Genetic diversity of *Arum L.* based on plastid marker

Leila Joudi Ghezeljeh Meidan[1]* and Mostafa Assadi[2]

1. Department of Agriculture and Animal Science, Islamic Azad University, Shabestar Branch, Shabestar, Iran
2. Research Institute of Forest and Rangelands, National Botanical Garden of Iran, Tehran, Iran

Abstract

TrnL-F region including intron trnL (UAA) and trnL (UAA) - trn (GAA) spacer in the large single-copy region of the chloroplast genome is widely used to infer phylogenetic relationships in plants. In this study, we obtained the trnL-F sequences from 8 samples of *Arum* L. in Iran. Phylogenetic analyses were conducted by the Bayesian inference, maximum parsimony, and maximum likelihood methods. The cladistics analysis of phylogenetic relationships indicated that all species constituted a monophyletic group within the Arae clade. *Eminium spiculatum* were considered as outgroups. *Biarum* and *Arum* species were placed in a monophyletic clade. *A. conophalloides* Schott. and *A. virescense* Stapf. were placed in one clade. *A. kotschyi* Boiss. and *A. korolkowii* L. formed a sister group. *Arum giganteum* Ghahreman. has been introduced as a new species in Iran, but in *Arum* monograph it is mentioned as potential equivalent to *Arum rupicola* Boiss. Molecular studies in this research can separate these two species and confirm previous studies.

Keywords: *Arum*; Iran; phylogeny; trnL-F region; plastid marker

Joudi Ghezeljeh Meidan, L. and **M. Assadi.** 2018. 'Genetic diversity of *Arum L.* based on plastid marker'. *Iranian Journal of Plant Physiology* 8 (3), 2479- 2485.

Introduction

Araceae is a monocotyledonous plant family of about 4000 species currently comprising 117 genera. The family is dominated by species occurring in the tropical regions of all continents (Grayum, 1990; Mayo et al., 1997; Cabrera et al., 2008). They are the most diverse in the tropics, and have a large variety of life forms from epiphytic to aquatic (Espindola et al., 2010). There is no doubt that the members of family Araceae had geographically great expansion in the Cretaceous (Friis et al., 2010). The nature of their fossils not only allows the strong calibration of DNA substitution rates, but also gives information about the previous ranges of certain clades (Nauheimer et al., 2012). The unique structure of the inflorescence is one of the most significant features of this family with small flowers emerging from the fleshy axis (spadix) subtended by a modified leaf (spathe) (Boyce, 1988). Cabrera et al. (2008) effectively settled the long-standing question of the relationships of the Araceae subfamily using a matrix of 102 aroid genera and 5188 aligned base pairs of chloroplast DNA. In order to convert such phylogeny information into a formal classification, it is ideal to compare and contrast them with phenotypic data so as to highlight the clades that are supported by distinctive morphological or anatomical synapomorphies and those supported by molecular synapomorphies. For example, Keating

*Corresponding author
E-mail address: Joudi.leila@iaushab.ac.ir
Received: January, 2018
Accepted: April, 2018

(2002) interpret the morphological and anatomical data by using the phylogenetic characteristics of the family presented by French et al. (1995), leading him to propose a new formal classification of the family. Bogner and Petersen (2007) identified an updated interpretation of the classification of Mayo et al. (1997), which resulted from the comparison of morpho-anatomical data with French et al. (1995) molecular tree.

Mayo et al. (1997) discussed the most detailed modern taxonomy of the Araceae as a distinct family, excluding the duckweeds (the former Lemnaceae, now Araceae subfamily Lemnoideae). Cusimano et al. (2010) examined 113 aroid genera and 4494 aligned nucleotides which resulted from adding 11 genera to the 2008 molecular matrix presented by Cabrera et al. including sequences of six chloroplast DNA regions; *rbcL, matK,* partial *trnK* intron, partial *trnA-Leu* gene, *trnL – trnF* spacer, and partial *trnA-Phe* gene. They also investigated 81 morphological characteristics with regard to the molecular phylogeny, utilizing a developed version of the 1997 morpho-anatomical data set.

