International Journal of Scientific Reports
Khandhar DB et al. Int J Sci Rep. 2022 Jan;8(1):7-22
http://www.sci-rep.com
pISSN 2454-2156 | eISSN 2454-2164

DOI: https://dx.doi.org/10.18203/issn.2454-2156.IntJSciRep20214932

Original Research Article

Authentication and identification markers for medicinal plants: I *Aristolochia bracteolata*

Dharam B. Khandhar, Pritesh P. Bhatt, Vrinda S. Thaker*

Department of Biosciences, Saurashtra University, Rajkot, Gujarat, India

Received: 15 September 2021
Revised: 09 December 2021
Accepted: 16 December 2021

***Correspondence:**
Dr. Vrinda S. Thaker,
E-mail: thakervs@gmail.com

ABSTRACT

Background: Aristolochiaceae is a unique plant family that contains aristolochic acids (AAs), with around 600 species.
Methods: The entire chloroplast genome of *Aristolochia bracteolata* was sequenced in this work, and fourteen chloroplast (cp) genomes were retrieved from the NCBI database.
Results: We also analyzed six types of microsatellite markers among these species and found some different markers for each species. A molecular phylogeny based on 7 barcode regions i.e. *matK, atpF* to *atpH*, *psbK* to *psbI*, *rbcL, rpoB, rpoC2* and *rpoC1* is proposed for *Aristolochia* and 16 species from Piperales, *Arabidopsis thaliana* is taken as an outgroup. Phylogenetic relationships using concatenated protein-coding genes from Chloroplast genomes of Piperales using the maximum likelihood method is also discussed.
Conclusions: In the present study we are providing some unique markers and phylogenetic relationships among Piperales members which will help in identification, authentication, to prevent adulteration and further investigation of these medicinal plants.

Keywords: *Aristolochia*, Chloroplast genome, Markers, Barcode, Phylogeny

INTRODUCTION

Aristolochiaceae is a plant family with around 600 species that contain aristolochic acids (AAs) and their derivatives, which are frequently utilized in the treatment of liver cancer.[1] Aristolochiaceae is a member of the class Piperales, divided into two subfamilies, Asaroideae and Aristolochioideae. Asaroideae occur chiefly in northern temperate regions with a centre of diversity ranging from Mali to Northern Nigeria, and in tropical East Africa and India.[2]

In various areas of the world, *Aristolochia* species have been utilized medicinally in numerous traditional medicine formulations and nutritional supplements. The main secondary compounds are aristolochic acids, aristolactams, benzylisoquinoline alkaloids, and mono-, sesqui-, di- and triterpenes, including derived diterpene labdanoic acids (LAs).[3] Sesqui- and diterpenes are the most abundant chemical compounds, and 24 types of sesquiterpenes have been described so far, with the major types including cadinanes, aristolanes, germacrenes and bicyclogermacrenes.[3,4] The physiological active components came from a methanolic extract of *A. bracteolata* plant sections. It also possesses a potent antiallergic activity and antibacterial and antifungal activities are also found.[5] The plant is used to cure scabies in India and leg itch in Ethiopia's Ogaden region. The root is bitter, mixed with lime juice are taken for snakebite and scorpion stings in Nigeria. In East of Lake Chad the root is also applied to scorpion stings.[6] *A. indica* and *A. tagala* were also employed in India as antirheumatics, emmenagogues, and snakebite treatments.[7] It is also used in the treatment of tumors and malaria fevers, but its application as an antimalarial is not suggested in its crude form.[8]

To avoid adulterations in the herbal formulations identification of molecular markers for herbs are imperative.[9] Many genes in the chloroplast genome are used for barcoding of the medicinal plants.[10] Approximately 100-150 genes are present on the chloroplast genome of the higher plants and algae.[11] From these concatenated protein-coding sequences of chloroplast genome often used as a tool for species authentication.[12] In addition chloroplast genome-based SSR markers are also a useful tool for such investigations.[13] Simple sequence repeats (SSRs), also known as microsatellites, are repeated sequences of 1–6 bp in length that have been identified in abundance in both coding and non-coding regions of prokaryotic and eukaryotic genomes.[14] Due to their co-dominant inheritance, high polymorphism, reproducibility and transferability SSRs are widely used in many laboratories.[15] Traditional approaches for screening polymorphic SSR (polySSR) markers and determining their suitability for genetic research are time-consuming and labor-intensive. With the application of NGS technologies, genomic resources have greatly increased in the last decade. Today in many studies phylogeny using concatenated protein-coding genes gives more accurate results instead of one or more barcode genes.[16-19]

This study presents the complete chloroplast genome of *A. bracteolata*, adding to the number of chloroplast genomes available for the family Aristolochiaceae. Its comparison with available Piperales species and its phylogenetic relationship is discussed.

METHODS

Extraction of cpDNA from plant material, and sequencing

Fresh leaves of *A. bracteolata* was collected from field 22.29º N, 70.74º E and identified. Leaves were kept in dark for 48ºC to decrease starch level. Chloroplast DNA was extracted from *A. bracteolata* using a standard protocol.[20] Chloroplast DNA quality and quantity were assessed based on spectrophotometry and purity was checked on 1% agarose gel electrophoresis. This DNA was subjected to high throughput sequencing using an ion torrent genome machine and reads were obtained by torrent suit software version v3.0.

Chloroplast genome assembly and annotation

Obtained reads were checked for quality and remaining adapter contamination. These reads were *de novo* assembled using CLC genomic workbench v12. Reference-based assembly was prepared with its closest relative *A. tagala* (NC_041455.1). This reference-based assembly was compared with *de novo* assembly and corrected. The complete chloroplast genome of *A. bracteolata* was annotated using two servers CpGAVAS and dual organellar genome annotator DOGMA.[21,22] Further tRNA were confirmed using tRNAscan-*SE* web *server*.[23] The chloroplast genome map was created using the default settings of the organelle genome DRAW (OGDRAW). The entire cp genome sequences of *A. bracteolata* have been deposited in GenBank, and accession numbers have been assigned to them as a MT 130705.

Data collection

Complete Piperales chloroplast sequences of *A. contorta* (NC_036152.1), *A. debilis* (NC_036153.1), *A. kaempferi* (NC_041452.1), *A. kunmingensis* (NC_041451.1), *A. macrophylla* (NC_041453.1), *A. mollissima* (NC_041457.1), *A. moupinensis* (NC_041454.1), *A. tagala* (NC_041455.1), *A. tubiflora* (NC_041456.1), *Asarum sieboldii* (NC_037190.1), *Piper cenocladum* (NC_008457.1), *P. kadsura* (NC_027941.1), *P. laetispicum* (NC_042254.1), and *Saruma henryi* (NC_039933.1) were retrieved from NCBI organelle database for comparative analysis.

Simple sequence repeats and phylogenetic analyses

With the default parameters of MISA and without primer synthesis mode, ChloroMitoSSRDB was used to detect microsatellites (mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats).[24] To estimate phylogenetic relationships within the Aristolochiaceae, plastomes of *A. bracteolata* and 16 Piperales species were selected. *Arabidopsis thaliana* was selected as an outgroup. Common protein-coding genes were identified and concatenated protein-coding sequence was prepared. Multiple sequence alignment of these concatenated sequences was performed using CLC genomic workbench v12. This alignment was exported to MEGA X and a phylogenetic tree was prepared using the maximum likelihood method using 1000 bootstrap replicates.

RESULTS

Phylogenetic analyses

Seven barcode genes were selected for the phylogenetic study of *A. bracteolata* with other available Piperales species. In these two intergenic regions, *atpF-atpH* and *psbK-psbI* were selected. In both intergenic regions, *A. sieboldii* is coming as an outgroup with elevated branch length, while *Arabidopsis thaliana* is coming with *Aristolochia* clade, which is misplaced. Therefore, these two intergenic regions are not reliable for the phylogeny of Piperales (Figure 2a and b). In five barcode genes (*matK*, *rbcL*, *rpoB*, *rpoC2* and *rpoC1*), *A. thaliana* is forming a proper outgroup. In these all *A. bracteolata* is forming a clade with *A. tagala*. Five barcode genes are showing *Saruma henryi* and *Asarum sieboldii* in a distinct clade (Figure 2 c-g), while two intergenic regions *atpF-atpH* and *psbK-psbI* are showing *Asarum sieboldii* in a different clade. The two barcode genes *rbcL* and *rpoC2* provided good bootstrap support than the other (Figure 2d and f), but these all barcodes are unable to provide accurate phylogeny for all the species. Therefore, we have gone for phylogeny using all concatenated protein-coding genes.

