A peer-reviewed version of this preprint was published in PeerJ on 7 November 2017.

View the peer-reviewed version (peerj.com/articles/4018), which is the preferred citable publication unless you specifically need to cite this preprint.

Kong H, Liu W, Yao G, Gong W. 2017. A comparison of chloroplast genome sequences in *Aconitum* (Ranunculaceae): a traditional herbal medicinal genus. PeerJ 5:e4018 https://doi.org/10.7717/peerj.4018

A comparison of chloroplast genome sequences in *Aconitum* (Ranunculaceae): a traditional herbal medicinal genus

Hanghui Kong [1, 2] , **Wanzhen Liu** [3] , **Gang Yao** [4] , **Wei Gong** Corresp. [3]

[1] Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China
[2] Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China
[3] College of Life Sciences, South China Agricultural University, Guangzhou, China
[4] College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou, China

Corresponding Author: Wei Gong
Email address: wgong@scau.edu.cn

The herbal medicinal genus *Aconitum* L., belonging to the Ranunculaceae family, represents the earliest diverging lineage within the eudicots. It currently comprises of two subgenera, *A*. subgenus *Lycoctonum* and *A*. subg. *Aconitum*. The complete chloroplast (cp) genome sequences were characterized in three species: *A. angustius*, *A. finetianum*, and *A. sinomontanum* in subg. *Lycoctonum* and compared to other *Aconitum* species to clarify their phylogenetic relationship and provide molecular information for utilization of *Aconitum* species particularly in Eastern Asia. The length of the chloroplast genome sequences were 156,109 bp in *A. angustius*, 155,625 bp in *A. finetianum* and 157,215 bp in *A. sinomontanum*, with each species possessing 126 genes with 84 protein coding genes (PCGs). While genomic rearrangements were absent, structural variation was detected in the LSC/IR/SSC boundaries. Five pseudogenes were identified, among which *Ψrps*19 and *Ψycf*1 were in the LSC/IR/SSC boundaries, *Ψrps*16 and *Ψinf*A in the LSC region, and *Ψycf*15 in the IRb region. The nucleotide variability (*Pi*) of *Aconitum* was estimated to be 0.00549, with comparably higher variations in the LSC and SSC than the IR regions. Eight intergenic regions were revealed to be highly variable and a total of 58 – 62 simple sequence repeats (SSRs) were detected in all three species. More than 80% of SSRs were present in the LSC region. Altogether, 64.41% and 46.81% of SSRs are mononucleotides in subg. *Lycoctonum* and subg. *Aconitum*, respectively, while a higher percentage of di-, tri-, tetra-, and penta-SSRs were present in subg. *Aconitum*. Most species of subg. *Aconitum* in Eastern Asia were first used for phylogenetic analyses. The availability of the complete cp genome sequences of these species in subg. *Lycoctonum* will benefit future phylogenetic analyses and aid in germplasm utilization in *Aconitum* species.

A comparison of chloroplast genome sequences in *Aconitum* (Ranunculaceae): a traditional herbal medicinal genus

Hanghui Kong[1, 4*], Wanzhen Liu [2*], Gang Yao[3], Wei Gong[2#]

[1] *Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China*

[2] *College of Life Sciences, South China Agricultural University, Guangzhou, 510614, China*

[3] *College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou, 510614, China*

[4] *Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China*

Corresponding author

*These authors contributed equally to this work

Corresponding Author: Wei Gong

College of Life Sciences, South China Agricultural University, Guangzhou 510614, China

E-mail: wgong@scau.edu.cn

Tel: +86 20-37088127; Fax: +86 20-37088127

Abstract

The herbal medicinal genus *Aconitum* L., belonging to the Ranunculaceae family, represents the earliest diverging lineage within the eudicots. It currently comprises of two subgenera, *A.* subgenus *Lycoctonum* and *A.* subg. *Aconitum*. The complete chloroplast (cp) genome sequences were characterized in three species: *A. angustius*, *A. finetianum*, and *A. sinomontanum* in subg. *Lycoctonum* and compared to other *Aconitum* species to clarify their phylogenetic relationship and provide molecular information for utilization of *Aconitum* species particularly in Eastern Asia. The length of the chloroplast genome sequences were 156,109 bp in *A. angustius*, 155,625 bp in *A. finetianum* and 157,215 bp in *A. sinomontanum*, with each species possessing 126 genes with 84 protein coding genes (PCGs). While genomic rearrangements were absent, structural variation was detected in the LSC/IR/SSC boundaries. Five pseudogenes were identified, among which *Ψrps*19 and *Ψycf*1 were in the LSC/IR/SSC boundaries, *Ψrps*16 and *Ψinf*A in the LSC region, and *Ψycf*15 in the IRb region. The nucleotide variability (*Pi*) of *Aconitum* was estimated to be 0.00549, with comparably higher variations in the LSC and SSC than the IR regions. Eight intergenic regions were revealed to be highly variable and a total of 58 – 62 simple sequence repeats (SSRs) were detected in all three species. More than 80% of SSRs were present in the LSC region. Altogether, 64.41% and 46.81% of SSRs are mononucleotides in subg. *Lycoctonum* and subg. *Aconitum*, respectively, while a higher percentage of di-, tri-, tetra-, and penta- SSRs were present in subg. *Aconitum*. Most species of subg. *Aconitum* in Eastern Asia were first used for phylogenetic analyses. The availability of the complete cp genome sequences of these species in subg. *Lycoctonum* will benefit future phylogenetic analyses and aid in germplasm utilization in *Aconitum* species.

INTRODUCTION

The chloroplast (cp) is an intracellular organelle that plays an important role in the process of photosynthesis and it is widely present in algae and plants (*Neuhaus & Emes, 2000; Inoue, 2011*). The cp genome in angiosperms is a circular DNA molecule with a typically quadripartite structure, consisting of two copies of a large inverted repeat (IR) region that separates a large-single-copy (LSC) region from a small-single-copy (SSC) region (*Raubeson & Jansen, 2005*; *Yang et al., 2010; Gree, 2011*; *Wicke et al., 2011*). Although highly conserved among plants, some differences in gene synteny, copy number and pseudogenes have been observed in cp genome structures (*Shradha et al., 2010*; *Lei et al., 2016*; *Ivanova et al., 2017*). A complete cp genome is valuable for plant taxonomical analyses, phylogenetic reconstructions, speciation processes, and biogeographical inferences at different taxonomic levels. The cp genome is useful in investigating the maternal origin in plants, especially those with polyploid species, due to their haploid maternal inheritance and high conservation in gene content and genome structure (*Birky, 1995*; *Soltis & Soltis, 2000*; *Song et al, 2002*). High-throughput sequencing technologies have enabled a rapid increase in the completion of cp genomes and have shifted the study of phylogenetics to phylogenomics. Highly informative universal markers based on indels, substitutions, and inversions of the cp genome have been further developed for various molecular studies in plants.

The genus *Aconitum* L. belongs to the tribe Delphinieae in the Ranunculaceae family and represents one of the earliest diverging lineages within the eudicots APG IV (*Wang et al., 2009*; *Sun et al., 2011*; *The Angiosperm Phylogeny Group, 2016*). It is currently divided into two subgenera, *A*. subgenus *Lycoctonum* and *A*. subgenus *Aconitum*, comprising about more than 400 species throughout Eurasia and North America with its diversification center in Eastern Asia

(*Utelli, Roy & Baltisberger*; 2000; *Jabbour & Renner, 2012*; *Wang et al., 2013*). Polyploid species were identified in both subgenera, particularly in subg. *Lycoctonum*. One of the tetraploid species in subg. *Lycoctonum* is *A. angustius* (2n = 4x = 32), which possesses heterologous chromosomes and is hypothesized to be a hybrid of *A. finetianum* (2n = 2x = 16) and *A. sinomontanum* (2n = 2x = 16) (*Gao, 2009*; *Kong et al., 2017b*). The three species display intermediate morphological characteristics and overlapping geographical distributions (*Shang & Lee, 1984*; *Yuan & Yang, 2006*; *Gao, 2009*; *Gao, Ren & Yang, 2012*). Based on previous morphological analysis and phylogenetic inference, *A. finetianum* was inferred to be the putative maternal progenitor of *A. angustius* (*Gao, 2009*; *Kong et al., 2017b*).

