations is the peculiar structure of the IR regions.

Recently, five chloroplast genomes from species of genus Dicliptera (Acanthaceae) were sequenced and analyzed (Huang et al. 2020). This study is the first systematic analysis of the completed chloroplast genomes of a genus in Acanthaceae’s large family. We had reported the first chloroplast genomes of genus Strobilanthes of the family Acanthaceae (Chen et al. 2018), and find its Inverted Repeat Region (~17 kb) is much shorter than that of most other angiosperms (~20–28 kb) (Chumley et al. 2006). However, the IR region of the genus Dicliptera is ~25kb, which means the IR contraction is not the characteristics of family Acanthaceae. Here we find that the IR region of S. bantonensis is ~17 kb, similar to that of the S. cusia, suggesting IR contraction may be a characteristic feature of the genus Strobilanthes.

The phylogenetic tree we have constructed is basically consistent with previous research (Ding et al. 2016; Li et al. 2019; Yaradua et al. 2019). Surprisingly, although Echinacanthus was a sister group to Strobilanthes, E. attenuates appears to be closer to Strobilanthes with high bootstrap values than other Echinacanthus species. This is consistent with the previous study (Gao et al. 2018), and this implies that Echinacanthus may not be a monophyletic group. It seems that E. attenuates should be classified as a species of Strobilanthes, and Echinacanthus and Strobilanthes may be more closely related than we previously thought.

The over 400 Strobilanthes species represent a rich source for bioprospecting. Correct identification of these species is a critical step. However, the Strobilanthes species are very similar in morphological characters and highly homoplastic, thus causing difficulty in further classification and differentiation. Different methods have been used to discriminate plants from the Strobilanthes genus, including morphological investigation, chemical analysis, and gene identification (Ni et al. 2012). DNA barcoding is one of the most promising methods. However, the single locus barcode does not have as much genetic information as the chloroplast genome for species discrimination. We are currently collecting a large number of Strobilanthes samples to test molecular markers predicted from this study.

Authors’ contributions
W.W. and C.L. conceived and designed this study. Y.L. and M.S.L. collected the samples. J.J. extracted DNA for next-generation sequencing. M.S.L. assembled the complete chloroplast genome. W.W. and J.J.L. annotated and analyzed the chloroplast genome and wrote the manuscript. All authors have read the work and agreed with its contents.

Funding
This work was supported by Guangxi Science and Technology Program [AD17292003], Guangxi R&D Project of Medical and Health Technology [S201529], National Science & Technology Fundamental Resources Investigation Program of China [2018FY100705], and The National Mega-Project for Innovative Drugs of China [2019ZX09735-002]. The funders were not involved in the study design, data collection, and analysis, decision to publish, or manuscript preparation. We also thank Prof. Yunfei Deng for the help of taxonomy identification of S. bantonensis.

ORCID
Chang Liu http://orcid.org/0000-0003-3879-7302

Data availability statement
The scientific name of the organism is Strobilanthes bantonensis Lindau. The specimen was collected in Nianjing, Napo, Baise, Guangxi, China (Geospatial coordinate: E105.880213, N23.939937). The specimen and its DNA are stored in the Herbarium of the Guangxi botanical garden of medicinal plants (#451026101000LMS1). The data that support the findings of this study are openly available in NCBI at https://www.ncbi.nlm.nih.gov/nuccore/MT576695.

References
Akowuah GA, Chin JH, Yeong SW, Quah SY, Ahmad M. 2020. In-vitro CYP3A4, CYP2E1 and UGT Activity in Human Liver Microsomes by Strobilanthes crispus Leaf Extracts. NPJ. 10(2):104–112.
Amiryousefi A, Hyvo宁en J, Poczai P. 2018. iScope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics. 34(17):3030–3031.
Asaf S, Khan AL, Khan MA, Waqas M, Kang SM, Yun BW, Lee UJ. 2017. Chloroplast genomes of Arabidopsis halleri spp. gemmifera and Arabidopsis lyrata ssp. petraea: structures and comparative analysis. Sci Rep. 7(1):7556.
Beier S, Thiel T, Munch T, Scholz U, Mascher M. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics. 33(16):2583–2585.
Chen H, Shao J, Zhang H, Jiang M, Huang L, Zhang Z, Yang D, He M, Ronaghi M, Luo X. 2018. Sequencing and analysis of Strobilanthes cusia (Nees) Kunze chloroplast genome revealed the rare simultaneous contraction and expansion of the inverted repeat region in angiosperm. Front Plant Sci. 9:324.
Chia-Lin L, Chien-Ming W, Hao-Chun H, Hung-Rong Y, Ying-Chyi S, Sheng-Jie Y, Yang-Chang W. 2019. Indole alkaloids indigodoles A-C from aerial parts of Strobilanthes cusia in the traditional Chinese medicine Qing Dai have anti-IL-17 properties. Phytochemistry. 162: 39–46.
Chin JE, Ikhutfar MRN, Kho CH, Ho KL, Cheah YK. 2017. Analysis of chemical constituents, antimicrobial and anticancer activities of dichloromethane extracts of Sordariomyces sp. endophytic fungi isolated from Strobilanthes crispus. World J Microbiol Biotechnol. 33(1):5.
Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK. 2006. 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. 23(11):2175–2190.
Deng Y, J-Q H, Daniel TF, Wood J, Wood JRL. 2011. Acanthaceae. In: Flora of China, Vol. 19. Beijing: Science Press.
Diercksens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.
Ding P, Shao Y, Li Q, Gao J, Zhang R, Lai X, Wang D, Zhang H. 2016. The complete chloroplast genome sequence of the medicinal plant Andrographis paniculata. Mitochondrial DNA A DNA Mapp Seq Anal. 27(4):2347–2348.
Gao C, Deng Y, Wang J. 2018. The complete chloroplast genomes of Echinacanthus species (Acanthaceae): phylogenetic relationships, adaptive evolution, and screening of molecular markers. Front Plant Sci. 9:1989.
Goulding SE, Olmstead RG, Morden CW, Wolfe KH. 1996. Ebb and flow of the chloroplast inverted repeat. Mol Gen Genet. 252(1-2):195–206.
