Reconstruction of the evolutionary biogeography reveals the origin and diversification of Arisaema (Araceae)

Van Tien Tran1,2, Tuan Phuong Nguyen Hoang3, Thomas B. Croat4, Zhengxu Ma5, Bing Liu6, Van Du Nguyen1,7 and Chi Toan Le3*

Received: August 05, 2021
Accepted: January 21, 2022

ABSTRACT
Explaining where plants originated and how plants migrated between eastern Asia and other Northern Hemisphere regions is still challenging. The genus Arisaema is considered a good example to illuminate this process, due to its wide distribution in both the tropical and temperate zones. Based on comprehensive taxon sampling, the historical biogeography of Arisaema was conducted using the molecular data from four chloroplast DNA regions. The phylogeny of Arisaema was studied by using ML and BI methods. Divergence times of Arisaema species were estimated by the uncorrelated lognormal Bayesian method. The ancestral area reconstruction of Arisaema was performed with BioGeoBEARS and Bayes-DIVA. Our results indicated that the crown group of Arisaema originated in Southeast Asia ca. 31.28 Ma. After the initial origin, Arisaema was inferred to have become widespread in Southeast Asia. Members in Himalayas, Africa and North China, Japan, the Korean Peninsula region (NJK region) were derived from migrants originating in Southeast Asia. American members were derived from migrants originating in the NJK region. Arisaema originated in Southeast Asia during the Oligocene with the development of tropical rainforest and a warm, moist climate. During the Oligocene to Miocene, Southeast Asia was the original center for the diversification of Arisaema in the NJK region.

Keywords: Arisaema, biogeography, origin, Eocene, Southeast Asia
Introduction

The genus *Arisaema* is one of the largest genera in the family Araceae in the World with about 200 species (Ohi-Toma et al. 2016). It is distributed in the temperate and warm-temperate areas from the East to South-East of North America to northern Mexico, from tropical East and North-East Africa to the Arabian Peninsula, and from temperate East Asia to South and South-East Asia and Malesia (Govaerts et al. 2002; Gusman & Gusman 2006; Murata 2011; Ohi-Toma et al. 2016). In the traditional medicine of India, China and Vietnam, some species of *Arisaema* are used as local medicine (Li et al. 2010). The young leaves and inflorescences of some species are used as vegetable or food plants (Gusman & Gusman 2006).

In the family Araceae, the genus *Arisaema* was placed in the tribe Arisaematae, subfamily Aroideae (Mayo et al. 1997; Renner & Zhang 2004). *Arisaema* is characterised by being seasonal or evergreen, monosexual or bisexual plants; having tuberous or rhizomatous stem, with or without pseudostem; leaf blades diverse from trifoliolate, pedate to radiate. Inflorescence solitary, bearing spadix monosexual or bisexual with the appendix morphology very diverse; male flowers with 2-5 stamens; female flowers with unilocular ovary containing several basal ovules (Fig. 1).

Since being established in 1831, there have been a wide variety of studies published within the genus *Arisaema* and several systems of classification proposed by Schott (1860), Engler (1920), Nakai (1950), Hara (1971), Murata (1984, 1991), Gusman & Gusman (2002, 2006), Li et al. (2010), Murata (2011), Murata et al. (2013) and Ohi-Toma et al. (2016). All of the above studies focused on systematics and have provided a foundation for investigating the phylogeny and historical biogeography of the genus *Arisaema*. Renner et al. (2004) conducted a phylogenetic study based on 81 individuals of *Arisaema*. Results of this study supported the monophyly of this genus, but several polytomies were found in the phylogenetic trees of *Arisaema*. Additionally, the study also estimated the divergence time of the genus, provided the main biogeographical event of *Arisaema*, and considered floristic links among East Africa, Arabia, the Himalayan region, China, and North America. However, this study used only one calibration point (the fossil similar to *A. triphyllum* at 18 million years ago) and penalized likelihood. The resulting time estimates are not sufficiently reliable to clarify divergence time of *Arisaema*, and the biogeography of *Arisaema* was not discussed in detail.