The genus *Aconitum* is known as a taxonomically and phylogenetically challenging taxon. Early divergence between subg. *Lycoctonum* and subg. *Aconitum* in Europe was suggested based on *trn*H-*psb*A and ITS (*Utelli, Roy & Baltisberger*; 2000). Although high morphological variability within and among populations was detected due to recent speciation, the morphological characteristics are poor indicators of relatedness. *Jabbour & Renner* (2012) conducted a phylogenetic reconstruction focusing on Delphineae based on *trn*L-F and ITS that suggested *Aconitum* was monophyletic clade and a sister group of *Delphinium*. However, few species from Eastern Asia were used, which may have affected the previous phylogenetic analysis. Most recently, phylogenetic inferences of polyploid species relationships in subg. *Lycoctonum* were made using four cpDNA intergenic regions (*ndh*F-*trn*L, *psb*A-*trn*H, *psb*D-*trn*T, and *trn*T-L) and two nrDNA regions (ITS and ETS) (*Kong et al., 2017b*), *Aconitum finetianum* was inferred as the maternal progenitor of *A. angustius*. With the same cpDNA intergenic regions, taxonomical revision has been conducted based on phylogenetic analyses of subg. *Lycoctonum* by *Hong et al.* (2017), yet phylogenetic information at the genomics level has been

absent.

Although some *Aconitum* species are highly toxic because of aconite alkaloid, many species are essential in the formulation of traditional herbal medicine in Asia (*Zhao et al., 2010*; *Semenov et al., 2016*; *Wada et al., 2016*; *Liang et al., 2017*). The current state of *Aconitum* phylogenetics lacks molecular information of some species in Eastern Asia, and thus inhibits identification and germplasm utilization of this genus. In this study, we report the complete cp genome sequences of three species in subg. *Lycoctonum*; we established and characterized the organization of the cp genome sequences of tetraploid *A. angustius* as well as diploid *A. finetianum* and *A. sinomontanum*. We further compared the structure, gene arrangement and microsatellite repeats (SSRs) with the related species in both subgenera of *Aconitum*. Altogether, 14 species and 2 varieties from *Aconitum* were used for phylogenetic reconstruction at the genomic level. Seven previously unanalyzed species from the subg. *Aconitum* in Eastern Asia were investigated for phylogenetic relationships, and the maternal origin of *A. finetianum* was explored in the tetraploid, *A. angustius*. Our results provide cp genomic information for taxonomical identification, phylogenetic inference, or the population history of *Aconitum* or Ranunculaceae, which can also aid in the utilization of the genetic resources of *Aconitum* as a traditional herbal medicine.

MATERIALS AND METHODS

Plant samples and DNA extraction

Fresh leaves were collected from *A. angustius*, *A. finetianum* and *A. sinomontanum* growing in the greenhouse of South China Botanical Garden, Chinese Academy of Sciences. Total genomic DNA was extracted from the fresh leaves of *A. angustius*, *A. finetianum* and *A. sinomontanum*

using the modified CTAB method (*Dolye & Dolye, 1987*). The DNA concentration was quantified using a Nanodrop spectrophotometer (Thermo Scientific, Carlsbad, CA, USA), and a final DNA concentration of >30 ng/μL was used for Illumina sequencing.

Chloroplast genome sequencing, assembly and annotation

We sequenced the complete cp genome of *A. angustius*, *A. finetianum* and *A. sinomontanum* with an Illumina HiSeq 2000 at Beijing Genomics Institute (BGI) in Wuhan, China. Genomic DNA was fragmented randomly and then the required length of DNA fragments was obtained by electrophoresis. Adapters were ligated to DNA fragments followed by cluster preparation and sequencing. A paired-end library was constructed with 270 bp insert size, and then 150 bp paired reads were sequenced using an Illumina HiSeq 2000.

We assembled the cp genomes using Geneious 9.1.4 (Biomatters Ltd., Auckland, New Zealand) with BLAST and map reference tools, respectively. Using the program DOGMA (http://dogma.ccbb.utexas.edu/) (*Wyman, Jansen & Boore, 2004*) and Geneious, annotation was performed in comparison with the cp genomes of *A. reclinatum* (MF186593) (*Kong et al., 2017a*), *A. barbatum* var. *puberulum* (KC844054) (*Chen et al., 2015*), and *A. barbatum* var. *hispidum* (KT820664) in subg. *Lycoctonum* as well as 10 species from the subg. *Aconitum* (*Choi et al., 2016; Kim et al., unpublished; Lim et al., 2017; Yang, unpublished; Yang et al., unpublished*) (Table 1). Altogether, 14 species and 2 varieties in both subgenera of *Aconitum* were used for annotation (Table 1). Among those species, *A. angustius*, *A. finetianum*, *A. sinomontanum*, *A. barbatum* var. *hispidum*, and *A. barbatum* var. *puberulum* were collected from China (*Chen et al., 2015*), *A. reclinatum* came from the United States (*Kong et al., 2017a*), while the remaining species were all sampled from Korea (*Choi et al., 2016; Kim et al., unpublished;*

139 *Lim et al., 2017; Yang, unpublished; Yang et al., unpublished*). Chloroplast genome sequences of
140 *Aconitum* species from Europe were not available in GenBank.

141 The annotation of tRNA genes were confirmed using the ARAGORN program (*Laslett &*
142 *Canback, 2004*), and then manually adjusted using the program Geneious. Contigs with BLAST
143 hits to consensus sequence from the "map to reference function" were assembled manually to
144 construct complete chloroplast genomes. Finally, the circular genome maps of the three species
145 were illustrated using the Organellar Genome DRAW tool (OGDRAW, http://ogdraw.mpimp-
146 golm.mpg.de/) (*Lohse et al., 2013*). The annotated chloroplast genomic sequences of *A.*
147 *angustius*, *A. finetianum* and *A. sinomontanum* have been submitted to GenBank (Accession
148 Number: MF155664, MF155665 and MF155666).

149

150 **Genome comparison and divergence hotspot**

151 The cp genome sequences from the finalized data set were aligned with MAFFT v7.0.0 (*Katoh &*
152 *Standley, 2012*) and adjusted manually when necessary. Based on many other cp genome studies,
153 the IRs expansion/contraction could lead to changes in the structure of the cp genome, leading to
154 the length variation of angiosperm cp genomes and contributing to the formation of pseudogenes
155 (*Kim & Lee, 2004*; *Nazareno, Carlsen & Lohmann, 2015*; *Ivanova et al., 2017*). Therefore, we
156 conducted comparative analysis to detect the variation in the LSC/IR/SSC boundaries among the
157 species/varieties. Comparative analysis of the nucleotide diversity (*Pi*) among the complete cp
158 genomes of *Aconitum* was performed based on a sliding window analysis using DnaSP 5.10
159 (*Librado & Rozas, 2009*). The window length was 600 bp and step size was 200 bp. To test and
160 visualize the presence of genome rearrangement and inversions, gene synteny was performed
161 using MAUVE as implemented in Geneious with default settings based on 14 species and 2

varieties in both subgenera.

Simple sequence repeats analysis

MISA (http://pgrc.ipk-gatersleben.de/misa/misa.html) (*Thiel et al., 2003*) is a tool for the identification and location of perfect microsatellites and compound microsatellites (two individual microsatellites, disrupted by a certain number of bases). We used MISA to search for potential simple sequence repeats (SSRs) loci in the cp genomes of the three species. The minimum number (thresholds) of SSRs was set as 10, 5, 4, 3, and 3 for mono-, di-, tri-, tetra-, and penta-nucleotides SSRs, respectively. All of the repeats found were manually verified and the redundant ones were removed.