One of the most effective phylogenetic tools in utilizing the genomic region, which has proven as a useful character for the phylogenetic data analysis in angiosperms, is the intron trnL (UAA) and trL (UAA) - trn (GAA) spacer in the large single-copy region of the chloroplast genome. Araceae has three genera in Iran, consisting of *Arum* L., *Biarum* Schott. and *Eminium* (Blume) Schott. Six species of *Arum* are *Arum maculatum* L., *A. virescense* Stapf, *A. conophalloides* schott., *A. kotschyi* Boiss. and Hohen., *A. korolkowii* L., and *A. giganteum* Ghahreman.

In this research, we aimed to provide a phylogenetic analysis for *Arum* species based on the sequences of intron trnL (UAA) and trnL (UAA) - trn (GAA) spacer (trnL-F) which is suitable to display relationships at the infrageneric level. This would allow us to assess the validity of the current classification, and identify the phylogeny and taxonomy relationships within different species of *Arum* in Iran.

Materials and Methods

Taxon sampling

Plant material of 10 specimens, including 2 genera and 8 species of Araceae were collected from different localities across Iran, and some of them were chosen from a herbaria in Iran (TARI = Herbarium of Research Institute of Forests and Rangelands) (Table1). Table 2 lists all taxa used in this study, and summarizes sources, voucher specimen data, and GenBank accession numbers.

Molecular analysis

Phylogenetic reconstructions were performed for 8 samples of *Arum* in six regions of Iran. In this study, we used the trnL-F sequence of 10 species of *Arum, Biarum and Eminium* from GenBank. The list of non-Iranian taxa used in our analysis along with GenBank accession numbers are indicated in Table 2. Also, we used the trnL-F

Table 1
List of taxa investigated and voucher specimens. Note: TARI= Herbarium of Research Institute of Forests and Rangelands

Species	Herbarium Locality
Arum maculatum 69216	Iran, Prov. N. Mazandara,. Assadi & Shahsavari. 1991.
A. maculatum 19123	Iran, Prov. N. Gilan. Joudi. 2015.
A.kotschyi 69151	Iran, Prov. N. Gorgan. Assadi & Shahsavari. 1991.
A.virescens 60066	Iran, Prov. N.Gilan. Assadi & Shah Mohammadi. 1987.
A.virescens 64555	Iran, Prov. Hamadan. Mozaffarian. 1988.
A.conophalloides 22222	Iran, Prov. Hamadan. Joudi. 2015.
A.giganteum 68100	Iran, Prov. Isfahan. Hamzehee. 1990.
A.korolkowi 50656	Iran, Prov. Khorasan. Assadi & Maasoumi.1984.
Biarum platyspatum 33333	Iran, Prov. Lordgan. Assadi & Maasoumi.1991.
B.straussi 16447	Iran, Prov. Lorestan. Wendelbo & Assadi. 1978.

sequence of *Eminium* as outgroup. For amplification of the trnL-F gene the forward primer was trnL -FM 5'- TAC GAC GAT CTY TCT AAA CAA GC-3' and trnF -RM 5'- GGA AAG ATT GCT CAA ATA CCA G -3. , using the following PCR protocol: 35 cycles of 1 min denaturation (94° C), 1 min annealing (50° C), and 72° C for 2 min, followed by 5 min final extension at 72° C for the completion of primer extension.

Forward and reverse sequences were visually compared and edited, and then they were initially aligned, using Sequencer 4 software (Gene Codes Corporation, Ann Arbor, MI USA). In addition to our sequences, 8 trnL-F sequences from other taxa were taken from GenBank (Table 2). All trnL-F sequences were assembled and aligned using MacClade 4 (Maddison and Maddison, 2010).