The phylogenetic relationships of Piperales were constructed based on 18 species (Annexure Table 2), in this *Arabidopsis thaliana* was taken as an outgroup. The phylogenetic tree was prepared using the maximum likelihood method using 1000 bootstrap replicates. Our phylogeny using concatenated protein-coding genes provides more resolution to *Aristolochia* and *Piperales* clades with 100% bootstrap support and all clades in the phylogeny are well supported. In phylogeny 10 *Aristolochia* species formed one clade while 7 species of the Piperaceae family formed another clade and *A. thaliana* come as an outgroup (Figure 3). The phylogenetic analyses revealed that this *A. bracteolata* was closely related to *A. tagala* and three Piperaceae species are also a close relative of the Aristolochiaceae family.

Repeat structure and simple sequence repeats analyses

The ChloroMitoSSRDB 2.0 server was used to evaluate repeats in 15 cp genomes, including 10 species of *Aristolochia*, *A. bracteolata*, *A. kaempferi*, *A. contorta*, *A. mollissima*, *A. debilis*, *A. moupinensis*, *A. kunmingensis*, *A. tubiflora*, *A. macrophylla*, *A. tagala* while including 5 species of piper, *Asarum sieboldii*, *P. cenocladum*, *P. kadsura*, *P. laetispicum* and *Saruma henryi*. The maximum number of microsatellite repeats 167 were found in *Asarum sieboldii* while minimum number of repeats 25 were found in *Piper cenocladum* (Table 1). The longest repeats, with a length of 560 bp, were detected in *Asarum sieboldii* (Annexure Table 1). The total number of SSRs in the 15 species cp genomes were also determined. From analyzed types of repeats, mononucleotide repeats found maximum in *A. tagala* with 47 (50%) and minimum in *Piper cenocladum* which is 11 (44%). Among analyzed six types of repeats mononucleotides were found to be highest while hexanucleotides were lowest in all 15 species. Most diverse types of repeats were observed in *Asarum sieboldii*. Repeats were observed in 3 groups based on overall similarity. The first group consists of *A. bracteolata*, *A. contorta*, *A. debilis*, *A. kunmingensis*, *A. macrophylla*, *A. tagala* and *A. tubiflora* while second group consist of *A. debilis*, *A. mollissima*, *A. moupinensis* and *Saruma henryi* (Figure 1). The third group consists of *P. cenocladum*, *P. kadsura* and *P. laetispicum*.

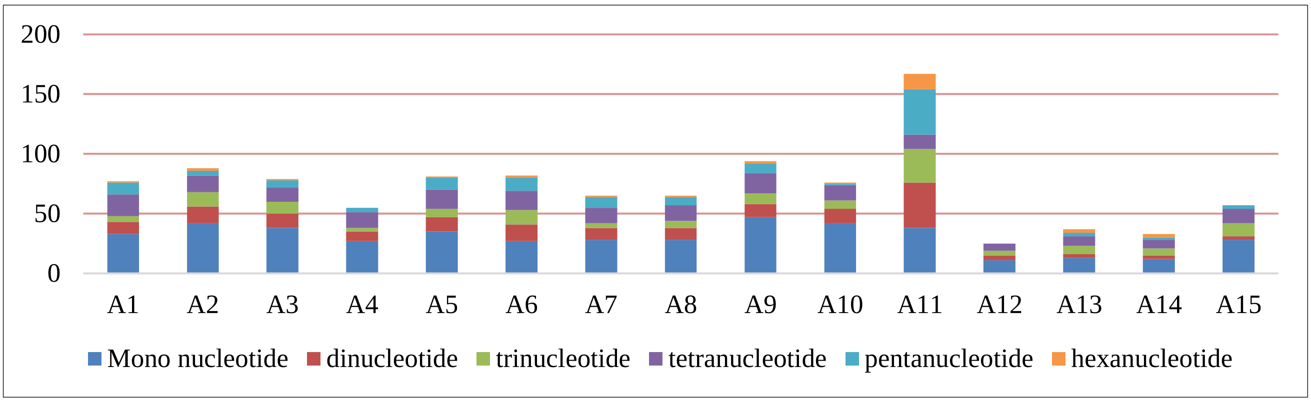

Figure 1: Distribution of number repeat chloroplast genome in fifteen species.

Table 1: Distribution of SSR in chloroplast genome in fifteen species.

Unit size	**Number of SSRs**														
	A. bracteolata	***A. contorta***	***A. debilis***	***A. kaempferi***	***A. kunmingensis***	***A. macrophylla***	***A. mollissima***	***A. moupinensis***	***A. tagala***	***A. tubiflora***	***Asarum sieboldii***	***P. cenocladum***	***P. kadsura***	***P. laetispicum***	***Saruma henryi***
Mononucleotide	33	42	38	27	35	27	28	28	47	42	38	11	13	12	28
Dinucleotide	10	14	12	8	12	14	10	10	11	12	38	4	3	3	3
Trinucleotide	5	12	10	3	7	12	4	6	9	7	28	4	7	6	11
Tetranucleotide	18	14	12	13	16	16	13	13	17	13	12	6	8	7	12

Continued.

Unit size	**Number of SSRs**														
	A. bracteolata	***A. contorta***	***A. debilis***	***A. kaempferi***	***A. kunmingensis***	***A. macrophylla***	***A. mollissima***	***A. moupinensis***	***A. tagala***	***A. tubiflora***	***Asarum sieboldii***	***P. cenocladum***	***P. kadsura***	***P. laetispicum***	***Saruma henryi***
Pentanucleotide	10	4	6	4	10	11	9	7	8	1	38	0	3	2	3
Hexanucleotide	1	2	1	0	1	2	1	1	2	1	13	0	3	3	0
Total number of identified SSRs	77	88	79	55	81	82	65	65	94	76	167	25	37	33	57

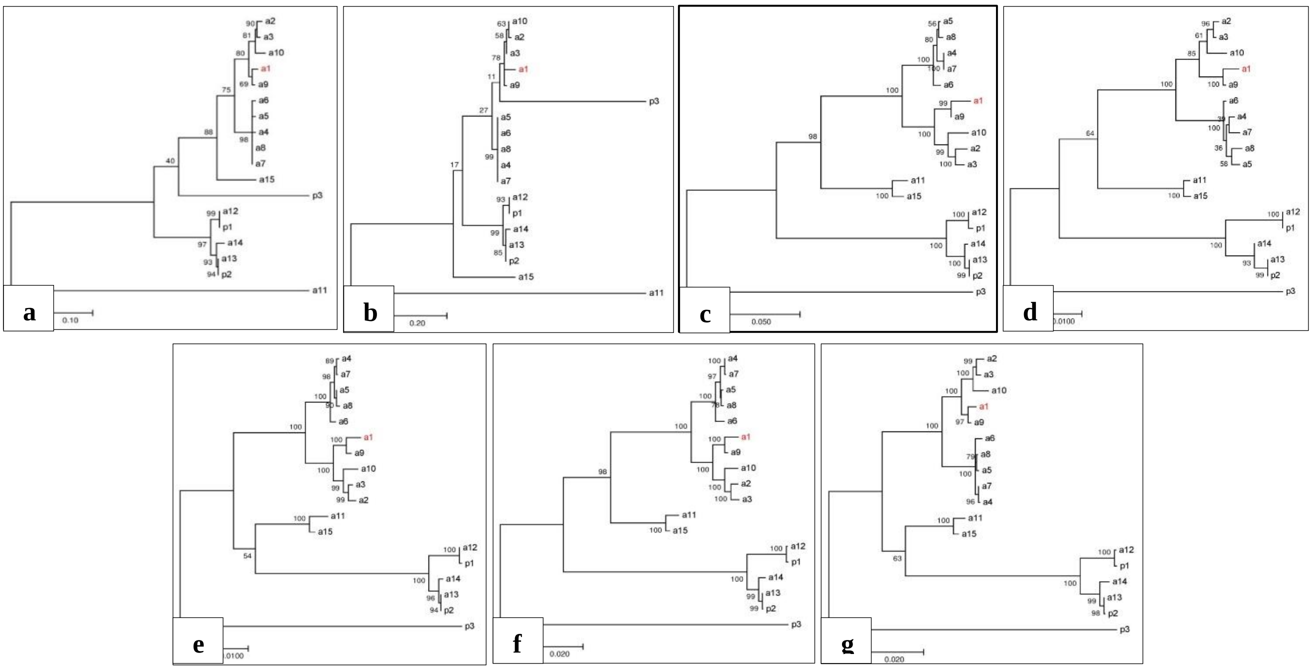

Figure 2: Phylogenetic relationship among 17 selected Piperales species and *A. thaliana* as an outgroup based on maximum likelihood method with 1000 bootstrap support. Here barcode genes are (a) *atpF-atpH*, (b) *psbK–psbI*, (c) *matK*, (d) *rbcL*, (e) *rpoB*, (f) *rpoC2*, and (g) *rpoC1*.