Ohi-Toma et al. (2016) introduced a new phylogenetic model using maximum parsimony and maximum likelihood methods to analyze 124 species of *Arisaema*. This work showed that the genus *Arisaema* includes 15 sections: *Anomala*, *Arisaema*, *Attenuata*, *Clavata*, *Decipientia*, *Dochafa*, *Fimbriata*, *Flagellarisaema*, *Frenchetiana*, *Nepenthoidia*, *Odorata*, *Pistillata*, *Sinarisaema*, *Tenuipistillata* and *Tortuosa*. Fourteen of these sections were recognized by Murata (2011, 2013), and section *Fimbriata* was reconfirmed as distinct from other sections in the genus. The updated phylogeny of *Arisaema* helped to clarify the relationship between phylogeny and geographical distribution for some sections in the genus, but this study did not focus on the evolutionary history of the genus.

Nauheimer et al. (2012) studied the global history of the family Araceae. This work included ten species of *Arisaema*, well calibrated with several fossils of Araceae, and indicated that Araceae began to diversify in the Early Cretaceous, and that all eight subfamilies existed before the K/T boundary. Early lineages persisted in Laurasia, and there were several relatively recent entries from Laurasia into Africa, South America, South-East Asia and Australia. *Arisaema* was estimated to have appeared during the Oligocene, after splitting from the ancestor of the group along with several other genera at the crown of Araceae. However, due to limitations of *Arisaema* samples, the results of divergence times and biogeography of *Arisaema* were not mentioned. Thus, the questions on historical biogeography and distribution hypothesis of this genus have not been clarified.

Data are now available to formulate a hypothesis for the relationships and distribution hypothesis of *Arisaema* species. Here we report phylogenetic and biogeographical analyses based mainly on published molecular data from chloroplast DNA regions from 127 of the 200 species of *Arisaema*. Our major aims are to reconstruct the biogeographical history of *Arisaema* reconciled with present day distributions, fossil data, and geological information. We explored the following questions: (1) How old is *Arisaema*? (2) What is the ancestral area of *Arisaema* and when were the major biomes colonized? (3) What is the best explanation for *Arisaema*’s large number of species and wide distribution?

Materials and methods

Sampling, DNA extraction, amplification and sequencing

Four chloroplast DNA regions (*trnL–trnF*, *rpl20–rps12*, *psbB–psbH* and *rps2–rpoC2*) were studied. The DNA sequences from *Arisaema* species were last downloaded on 12 September 2020. Duplicate and uncertain sequences were excluded. Two new sequences were obtained for an undescribed species recently found in Vietnam. In total, our data were obtained for 129 species including 127 identified species of *Arisaema* for all 15 sections of the genus that were recognized by Ohi-Toma et al. (2016). Thus, the molecular data in this study represents most geographical and taxonomic diversity in the genus with approximately 84% of *Arisaema* species. Voucher information and GenBank accession numbers are presented in Tab. S1.
Reconstruction of the evolutionary biogeography reveals the origin and diversification of *Arisaema* (Araceae)

For the sample from Vietnam, we extracted genomic DNA from silica gel dried leaves using the CTAB procedure (Doyle & Doyle 1987). Polymerase chain reactions and sequencing were performed using the primers presented in Ohi-Toma et al. (2016). PCR products were separated and visualized using an ABI3730 automated sequencer (Applied Biosystems, USA). All sequences were aligned in Geneious v.8.0.5 (Kearse et al. 2012).