Maximum parsimony analysis (MP)

Parsimony analysis was implemented by employing PAUP version 4.0 (Swofford, 2002) using the following criteria: 100 heuristic search replicates, random stepwise addition of taxa, and tree-bisection reconnection (TBR) branch swapping. These parsimonious trees were used to calculate the consensus tree. Bootstrap analysis (BS) were applied to determine the clade support. BS of clades was calculated using PAUP with 100 replicates of heuristic searches, and randomly stepwise addition of taxa. Clades with a bootstrap value of 70% or more were considered as well supported clades.

Bayesian analysis (BA)

The BA analyses of the trnL-F datasets were performed using MrBayes version 3.0 (Huelsenbeck and Ronquist, 2001). In order to find the appropriate model of DNA substitution, the maximum likelihood criteria for datasets were determined by the Akaike information criterion (AIC; Akaike 1974) as implemented in the software Model Test version 3.7 (Posada et al., 1998).

Modeltest

Table 2
List of non-Iranian taxa used in the analysis and their GenBank accession numbers (for trnL-F marker)

Species	*trn GenBank accession number*
Arum italicum voucher Barabe	AY*555183*
A. maculatum	GU*067633*
Biarum tenufolium	EU*193355*
B. tenufolium	EU*193356*
B. davisii subsp marmarisense.	EU*193348*
B. arundanum	EU*193346*
B. dispar	EU*193352*
Eminium spiculatum	AM*932360*

For the maximum likelihood (ML) and MrBayes analyses (MB), the best fit of DNA substitution model should be found. The Akaike information criterion and hierarchical likelihood ratio test (hLRT) were calculated based on the log likelihood scores of 56 models using the Modeltest 3.7 (Posada et al. 1998). In general, AIC was chosen (Posada 2008). For trnL-F dataset of this paper, likelihood settings from the best-fit model (TIM + G) were selected by AIC in the Modeltest 3.7 with the nucleotide frequencies A = 0.3907, C = 0.1523, G = 0.1015, T = 0.3555, a gamma shape parameter of 0.6567, and an assumed proportion of invariable sites of 0.

Maximum likelihood

Maximum search was performed on the basis of the results of the Modeltest in PAUP. The parameters of the best model, such as the base frequency, the mean relative substitution rates, proportion of invariable sites, and gamma distribution shape were employed. The heuristic search and bootstrap were implemented as in parsimony analysis in PAUP above mentioned.

Bayesian inference

Bayesian inference of the phylogenetic trees was analyzed by some parameters of the Modeltest, and was included in the analysis. The option was set up using 1,000,000 generations of

Markov Chain Monte Carlo (MCMC) searches and a sample frequency of 1000. Saturation was reached after a burn-in of 1000 generations. The clade support was assessed using Bayesian posterior probabilities employing the MrBayes version 3.0 (Huelsenbeck and Ronquist, 2001).

Results

Molecular phylogeny

The data set of the trnL-F region included 611 characters with 128 constant positions within the ingroup while 463 characters were parsimony informative. The Bayesian 50% majority-rule consensus tree for trnL-F contained 11 internal nodes with a posterior probability (PP) of 1.0. Strict consensus phylogeny trees with 678 steps resulted in the consistency index (CI) of 0.872 and retention index (RI) of 0.960.

Phylogenetic analyses of trnL-F sequences by Bayesian Method

The result of phylogenetic analysis of trnL-F by Bayesian method is shown in Fig. I. *Eminium spiculatum* was introduced as outgroup. Our results indicate that Iranian *Arum* and *Biarum* placed in monophyletic clade as ingroup (species from GenBank fixed as polytomic). There are 2 main clades: clade A and B (PP=0.7). Clade A includes *B. Platyspathum* 33333 and *B. Strussi* 16447 by 0.99 support from Iran that form a sister group. Clade B with *Arum* species is divided into two subclades: BI and BII (PP=0/76). All Iranian species belong to clade BI by PP=0.87. *A. conophalloides* 22222 and *A. virescense* 60066 formed a sister group. *A. giganteum* 68100 is located at a distance from other species and showed a separate evolution. Clade BII had two species of GenBank: *A. maculatum* and *A. Italicum* (PP=0.83).