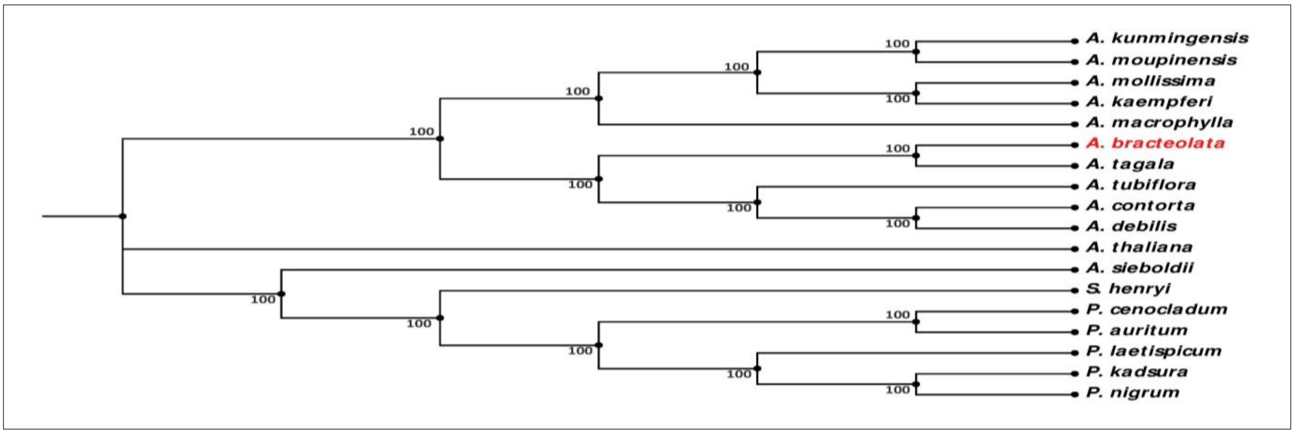

Figure 3: Phylogenetic tree of *A. bracteolata* with eighteen complete chloroplast genomes by maximum likelihood method.

DISCUSSION

Aristolochia is the biggest genus in the Aristolochiaceae family, with about 600 species.[25] Previous investigations used one or more genes to undertake phylogenetic analyses; however, these analyses were limited to Aristolochiaceae species and contained few species from other families.[3,26,27] Chloroplast genomes include a wealth of information that may be used in evolutionary, taxonomic, and phylogenetic investigations.[28-30] Phylogenetic analyses employing one or several genes have been performed in previous studies however, these analyses were restricted to the species of Aristolochiaceae or limited to phylogeny by one gene.[3,26,27] Two barcode genes *matK* and *rbcL* were previously used to resolve different plant species.[31,32] In comparison to other barcodes, *matK* and *rbcL* have a larger number of parsimony-informative characters.[1] Both the *rbcL* and *matK* phylogenies clearly distinguished *Aristolochia* from Piper, indicating a monophyletic origin with solid evidence.[33] Bootstrap support more than 70% for any clade is considered reliable.[34]

SSRs are commonly utilized as molecular markers in species identification, population genetics, and phylogenetic studies, Because of their high amount of polymorphism.[35,36] As a result, SSRs are widely used as a molecular marker for breeding, population genetics, genetic linkage map construction, and gene mapping.[37-39] Phylogenetic connections at practically any taxonomic level have been effectively resolved using complete chloroplast genomes and protein-coding genes in the previous decade.[28,40]

This finding (inferred from chloroplast genome data) had strong support values, indicating that the chloroplast genome was capable of efficiently resolving this family's phylogenetic locations and relations. However, in order to correctly depict the development of the Aristolochiaceae family, more species must be used to evaluate the phylogeny. This research will also serve as a guide for identifying species in the *Aristolochia* and other genus utilising the chloroplast genome.

CONCLUSION

The phylogenetic analyses illustrated that these *Aristolochia* species were positioned close to three species from the family Piperaceae and phylogenetically closer to *A. tagala*. The results provided the basis for the study of the evolutionary history of *A. bracteolata*. We also discovered SSR sites, which will be used to create tools for further research into *Aristolochia* species. All the data presented will facilitate the identification, authentication, to prevent adulteration and further investigation of these medicinal plants.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the institutional ethics committee

REFERENCES

1. Li X, Zuo Y, Zhu X, Liao S, Ma J. Complete chloroplast genomes and comparative analysis of sequences evolution among seven Aristolochia (Aristolochiaceae) medicinal species. Int J Mol Sci. 2019(5);20:1045.
2. Neinhuis C, Wanke S, Hilu K, Müller K, Borsch T. Phylogeny of Aristolochiaceae based on parsimony, likelihood, and Bayesian analyses of trn L-trn F sequences. Plant Syst Evol. 2005;250(1):7-26.
3. Silva-Brandão K, Solferini V, Trigo J. Chemical and phylogenetic relationships among Aristolochia L. (Aristolochiaceae) from southeastern Brazil. Biochem Syst Ecol. 2006;34(4):291-302.
4. Wu T, Damu A, Su C, Kuo P. Terpenoids of Aristolochia and their biological activities. Nat Prod Rep. 2004;21(5):594-624.
5. Das T, Latha R, Agastian P. Evaluation of Aristolochia bracteolata Linn. for antimicrobial activity, α-glucosidase inhibition, and its phytochemical constituents. Evaluation. 2016;9(1).
6. Gbadamosi I, Egunyomi A. In–Vitro Propagation and Antimycotic Potential of Extracts and Essential Oil of Roots of Aristolochia Bracteolata Linn.(Aristolochiaceae). Afr J Tradit Complement Altern Med. 2012;9(1):50-5.
7. Dechbumroong P, Aumnouypol S, Denduangboripant J, Sukrong S. DNA barcoding of Aristolochia plants and development of species-specific multiplex PCR to aid HPTLC in ascertainment of Aristolochia herbal materials. PloS One. 2018;13(8):e0202625.
8. Suliman Mohamed M, Timan Idriss M, Khedr I, Abd AlGadir H, Takeshita S, Shah M et al. Activity of Aristolochia bracteolata against Moraxella catarrhalis. Int J Bacteriol. 2014.
9. Sharma A, Namdeo A, Mahadik K. Molecular markers: New prospects in plant genome analysis. Pharmacogn Rev. 2008;2(3):23.
10. Hollingsworth P, Graham S, Little D. Choosing and using a plant DNA barcode. PloS one. 2011;6(5):e19254.
11. Ramundo S, Rochaix J. Controlling expression of genes in the unicellular alga Chlamydomonas reinhardtii with a vitamin-repressible riboswitch. Method Enzymol. 2015;550:267-81.
12. Wu M, Li Q, Xu J, Li X. Complete chloroplast genome of the medicinal plant Amomum compactum: gene organization, comparative analysis, and phylogenetic relationships within Zingiberales. Chin Med J. 2018;13(1):1-12.
13. Ganie S, Upadhyay P, Das S, Sharma M. Authentication of medicinal plants by DNA markers. Plant Gene. 2015;4:83-99.
14. Li Y, Korol A, Fahima T, Beiles A, Nevo E. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol. 2002;11(12):2453-65.