Phylogenetic analysis

Four species and two varieties in subg. *Lycoctonum* and 10 species in subg. *Aconitum* were used for phylogenetic reconstruction, with *Megaleranthis saniculifolia and Clematis terniflora*v as the outgroup. Except for *A. kusnezoffii*, *A. volubile*, and *A. ciliare*, the remaining seven species in subg. *Aconitum* from Korea were first used for phylogenetic analysis. The complete cp genome sequences and PCGs were used for the phylogenetic reconstruction of *Aconitum* species in Eastern Asia. Three different methods including Bayesian Inference (BI), Maximum Parsimony (MP), and Maximum Likelihood (ML) were employed. In all analyses, gaps were treated as missing.

Bayesian Inference (BI) of the phylogenies was performed using MrBayes v.3.2 (*Huelsenbeck & Ronquist, 2001*; *Ronquist & Huelsenbeck, 2003*). The best model was determined for each sequence partition, after comparisons among 24 models of nucleotide

substitution using Modeltest v.3.7 (*Posada & Crandall, 1998*). We performed MP using PAUP* v.4.0b10 (*Swofford, 2002*). We calculated the bootstrap values with 1000 bootstrap replicates, each with 10 random sequence addition replicates holding a single tree for each run. We conducted ML using RAxML (*Stamatakis, 2006*) and the RAxML graphical interface (raxmlGUI v.1.3 (*Silvestro & Michalak, 2012*) with 1000 rapid bootstrap replicates. The general time-reversible (GTR) model was chosen with a gamma model for the rate of heterogeneity.

RESULTS AND DISCUSSION

Genome Organization and Features

Using the Illumina HiSeq 2000 sequencing platform, a total number of 2x150 bp pair-end reads ranging from 9,879,068 to 27,530,148 bp were produced for the three species in subg. *Lycoctonum*. Altogether, 1,270 Mb of clean data were produced for *A. angustius*, 3,586 Mb for *A. finetianum,* and 3,590 Mb for *A. sinomontanum*. The assembly generated an average of 6,713 contigs with a N50 length of 732 bp for *A. angustius*, an average of 6,201 contigs with a N50 length of 801 bp bp for *A. finetianum*, and an average of 6,999 contigs with a N50 length of 769 bp for *A. sinomontanum*. Scaffolds from the assembly with k-mer values of 35 to 149 were matched to reference cp genome sequences, which were used to determine the relative position and direction respectively. We generated a new draft chloroplast genome by manually identifying the overlapping regions. To further refine the draft genome, the quality and coverage of each was double-checked by remapping reads. The complete cp genome sequences of the three species with full annotations were deposited into GenBank.

The size of the cp genomes was 156,109 bp for *A. angustius*, 155,625 bp for *A. finetianum* and 157,215 bp for *A. sinomontanum* (Table 1). The chloroplast genomes displayed a typical

quadripartite structure, including a pair of IRs (25,927-26,225 bp) separated by LSC (86,664-88,074 bp) and SSC (16,914-17,107 bp) regions (Fig. 1 and Table 1). The GC content of the three cp genomes was 38.00%, demonstrating congruence with other *Aconitum* species (38.00% or 38.10%) (Table 1).

When duplicated genes in the IR regions were counted only once, each of the three cp genomes encode 126 predicted functional genes, including 84 PCGs, 38 tRNA genes, and four rRNA genes. The remaining non-coding regions include introns, intergenic spacers, and pseudogenes. Altogether 18 genes were duplicated in the IR regions, including seven PCGs, seven tRNA genes, and four rRNA genes (Fig. 1; Table S1). Each of the thirteen genes (eight PCGs and five tRNA genes) contained one interval, and three PCGs (*clp*P, *ycf*3 and *rps*12) had two intervals each (Table S1). The maturase K (*mat*K) gene in the cp genomes of the three species is located within *trn*K intron, which is similar in most of the other plants species (*Kong & Yang, 2017*). In the IR regions, the four rRNA genes and two tRNA genes (*trn*I and *trn*A) are clustered as 16S-*trn*I-*trn*A-23S-4.5S-5S. This has also been reported in the cp genomes of *A. barbatum* var. *hispidum*, *A. barbatum* var. *puberulum*, and many other plant species (*Mardanov et al., 2008*; *Wu et al., 2014*; *Chen et al., 2015*).

Comparative analysis of genomic structure

Synteny analysis identified a lack of genome rearrangement and inversions in the cp genome sequences of the *Aconitum* species. No gene rearrangement and inversion events were detected (Fig. S1). Genomic structure, including gene number and gene order, is highly conserved among the *Aconitum* species; however, structural variation was still present in the LSC/IR/SSC boundaries (Fig. 2). The genes *rps*19-*rp*l2-*trn*H and *ycf*1-*ndh*F were located between the

junction of the LSC/IR and SSC/IR regions. The *rps*19 gene crosses the LSC/IRa junction region in *A. sinomontanum*, *A. barbatum* var. *puberulum* and *A. barbatum* var. *hispidum* of subg. *Lycoctonum*, as well as in *A. jaluense*, *A. volubile*, *A. carmichaelii*, *A. kusnezoffii* and *A. monanthum* of subg. *Aconitum*. As a result, the *rps*19 gene has apparently lost its protein-coding ability due to being partially duplicated in the IRb region, thus a producing pseudogenized *Ψrps*19 gene. The same was found with the *ycf*1 gene, as the IRb/SSC junction region is located within the *ycf*1 CDS region and only a partial gene is duplicated in the IRa region, resulting in a pseudogene. This is a general structure among the dicots. The *Ψycf*1 pseudogene in the IR region was 1,279 bp for two varieties in subg. *Lycoctonum* and seven species in subg. *Aconitum*. However, length variation was present in the IR of the remaining six species: 1,292 bp in *A. angustius*, *A sinomontanum*, and *A. reclinatum*; 1,165 bp in *A. finetianum*; 1,274 bp in *A. chiisanense*; 1,356 bp in *A. volubile*; and 1,263 bp in *A. carmichaelii* (Fig. 2; Table 2).

Three pseudogenes, *Ψycf*15, *Ψrps*16, and *Ψinf*A, were identified in the gene annotations (Table 2). The *Ψycf*15 gene is pseudolized in *A. austrokoreense* and *A. chiisanense* with four base insertions and pseudolized in *A. monanthum* with a one base insertion, contributing to several internal stop codons. The *Ψinf*A region is pseudogenized with two nonsynonymous substitutions producing internal stop codons in all of the members of subg. *Lycoctonum*. This pseudogenized *Ψinf*A gene has also been found in other angiosperm chloroplast genomes (*Raman & Park, 2015*; *Lu, Li & Qiu, 2017*). The gene *rps*16 encodes the ribosomal protein S16 and is present in the cp genome of most if the higher plants. However, *rps*16 has been functionally lost in various plant species (*Shradha et al, 2010*). A pseudogene *Ψrps*16 was also present in the cp genomes of *A. angustius*, *A. finetianum* and *A. reclinatum* in subg. *Lycoctonum* as well as in the nine species in subg. *Aconitum* due to the loss of one CDS region (Table 2). As

has been revealed in other studies, the functional loss of the *rps*16 gene might be compensated by the dual targeting of the nuclear *rps*16 gene product (*Keller et al., 2017*).

Sequence divergence among the species in *Aconitum*

The average nucleotide variability (*Pi*) values were estimated to be 0.00549, ranging from 0 to 0.03856, based on the comparative analysis of cp genome sequences in *Aconitum* species. The highest variation was found in the LSC and SSC regions, with an average *Pi* = 0.007140 and 0.008368, respectively. The IR regions had a much lower nucleotide diversity with *Pi* = 0.001079 and 0.001459. Eight intergenic regions (*trn*H-*psb*A, *trn*K-*rps*16, *trn*D-*trn*Y, *trn*Y-*trn*E, *trn*E-*trn*T, *trn*T-*trn*L, *rpl*12-*clp*P and *trn*H-*trn*R) were highly variable, with *Pi* value ~ 0.023 (Fig. 3). The former eight loci are present in the LSC, while the pseudogene *Ψycf*1 is in the SSC region. The single-copy regions have been demonstrated to be highly variable with loci clustered in 'hot spots' (*Kong & Yang, 2017*). Among the eight intergenic regions, *trn*H-*psb*A and *trn*T-*trn*L are variable and useful for phylogenetic reconstruction in the subg. *Lycoctonum* (*Utelli, Roy & Baltisberger, 2000*; *Kong et al., 2017b*). However, the other intergenic regions, even with higher nucleotide variability, have never been involved in the phylogenetic analysis for the genus *Aconitum*. The highly variable loci detected in the current study may provide a basis for further phylogenetic characterization of this genus. The observed divergence hotspot regions provide abundant information for marker development in phylogenetic analysis or conservation genetics of *Aconitum*.

