Journal of Asia-Pacific Biodiversity 9 (2016) 245–249

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Journal of Asia-Pacific Biodiversity

journal homepage: http://www.elsevier.com/locate/japb

Journal of Asia-Pacific Biodiversity

Short communication

The complete chloroplast genome sequence of *Abies nephrolepis* (Pinaceae: Abietoideae)

Dong-Keun Yi, Kyung Choi, Minjung Joo, Jong Cheol Yang, Feruza U. Mustafina, Jun-Soo Han, Dong Chan Son, Kae Sun Chang, Chang Ho Shin, You-Mi Lee*

Division of Forest Biodiversity and Herbarium, Korea National Arboretum, Pocheon 11186, South Korea

ARTICLE INFO

Article history:
Received 3 March 2016
Received in revised form
15 March 2016
Accepted 16 March 2016
Available online 31 March 2016

Keywords:
Abies
chloroplast genome
Khinghan fir
NGS sequencing

ABSTRACT

The plant chloroplast (cp) genome has maintained a relatively conserved structure and gene content throughout evolution. Cp genome sequences have been used widely for resolving evolutionary and phylogenetic issues at various taxonomic levels of plants. Here, we report the complete cp genome of *Abies nephrolepis*. The *A. nephrolepis* cp genome is 121,336 base pairs (bp) in length including a pair of short inverted repeat regions (IRa and IRb) of 139 bp each separated by a small single copy (SSC) region of 54,323 bp (SSC) and a large single copy region of 66,735 bp (LSC). It contains 114 genes, 68 of which are protein coding genes, 35 tRNA and four rRNA genes, six open reading frames, and one pseudogene. Seventeen repeat units and 64 simple sequence repeats (SSR) have been detected in *A. nephrolepis* cp genome. Large IR sequences locate in 42-kb inversion points (1186 bp). The *A. nephrolepis* cp genome is identical to *Abies koreana*'s which is closely related to taxa. Pairwise comparison between two cp genomes revealed 140 polymorphic sites in each. Complete cp genome sequence of *A. nephrolepis* has a significant potential to provide information on the evolutionary pattern of Abietoideae and valuable data for development of DNA markers for easy identification and classification.

Introduction

The chloroplasts (cp) are specialized intracellular organelles for photosynthesis. The cp genomes have maintained a relatively conserved structure and gene content throughout evolution (Shinozaki et al 1986; Douglas 1998). The complete cp genomes of more than 830 accessions have been reported from seed plants until now National Center for Biotechnology Information (NCBI) Organelle Genome Resources, http://www.ncbi.nlm.nih.gov/genomes/). The majority of spermatophytes contain 90–110 unique genes within 120–170 kb of circular chromosome (Sugiura 1992). The gene contents and the polycistronic transcription units of cp genome are generally conserved among seed plant species with the exception of nonphotosynthetic parasitic plants. The gene order of the cp genome is also relatively conserved; however, it was frequently modified by inversion mutations. Various structural modifications including loss of inverted repeat (IR) or short IR regions, rearrangement of gene order and loss of some genes are reported for gymnosperm cp genomes (Wu et al 2009; Lin et al 2010; Wu et al 2011; Wu and Chaw 2014).

The cp genome sequences have been widely used for resolving evolutionary and phylogenetic issues at various taxonomic levels of plants (Raubeson and Jansen 1992; Downie et al 1996; Jansen et al 2007). Phylogenetic analyses and taxonomic systematizations are performed by a number of cp markers (Kress and Erickson 2007; Erickson et al 2008; Group CPW 2009; Hollingsworth et al 2011). Complete cp genome sequencing provides valuable information for determination of suitable cp DNA markers for plant species classification. Moreover, acquisition of genome sequence nowadays does not demand much time and funds due to availability and accessibility of sequencing technologies.

Firs (genus *Abies*, family Pinaceae) include about 48 species of evergreen coniferous trees (Suyama et al 2000; Farjon 2010; Xiang et al 2015). *Abies nephrolepis* (Traits. ex Maxim.) Maxim. is

* Corresponding author. Tel.: +82 0315401001.
E-mail address: sl1962@korea.kr (Y.-M. Lee).