15. Zane L, Bargelloni L, Patarnello T. Strategies for microsatellite isolation: a review. Mol Ecol. 2002;11(1):1-16.
16. Mekvipad N, Satjarak A. Evolution of organellar genes of chlorophyte algae: Relevance to phylogenetic inference. PloS One. 2019;14(5):e0216608.
17. Ren X, Xin G, Jia G, Zhang X, Liu H, Yang C et al. Characterization of the complete chloroplast genome sequence of Tapiscia sinensis (Tapisciaceae). Conserv Genet Resour. 2018;10(4):765-68.
18. Somaratne Y, Guan D, Wang W, Zhao L, Xu S. The Complete chloroplast genomes of two Lespedeza species: insights into codon usage bias, rRNA editing sites, and phylogenetic relationships in Desmodieae (Fabaceae: Papilionoideae). Plants. 2020;9(1):51.
19. Huang Y, Wang J, Yang Y, Fan C, Chen J. Phylogenomic analysis and dynamic evolution of chloroplast genomes in Salicaceae. Front Plant Sci. 2017;8:1050.
20. Bhatt P, Thaker V. A comparative study on 193 plastomes of Poaceae for validity and implications of individual barcode genes and concatenated protein coding sequences with selected plastomes of grasses from the desert of India. Meta Gene. 2021;100921.
21. Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC genomics. 2012;13(1):1-7.
22. Wyman S, Jansen R, Boore J. Automatic annotation of organellar genomes with DOGMA. Bioinformatics. 2004;20(17):3252-5.
23. Lowe T, Chan P. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. 2016;44(1):54-7.
24. Sablok G, Padma Raju G, Mudunuri S, Prabha R, Singh D, Baev V et al. ChloroMitoSSRDB 2.00: more genomes, more repeats, unifying SSRs search patterns and on-the-fly repeat detection. Database. 2015.
25. The Editorial Committee of Flora of China. Flora of China; Science Press: Beijing, China; Missouri Botanical Garden Press: St. Louis, MO, USA. 2003;5:246-69.
26. Murata J, Ohi T, Wu S, Darnaedi D, Sugawara T, Nakanishi T, Murata H. Molecular phylogeny of Aristolochia (Aristolochiaceae) inferred from matK sequences. Acta phytotaxonomica et geobotanica. 2001;52(1):75-83.
27. Ohi-Toma T, Sugawara T, Murata H, Wanke S, Neinhuis C, Murata J. Molecular phylogeny of Aristolochia sensu lato (Aristolochiaceae) based on sequences of rbcL, matK, and phyA genes, with special reference to differentiation of chromosome numbers. Syst Bot. 2006;31(3):481-92.
28. Qian J, Song J, Gao H, Zhu Y, Xu J, Pang X, Yao H, Sun C, Li X, Li C, Liu J. The complete chloroplast genome sequence of the medicinal plant Salvia miltiorrhiza. PloS One. 2013;8(2):e57607.
29. Jansen K, Cai Z, Raubeson A, Daniell H, Depamphilis W, Leebens-Mack J, Müller F, Guisinger-Bellian M, Haberle C, Hansen K, Chumley W. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. PNAS. 2007;104(49):19369-74.
30. Moore J, Bell D, Soltis S, Soltis E. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. PNAS. 2007;104(49):19363-8.
31. Lin X, Dong J, Yang Q, Zhou W, Wang Y, Zhang Y, Ahmad M, Sun Y, Wang Y, Ling J. Identification of three seagrass species in coral reef ecosystem by using multiple genes of DNA barcoding. Ecotoxicology. 2021;30(5):919-28.
32. Amandita F, Rembold K, Vornam B, Rahayu S, Siregar I, Kreft H, Finkeldey R. DNA barcoding of flowering plants in Sumatra, Indonesia. Ecol Evol. 2019;9(4):1858-68.
33. Murali S, Rashmi T, Francis M. Phylogenetic relationship of selected Aristolochia spp. with different generic segregates inferred using rbcL and matK genes. J Med Plants. 2016;4:99-103.
34. Hillis D, Bull J. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Bot. 1993;42(2):182-92.
35. Jiao Y, Jia M, Li W, Chai L, Jia J, Chen Z, Wang Y, Chai Y, van de Weg E, Gao S. Development of simple sequence repeat (SSR) markers from a genome survey of Chinese bayberry (Myrica rubra). BMC genomics. 2012;13(1):1-16.
36. Xue J, Wang S, Zhou L. Polymorphic chloroplast microsatellite loci in Nelumbo (Nelumbonaceae). Am J Bot. 2012;99(6):240-4.
37. Gupta P, Varshney R. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica. 2000;113(3):163-85.
38. Kim K, Ratcliffe S, French B, Liu L, Sappington T. Utility of EST-derived SSRs as population genetics markers in a beetle. J Hered. 2008;99(2):112-24.
39. Liu S, Yin X, Lu J, Liu C, Bi C, Zhu H et al. The first genetic linkage map of Ricinus communis L. based on genome-SSR markers. Ind Crops Prod. 2016;89:103-8.
40. Li X, Zhang C, Qiao Q, Ren Z, Zhao J, Yonezawa T, Hasegawa M, Crabbe C, Li J. Zhong Y. Complete chloroplast genome sequence of holoparasite Cistanche deserticola (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host Haloxylon ammodendron (Chenopodiaceae). PloS One. 2013;8(3):e58747.

Cite this article as: Khandhar DB, Bhatt PP, Thaker VS. Authentication and identification markers for medicinal plants: I *Aristolochia bracteolata*. Int J Sci Rep 2022;8(1):7-22.

ANNEXURE

Table 1: Similar types and size remove of SSRs in the *A. bracteolata, A. contorta, A. debilis, A. kaempferi, A. kunmingensis, A. macrophylla, A. mollissima, A. moupinensis, A. tagala, A. tubiflora, Asarum sieboldii, P. cenocladum, P. kadsura, P. laetispicum* and *Saruma henryi* chloroplast genomes.

Chloroplast genomes	Size	Gene name	Start	End
SSR *A. bracteolata*				
(A) 13	13	-	107	119
(AT) 7tattttatttt (TTA) 4	37	*trnK-UUU*	1940	1976
(AATAT) 3ttattaaatattaatattaaaataatatata (AT) 10	66	*trnK-UUU – rps16*	4398	4463
(ATATA) 3	15	*trnK-UUU – rps16*	4574	4588
(TTTG) 3	12	*trnK-UUU – rps16*	5081	5092
(A) 14	14	*rps16 – trnQ-UCG*	6528	6541
(A) 25	25	*rps16 – trnQ-UCG*	6685	6709
(AATA) 3 (AT) 6*aaataaaaaaaaaaataaataataataaataaaaaatatattctattttaatatttttaatatt (TA) 6	96	*atpH – atpL*	14334	14429
(ATTAT) 4atgggtccctgtacaaaatatggccctgtacaaatatgatccagacac (AT) 7	82	*psbZ – trnG-GCC*	38936	39017
(AT) 6	12	*ycf3 – trnS-GGA*	47988	47999
(CATAT) 3	15	*psaL – ycf4*	62896	62910
(AAAT) 3aaaat (ATAA) 3	29	*rpl33 – rps18*	71568	71596
(T) 14	14	*rps18 – rpl20*	72211	72224
(TAATT) 3	15	*rps18 – rpl20*	72348	72362
(TATAT) 3	15	*clpP*	75444	75458
(T) 16	16	*rps11 – rpl36*	83463	83478
(AT) 7ctatatatttcta (T) 17	44	*rpl16*	85779	85822
(A) 16gaagtagtaaataaattatgtattctatacaccagacgaatcagtgggttacctga (ATTT) 3	84	*psaC – ndhE*	120962	121045
(ATAGA) 3aggatatccatgtttatttata (TCTAT) 3attacgactatttattcaacaaat (TTGA) 3	88	*psaC – ndhE*	121212	121299
(T) 13	13	*ndhE – ndhG*	121782	121794
SSR *A. contorta*				
(CATATA) 3	18	*trnH-GUG – psbA*	110	127
(A) 12	12	*psbA – trnK-UUU*	1555	1566
(TA) 9	18	*trnK-UUU*	3961	3978
(ATA) 4	12	*trnK-UUU – rps16*	4409	4420
(AT) 7	14	*trnK-UUU – rps16*	4578	4591
(TTTA) 3	12	*trnG-UCC*	9853	9864
(A) 17	17	*trnG-UCC – trnR-UCU*	10233	10249
(ATT) 4	12	*atpF*	13000	13011
©14(T) 14	28	*petN – psbM*	30575	30602
(ATATA) 3tatgatacgg (TAATA) 3	40	*trnT-GGU – psbD*	34731	34770
(T) 15ggactcatatttaattttaatgtgtctcacagtcggaattccttcgggggaggggtcacttcttgaatg (A) 14	98	*psbZ – trnG-GCC*	38602	38699
(TTA) 4taatatactattatctattataatataattatactatttgtttatttattat (TA) 8a (AT) 6aatatttataatatttttatataatatatattaatataataat (ATA) 4tatagatatagtttat (TA) 9	182	*trnT-UGU – trnL-UAA*	49428	49609
(AT) 7tcatattcaa (AT) 10	44	*ndhC – trnV-UAC*	53971	54014
(TTA) 4	12	*ndhC – trnV-UAC*	54606	54617
(T) 14	14	*trnV-UAC – trnM-CAU*	56282	56295

Continued.