Phylogenetic analyses of trnL-F sequences by Maximum Parsimony

In the maximum parsimony analysis for trnL-F marker, *E. spiculatum* was introduced as an outgroup with bootstrap of 100. The ingroup in

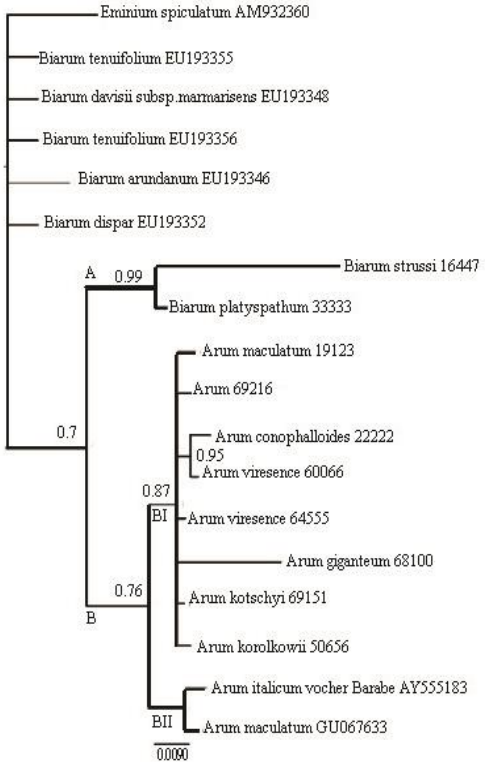

Fig. I. The parsimony strict consensus tree based on concatenated plastid data, (numbers above branches are maximum likelihood percentage bootstrap values).

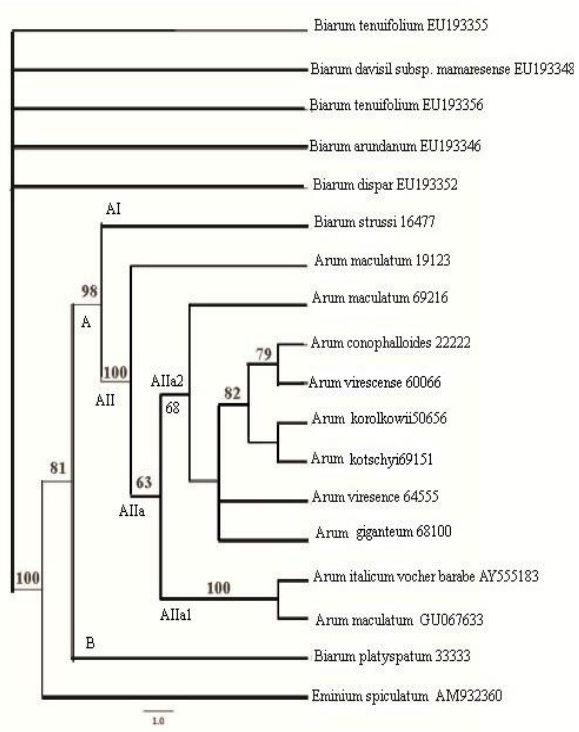

Fig. II. Bayesian consensus tree based on concatenated plastid data (numbers on branches are Bayesian posterior probabilities)

this analysis is related to the *Arum* and *Biarum* genus with a BPP = 81 including clades A and B. Clade B dedicated *B. platyspathum* 33333. Clade A is divided into 2 subclades (BPP = 98), which include the AI subclade with *B. straussi* 16447 and the subclade AII with the genus *Arum* species. In the AII subclade with bootstrap 100, the species of *A. maculatum* 19123 is completely separate from others. In the case of the rest of the species, there is a group called AIIa where bootstrap is 63. AIIa has 2 clusters: AIIa1 include the GenBank and AIIa2 comprises the species studied. In AIIa2, the species *A. conophalloides* 22222 and *A. giganteum* 68100 form a sister group and the species *A. korolkowii* and *A. kotschyi* form the sister group, and in the latter these four species will be monophyletic with bootstrap 82. The *A. virescense* is independently located in the AIIa2 cluster (Fig. II).