**Figure 1.** Diversity of morphological characters in the genus *Arisaema*. A – B: Diversity of habits and leaf blades; D – I: Diversity of inflorescences; J – O: Diversity of spadices. A. *A. menglaense*; B. *Arisaema* sp.; C. *A. rhizomatum*; D. *A. pierreanum*; E. *A. lidaense*; F. *A. averyanovii*; G. *A. franchetianum*; H. *A. rhizomatum*; I. *A. langbianense*; J & K. *A. menglaense*; L. *A. pingbianense*; M. *A. rhizomatum*; N. *A. rostratum*; O. *A. pierreanum* (Photograph: V. D. Nguyen).
Phylogenetic analyses

The phylogenetic analyses of *Arisaema* were conducted using maximum likelihood (ML) and Bayesian inference (BI) methods. The GTR + I + G substitution model was determined as a best-fitting model for the combined datasets by the Akaike Information Criterion (AIC) as implemented in jModelTest v.2.1.6 (Darriba et al. 2012) available in the CIPRES Science Gateway (Miller et al. 2010).

The ML analyses were performed using RAxML v.8.2.8 (Stamatakis 2014) available in the CIPRES Science Gateway (Miller et al. 2010), with the GTR + I + G substitution model and 1000 bootstrap replicates for statistical support. The Bayesian analysis was performed in MrBayes v.3.1.2 (Ronquist & Huelsenbeck 2003) on the CIPRES Science Gateway (Miller et al. 2010) using the GTR + I + G substitution model. The Markov chain Monte Carlo (MCMC) algorithm was run for 10 million generations with a total of four chains, starting from a random tree, and trees were sampled every 1000 generations. The program Tracer v.1.6 (Rambaut & Drummond 2007) was used to check that effective sample sizes (ESSs) were obtained for all relevant parameters assessed (> 200). With the first 25% of sampled generations discarded as burn-in, the 50% majority-rule consensus tree and Bayesian posterior probabilities (PP) were obtained using the remaining trees.

Divergence time estimation

The divergence times of *Arisaema* were estimated by using the uncorrelated lognormal Bayesian method in BEAST v.1.8.0 (Drummond et al. 2012). All dating runs relied on a Yule process tree prior. A lognormal distribution was used for the calibration points (Deng et al. 2014; Liu et al. 2018), and a normal distribution was used for the secondary calibration points (Lu et al. 2013; Gonçalves et al. 2020). We conducted two separate MCMC runs of 100 million generations with samples taken every 10,000 generations. Tracer was used to check if ESSs for all relevant parameters assessed (> 200). With the first 25% of sampled generations discarded as burn-in, the 50% majority-rule consensus tree and Bayesian posterior probabilities (PP) were obtained using the remaining trees.

Ancestral area reconstruction

The ancestral area reconstruction of *Arisaema* was conducted using (1) BioGeoBEARS (Matzke 2013) implemented in R (R Development Core Team 2016), and (2) a Bayesian approach to dispersal-vicariance analysis (Bayes-DIVA; Nylander et al. 2008) implemented in RASP v.3.2 (Yu et al. 2011) using datasets from the BEAST analysis. For the BioGeoBEARS, two likelihood-based models were implemented: Dispersal-Extinction-Cladogenesis (DEC) (Ree & Smith 2008) and the likelihood version of dispersal-vicariance (DIVA) (Ronquist 1997; herein DIVALIKE). In the two models, an additional “\( j \)” parameter (founder event/jump speciation) was added, which allows descendant lineages to have a different area from the direct ancestor (Matzke 2013; Vasconcelos et al. 2017; Liu et al. 2018). In total, four models resulted in the BioGeoBEARS analysis. In addition, we estimated the number and type of biogeographical events by biogeographical stochastic mapping (BSM) implemented in ‘BioGeoBEARS’ (Matzke 2015) under the best-fit biogeographical model (see results). According to Dupin et al. (2017), we divided the biogeographical events under the models including within-area speciation, vicariance, and dispersal events (range expansions and founder events). Frequency of events was estimated from the mean and standard deviation of event counts from 100 BSMs.

The Bayes-DIVA method allows minimal phylogenetic uncertainty by utilizing the prior distribution of trees resulting from BEAST and generating credibility support values for alternative phylogenetic relationships (Nylander et al. 2008; Yu et al. 2011). For the Bayes-DIVA analysis, we performed based on 10,000 trees from the BEAST analysis and computed a condensed tree as the final representative tree with the first 2500 trees discarded as burn-in.