Chloroplast genomes	Size	Gene name	Start	End
(ATTA) 4	16	*accD – psaL*	63256	63271
(A) 13tgcatatc (AT) 7	35	*psaL – ycf4*	64015	64049
(TAT) 4attatgtaattatgtatgatc (A) 13	46	*petA – psbJ*	67825	67870
(ATAA) 3attaaaattataatataaaattataaaataaaaaatataattatataattaaattaataatattatattatatattaatttaat (TA) 7aat (TA) 9atataataatata (TTCTAT) 3	162	*rpl33 – rps18*	72707	72868
(A) 12tataatttttatttaatttaatattaataatttaatattaaataaatattatattaatttattttataatttaatataatattaatatat (ATA) 4aaatggggggagagagggagagaaatc (T) 14	157	*rps18 – rpl20*	73458	73614
(A) 12gaaattttactgatatcgatattaagtatcaaataaatagtaatagtaataaaaaatacgtagtaaatgg (ATAA) 3	94	*rpl20 – rps12*	74292	74385
(TAT) 4	12	*ndhF – rpl32*	118540	118551
(ACATA) 3	15	*rpl32 – trnL-UAG*	119344	119358
(A) 12	12	*rpl32 – trnL-UAG*	119649	119660
SSR *A. debilis*				
(T) 14cattgaataaattaacttttcttcttgtgccgtagtgaattttgtgaacccatcatttaa (ATG) 4	86	*trnK-UUU – rps16*	4728	4813
(TATTA) 3	15	*trnS-GCU – trnG-UCC*	8349	8363
(T) 12	12	*trnS-GCU – trnG-UCC*	8682	8693
(TATAA) 3atatataaaataaataatttatatatattctattttaatttttaattta (AT) 6	76	*atpH – atpL*	14143	14218
(T) 12	12	*rpoB*	26621	26632
(T) 12ca (T) 12	26	*trnC-GCA – petN*	29448	29473
©13	13	*petN – psbM*	30202	30214
(ATA) 4	12	*petN – psbM*	30643	30654
(T) 14	14	*trnT-GGU – psbD*	34255	34268
(T) 13attcatagattagattatatctatttctatatttatatttctttctatttc (TA) 7	78	*psbC – trnG-UGA*	37750	37827
(AT) 11g (TA) 6	35	*psbZ – trnG-GCC*	38925	38959
(AATCA) 3	15	*ycf3*	46967	46981
(ATA) 4taaatataatattaatatataaatataatattaataatattaatt (ATA) 4	69	*trnT-UGU – trnL-UAA*	49660	49728
(A) 16	16	*trnL-UAA*	50367	50382
(AT) 6	12	*ndhC – trnV-UAC*	54289	54300
(TTAAT) 3	15	*rbcL – accD*	60915	60929
(TAT) 4gtatgatcc (A) 12	33	*petA – psbJ*	67863	67895
(TTTTA) 3	15	*petA – psbJ*	68093	68107
(A) 17	17	*psbE – petL*	70523	70539
(TAAC) 3	12	*trnP-UGG – psaJ*	71759	71770
(TA) 6attataataaattctatttctattataatagaatatttaatttagatagtttatctataaactaaacata (TAT) 4	94	*rpl33 – rps18*	72831	72924
(A) 16	16	*rps18 – rpl20*	73456	73473
(TA) 7	14	*trnL-CAA – ndhB*	99799	99812
(TCTA) 3	12	*psaC – ndhE*	122789	122800
(TTGA) 3	12	*ndhE*	122914	122925
(A) 16	16	*rps15 – ycf1*	129152	129167
(AGTATC) 3	18	*ycf1*	132810	132827
(AT) 7	14	*ndhB – trnL-CAA*	149590	149603
SSR *A. kaempferi*				
(TAATA) 3 (ATA) 5* (AT) 13*	47	*trnK-UUU – rps16*	4169	4215

Continued.

Chloroplast genomes	Size	Gene name	Start	End
(CCTT) 4atcttatatatatgcttatatatatgatatatatgatatatgcttatatatatgct (TA) 6tggaaatggaaggataggta (T) 12	116	*rps16 – trnQ-UUG*	6028	6143
(TATT) 5tatctattattaatat (AATA) 3	48	*trnT-UGU – trnL-UAA*	49375	49422
(T) 13	13	*rpl16*	87182	87194
SSR *A. kunmingensis*				
(AT) 8aaatataataataaatataatataaatatataaaatagaatataaaataatat aattctaatataaat (ATA) 4tattatata (AT) 6	117	*trnK-UUU – rps16*	4250	4366
(ATT) 4	12	*trnG-UCC*	9071	9082
(CATAT) 3ttataatatatataatc (CATAT) 3ttatataatatata (AT) 8	77	*trnG-UCC – trnR-UCU*	9742	9818
(TTTCAT) 3ttaaaatttaat (A) 12	42	*atpF*	12913	12954
(T) 12cttaatttccttatttaaaatttaaataaagaaatggaatctattctaat (TTTA) 4	78	*atpF – atpH*	13234	13311
(G) 19cccccccataaaaggcgggaatgataggccattgcgcaa (G) 27	85	*atpF – atpH*	13427	13511
(AT) 7tatatataatatatataaaataaataatataataatattaaattaat (ATA) 4	73	*atpH – atpL*	14356	14428
(A) 12	12	*rpoB – trnC-GCA*	28595	28606
(TATAT) 3	15	*psbM – trnD-GUC*	31611	31625
(TTATT) 3	15	*trnS-GGA – rps4*	48559	48573
(TATT) 3tatc (TATT) 3tatctattattaatat (AATA) 3aagat (AATA) 3aagagtagtttgttt (TTA) 4tatatattatatatattattatata (TAT) 4atatattaaattaaatatatat (TA) 6aatacaatatcaatattatattaataatacaatat (TTTAA) 3attaaaattaaatagaatattaattatgacaacattaacattaatattaataatt acttaatattaataaaagaaact (A) 12	311	*trnT-UGU – trnL-UAA*	49683	49993
(A) 16gattaaaaaaaaaaagaaaataaaaatatta (TATAT) 3	62	*ndhC – trnV-UAC*	54307	54368
(ATA) 4	12	*atpB*	58742	58753
(TTA) 4	12	*rps18*	73202	73213
(ATATA) 3	15	*clpP*	74990	75004
(T) 12	12	*ndhF*	117249	117260
SSR *A. macrophylla*				
(AATATA) 3aaataataatataaat (ATAA) 3	46	*trnK-UUU – rps16*	4116	4161
(TAATAT) 3aatatataat (A) 15	43	*trnK-UUU – rps16*	4309	4351
(TA) 9tgct (TA) 6tggaaatggaaggataggtattttttatgct (TA) 6	77	*rps-16 – trnQ-UUG*	6222	6298
(AT) 6aatatat (ATAAA) 3taataatataaattataataaatataaaa (AATAT) 3	78	*atpH – atpL*	14296	14373
(ATAAG) 3	15	*rpoB – trnC-GCA*	28140	28154
(AT) 7	14	*psbM – trnD-GUC*	31486	31499
(AT) 10c (AT) 9	39	*psbZ – trnG-GCC*	38937	38975
(AATAT) 3ataataacaatataatatataataaaataataaaaataaaat (AATA) 3ttaatataat (TAATA) 3attaata (AT) 7	115	*ycf3*	46940	47054
(TATT) 3tatctattattaatat (AATA) 3aagagtagtttgttt (TTA) 4taattatatatttata (TAT) 4atttatattaattatat (TA) 8	128	*trnT-UGU – trnL-UAA*	49648	49775
(TAA) 5c (ATT) 6aatgttgttatg (TAA) 5tattaatataatataattatatattatattattta (TAT) 4taattctattatgattatgtaatattagtaatattaattactt (AATAA) 3act (A) 14	183	*trnT-UGU – trnL-UAA*	49878	50060
(TTTAA) 3tttctgaa (T) 13	36	*trnF-GAA – ndhJ*	52075	52110
(A) 12gaattaaagactatttgaaataaaataatattaaaataatatgaatattattt ctattcataaaatcat (A)	168	*ndhC – trnV-UAC*	54390	54557

Continued.