Peer review under responsibility of National Science Museum of Korea (NSMK) and Korea National Arboretum (KNA).

http://dx.doi.org/10.1016/j.japb.2016.03.014

commonly known as the Khinghan fir and is an important timber tree in Northeast Asia. *A. nephrolepis* shows close relationships with *Abies koreana*, *Abies sachalinensis*, *Abies sibirica*, and *Abies veitchii*. Complete cp genome sequence of *Abies koreana* has already been published (Yi et al 2015). Thus, analysis of cp genome structure of *A. nephrolepis* and close related species, as well as sequence comparison, will contribute to a better understanding of the evolutionary mode of the cp genome and marker development for using in *Abies* genus classification.

Materials and methods

Plants materials and DNA extraction

Approximately 5 g of fresh leaves of *Abies nephrolepis* (Trautv. ex Maxim.) Maxim. were collected from a single individual in the natural forest habitat at the southeast part of South Korea. Voucher specimen and DNA sample (KHB1343154) were deposited in Korea National Arboretum (KH). Total genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle 1987). Extracted DNA was quantified using NanoDrop spectrophotometers (Nanodrop 2000; Thermo Scientific, Wilmington, DE, USA).

Sequencing and chloroplast genome assembly

Chloroplast genome sequences were analyzed using Illumina Hiseq 2000 (Illumina, San Diego, CA, USA). A total number of 35,665,348 reads were analyzed to generate 3,602,200,148 base pairs (bp) of sequence. Acquisition of precise sequence data allowed low quality reads ($<$ Q20) to be filtered out from raw data. The filtered sequences were assembled using the Bowtie2 software (ver. 2.2.3, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; Langmead and Salzberg 2012) with reference sequence of *A. koreana*. A total number of 433,280 reads were mapped in reference sequence with an average coverage of 346.7X. Gaps or uncertain sites were filled and resequenced via direct sequencing of Polymerase Chain Reaction (PCR) products amplified using primers designed from assembled sequences and previously published primers (Tsumura et al 2000). The sequenced fragments were assembled using Geneious 8.0.3 (Biomatters, Auckland, NZ, http://www.geneious.com; Kearse et al 2012).

Chloroplast gene annotation

Gene annotations were performed using the Basic Local Alignment Search Tool (BLAST; BLASTN, PHI-BLAST and BLASTX), Open Reading Frame (ORF) finder program from the National Center for Biotechnology Information and DOGMA (Wyman et al 2004). The nomenclature of cp genes was used according to Chloroplast Genome Database (http://chloroplast.cbio.psu.edu) and previous published cp genomes (Cui et al 2006). A circular map of cp genomes was drawn by OGDRAW (http://ogdraw.mpimp-golm.mpg.de/; Lohse et al 2007).

Sequence analyses

A-T contents were evaluated by MEGA6 (version 6.06, http://www.megasoftware.net; Tamura et al 2013). Repeating sequences were analyzed by using REPuter (https://bibiserv2.cebitec.uni-bielefeld.de/reputer; Kurtz et al 2001) and Tandem Repeats Finder, ver. 4.07b (https://tandem.bu.edu/trf/trf.html; Benson 1999). Simple sequence repeat (SSR) loci were identified by Simple Sequence Repeats Extractor (http://www.aridolan.com/ssr/ssr.aspx). All SSR regions were PCR amplified and resequenced manually to mitigate errors from the Next Generation Sequencing (NGS) sequencing procedure. Protein coding genes sequences were aligned using MUSCLE (Edgar 2004) and adjusted by hand. Nucleotide diversity and ratio of divergence at nonsynonymous and synonymous sites (dN/dS) value were analyzed using DnaSP (version 5.10.01, http://www.ub.edu/dnasp; Librado and Rozas 2009) and MEGA6. Mauve whole genome aligner (Darling et al 2004) was used for alignment of complete cp genomes.