Characterization of simple sequence repeats

MISA was used to identify SSRs with minimum a of 10 bp repeats among the three species. In *A.*

angustius, 60 SSRs were found, while 62 SSRs were found in *A. finetianum*, and 58 in *A. sinomontanum*. This result is comparable with *A. reclinatum* (61 SSRs), *A. barbatum* var. *hispidum* (53 SSRs), and *A. barbatum* var. *puberulum* (57 SSRs). An average of 59 SSRs were identified in subg. *Lycoctonum*, which is relatively higher than that of subg. *Aconitum* (47). In both subgenera, most SSRs are in the LSC regions, accounting for an average of 85.31% and 80.85% in subg. *Lycoctonum* and subg. *Aconitum*, respectively. Among all of the SSRs, the mononucleotide A/T repeat units occupied the highest proportion, with 64.41% and 46.82% of the total SSRs in subg. *Lycoctonum* and subg. *Aconitum*, respectively. Although few SSRs were detected in subg. *Aconitum*, a higher proportion of di-, tri-, tetra- and penta-nucleotide repeats were detected (Table 3). The SSRs have a remarkably high A/T content with only seven SSRs, namely $(ATCT)_3$, $(TTCT)_3$, $(CTTT)_3$, $(TAAAG)_3$, $(TTTC)_3$, $(ATAC)_3$ and $(CATT)_3$, that contain one C or G nucleotide.

A total of 11 cp SSR loci were shared among the cp genomes of tetraploid *A. angustius* and diploid *A. finetianum*. No common cp SSRs were specifically found between *A. angustius* and *A. sinomontanum*. This result provides evidence of the maternal origin of the tetraploid *A. angustius* from diploid *A. finetianum*, which is consistent with previous research (*Gao, 2009*; *Kong et al., 2017b*). Among the three species, the highest number of unique SSRs loci were present in *A. sinomontanum* (11) followed by *A. angustius* (7), *A. finetianum* (6), and *A. reclinatum* (5).

Phylogenetic analyses

In the present study, three phylogenetic methods (BI, MP and ML) resulted in identical phylogenetic trees within each data set. Different analyses based on the two datasets generated largely congruent topologies (Fig. 4). The total aligned length with parsimony informative loci

was 178,392 bp with 4,342 for the complete cp genome sequences, and 106,535 bp with 3,164 for PCGs, respectively. All of the phylogenetic trees support that *Aconitum* comprises two monophyletic subgenera. High Bayesian posterior probabilities and bootstrap values were detected at most nodes, particularly based on the complete cp genomes (Fig. 4A).

The phylogenetic relationship of Korean species in subg. *Aconitum* was investigated for the first time. The monophyletic clade was formed by *A. ciliare*, *A. carmichaelii*, *A. japonicum* subsp. *napiforme* and *A. kusnezoffii*, with strong support values (Fig. 4). The clade comprised of *A. jaluense* subsp. *jaluense* and *A. volubile* exhibited moderate-to-high support, forming a monophyletic sister group. The positions of the four species *A. ciliare*, *A. carmichaelii*, *A. austrokoreense*, and *A. chiisanense*, demonstrated inconsistencies based on the two data sets. Obviously, these species received stronger support based on the sequences of the complete cp genome rather than PCGs, indicating that whole genomes are more efficient in determining phylogenetic relatedness in *Aconitum* than PCGs alone.

Based on the phylogenetic tree, the tetraploid *A. angustius* was always closely related with diploid *A. finetianum*, which further supports previous research (*Kong et al., 2017b*). The two species co-occur on several mountains in southeast China and even grow very closely within a community (*Yuan & Yang, 2006*). They show similar morphological characteristics in having 3-part leaves, the cylindric upper sepals and retrosely pubescent pedicels, resulting in common misidentification (*Gao, Ren & Yang, 2012*). *Aconitum finetianum* is the most likely maternal progenitor of *A. angustius* based on both molecular phylogenetic and morphological evidence (*Kong et al., 2017*); therefore, it is reasonable to see that the two species have a close phylogenetic relationship.

The five pseudogenes exhibit different evolutionary histories from each other. Concerning

the evolution of *Ψycf15*, it occurs in only three species *A. monanthum, A. austrokoreense*, and *A. chiisanense* of subgen. *Aconitum*, which was probably pseudogenized once in each species independently and subsequently restored to a functional copy. We propose that *Ψrps16* was pseudogenized during the divergence between the two subgenera and restored to a functional copy within the *A. sinomontanum-A. barbatum* clade of subgen. *Lycoctonum*. With respect to *Ψrps19*, it appears to have been pseudogenized multiple times independently in phylogenetically distant species of the two subgenera. *Ψycf1* is commonly found among cp genomes of plant species. Within *Aconitum*, *Ψycf1* exhibits length variation and multiple convergent mutation events, which are not consistent with the phylogenetic relationships of the genus. Only *ΨinfA* shows an evolutionary history congruent with the phylogeny of *Aconitum* (Fig. 4B; Table 2). Overall, our results show that similarities among pseudo-gene sequences do not necessarily predict phylogenetic relationships among species.

Acknowledgement

We would like to provide a special thank to Dr. Tongjian Liu for his assistance in lab work and data analyses. We thank Dr. AJ Harris and LetPub (www.letpub.com) for their linguistic assistance during the preparation of this manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences: GenBank accession number: MF155664, MF155665 and MF155666.

Data Availability

The following information was supplied regarding data availability: The raw data can be found in https://doi.org/10.6084/m9.figshare.5092414.v1, https://doi.org/10.6084/m9.figshare.5092420.v1 and with the GenBank accession numbers in Table 1.

Supplemental Information

Supplemental information for this article can be found online.

REFERENCES

Birky CW. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America* **92(25):**11331–11338 DOI 10.1073/pnas.92.25.11331.

Chen XC, Li QS, Li Y, Qian J, Han JP. 2015. Chloroplast genome of *Aconitum barbatum* var. *puberulum* (Ranunculaceae) derived from CCS reads using the PacBio *RS* platform. *Frontiers in Plant Science* **6(42):**1–9 DOI 10.3389/fpls.2015.00042.

Choi JE, Kim GB, Lim CE, Yu HJ, Mun JH. 2016. The complete chloroplast genome of *Aconitum austrokoreense* Koidz. (Ranunculaceae), an endangered endemic species in Korea. *Mitochondrial DNA B Resour* **1(1):** 688–689 DOI 10.1080/23802359.2016.1219644

Doyle JJ, Doyle JL, Rausher J, Brown AHD. 2004. Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). *New Phytologist* **161:**121–132 DOI 10.1046/j.1469-8137.2003.00949.x.

Gao Q. 2009. Taxonomic revision of *Aconitum* L. subgenus *Lycoctonum* (DC.) *Peterm.* (Ranunculaceae) from China. D. Phil. Thesis, Institute of Botany, the Chinese Academy

of Sciences, Beijing, China

Gao Q, Ren C, Yang QE. 2012. Taxonomic status and distributional range of *Aconitum angustius* (Ranunculaceae) based on cytological evidence. *Nordic Journal of Botany* **30:**1–13 DOI 10.1111/j.1756-1051.2012.01506.x.

Green BR. 2011. Chloroplast genomes of photosynthetic eukaryotes. *the Plant Journal* **66:**34–44 DOI 10.1111/j.1365-313X.2011.04541.x.