Chloroplast genomes	Size	Gene name	Start	End
17gattaaaattaaatataaatatataacataacaatgttatattaatattaatat (ATTA) 4				
(ATA) 5aataataatat (ATATA) 3	41	*ndhC – trnV-UAC*	55842	55882
(AATA) 4aaatctaattaaataatttaaaatgaaattagaatttaattaatataattataattaaatta (ATAAT) 3tattaaataaatattaaatattttaaattctaaaaatgaaaataaaaaataaacaatatattatattgttatacatatattatattgtta (TAT) 4attatttatattatt (ATA) 4	222	*rpl33 – rps18*	72870	73091
(TTTC) 3	12	*rpl20 – rps12*	74163	74174
(T) 12	12	*clpP – psbB*	77508	77519
(T) 15	15	*rpl16 – rps3*	88679	88093
SSR *A. mollissima*				
(ATA) 5taat (TAATA) 3 (AT) 7	48	*trnK-UUU – rps16*	4146	4193
(TA) 6tggaaatggaaggataggta (T) 17ctggagaagtctcagcaaggatttaacatgatcc (TATAT) 3	98	*rps-16 – trnQ-UUG*	6088	6185
(AT) 6	12	*trnG-UCC – trnR-UCU*	9617	9628
(T) 17	17	*atpF – atpH*	13024	13040
(TTA) 4	12	*trnt-GGU – psbD*	33693	33704
(TATT) 5tatctattattaatat (AATA) 3aagagtcgtttgtttttattattatattattatttttattattatatatattatatatattatatatattaaatta (AATAT) 3taattatgacaacattaatattatattaataattacttaatattaataaaataaact (A) 16	212	*trnT-UGU – trnL-UAA*	49410	49621
(A) 19g (ATAAA) 3	35	*ndhC – trnV-UAC*	53909	53943
(TATAT) 3	15	*psaJ – rpl33*	71702	71716
(T) 17	17	*rpl32 – trnL-UAG*	118292	118308
(TATC) 3catattctatccatatatatatagaatatatagaaggatatctgtgtttatttatatctatatcattacgactacgattatttattcaacaaat (TTGA) 3	118	*psaC – ndhE*	122140	122257
(TTTATT) 3	18	*ndhE – ndhG*	122755	122772
SSR *A. moupinensis*				
(A) 18	18	*matK*	3405	3422
(TAATA) 3ataaaaaaaaatatataaaatataattataatattaat (ATA) 4ttaattagataattatattattataatataata (AT) 11	120	*trnK-UUU – rps16*	4200	4319
(CATAT) 3ttataatatatataatatta (AT) 6	47	*trnG-UCC – trnR-UCU*	9714	9760
(AT) 7aaaataaataataatattaa (TAT) 4aattataataaataaaattataataaataaaa (AATAT) 3	93	*atpH – atpL*	14276	14368
(AAAAAG) 3	18	*psbM – trnD-GUC*	30398	30415
(ATA) 5taataag (TATAA) 3	37	*ycf3*	46364	46400
(TATT) 3tatctattattaatatttaatat (AATA) 3aagat (AATA) 3aagagtagtttgtttttattattcttatatatattatataa (ATTAT) 3atatt (TA) 7aataatatataaatacaatatcaatattatat (TAA) 4tacaatatttaatttaaatttaaattaaatagaatattaattatgacaacattaacattaatattaataattacttaatattaataaaataaact (A) 13	291	*trnT-UGU – trnL-UAA*	49130	49420
(T) 12	12	*ndhA*	124832	124843
SSR *A. tagala*				
(A) 14	14	*trnH-GUG – psbA*	95	108
(AAAT) 3	12	*psbA – trnK-UUU*	1544	1555
(TTAAA) 3tattaaattaataaaaattaaataatattaatttaatataattataataaatta (AAT) 4aaaaaataaaaaataaataa (AAT) 4	113	*trnK-UUU*	3933	4045
(AATAT) 3ttattaaatataatattttattaaatataatataaatataaaaatttattaaatata	158	*trnK-UUU – rps16*	4412	4569

Continued.

Chloroplast genomes	Size	Gene name	Start	End
atataaatataaaaatataatataaatataaaaataaaata (TTTAAT) 3ataatt (ATA) 4 (AT) 6*				
(TTTG) 3ttagggaagtacatcc (A) 13	41	*trnK-UUU – rps16*	5096	5136
(T) 14	14	*rps16 – trnQ-UCC*	6461	6474
(A) 24	24	*rps16 – trnQ-UUC*	6716	6739
(AT) 7aaagaaagataaag (A) 15	43	*psbM – trnD-GUC*	31656	31698
(A) 18taaagaaaaaaaaagttgtttgtagggagaggggtcacttcttgaatg (A) 16tgaataatctaatgttatgttaaaataaaataattggaatccattatattatatggccctgtacaaaatatggccctgtacaaatatgatccagacac (AT) 6	192	*psbZ – trnG-GCC*	38878	39069
(TTTA) 3actatttgtttatttattattaatataataaatagaaattaagaaatagtagtttgtttttatta (ATATT) 3atttaatatatatataat (ATTAA) 3taaaaataaattatttt (TAAA) 3	154	*trnT-UGU – trnL-UAA*	49749	49902
(AT) 7caaatttct (AATAA) 3taaaaaaatattattaattattatattataatataattaaaataataaataatattaataata (AAT) 4attaatataaataatataatatatagaatatagaaatatatagaataaaataaggaaaaaataataggcctgacttgggttaattttctc (T) 12	216	*ndhC – trnV-UAC*	54224	54439
(A) 13	13	*psaL – ycf4*	63971	63983
(ATAA) 3aatct (AATTA) 3aaaattaataattattaataattatattataatgttatatttagatattatatataaaaataataatt (ATTATA) 3	119	*rpl33 – rps18*	72663	72781
(AAT) 6gggggggggagataaata (T) 17	53	*rps18 – rpl20*	73269	73321
(TA) 7	14	*rpl16*	86870	86883
(A) 16	16	*psaC – ndhE*	122045	122060
(AT) 6agaaggatatccatgtttatttatatctattctatatcattacgactacgactatttattcaacaaat (TTGA) 3	92	*psaC – ndhE*	122287	122378
SSR *A. tubiflora*				
(TA) 6	12	*psbA – trnK-UUU*	1373	1384
(T) 16	16	*trnK-UUU*	1836	1851
(ATA) 4	12	*trnK-UUU*	3916	3927
(TA) 7	14	*rps16 – trnQ-UUG*	6486	6499
(A) 19tggaatcacattgtatccaatattcccccactttta(G) 12	67	*rps16 – trnQ-UUG*	6789	6855
(T) 12	12	*rpoB – trnC-GCA*	28318	28329
(T) 16ctgtttcaatatcaat (A) 13	45	*rpoB – trnC-GCA*	28474	28518
©12 (A) 13	25	*rpoB – trnC-GCA*	28647	28671
(AATAA) 3	15	*psbM – trnD-GUC*	31864	31878
(AT) 6taaatatatttaaattaaatattaaataaa (AT) 6tataaatataata (TAT) 4	79	*ndhC – trnV-UAC*	54269	54347
(A) 14 (AAT) 4*	24	*ndhC – trnV-UAC*	54577	54600
(T) 16	16	*petG – trnW-CCA*	71271	71286
(T) 13	13	*trnP-UGG – psaJ*	71761	71773
(ATAA) 4	16	*rpl33 – rps18*	72925	72940
(AAT) 4	12	*rps18 – rpl20*	73628	73639
(T) 14caatttggtcgatttcatgccttccgccaataccaaatatttgatgtattcatcatattactttattgattggcagtttgagatggcttata (TTAT) 3	118	*clpP*	76791	76908
©16	16	*rps11 – rpl36*	84983	84998
(TG) 6 (T) 18	30	*rpl32 – trnL-UAG*	119179	119208
SSR *Asarum sieboldii*				
(AT) 6aaatatataaatataaattttg (AT) 6 (ATATAA) 4*	66	-	5	70
(A) 14	14	-	354	367

Continued.