Discussion

We provided phylogenetic analysis of the *Arum* from Iran. Phylogenetic relationships between the Araceae genera in this study intensely organized the species of *Arum*, *Biarum* into a supported monophyletic group. Results of cladistic analysis of the phylogenetic relationships among *Arum* species showed its monophyletic origin, and *Biarum* was its sister group. According to earlier studies, all species examined in our research were also nested inside Arae clade within the Mediterranian clade, comprising *Arum*, *Dracunculus*, *Biarum, Eminium* and *Helicodiceros*, which diverged allopathically in a region encompassing tropical Asia and Anatolia during the late Eocene (Mansion et al., 2008). Based on Espindola et al., the center of origin of *Arum* species was the European–Aegean region, and that major diversification happened during the last 10 Myr.

In Iran, *A. korolkowii* and *A. kotschyi* are similar in their morphological characteristics (Joudi et al., 2016) and molecular analyses. They were placed in a unique clade, and the monophyly of the clade was shown. Linz et al. (2010) demonstrated that *Arum rupicola*, *A. jacquemontii*, and *A. korolkowii* form a sister group and A*rum jacquemontii* and *A. korolkowii* have the most eastern area of distribution, reaching into Central Asia.

A. maculatum species was placed in separated clade. In another research conducted by Espindola et al. (2011), *A. maculatum* was placed in different positions. Also, *A. maculatum* was placed in different locations based on trnK, trnL, ndh, and trnT molecular markers (Espindola et al., 2010), which confirm our results, indicating that *A. maculatum* was placed in various positions.

Arum conophaloides 59955 and *A. virescense* formed a monophyletic clade, which was supported by morphometric research findings (Joudi et al., 2016). *A. giganteum* is widely distributed in Iran. Studies on morphological differences have separated this species confirming earlier findings by identifying the synapomorphies (Joudi et al., 2016). *A. giganteum* has been introduced as a new species in Iran (Ghahreman, 1983) but in *Arum* monograph it is mentioned as potential equivalent to *Arum rupicola* Boiss. Molecular studies in this research can separate these two species and confirm before studies.

Conclusion

trnL-F region including intron trnL (UAA) and trnL (UAA) - trn (GAA) spacer is a widely used molecular marker for the reconstruction of evolutionary patterns in plant. This research indicated that this marker can potentially be a valuable tool to identify the species of *Arum*. Based on the results of cladistics analysis, the phylogenetic relationships among Arum species from Araceae family have confirmed the morphological investigations at high levels. Furthermore, *Arum giganteum* has been introduced as a new species in Iran with its exclusive morphological features such as the whole size of the plant, size of the leaves, and inflorescence that could not be seen in other *Arum* species.

However, trn marker could identify the relationships between some of the major groups, although some groups remained unresolved. According to our findings, we assume that further research regarding the insufficiently resolved nodes within the *Arum* genus would provide important insights into the relations between the species of *Arum*. The combined approaches,

including the application of different markers, would increase the resolutions and support the *Arum* clades.

Acknowledgements

The authors would like to express their thanks to Islamic Azad University, Science and Research Branch, Tehran, Iran and Research Institute of Forest and Rangelands, National Botanical Garden of Iran for providing the facilities necessary to carry out the present study.

References

Akaike, H. 1974. 'A new look at the statistical model identification'. *IEEE Transactions on Automatic Control* ,19, 718-723.

Bogner, J., K.R. Johnson, Z. Kvacek and **G.R. Upchurch.** 2007. 'New fossil leaves of Araceae from the Late Cretaceous and Paleogene of western North America'. *Zitteliana,* 47: 133–147.

Boyce, P. C. 1988. 'Plant portrait: *Arum pictum*'. *Kew Magazine 5:72–76.*

Boyce, P.C. 1993. 'The genus Arum. A Kew magazine monograph'. Royal Botanic Gardens, Kew, London. Norwich: HMSO Books.

Boyce, P.C. 2006. 'Arum – a decade of change'. *Aroideana. 29: 132–9.*

Boyce, P.C., and **T.B. Croat.** 2011. 'The Uberlist of Araceae, totals for published and estimated number of species in Aroid genera http://www.aroid.org/genera/111109uberlist.