Hong Y, Luo Y, Gao Q, Ren C, Yuan Q, Yang QE. 2017. Phylogeny and reclassification of *Aconitum* subgenus *Lycoctonum* (Ranunculaceae). *PLoS ONE* **12(1):**e0171038 DOI 10.1371/journal.pone.0171038.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:**754–755 DOI 10.1093/bioinformatics/17.8.754.

Inoue K. 2011. Emerging roles of the chloroplast outer envelope membrane. *Trends in Plant Science* **16(10):**550–557 DOI 10.1016/j.tplants.2011.06.005.

Ivanova Z, Sablok, G, Daskalova E, Zahmanova G, Apostolova E, Yahubyan G, Baev V. 2017. Chloroplast genome analysis of resurrection Tertiary relic *Haberlea rhodopensis* highlights genes important for desiccation stress response. *Frontiers of Plant Sciences* **8:**204 DOI 10.3389/fpls.2017.00204.

Jabbour F, Renner SS. 2012. A phylogeny of *Delphinieae* (Ranunculaceae) shows that *Aconitum* is nested within *Delphinium* and that Late Miocene transitions to long life cycles in the Himalayas and Southwest China coincide with bursts in diversification. *Molecular Phylogenetics and Evolution* **62:**928–942 DOI 10.1016/j.ympev.2011.12.005.

Katoh K, Frith MC. 2012. Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics* **28(23):**3144–3146 DOI

10.1093/bioinformatics/bts578.

Kearse M, Moir R, Wilson, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:**1647–1649 DOI 10.1093/bioinformatics/bts199.

Keller J, Rousseau-Gueutin M, Martin GE, Morice J, Boutte J, Coissac E, Ourari M, Aïnouche M, Salmon A, Cabello-Hurtado F, Aïnouche A. 2017. The evolutionary fate of the chloroplast and nuclear *rps16* genes as revealed through the sequencing and comparative analyses of four novel legume chloroplast genomes from *Lupinus*. *DNA Research* in press DOI 10.1093/dnares/dsx006.

Kim KJ, Lee HL. 2004. Complete chloroplast genome sequences from Korean ginseng (Panax *schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Research* **11:**247–261 DOI 10.1093/dnares/11.4.247.

Kim GB, Lim CE, Mun JH. Unpublished. Complete chloroplast genomes of Aconitum species from Korea.

Kong WQ, Yang JH. 2017. The complete chloroplast genome sequence of *Morus cathayana* and *Morus multicaulis*, and comparative analysis within genus *Morus* L. *PeerJ* **5:**e3037 DOI 10.7717/peerj.3037.

Kong HH, Liu WZ, Yao G, Gong W. 2017a. Characterization of the whole chloroplast genome of a rare and endangered species *Aconitum reclinatum* (Ranunculaceae) in the United States. *Conservation Genetics Resources* in press DOI 10.1007/s12686-017-0789-y

Kong HH, Zhang Y, Hong Y, Barker MS. 2017b. Multilocus phylogenetic reconstruction

informing polyploid relationships of *Aconitum* subgenus *Lycoctonum* (Ranunculaceae) in China. *Plant Systematics and Evolution* **303:**727–744 DOI 10.1007/s00606-017-1406-y.

Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* **32:**11–16 DOI 10.1093/nar/gkh152.

Lei WJ, Ni DP, Wang YJ, Shao JJ, Wang XC, Yang D, Wang JS, Chen HM, Liu C. 2016. Intraspecific and heteroplasmic variations, gene losses and inversions in the chloroplast genome of *Astragalus membranaceus*. *Scientific Reports* **6:**21669 DOI 10.1038/srep21669.

Liang X, Chen L, Song L, Fei W, He M, He C, Yin Z. 2017. Diterpenoid alkaloids from the root of *Aconitum sinchiangense* W. T. Wang with their antitumor and antibacterial activities. *Natural Product Research* **11:**1–8 DOI 10.1080/14786419.2016.1272113

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:**1451–1452 DOI 10.1093/bioinformatics/btp187.

Lim CE, Kim GB, Baek S, Han SM, Yu HJ, Mun JH. 2017. The complete chloroplast genome of *Aconitum chiisanense* Nakai (Ranunculaceae). *Mitochondrial DNA A* **28(1):**75–76 DOI: 10.3109/19401736.2015.1110805.

Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW — a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research* **41(W1):**W575–W581 DOI 10.1093/nar/gkt289.

Lu RS, Li P, Qiu YX. 2017. The Complete Chloroplast Genomes of Three *Cardiocrinum* (Liliaceae) Species: Comparative Genomic and Phylogenetic Analyses. *Frontiers of*

Plant Sciences **7:**2054 DOI 10.3389/fpls.2016.02054

Mardanov AV, Ravin NV, Kuznetsov BB, Samigullin TH, Antonov AS, Kolganova TV, Skyabin KG. 2008. Complete sequence of the duckweed (*Lemna minor*) chloroplast genome: structural organization and phylogenetic relationships to other angiosperms. *Journal of Molecular Evolution* **66:**555–564 DOI 10.1007/s00239-008-9091-7

Nazareno AG, Carlsen M, Lohmann, LG. 2015. Complete chloroplast genome of *Tanaecium tetragonolobum*: the first Bignoniaceae plastome. *PLoS ONE* **10:**e129930 DOI 10.1371/journal.pone.0129930.

Neuhaus H, Emes M. 2000. Nonphotosynthetic metabolism in plastids. *Annual Review of Plant Biology* **51:**111–140 DOI 10.1146/annurev.arplant.51.1.111.

Posada D, Crandall KA. **1998.** Modeltest: testing the model of DNA substitution. *Bioinformatics* **14:**817–818 DOI 10.1093/bioinformatics/14.9.817.

Raubeson LA, Jansen RK. 2005. Chloroplast genomes of plants. In: Henry RJ, ed. *Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants*. Cambridge: CABI Press, 45–68.

Roman G, Park S. 2015. Analysis of the Complete Chloroplast Genome of a Medicinal Plant, *Dianthus superbus* var. *longicalyncinus*, from a Comparative Genomics Perspective. **10(10):** e0141329 DOI 10.1371/journal.pone.0141329.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:**1572–1574 DOI 10.1093/bioinformatics/btg180.

Roy S, Ueda M, Kadowaki K, Tsutsumi N. 2010. Different status of the gene for ribosomal protein S16 in the chloroplast genome during evolution of the genus *Arabidopsis* and closely related species. *Genes and Genetic Systems* **85:**319–326 DOI

10.1266/ggs.85.319.

Semenov AA, Enikeev AG, Snetkova LV, Permyakov AV, Sokolova NA, Dudareva LV. 2016. Ortho-phthalic acid esters in lipophilic extract from the cell culture of *Aconitum baicalense* Turcz ex Rapaics 1907. *Doklady Biochemistry and Biophysics* **471:**421–422 DOI 10.1134/S1607672916060120.

Shang XM, Lee CL. 1984. Chromosome studies of ten species of *Aconitum* in China. *Acta Phytotaxonomica Sinica* **22:**378–385

Silvestro D, Michalak I. 2012. raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution* **12:**335–337 DOI 10.1007/s13127-011-0056-0.

Soltis PS, Soltis DE. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences of the United States of America* **97(13):**7051–7057 DOI 10.1073/pnas.97.13.7051.

Song BH, Wang XQ, Wang XR, Sun LJ, Hong DY, Peng PH. 2002. Maternal lineages of *Pinus densata*, a diploid hybrid. *Molecular Ecology* **11:**1057–1063 DOI 10.1046/j.1365-294X.2002.01502.x.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:**2688–2690

Sun G, Dilcher DL, Wang HS, Chen ZD. 2011. A eudicot from the Early Cretaceous of China. *Nature* **471:**625–628 DOI 10.1038/nature09811.

Swofford DL. 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods). Version 4.0b10, Sinauer Associates, Sunderland, Massachusets, USA

The Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal*

of the Linnean Society **181:**1–20 DOI 10.1111/boj.12385.