Results

General features of the A. nephrolepis cp genome

The complete cp genome of *A. nephrolepis* is 121,336 bp in length (Figure 1). Its sequence is deposited in GenBank with the accession number KT834974. The *A. nephrolepis* cp genome includes a pair of short inverted repeat regions (IRa and IRb) of 139 bp each separated by a small single copy region of 54,323 bp (SSC) and a large single copy region of 66,735 bp (LSC). It contains 114 individual genes including 68 protein encoding genes, 35 tRNA and four rRNA genes, six open reading frames, and one pseudogene. *A. nephrolepis* is identical to *A. koreana* in cp genome content and structure (Yi et al 2015). Fifty-three protein coding and 16 tRNA genes, three open reading frames, and one pseudogene are located in the LSC region, while 15 protein coding and 17 tRNA genes, four rRNA, and three open reading frames are sited in the SSC region. Only one tRNA gene (*trn*I-CAU) is duplicated and found in IR regions. The cp genome of *A. nephrolepis* lacks NADH dehydrogenase genes (*ndh* genes) as reported for the cp genome of other representatives of the family Pinaceae (Lin et al 2010; Wu et al 2011). Twelve genes contain one intron, *rps*12 and *ycf*3 genes have two introns. The major part (56.0%) of the *A. nephrolepis* cp genome consists of gene-coding regions (50.1% protein and 5.9% RNA coding regions), whereas the intergenic spacers (including 15 introns) comprise 44.0%. The overall A-T content of *A. nephrolepis* cp genome is 61.7%, which is slightly lower than in the cp genome of other conifers and generally in angiosperms (Shinozaki et al 1986; Wu et al 2009; Lin et al 2010; Wu et al 2011; Yi and Kim 2012; Wu and Chaw 2014). The A-T content in noncoding regions (64.6%) is higher than in coding regions (59.5%). IR regions contain 58.3% of A-T while LSC and SSC regions have 62.6% and 60.8%, respectively (Table 1).

Examination of repeat unit and SSRs in A. nephrolepis cp genome

Seventeen repeat units of at least 22 bp length are identified in the *A. nephrolepis* cp genome. One 1186-bp length repeat is novel for the IR region. It is described for cp genomes for some representative of *Abies* and *Tsuga* genera (Tsumura et al 2000). The other 16 repeat units include six direct, six dispersed and four palindromic repeats (Table 2). Sixty-four SSRs with the length longer than 10 bp are identified. Thirty-six mononucleotide SSRs contain 3 Cs, 1 Gs, and 32 As or Ts. Fourteen dinucleotide SSRs contain ATs or TAs and none of GAs. The other SSRs are identified as dinucleotides or trinucleotides, 10 tetranucleotides (3 repeats were duplicated in di-nucleotide) and two pentanucleotides (Table S1).

Comparative analysis of cp genomes

A. nephrolepis is identical to *A. koreana* in cp genome structure, gene order, and its content. Pairwise sequence comparison between their cp genomes revealed high identity (99.9%) and