The biogeographical data for species within *Arisaema* were compiled from the distribution information that
describes the extant species of the genus and their relatives in the literatures (Hara 1971; Gusman 2006; Li et al. 2010), and herbarium specimens. Five biogeographical areas were defined based on the geographic barriers and floristic divisions of Arisaema (Hara 1971; McLoughlin 2001; Gusman 2006; Li et al. 2010): A: Southeast (SE) Asia (including mainland of Southeast Asia, Indochina and Malesia, but excluding New Guinea); B: Africa (including the coastal area of the Arabian Peninsula and East Africa); C: America (including North America, Mexico); D: North China, Japan, the Korean Peninsula region (NJK region); E: Himalayas (including Pakistan, India, Himalaya region, and SW China).

Arisaema is distributed in different areas in three continents (Asia, Africa, and America). There are multiple differences in climatic and geographic barriers among these continents. The morphology and habitat of Arisaema species in Southeast Asia and NJK region or Himalayas are different. Murata (1984) recognized that in Arisaema, tuberous plants with the spilodistichous leaf arrangement are always deciduous and are better adapted to winter dormancy in cool climates. Those characters mostly belong to species distributed in NJK region, America and Himalayas. Murata (1990) also noted that plants with rhizomatous stem and quincuncial leaf arrangement are evergreen plants, better adapted to non-seasonal growth habits under tropical climate. Those characters are found in members mainly distributed in low elevations of Southern China to SE Asia. Therefore, we defined Southeast Asia, NJK region (North China, Japan, and the Korean Peninsula region) and Himalayas as three separate biogeographical areas.

Results

Phylogenetic relationships

The lengths of the four plastid data sets: trnLF, rpl20–rps12, psbB–psbH and rps2–poC2 were 472, 1079, 736, 566 bps respectively. The combined dataset resulted in a matrix with 2853 characters. The topology from ML and BI analyses of the combined dataset were highly congruent, only differing in low resolution nodes. We thus present the BI tree with BS and PP values in Fig. 2.

The phylogenetic results reported here are more advanced compared to other previous analyses (Renner et al. 2004; Ohi-Toma et al. 2016) based on the robust dataset. Our phylogenetic analyses supported Arisaema as monophyletic, with 15 sections being recognized within the genus (Fig. 2). Polytomies are distributed in the phylogenetic tree, however, placement and support values for most clades are improved. The basal of the phylogenetic trees includes eight sections Tenuipistillia, Dochafa, Tortuosa, Arisaema, Fimbriata, Attenuata, Odorata, and Anomala. However, the positions of the seven sections are not stable. The three sections Tenuipistillia, Dochafa, and Tortuosa weakly formed a clade. Section Arisaema was well supported as a single clade (BS: 98 %, PP: 1.0) (Fig. 2). While, four sections Fimbriata, Attenuata, Odorata, and Anomala weakly formed a clade, with Fimbriata poorly supported as sister to the three rest sections (Fig. 2). The middle portion of the Arisaema phylogenetic tree includes the remaining seven sections Clavata, Flagellarisaea, Nepenthoidae, Decipientia, Pistillata, Franchetiana and Sinarisaema. The sections Nepenthoidae and Pistillata are not supported as monophyletic. This situation was also seen in Ohi-Toma et al. (2016).

In the section Anomala, our taxon which was collected in Tuyen Quang province in northern Vietnam (Arisaema sp. VN25) was placed closely relative to A. petelotii, A. hainanense and A. omkoiense, and was strongly supported by the molecular data (Fig. 2). However, our species from Tuyen Quang province is close to A. grapsospadix and A. balansae morphologically. It can be distinguished from A. grapsospadix by having no pseudostem, a spathe without a white area at the base of spathe lamina and from A. balansae by the spathe lacking an auricle at the tube mouth and appendix covered by short neuters. Thus, it should be described as a new species based on the evidence here for future studies.