Thiel T, Michalek W, Varshney R, Graner A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeumvulgare* L.). *Theoretical Application of Genetics* **106:**411–422 DOI 10.1007/s00122-002-1031-0.

Utelli AB, Roy BA, Baltisberger M. 2000. Molecular and morphological analyses of European *Aconitum* species (Ranunculaceae) *Plant Systematics and Evolution* **224:**195–212 DOI 10.1007/BF00986343.

Wada K, Takeda K, Haraguchi M, Abe Y, Kuwahara N, Suzuki S, Terui A, Masaka T, Munakata N, Uchida M, Nunokawa M, Yamashita H, Goto M, Lee KH. 2016. Four new diterpenoid alkaloids from *Aconitum japonicum*. *Planta Medica* **81(S01):**S1–S381 DOI 10.1055/s-0036-1596757

Wang W, Liu Y, Yu SX, Gao TG, Chen ZD. 2013. *Gymnaconitum*, a new genus of Ranunculaceae endemic to the Qinghai-Tibetan Plateau. *Taxon* **62:**713–722 DOI 10.12705/624.10.

Wang W, Lu AM, Ren Y, Endress ME, Chen ZD. 2009. Phylogeny and classification of Ranunculales: Evidence from four molecular loci and morphological data. *Perspectives in Plant Ecology Evolution and Systematics* **11:**81–110 DOI 10.1016/j.ppees.2009.01.001.

Wicke S, Schneeweiss GM, dePamphilis, CW, Müller, KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology* **76:**273–297 DOI 10.1007/s11103-011-9762-4.

Wu Z, Gui S, Quan Z, Pan L, Wang S, Ke W, Liang D, Ding Y. 2014. A precise chloroplast

genome of *Nelumbo nucifera* (Nelumbonaceae) evaluated with Sanger, Illumina MipSeq, and PacBio RS II sequencing platforms: insight into the plastid evolution of basal eudicots. *BMC Plant Biology* **14:**289 DOI 10.1186/s12870-014-0289-0.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **20(17):**3252–3255 DOI 10.1093/bioinformatics/bth352.

Yang M, Zhang XW, Liu GM, Yin YX, Chen KF, Yun QZ, Zhao DJ, Al-Mssallem IS, Yu J. 2010. The Complete Chloroplast Genome Sequence of Date Palm (*Phoenix dactylifera* L.). *PLoS ONE* **5(9):**e12762 DOI 10.1371/journal.pone.0012762.

Yuan Q, Yang QE. 2006. Polyploidy in *Aconitum* subgenus *Lycoctonum* (Ranunculaceae). *Botanical Journal of the Linnean Society* **150:**343–353 DOI 10.1111/j.1095-8339.2006.00468.x.

Zhao Y, Bu G, Zhou Y, Lv L, Yan G, Chen B, Wang L, Cen X. 2010. Mechanism study of *Aconitum*-induced neurotoxicity in PC12 cells: involvement of dopamine release and oxidative damage. *Neurotoxicology* **31:**752–757 DOI 10.1016/j.neuro.2010.06.005.

Figure 1(on next page)

Figure 1. The gene maps of *Aconitum angustius*, *A. finetianum*, and *A. sinomontanum*.

The genes lying inside and outside the circles are transcribed in the clockwise and counterclockwise directions, respectively. Different colors denote the genes belonging to different functional groups. The thicknesses indicate the extent of the inverted repeats (IRa and IRb) that separate the small single-copy (SSC) region from the large single-copy (LSC) region. The dark gray in the inner circle corresponds to GC content, and the light gray to AT content.

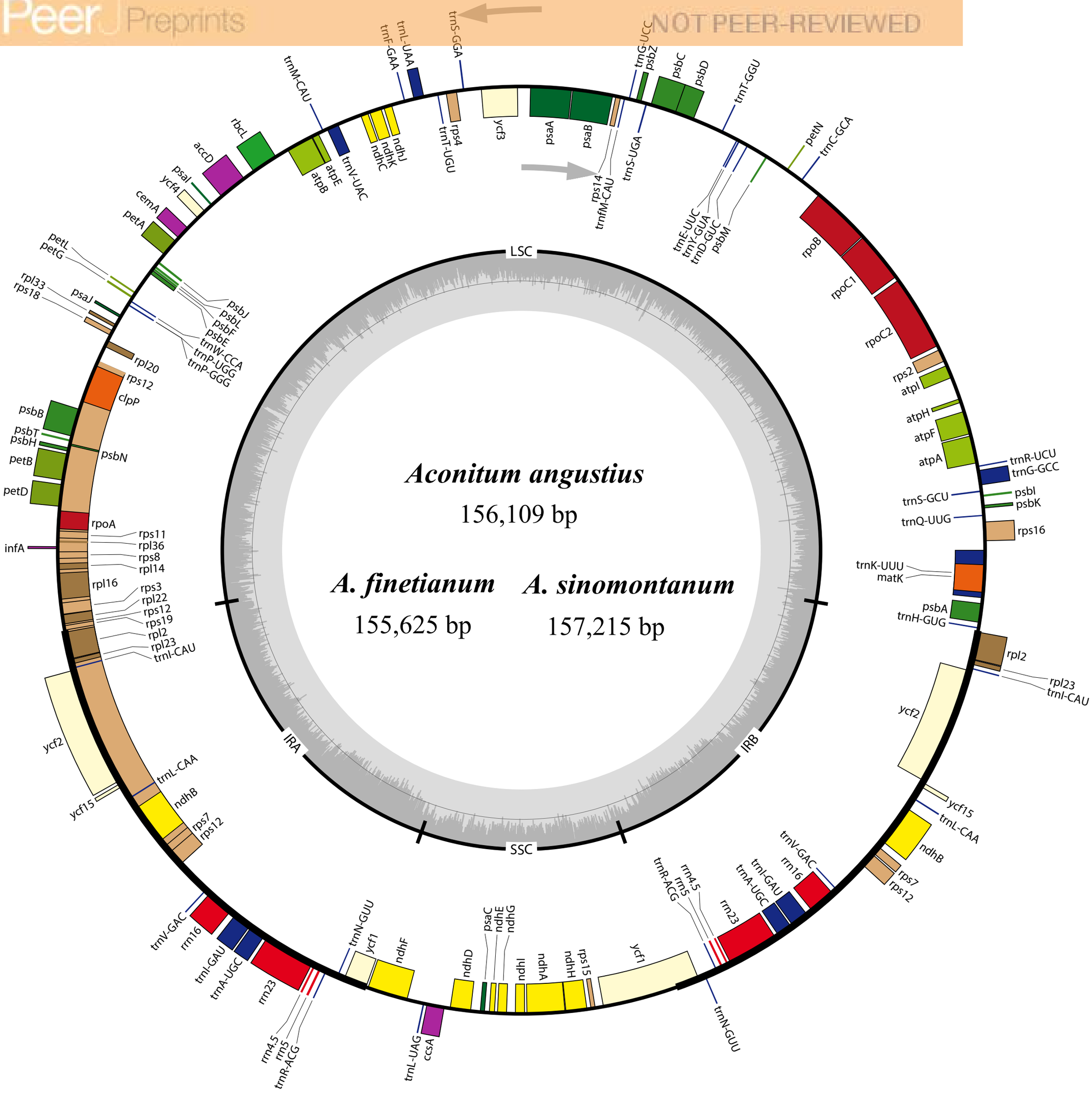
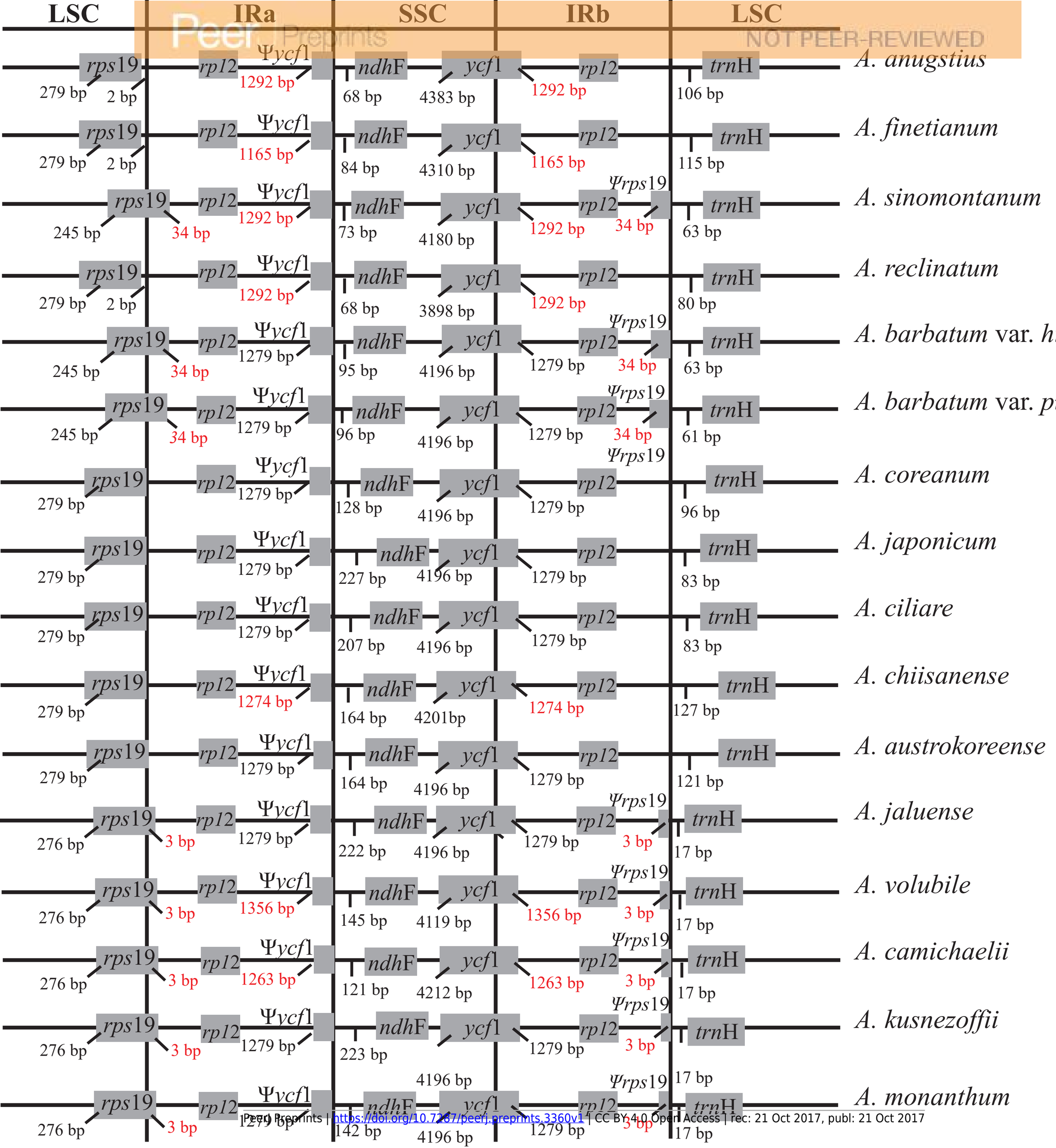