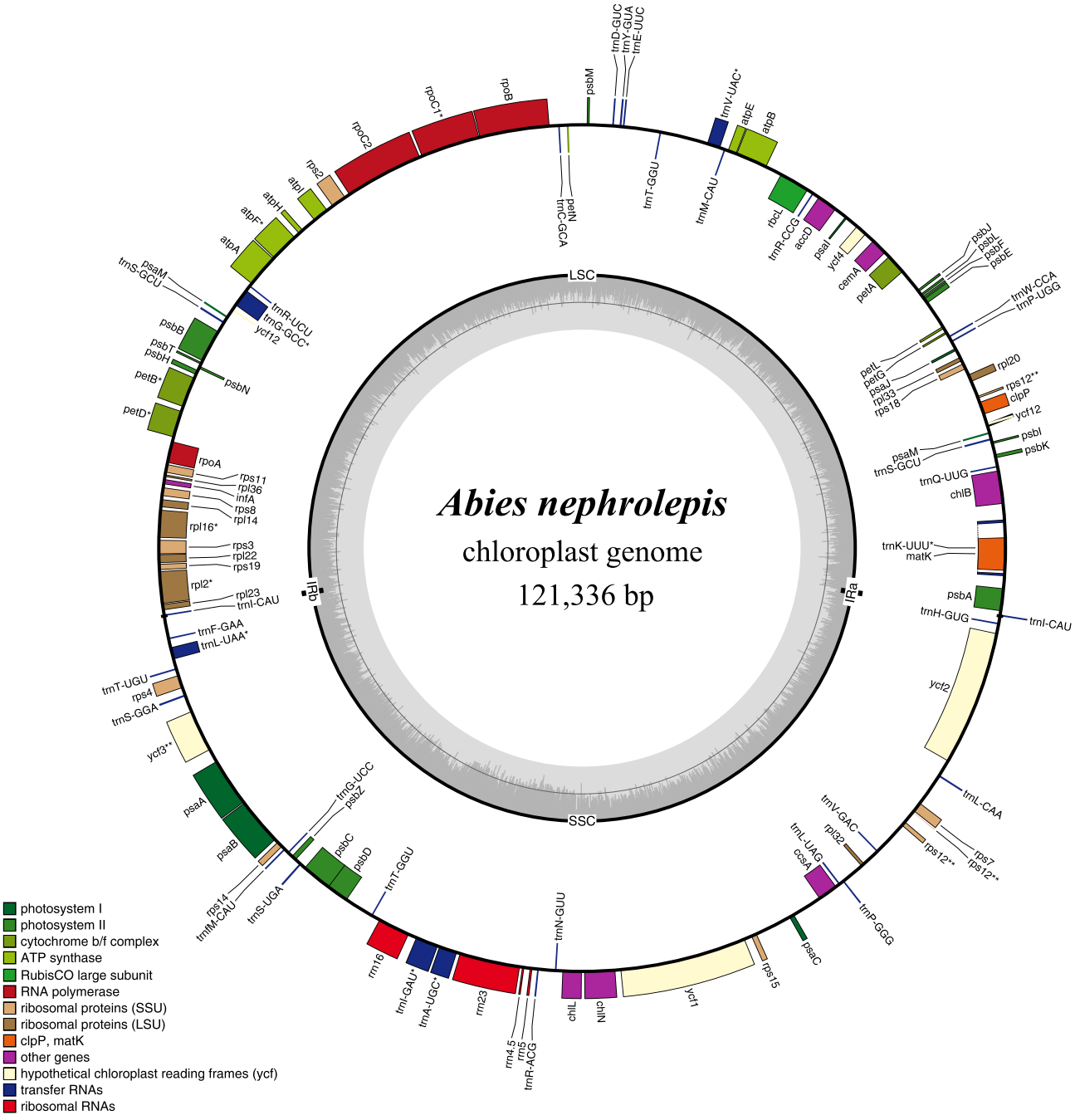

Figure 1. The complete chloroplast genome map of *Abies nephrolepis*. A pair of thin blocks inside the circle represents inverted repeats (IRa and IRb; 139 bp each), which separate the large single copy region (LSC; 66,735 bp) from the small single copy region (SSC; 54,323 bp). Genes drawn inside the circle are transcribed clockwise, while those drawn outside the circle are transcribed counterclockwise. Intron-containing genes are marked by asterisks.

defined 140 polymorphic sites. Among these 140 polymorphic sites, 108 are characterized by indel polymorphisms and the others are identified as SNPs. Seven polymorphic sites locate in protein coding regions (one in *rpl*14 and six in *ycf*2 gene) and the other 132 polymorphic sites locate in intergenic spacer regions. Whole cp genome alignment with published cp genomes of two Abietoideae representatives (*Cedrus deodara* and *Keteleeria davidiana*) revealed 42-kb inversion in cp genomes of *A. nephrolepis* and *A. koreana* (Lin et al 2010; Wu et al 2011). Base substitution analysis among three cp genomes of *A. nephrolepis*, *C. deodara*, and *K. davidiana* was performed. The *psa*M gene shows highest divergence (18.77%), followed by *psb*M (8.75%), *psa*J (7.83%), *ycf*2 (7.12%), *psa*I (6.01%), *ycf*1 (5.98%), *cem*A (4.79%), *mat*K (4.68%), *ycf*4 (4.56%), and *acc*D (4.42%) genes. High divergent genes are located in LSC regions with the exception of the genes with unknown function (*ycf*1 and *ycf*2; Table S2).