Chloroplast genomes	Size	Gene name	Start	End
(AATTAT) 3	18	*psbM – petN*	3366	3383
(TTTA) 3	12	*trnC-GCA – rpoB*	4239	4250
(A) 15	15	*trnC-GCA – rpoB*	4466	4480
(TTAAT) 3	15	*trnC-GCA – rpoB*	5353	5367
(ATATT) 3	15	*atpL – atpH*	18062	18076
(AAAAAT) 3 (AAAT) 4*agattcctattaaat (TTATA) 3attatattatataattatataaatattatattatataatt (ATATAA) 3t (ATATA) 3taatatt (ATATA) 3 (TAATA) 3	171	*atpH – atpF*	19297	19467
(AT) 6aaaaatataaaa (AT) 7a (AT) 6aatataaatatataaatataataaatatataaatataaattatataattatataaatatataaattatataaataaat (TA) 6aataaattatatattatataaataaattatataaataaatt (ATATAA) 3ataaatatattaaatatataaatatataatatatatataatatatttataatatattaaatatataatatattaaatattcaatattaaatatata (AT) 6aa (AT) 6aaa (AT) 6a (AT) 6aatatatatataaatatatataaa (AT) 6a (AT) 6aatatatatataaatatatatata (AT) 6aatatatatataa (AT) 6aaatataaataaaatataataaaatataaataaaatataaataaaatataataaaatataaatataaataaaatatagtaaaat (ATAA) 4	560	*atpH – atpF*	19602	20161
(T) 12	12	*atpF – atpA*	21584	21595
(A) 14	14	*trnG-UCC*	23938	23951
(A) 12ttatgcttcgcgattccataatccaatttttgattcaattttaaattgacct (TTAGA) 4	84	*trnG-UCC*	24054	24137
(AATTAT) 3	18	*trnG-UCC – trnS-GCU*	24948	24965
(AT) 13	26	*trnG-UCC – trnS-GCU*	25146	25171
(T) 16	16	*trnQ-UUG – rps16*	26942	26957
(T) 12	12	*rps16*	29042	29053
(T) 16	16	*rps16 – trnK-UUU*	29836	29851
(A) 18	18	*rps16 – trnK-UUU*	30054	30071
(TTTTTA) 3	18	*rps16 – trnK-UUU*	30270	30287
(AT) 6	12	*trnK-UUU – psbA*	33175	33186
(AT) 6aaatataaaaaatatataaatatataaatatatataaaaatagaaatttataaaatttttataaatttc (TAT) 4	94	*psbA – trnH – GUG*	35128	35221
(AT) 6	12	*psbA – trnH – GUG*	35326	35337
(TA) 6	12	*psbA – trnH – GUG*	35445	35456
(TA) 7ttatt (TA) 7	33	*psbA – trnH – GUG*	35716	35748
(ATTTAT) 3	18	*psbA – trnH – GUG*	36058	36075
(AAT) 6aaaataatataataaaaaaaataattaataataaatagaattctataataaatagaattataataa (AAT) 4	96	*trnH-GUG – trnT-GGU*	36656	36751
(TATAT) 3	15	*trnH-GUG – trnT-GGU*	37503	37517
(ATATA) 3	15	*trnH-GUG – trnT-GGU*	37668	37682
(ATA) 4aatataaattattctaattattagaattaatatattattataaata (TAT) 6aaatatattattataattaatataataaatataaattaatataattctaataaat (TAATA) 3aatttattagaattaatattaatataataaatataaattaatataataaat (TAATA) 3	212	*trnH-GUG – trnT-GGU*	37966	38177
(ATATA) 4	20	*trnH-GUG – trnT-GGU*	38376	38395
(TATATG) 3	18	*trnH-GUG – trnT-GGU*	38618	38635
(A) 12	12	*trnT-GGU – psbD*	38910	38921
(AT) 6a (AT) 6aatataga (AT) 6aatatttcttatatt (TA) 12	84	*trnT-GGU – psbD*	39811	39894
(AT) 7catatatatat (TA) 7	39	*trnT-GGU – psbD*	40024	40062
(TATT) 3	12	*psbC – trnS-UGA*	43158	43169

Continued.

Chloroplast genomes	Size	Gene name	Start	End
(A) 13	13	*trnG-GCC – trnfM-CAU*	44300	44312
(TAATA) 3ttaatataatat (TAATA) 3	42	*ycf3*	52162	52203
(AT) 6tata (AT) 6 (ATA) 4*tatatattatatat (ATA) 5taatat (ATA) 4tataataataattcataattatattattatattaataaaatattaataaataattatattattatattaattatattattatattaata (AAT) 5atataataagaattcataattatat (TATTA) 3attatattaataaataataatatataataatataaatataaataataatatatataataagaattcataattatat (TATTA) 3	318	*ycf3*	52343	52660
(ATATA) 3ataattataaataaaataattataa (ATAAT) 3	55	*ycf3*	52845	52899
(AATAT) 3	15	*ycf3*	53090	53104
(ATA) 4	12	*ycf3*	53242	53253
(ATA) 6aatataataatataaataaaataatataaat (AATA) 3	61	*ycf3*	53626	53686
(AAAT) 3	12	*rps4 – trnT-UGU*	55779	55790
(A) 12	12	*ndhC – trnV-UAC*	61080	61091
(ATGAA) 3ttatgaattattctatttatattta (TATAT) 3attattata (TAT) 6a (TAT) 5atat (TTA) 5	117	*ndhC – trnV-UAC*	61200	61316
(TTA) 4tatttatatattatg (TAT) 4	39	*ndhC – trnV-UAC*	61683	61721
(TATAT) 3attatatta (TATAT) 3att (TTATA) 3	57	*ndhC – trnV-UAC*	61836	61892
(TA) 6	12	*ndhC – trnV-UAC*	62054	62065
(TATAT) 3tttta (TATAT) 3taatatataatatattaattatatta (TATAT) 3attattatatataatttaaaatattatatattatttaaaataaaatatataattataattaataattatttaaaatatataatatattaattatatta (TATAT) 3attattttattatattattttattatattattttattatattattttatta (TATTT) 3ta (TATAT) 3	272	*ndhC – trnV-UAC*	62219	62490
(AT) 8aataatatataataa (AT) 10	51	*ndhC – trnV-UAC*	62739	62789
(TAATA) 3gaatataattat (TAATA) 3	42	*ndhC – trnV-UAC*	63136	63177
(ATAGAA) 3	18	*ndhC – trnV-UAC*	63999	64016
(T) 13	13	*trnM-CAU – atpE*	65846	65858
(AAAT) 3	12	*psaL – ycf4*	73232	73243
(AAG) 4	12	*ycf4 – petA*	74241	74252
(A) 12	12	*ycf4 – petA*	74917	74928
(A) 13	13	*ycf4 – petA*	75057	75069
(AATG) 3	12	*ycf4 – petA*	75732	75743
(CTT) 4	12	*psaj – rpl33*	81834	81845
(A) 12	12	*rps18 – rpl20*	82686	82697
(ATAA) 3 (TAAA) 3	24	*rpl20 – rps12*	83486	83509
(TAT) 4attattatta (TAT) 4aaatattcttattataaatataaaaaattattattataaatataaaaaattataaa (TAT) 4	102	*petD – rpoA*	92221	92322
(ATATAA) 3aaaaatattctattatttataaaattaaatttatataaatttataaaataaaatataaatataaaatattata (ATTAT) 3a (TAT) 4atattataattatataattatattataaata (TATAT) 3 (TAT) 4*	176	*petD – rpoA*	92521	92696
(TTGTT) 3	15	*ycf2 – trnL-CAA*	108764	108778
(A) 16	16	*rrn5 – trnR-ACG*	123268	123283
(AATACC) 4	24	*ycf1*	126504	126527
(GAAT) 3	12	*ycf1*	128299	128310
(A) 12	12	*ycf1*	128952	128963
(ATT) 4	12	*ycf1 – rps15*	130475	130486
(A) 12taaggctttgttctgtcttattgttatcacatccatttcccttcaattctattttcatattgttcatatatacatatattgttcatatatac (AT) 6aactataac (AT) 6	137	*ndhG – ndhE*	135813	135949
(AAT) 4	12	*ndhG – ndhE*	136114	136125

Continued.