Figure 2(on next page)

Figure 2. Comparison of the border positions of LSC, SSC and IR repeat regions among fourteen species and two varieties in *Aconitum*.

Genes are denoted by grey boxes and the gaps between the genes and the boundaries are indicated by the base lengths (bp). Extensions of the genes are also indicated above the boxes.

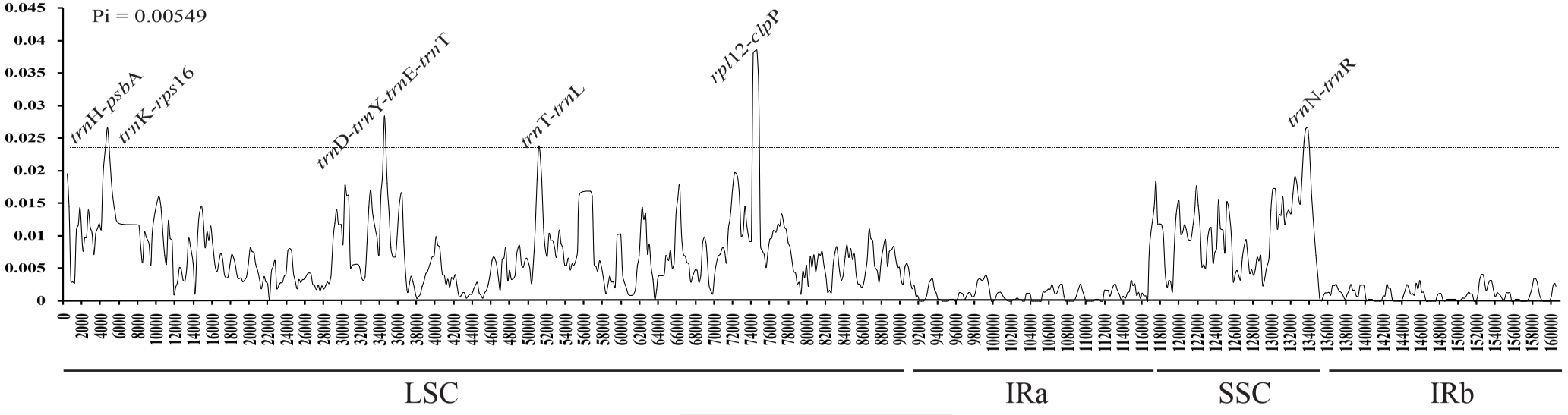


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3360v1 | CC BY 4.0 Open Access | rec: 21 Oct 2017, publ: 21 Oct 2017

Figure 3(on next page)

Figure 3. Sliding window analysis of the whole cp genome for fourteen species and two varieties in *Aconitum*.

X-axis: position of the midpoint of a window; *Y*-axis: nucleotide diversity (*Pi*) of each window.