Discussion

A. nephrolepis is an economically important species used for timber and as a garden tree. The length of complete cp genome sequence of *A. nephrolepis* is 121,336 bp and includes 114 individual genes which make up to 56.0% of the whole cp genome. Short IR sequences (1186bp) are identified in *A. nephrolepis* cp genome. They

Table 1. Genes of *Abies nephrolepis* cp genome (total 114 genes).

Group of gene		Name of gene*	No.
Protein genes	Photosynthesis		
	Photosystem I	*psa*A, *psa*B, *psa*C, *psa*I, *psa*J, *psa*M(x2)	7
	Photosystem II	*psb*A, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*I, *psb*J, *psb*K, *psb*L, *psb*M, *psb*N, *psb*T, *psb*Z	15
	Cytochrome	*pet*A, *pet*B*, *pet*D*, *pet*G, *pet*L, *pet*N	6
	ATP synthase	*atp*A, *atp*B, *atp*E, *atp*F*, *atp*H, *atp*I	6
	Rubisco	*rbcL*	1
	ATP-dependent protease subunit P	*clp*P	1
	Chloroplast envelope membrane protein	*cem*A	1
	Chlorophyll biosynthesis	*chl*B, *chl*L, *chl*N	3
Ribosomal proteins	large units	*rpl*2*, *rpl*14, *rpl*16*, *rpl*20, *rpl*22, *rpl*23, *rpl*32, *rpl*33, *rpl*36	9
	small units	*rps*2, *rps*3, *rps*4, *rps*7, *rps*8, *rps*11, *rps*12+, *rps*14, *rps*15, *rps*18, *rps*19	11
Transcription/ translation	RNA polymerase	*rpo*A, *rpo*B, *rpo*C1*, *rpo*C2	4
	Miscellaneous proteins	*acc*D, *ccs*A, *inf*A, *mat*K	4
	Hypothetical proteins & Conserved reading frame	*ycf*1, *ycf*2, *ycf*3**, *ycf*4, *ycf*12(x2)	6
RNA genes	Transfer RNAs	*trn*A-UGC*, *trn*C-GCA, *trn*D-GUC, *trn*E-UUC, *trn*F-GAA, *trn*fM-CAU, *trn*G-GCC, *trn*G-UCC*, ψ*trn*G-UCC, *trn*H-GUG, *trn*I-CAU(x2), *trn*I-GAU*, *trn*K-UUU*, *trn*L-CAA, *trn*L-UAA*, *trn*L-UAG, *trn*M-CAU, *trn*N-GUU, *trn*P-GGG, *trn*P-UGG, *trn*Q-UUG, *trn*R-ACG, *trn*R-CCG, *trn*R-UCU, *trn*S-GCU(x2), *trn*S-GGA, *trn*S-UGA, *trn*T-GGU(x2), *trn*T-UGU, *trn*V-GAC, *trn*V-UAC*, *trn*W-CCA, *trn*Y-GUA	36
	Ribosomal RNAs	*rrn*4.5, *rrn*5, *rrn*16, *rrn*23	4
Total			114

are presented by *trn*S-*psa*M-*ycf*12-ψ*trn*G and *trn*G-*ycf*12-*psa*M-*trn*S and found in other species of genus *Abies* and genus *Tsuga*. These repeat sequences are identical to those of *A. sachalinensis* and *A. veitchii* which show their close relationship with *A. nephrolepis* (Tsumura et al 2000). Whole genome alignment of cp genomes of *Cedrus deodara*, *Keteleeria davidiana*, and *A. nephrolepis* revealed 42-kb inversion points in cp genome of *A. nephrolepis* mentioned above and their location in 42-kb inversion points has been reported earlier for representatives of *Abies* and *Tsuga* genera. PCR analyses were performed with the use of earlier published primers for verifying 42-kb inversion arrangement (Tsumura et al 2000). The analyses revealed isomeric cp forms presented in each individual of *A. nephrolepis*. Previous studies have noted that these results are likely to emerge from PCR-mediated recombination which occurred during PCR amplification (Guo et al 2014). Therefore, different experimental approaches or explanatory models like sub-stoichiometric shifting are needed for understanding 42-kb inversion rearrangement in cpDNA and verifying the presence of isomeric cp forms (Yi et al 2013; Guo et al 2014).