Divergence times

The divergence time estimations for Arisaema are presented in Fig. 3. The stem age of Arisaema was estimated to be 33.27 Ma (95 % HPD: 29.82, 36.58 Ma). The group including sections Fimbriata, Odorata, Attenuata, and Anomala diverged early at the crown node of Arisaema 27.86 Ma (95 % HPD: 22.86, 32.99 Ma). The African member A. chiperianum diverged 8.83 Ma (95 % HPD: 2.52, 15.51 Ma) at node 6 during late Miocene (Fig. 3), while the American Arisaema are not monophyletic and diverged since the Mid Miocene to Pliocene (Fig. 3). A comparison dating from this study, Nauheimer et al. (2012) and Renner et al. (2004) is presented in Tab. S2.

Ancestral area reconstruction

The results of ancestral area reconstruction from BioGeoBEARS and Bayes-DIVA are highly congruent. However, the results from BioGeoBEARS showed higher resolution for more nodes than Bayes-DIVA. Among the four models of BioGeoBEARS, the results indicated higher log likelihood values for three parameters in comparison to two parameters (Tab. 1), thus indicating jump speciation (i.e., dispersal between non-adjacent areas) as an important pattern in range variation of Arisaema. Moreover, the BioGeoBEARS analyses showed DEC-j as the best-fit biogeographical model. Thus, we only present the reconstruction of BioGeoBEARS under the DEC-j model (Fig. 4). The result of ancestral area reconstructions from Bayes-DIVA is shown in Fig. S1. The number of nodes in Fig. 3 and Fig. 4 are consistent, and a summary of the divergence age
Van Tien Tran, Tuan Phuong Nguyen Hoang, Thomas B. Croat, et al.

Figure 2. Majority rule consensus tree of *Arisaema* based on the combined datasets of four DNA regions (*trnL–trnF*, *rpl20–rps12, psbB–psbH* and *rps2–rpoC2*). ML bootstrap values and posterior probabilities (PP) of the BI analysis are presented above the branches. “-” indicates the support values less than 50%. The circumscription of sections follows the classification of Ohi-Toma et al. (2016).

Reconstruction of the evolutionary biogeography reveals the origin and diversification of *Arisaema* (Araceae)

Figure 3. Maximum clade credibility tree inferred from BEAST based on the combined datasets of four DNA regions. The bars around node ages indicate 95% highest posterior density intervals. Node constraints are indicated with stars. Nodes of interests were marked as 0–18.
estimations and ancestral area reconstruction is presented in Tab. 2. Our biogeographical stochastic mapping (BSM) results under the best-fit biogeographical model DEC+j revealed that most biogeographical events comprise within-area speciation (74 %) and dispersals (23 %), with very few (3 %) vicariant events (Tab. S3).

Our results suggest that *Arisaema* most likely originated in Southeast Asia (area A) (node 0, Fig. 4), and subsequently diversified in Southeast Asia during the Oligocene (node 1, Fig. 4). The group including Southeast Asian and Himalayas members of four sections *Fimbriata*, *Anomala*, *Attenuata* and *Odorata* originated in Southeast Asia at node 1 during the Mid Oligocene. In this clade, *A. album*, the endemic species of India, diverged in the late Miocene at node 3.

Similarly, all remaining sections likely originated in Southeast Asia and then spread to other areas during the Oligocene to Miocene (Fig. 4). The dispersal events from Southeast Asia to Himalayas, Southeast Asia to NJK region or NJK region to America were presented multiple times within *Arisaema* (Fig. 4).

**Discussion**

**Origin of *Arisaema* in Southeast Asia**

The biogeographical analyses and divergence time estimations suggest that *Arisaema* originated in Southeast Asia *ca*. 33.27 Ma (95 % HPD: 29.82, 36.58 Ma) during the early Oligocene (node 0, Figs. 3, 4; Tab. 2), and soon after the initial origin *Arisaema* started to diverge at 31.28