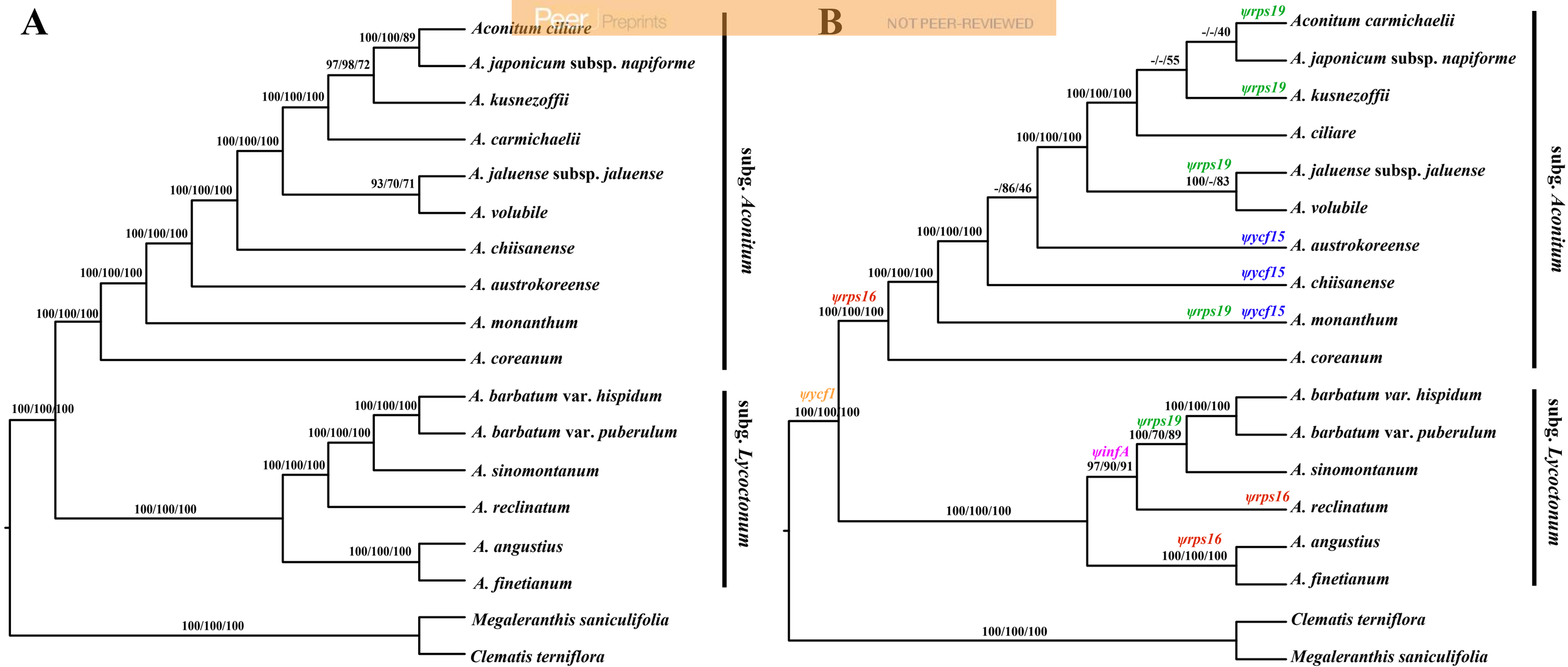


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3360v1 | CC BY 4.0 Open Access | rec: 21 Oct 2017, publ: 21 Oct 2017

Figure 4(on next page)

Figure 4. Phylogenetic relationship among *Aconitum* species.

Based on the two data sets of complete cp genome sequences (A) and PCGs (B), respectively, phylogenetic reconstruction was conducted using three methods: Bayesian Inference (BI), Maximum Parsimony (MP) and Maximum Likelihood (ML). Numbers above the branches represent BI posterior probabilities, MP and ML bootstrap values. The pseudogenes are indicated above the branches in different colors on the phylogenetic tree based on PCGs (B).



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3360v1 | CC BY 4.0 Open Access | rec: 21 Oct 2017, publ: 21 Oct 2017

Table 1(on next page)

Table 1. Summary of characteristics in chloroplast genome sequences of thirteen species and two varieties in *Aconitum*.

1 **Table 1 Summary of characteristics in chloroplast genome sequences of thirteen species and two varieties in *Aconitum*.**

	GenBank No.	Voucher Number/Herbarium	Total genome size (bp)	LSC (bp)	SSC (bp)	IR (bp)	Total number of genes	Protein-coding genes	tRNA genes	rRNA genes	GC content
subg. *Lycoctonum*											
A. angustius	MF155664	ZY37/IBSC	156,109	86,719	16,914	26,225	126	84	38	4	38%
A. finetianum	MF155665	ZY25/IBSC	155,625	86,664	17,107	25,927	126	84	38	4	38%
A. sinomontanum	MF155666	ZY46/IBSC	157,215	88,074	16,926	26,090	126	84	38	4	38%
A. reclinatum	MF186593	US17/IBSC	157,354	88,269	16,963	26,061	127	86	37	4	38%
A.barbatum var. *puberulum*	KC844054	Not provided/-	156,749	87,630	16,941	26,089	127	85	38	4	38%
A.barbatum var. *hispidum*	KT820664	VP0000486327/NIBR	156,782	87,661	16,987	26,067	127	85	38	4	38%
subg. *Aconitum*											
A. austrokoreense	KT820663	VP0000494173/NIBR	155,682	86,388	17,054	26,120	126	83	39	4	38.1%
A.carmichaelii	KX347251	ACAR20151205/-	155,737	86,330	17,021	26,193	124	83	37	4	38.1%
A. chiisanense	KT820665	VP0000494177/NIBR	155,934	86,559	17,085	26,145	125	82	39	4	38.1%
A. ciliare	KT820666	VP0000486323/NIBR	155,832	86,452	17,084	26,148	126	83	39	4	38.1%
A. coreanum	KT820667	VP0000486326/NIBR	157,029	87,622	17035	26,186	128	86	38	4	38.0%
A. jaluense	KT820669	VP0000494219/NIBR	155,926	86,406	17,090	26,215	126	83	39	4	38.1%
A. japonicum	KT820670	VP0000494223/NIBR	155,878	86,480	17,104	26,147	127	84	39	4	38.1%

A. kusnezoffii	KT820671	VP0000529885/NIBR	155,862	86,335	17,103	26,212	126	84	39	4	38.1%
A. monanthum	KT820672	VP0000529886/NIBR	155,688	86,292	16,996	26,200	125	82	39	4	38.1%
A. volubile	KU556690	MBC_KIOM-2015-73/KIOM	155,872	86,348	16,944	26,290	126	83	38	4	38.1%

2

Table 2(on next page)

Table 2. The distribution of the five pseudogenes in *Aconitum*.

Table 2 The distribution of the five pseudogenes in *Aconitum*.

Locations	**LSC**		**LSC/IRa**	**IRa**	**IRa/SSC**
Genes	*Ψrps*16	*Ψinf*A	*Ψrps*19	*Ψycf*15	*Ψycf*1
Aconitum* subg. *Lycoctonum					
A. angustius	+				**+/1292bp**
A. finetianum	+				**+/1165bp**
A. sinomontanum		+	+/34bp		**+/1292bp**
A. reclinatum	+	+			+/1292bp
A. barbatum var. *puberulum*		+	+/34bp		+/1279bp
A. barbatum var. *hispidum*		+	+/34bp		+/1279bp
Aconitum* subg. *Aconitum					
A. austrokoreense	+			+/4bp indel	+/1279bp
A. carmichaelii	+		+/3bp		**+/1263bp**
A. chiisanense	+			+/4bp indel	**+/1274bp**
A. ciliare	+				+/1279bp
A. coreanum	+				+/1279bp
A. jaluense	+		+/3bp		+/1279bp
A. japonicum	+				+/1279bp
A. kusnezoffii	+		+/3bp		+/1279bp
A. monanthum	+		+/3bp	+/1bp indel	+/1279bp
A. volubile	+		+/3bp		**+/1356bp**

3 +: indicating the presence of pseudogenes

Table 3(on next page)

Table 3. Number of chloroplast SSRs in different regions or different types present in *Aconitum* species.

1 **Table 3 Number of chloroplast SSRs in different regions or different types present in *Aconitum* species.**

Species						Number of SSRs in different regions			
	Homo (>10)	Di (>5)	Tri (>5)	Te(>3)	Pen (>3)	LSC	SSC	IR	Total
subg. *Lycoctonum*	38 (64.41%)	10 (16.95%)	3 (4.80%)	8(12.99%)	0 (0.00%)	50 (85.31%)	7 (11.02%)	2 (3.39%)	59
A. angustius	40	9	2	8	1	50	8	2	60
A. finetianum	42	9	2	8	1	51	9	2	62
A. sinomontanum	36	12	2	8	0	50	6	2	58
A. reclinatum	42	10	2	7	0	53	6	2	61
A.barbatum var. *puberulum*	36	10	2	8	0	49	5	2	56
A.barbatum var.*hispidum*	32	10	7	7	0	49	5	2	56
subg. *Aconitum*	22 (46.81%)	15 (31.91%)	1 (21.28%)	7 (14.89%)	1 (21.28%)	38 (80.85%)	7 (14.89%)	2 (4.36%)	47
A. austrokoreense	22	15	0	7	0	32	10	2	44
A.carmichaelii	21	16	1	7	0	37	6	2	45
A. chiisanense	21	16	1	7	2	39	6	2	47
A. ciliare	23	16	1	7	1	41	5	2	48
A. coreanum	39	14	1	7	1	50	10	2	62
A. jaluense	17	14	1	6	2	33	6	2	41
A. japonicum	20	16	1	7	1	37	6	2	46
A. volubile	17	15	1	6	1	35	3	2	40
A. kusnezoffii	19	16	1	7	1	37	5	2	44

A. monanthum	18	13	0	7	2	36	9	2	47

2