Table 2. Distribution of repeat units larger than 21 base pairs (bp) in *Abies nephrolepis* cp genome.

Repeat No.	Size (bp)	Repeat	Location	Repeat unit length
1	1186	Dispersed	trnS-psaM-ycf12-ΨtrnG, trnG-ycf12-psaM-trnS	
2	277	Direct	IGS(ycf2/trnH)	38
3	216	Direct	CDS(ycf1)	24
4	133	Direct	CDS(ycf1)	48
5	101	Dispersed	IGS(trnV/trnT), IGS(psbT/trnT)	
6	81	Direct	CDS(ycf1)	27
7	73	Dispersed	IGS(rpl32/trnV), CDS(ycf2)	
8	59	Direct	IGS(rbcL/atpB)	23
9	52	Dispersed	IGS(trnT/trnE), IGS(trnT/rrn16)	
10	46	Palindromic	IGS(trnR/trnG)	
11	42	Direct	CDS(ycf1)	21
12	29	Palindromic	CDS(rps14/trnfM)	
13	28	Palindromic	IGS(chlN/ycf1)	
14	24	Dispersed	IGS(trnT/trnE), IGS(trnT/rrn16)	
15	24	Palindromic	IGS(trnG/psbZ)	
16	24	Dispersed	IGS(trnT/trnE), IGS(trnT/rrn16)	
17	22	Dispersed	IGS(ycf3/psaA)	

The cp genomes of *C. deodara* and *K. davidiana* have a pair of *psb*I genes within novel IR regions whereas two species of genus *Abies* possess a single *psb*I gene located outside of the novel IR region. Expansion and contraction events in novel IR regions of the Abietoideae cp genome have the same pattern as in angiosperms cp genome (Yi and Kim 2012).

Sequence diversity analyses defined *psa*M as the most divergent gene for three species (*A. nephrolepis*, *C. deodara*, and *K. davidiana*). It is located in the 1186 bp IR region of the cp genome. Compared with two other genera, deletion of 6 bp size is observed for *A. nephrolepis* cp genome in this region.

Identity between *A. nephrolepis* and *A. koreana* in genes order and their content is affirmed by pairwise comparison which revealed 140 polymorphic sites in both cp genomes. They include 32 SNPs and 108 indel polymorphic sites, 28 of which are indel events, 48 which resulted from three indel events of SSR loci (39 bp, 5 bp and 4 bp), and the other 60 polymorphic sites are caused from indel of mononucleotide repeat. Using SNPs and indel polymorphic sites in identification of *A. koreana* and *A. nephrolepis* provides valuable data for markers development and their implementation in taxonomic classification.

Conclusion

The cp genome is a rich source of evolutionary data; its complete sequencing is valuable information for plant species taxonomic positioning. Complete cp genome sequence of *A. nephrolepis* has a significant importance for further developing of evolutionary pattern for Abietoideae. Furthermore, comparative analyses of cp genome sequences provide developing DNA markers for easy identification and classification.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.japb.2016.03.014.

References

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27:573–580.

Cui L, Veeraraghavan N, Richter A, et al. 2006. ChloroplastDB: the Chloroplast Genome Database. *Nucleic Acids Research* 34:D692–696.

Darling AC, Mau B, Blattner FR, et al. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Research* 14:1394–1403.

Douglas SE. 1998. Plastid evolution: origins, diversity, trends. *Current Opinion In Genetics & Development* 8:655–661.