Chloroplast genomes	Size	Gene name	Start	End
(A) 12	12	*ccsA – trnL-UAG*	139975	139986
(T) 13ctagctcattttttgatgattcagttctcatttatctgatttcatagatatgaaatgaaaaaa (AAAAT) 3	90	*trnL-UAG – rpl32*	140448	140537
(T) 15	15	*trnL-UAG – rpl32*	141116	141130
(TA) 12	24	*rpl32 – ndhF*	141913	141936
(AT) 12	24	*ndhF – rpl32*	147954	147977
(A) 15	15	*rpl32 – trnL-UAG*	148761	148775
(TATTT) 3ttttttcatttcatatctatgaaatcagataaatgagaactgaatcatcaaaaaatgagctag (A) 13	91	*rpl32 – trnL-UAG*	149353	149443
(T) 12	12	*trnL-UAG – ccsA*	149905	149916
(TAT) 4	12	*ndhE – ndhG*	153765	153776
(TA) 6tgttatagt (TA) 6tgtatatatgaacaatatatgtatatatgaacaatatgaaaatagaattgaagggaaatggatgtgataacaataagacagaacaaagcctta (T) 12	138	*ndhE – ndhG*	153941	154078
(CATT) 3	12	*ycf1*	161580	161591
(GTATTG) 4	24	*ycf1*	163359	163382
(AACAA) 3	15	*trnL-CAA – ycf2*	181113	181127
SSR *P. cenocladum*				
(A) 12	12	*trnK-UUU – rps16*	4798	4809
©13	13	*rps16*	5540	5552
(T) 12	12	*rps16 – trnQ-UUG*	6663	6674
(AT) 8	16	*rps16 – trnQ-UUG*	7193	7208
©17	17	*trnS-GCU – trnG-GCC*	9314	9330
(TTGG) 3	12	*ndhC – trnV-UAC*	53464	53475
(T) 12	12	*rps8 – rpl14*	83823	83834
(T) 12	12	*rps12*	111935	111946
(A) 14	14	*rps12*	114502	114515
(ATAA) 3aatattattaaaatataaatgaaaaaaaaaaagttttcatgaataatacgtaaagtgaagt (A) 12	85	*ccsA – ndhL*	127169	127253
(AT) 7	14	*trnL-UAG – rpl32*	129245	129258
(T) 14	14	*ndhF – ycf1*	133172	133185
(T) 14	14	*ndhF – ycf1*	133778	133791
(A) 12	12	*trnR-ACG – 5srRNA*	136347	136358
SSR *P. kadsura*				
(A) 13	13	*trnK-UUU*	1892	1904
(T) 14	14	*psbL – trnS-GCU*	8746	8759
(T) 12	12	*ndhC – trnV-UAC*	53513	53524
(ATTACT) 3	18	*petA – psbJ*	68102	68119
(ATA) 4 (A) 13*	24	*trnP-UGG – psaJ*	71413	71436
(ATAGT) 4	20	*psbH – petB*	79927	79946
(ACTTA) 3atttacatt (TACTTA) 3	42	*rps12 – trnV-GAC*	104193	104234
(T) 14ctt (A) 13	30	*rrn5 – trnR-ACG*	113389	113418
(TTAA) 3aaataattagataaaatattataaaaataaaaatataaataaatgaaaaaaaaaagttttcatgaataatatgtaaaatgaagt (A) 12	108	*ndhD – ccsA*	128573	128680
(A) 12	12	*rpl32 – ndhF*	131680	131691
(T) 13aag (A) 14	30	*trn-ACG – rrn5*	137168	137197
(TAAGTA) 3aatgtaaa (TTAAG) 3	41	*trnV-GAC – rps12*	146352	146392
SSR *P. laetispicum*				
(TATAA) 3	15	*trnK-UUU – rps16*	4991	5005
(ATTAA) 3	15	*atpF – atpH*	14146	14160
(T) 12	12	*petN – psbM*	30156	30167

Continued.

Chloroplast genomes	Size	Gene name	Start	End
(ATTAAT) 3	18	*petN – psbM*	30285	30302
(T) 12	12	*trnD-GUC – trnY-GUA*	32313	32324
(T) 12	12	*psbC – trnS-UGA*	37801	37812
(A) 12	12	*rpl14 – rpl16*	86027	86038
(TATTAG) 3	18	*rps12 – trnV-GAC*	104998	105015
(T) 14aag (A) 12	29	*rrn5 – trnR-ACG*	113563	113591
(T) 12ctt (A) 14	29	*trnR-ACG – rrn5*	137355	137383
(ACTAAT) 3	18	*trnV-GAC – rps12*	145930	145947
SSR *Saruma henryi*				
(ATA) 6taaataa (AAT) 5ata (AAT) 4ata (AAT) 4	70	*trnK-UUU – rps16*	2	71
(T) 12	12	*trnK-UUU – rps16*	4881	4892
(A) 16	16	*rps16*	5063	5078
(A) 15	15	*psbK – psbL*	5856	5870
(A) 14	14	*trnS-GCU – trnG-UCC*	8453	8466
(TA) 7	14	*trnS-GCU – trnG-UCC*	9326	9339
(CAA) 4	12	*trnS-GCU – trnG-UCC*	9570	9581
(T) 14	14	*trnG-UCC*	9916	9929
(T) 14	14	*trnG-UCC*	10545	10558
(T) 14	14	*trnG-UCC*	10661	10674
(T) 12	12	*atpA – atpF*	10925	10936
(A) 12	12	*atpF*	13101	13112
(TTAT) 3aatataaatttataaataggaatcc (ATTT) 4	53	*atpH – atpL*	14530	14582
(T) 12	12	*rpoC1*	15884	15895
(T) 12	12	*rpoB – trnC-GCA*	24371	24382
(A) 14t (ATAA) 3	27	*psbC – trnS-UGA*	29459	29485
(AGAT) 3	12	*pafL*	31671	31682
(TA) 9	18	*pafL – trnS-GGA*	34676	34693
(TATT) 4	16	*rps4 – trnD-GUC*	37737	37752
(AAT) 4	12	*pafL*	46710	46721
(A) 12	12	*pafL – trnS-GGA*	46978	46989
(AAAT) 3	12	*rps4 – trnT-UGU*	48741	48752
(TATTT) 3	15	*trnL-UAA*	50269	50283
(TAT) 4ttttattaatatctaatat (AATA) 3	43	*ndhC – trnV-UAC*	53770	53812
(ACAAT) 3	15	*ndhC – trnV-UAC*	54320	54334
(T) 15atttatttttatcttttccctattattctccccctagtcatacttgtttgta (T) 13	81	*trnM-CAU – atpE*	56273	56353
(A) 13	13	*atpB – rbcL*	58883	58895
(AAAT) 3	12	*psaL – pafll*	63646	63657
(AAG) 4	12	*pafll – cemA*	64654	64665
(T) 13	13	*petA – psbJ*	68086	68098
(CTT) 4	12	*psaJ – rpl33*	72210	72221
(ATAA) 3	12	*rpl20 – rps2*	73813	73824
(T) 12	12	*clpP*	76088	76099
(T) 16	16	*psbB – psbT*	78608	78623
(A) 21	21	*rrn5 – trnR-ACG*	112936	112956
(ATT) 4cctgatcggtaaaatccatatttgtggaaaat (A) 12	56	*ycf1 – rps15*	120056	120111
(T) 12	12	*ndhA*	122815	122826
(AT) 7	14	*ndhG – ndhE*	125483	125496
(ATAG) 3	12	*ndhE – psaC*	125973	125984
(A) 13	13	*ndhE – psaC*	126134	126146
(T) 15	15	*trnL-UAG – rpl32*	130146	130160
(AAAAG) 3	15	*rpl32 – ndhF*	130726	130740
(T) 20	20	*trnR-ACG – rrn5*	135603	135622

Table 2: Chloroplast genome sequences used for phylogenetics tree construction with their genebank ID.

Species code	Species name	Genebank ID
a1	*Aristolochia bracteolata*	MT_130705
a2	*Aristolochia contorta*	NC_036152.1
a3	*Aristolochia debilis*	NC_036153.1
a4	*Aristolochia kaempferi*	NC_041452.1
a5	*Aristolochia kunmingensis*	NC_041451.1
a6	*Aristolochia macrophylla*	NC_041453.1
a7	*Aristolochia mollissima*	NC_041457.1
a8	*Aristolochia moupinensis*	NC_041454.1
a9	*Aristolochia tagala*	NC_041455.1
a10	*Aristolochia tubiflora*	NC_041456.1
a11	*Asarum sieboldii*	NC_037190.1
a12	*Piper cenocladum*	NC_008457.1
a13	*Piper kadsura*	NC_027941.1
a14	*Piper laetispicum*	NC_042254.1
a15	*Saruma henryi*	NC_039933.1
p1	*Piper auritum*	NC_034697.1
p2	*Piper nigrum*	NC_034692.1
p3	*Arabidopsis thaliana*	NC_000932.1