Cabrera, L.I., G.A. Salazar, M.W. Chase, S.J. Mayo, J. Bogner and **P. Davila**. *2008*. 'Phylogenetic relationships of aroids and duckweeds (Araceae) inferred from coding and noncoding plastid DNA'. *American Journal of Botany 95: 1153–1165*.

Chouteau, M., M. Giberna and **D. Barabe.** *2008*. 'Relationships between floral characters, pollination mechanisms, life forms, and habitats in Araceae'. *Botanical Journal of the Linnean Society, 156: 29–42.*

Cusimano, N., J. Bogner, S. Mayo, P.C. Boyce, N. Wong, M. Hesse, W. Hetterscheid, R. Keating and **J. French.** 2011. 'Relationships within the Araceae: comparison of morphological patterns with molecular phylogenies'. *American Journal of Botany 98(4): 654–668.*

Cusimano, N., M.D. Barrett, W.L.A Hetterscheid and **S.A. Renner**. 2010. 'Phylogeny of the Areae (Araceae) implies that Typhonium, Sauromatum, and the Australian species of Typhonium are distinct clades'. *TAXON ,59 (2): 439–447.*

Espindola, A., S. Buerki, M. Bedalov, P. Kupfer and **N. Alvarez**. 2010. 'New insights into the phylogenetic and biogeography of arum (Araceae): unravelling its evolutionary history'. *Botanical journal of the linnean society*, *163, 14–32.*

Friis, E.M., K.R. Pedersen and **P.R. Crane**. 2010. 'Diversity in obscurity: fossil flowers and the early history of angiosperms'. Philosophical Transactions of the Royal Society B. *Biological Sciences*, 365: 369–382.

Gibernau, M., Ch. Favre, T. Talou and **Ch. Raynaud**.2004. 'Floral odor of *Arum italicum'. Aroideana, 27: 142-147*.

Huelsenbeck, J.P. and **F. Ronquist**. 2001.' Mr Bayes: Bayesian inference of phylogeny'. *Bioinformatics*, 17: 754-755.

Joudi, L., I. Mehregan, M. Assadi and **D. Farajzadeh**. 2016. 'Genetic diversity and classification of wild Arum (Araceae) species using morphological characters in Iran'. *Electronic Journal of Biology*, 12(3): 302-308.

Keating, R.C. 2000. 'Collenchyma in Araceae: Trends and relation to classification'. *Botanical Journal of the Linnean Society, 134:203 – 214*.

Linz, J., J. Stökl, I. Urru, T. Krügel and **M. Stensmyr**. 2010. 'Molecular phylogeny of the genus Arum (Araceae) inferred from multi-locus sequence data and AFLPs'. *TAXON*, 59 (2): 405–415.

Maddison, W.P. and **D.R. Maddison**. 2010. 'Mesquite (version 2.74): A modular system for evolutionary analysis. http:// mesquiteproject. Org/

Mansion, G., G. Rosenbaum, N. Schoenenberger, G. Bacchetta, J.A. Rosselló, and **E. Conti** 2008. 'Phylogenetic analysis informed by geological history supports

multiple, sequential invasions of the Mediterranean Basin by the angiosperm family Araceae'. *Systematic Biology*, 57(2): 269-285.

Mayo, S.J., **J. Bogner** and **P.C. Boyce**.1997. 'The genera of Araceae'. The Trustees, Royal Botanic Gardens, Kew, U.K.

Posada, D. 2008. Model Test: Phylogenetic Model Averaging. *Molecular Biology Evolution*, 25: 1253-1256.

Riedl, H. 1963. Flora Iranica 1.

Ronquist, F. and **J.P. Huelsenbeck**. 2003. *Bioinformatics,19: 1572–1574.*

Swofford, D.L. 2002. Phylogenetic analysis using parsimony (PAUP). Ver. 4. Sinauer Associated Sunderlandm Massachusetts.