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter No</th>
<th>LnL</th>
<th>d</th>
<th>e</th>
<th>j</th>
<th>AIC</th>
<th>AIC weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEC</td>
<td>2</td>
<td>−116.85</td>
<td>7.63×10^{-3}</td>
<td>7.51×10^{-2}</td>
<td>0</td>
<td>237.78</td>
<td>0.268</td>
</tr>
<tr>
<td>DEC+j</td>
<td>3</td>
<td>−115.90</td>
<td>7.13×10^{-3}</td>
<td>8.80×10^{-2}</td>
<td>0.000517</td>
<td>235.70</td>
<td>0.731</td>
</tr>
<tr>
<td>DIVALIKE</td>
<td>2</td>
<td>−120.55</td>
<td>9.44×10^{-3}</td>
<td>0.1377</td>
<td>0</td>
<td>258.90</td>
<td>0.0010</td>
</tr>
<tr>
<td>DIVALIKE+j</td>
<td>3</td>
<td>−126.45</td>
<td>4.68×10^{-3}</td>
<td>1.00×10^{-12}</td>
<td>0.00716</td>
<td>245.10</td>
<td>0.9989</td>
</tr>
</tbody>
</table>

**Table 1.** Comparison of the fit of different models of biogeographical range evolution and model specific estimates for different parameters (*d* = dispersal, *e* = extinction, *j* = weight of jump dispersal (founder speciation)).

<table>
<thead>
<tr>
<th>Node</th>
<th>Bayesian PP</th>
<th>Age estimates mean (95% HPD) [Ma]</th>
<th>BioGeoBEARS (DEC+j)</th>
<th>Bayes-DIVA (maximum probability)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0</td>
<td>33.27 (29.82, 36.58)</td>
<td>A</td>
<td>A (80.87)</td>
</tr>
<tr>
<td>1</td>
<td>1.0</td>
<td>31.28 (27.08, 35.45)</td>
<td>A</td>
<td>A (81.39)</td>
</tr>
<tr>
<td>2</td>
<td>&lt; 0.90</td>
<td>27.86 (22.86, 32.99)</td>
<td>A</td>
<td>A (100)</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>04.73 (04.22, 12.83)</td>
<td>A</td>
<td>AE (100)</td>
</tr>
<tr>
<td>4</td>
<td>&lt; 0.90</td>
<td>30.21</td>
<td>A</td>
<td>A (81.60)</td>
</tr>
<tr>
<td>5</td>
<td>1.0</td>
<td>21.28 (13.88, 28.91)</td>
<td>A</td>
<td>AE (65.85)</td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>08.83 (02.52, 15.51)</td>
<td>A</td>
<td>AB (50)</td>
</tr>
<tr>
<td>7</td>
<td>1.0</td>
<td>01.40 (00.01, 03.49)</td>
<td>AE</td>
<td>A (100)</td>
</tr>
<tr>
<td>8</td>
<td>1.0</td>
<td>25.99 (21.67, 30.57)</td>
<td>A</td>
<td>A (52)</td>
</tr>
<tr>
<td>9</td>
<td>&lt; 0.90</td>
<td>21.27 (14.41, 28.20)</td>
<td>A</td>
<td>A (38)</td>
</tr>
<tr>
<td>10</td>
<td>1.0</td>
<td>10.42 (04.94, 16.19)</td>
<td>AD</td>
<td>AD (32)</td>
</tr>
<tr>
<td>11</td>
<td>1.0</td>
<td>10.62 (04.89, 17.20)</td>
<td>D</td>
<td>A (19.58)</td>
</tr>
<tr>
<td>12</td>
<td>&lt; 0.90</td>
<td>04.12 (00.85, 08.17)</td>
<td>AC</td>
<td>AD (100)</td>
</tr>
<tr>
<td>13</td>
<td>&lt; 0.90</td>
<td>04.24</td>
<td>D</td>
<td>AD (100)</td>
</tr>
<tr>
<td>14</td>
<td>1.0</td>
<td>22.39 (18.78, 26.27)</td>
<td>A</td>
<td>AE (94.59)</td>
</tr>
<tr>
<td>15</td>
<td>0.99</td>
<td>18.78 (16.81, 21.17)</td>
<td>A</td>
<td>A (40)</td>
</tr>
<tr>
<td>16</td>
<td>1.0</td>
<td>06.07 (02.54, 10.11)</td>
<td>A</td>
<td>AD (26)</td>
</tr>
<tr>
<td>17</td>
<td>&lt; 0.90</td>
<td>18.36</td>
<td>A</td>
<td>AD (40.91)</td>
</tr>
<tr>
<td>18</td>
<td>0.99</td>
<td>17.45 (16.37, 18.82)</td>
<td>AD</td>
<td>CD (100)</td>
</tr>
</tbody>
</table>