Downie SR, Katz-Downie DS, Cho KJ. 1996. Phylogenetic analysis of Apiaceae subfamily apioideae using nucleotide sequences from the chloroplast rpoC1 intron. *Molecular Phylogenetics and Evolution* 6:1–18.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11–15.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792–1797.

Erickson DL, Spouge J, Resch A, et al. 2008. DNA barcoding in land plants: developing standards to quantify and maximize success. *Taxon* 57:1304–1316.

Farjon A. 2010, . *A Handbook of the World's Conifers (2 Vols.)*, Vol. 1. Leiden, Netherlands: Brill.

Group CPW. 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America* 106:12794–12797.

Guo W, Grewe F, Cobo-Clark A, et al. 2014. Predominant and substoichiometric isomers of the plastid genome coexist within *Juniperus* plants and have shifted multiple times during cupressophyte evolution. *Genome Biology and Evolution* 6:580–590.

Hollingsworth PM, Graham SW, Little DP. 2011. Choosing and using a plant DNA barcode. *PloS One* 6:e19254.

Jansen RK, Cai Z, Raubeson LA, et al. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences of the United States of America* 104:19369–19374.

Kearse M, Moir R, Wilson A, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649.

Kress WJ, Erickson DL. 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PloS One* 2:e508.

Kurtz S, Choudhuri JV, Ohlebusch E, et al. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* 29:4633–4642.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357–359.

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

Lin CP, Huang JP, Wu CS, et al. 2010. Comparative chloroplast genomics reveals the evolution of Pinaceae genera and subfamilies. *Genome Biology and Evolution* 2:504–517.

Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Current Genetics* 52:267–274.

Raubeson LA, Jansen RK. 1992. A rare chloroplast DNA structure mutation is shared by all conifers. *Biochemical Systematics and Ecology* 20:17–24.

Shinozaki K, Ohme M, Tanaka M, et al. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *The EMBO Journal* 5:2043–2049.

Sugiura M. 1992. The chloroplast genome. *Plant Molecular Biology* 19:149–168.

Suyama Y, Yoshimaru H, Tsumura Y. 2000. Molecular phylogenetic position of Japanese *Abies* (Pinaceae) based on chloroplast DNA sequences. *Molecular Phylogenetics and Evolution* 16:271–277.

Tamura K, Stecher G, Peterson D, et al. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30:2725–2729.

Tsumura Y, Suyama Y, Yoshimura K. 2000. Chloroplast DNA inversion polymorphism in populations of *Abies* and *Tsuga*. *Molecular Biology and Evolution* 17:1302–1312.

Wu CS, Chaw SM. 2014. Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): evolution towards shorter intergenic spacers. *Plant Biotechnology Journal* 12:344–353.

Wu CS, Lai YT, Lin CP, et al. 2009. Evolution of reduced and compact chloroplast genomes (cpDNAs) in gnetophytes: selection toward a lower-cost strategy. *Molecular Phylogenetics and Evolution* 52:115–124.

Wu CS, Lin CP, Hsu CY, et al. 2011. Comparative chloroplast genomes of Pinaceae: insights into the mechanism of diversified genomic organizations. *Genome Biology and Evolution* 3:309–319.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20:3252–3255.

Yi DK, Kim KJ. 2012. Complete chloroplast genome sequences of important oilseed crop *Sesamum indicum* L. *PloS One* 7:e35872.

Yi DK, Yang JC, So SK, et al. 2015. The complete plastid genome sequence of *Abies koreana* (Pinaceae: Abietoideae). *Mitochondrial DNA*; 2015:1–3. http://dx.doi.org/10.3109/19401736.2015.1025260.

Yi X, Gao L, Wang B, et al. 2013. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): evolutionary comparison of *cephalotaxus* chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms. *Genome Biology and Evolution* 5:688–698.

Xiang QP, Wei R, Shao YZ, et al. 2015. Phylogenetic relationships, possible ancient hybridization, and biogeographic history of *Abies* (Pinaceae) based on data from nuclear, plastid, and mitochondrial genomes. *Molecular Phylogenetics and Evolution* 82 (Pt A):1–14.