**Table 2.** Summary of supported clades, divergence age estimation (Ma) by BEAST and ancestral area reconstructions by BioGeoBEARS and Bayes-DIVA. Node numbers refer to Figs. 3, 4. Area abbreviations are as follows: A: Southeast (SE) Asia (including mainland of Southeast Asia, Indochina and Malesia, but excluding New Guinea); B: Africa (including the coastal area of the Arabian Peninsula and East Africa); C: America (including North America, Mexico); D: North China, Japan, the Korean Peninsula region (NJK region); E: Himalayas (including E Pakistan, N India, Himalaya region, and SW China).
Reconstruction of the evolutionary biogeography reveals the origin and diversification of Arisaema (Araceae)

Figure 4. Ancestral area reconstruction of Arisaema by BioGeoBEARS ($j = 0.000517$, LnL = $-115.90$). Geologic time scale is shown at the bottom. Area abbreviations are as follows: A: Southeast (SE) Asia (including mainland of Southeast Asia, Indochina and Malesia, but excluding New Guinea); B: Africa (including the coastal area of the Arabian Peninsula and East Africa); C: America (including North America, Mexico); D: North China, Japan, the Korean Peninsula region (NJK region); E: Himalayas (including E Pakistan, N India, Himalaya region, and SW China).
Arisaema was closely related to birds. and diversification of ca. 50 Ma (95 % HPD: 27.08, 35.45 Ma; node 1, Figs. 3, 4; Tab. 2). Forests (including tropical forests and temperate forests or mountainous forests at high elevation) were present throughout the Northern Hemisphere during the early Cenozoic, which allowed for intercontinental dispersal of plants (Wei et al. 2015; West et al. 2015). In fact, both the tropical forests and temperate forests (including mountainous forests at high elevation) can harbor Arisaema, with many species also occurring in the deciduous forests and subalpine forests in East Asia and East North America. Among the sections of Arisaema, only the sect. Anomala (ca. 25 spp.), sect. Attenuata (ca. 20 spp.) and sect. Fimbriata (ca. 2 spp.) are distributed in tropical forests, others (ca. 150 spp.) in temperate forests, or relatively higher elevation forests of tropical regions (i.e., the mountainous forests in Himalaya, 1000-3000 m) (Murata 1984, 1990; Mayo et al. 1997; Gusman & Gusman 2002; Govaerts et al. 2002; Murata et al. 2013; Ohi-Toma et al. 2016).

Based on our observation in the field and suggestion from Mayo et al. (1998), Arisaema and several genera of Araceae prefer wet to swampy conditions, or grow near open water. They prefer growing in woodlands or in shady places under trees in deciduous forests and rainforests. Tropical forests during the Paleocene/Eocene were located at middle paleolatitudes where the climate was more favorable (Morley 2000; 2003; Jablonski 2004; Schuettpelz & Pryer 2009; Couvreur et al. 2011). By late Eocene to early Oligocene, they retreated to lower latitudes (Contreras et al. 2013). Modern tropical forests are dominated by angiosperms in taxa such as Fabaceae (Lavin et al. 2005), Sapindales (Wang et al. 2009), Rubiaceae (Antonelli et al. 2009) and Fagaceae (Xiang et al. 2014). The woody families above appear to have diverged or radiated before the time of origin of Arisaema and attained a level of ecological prominence comparable to that of today (Crane & Herendeen 2006). During the late Eocene and early Oligocene, the climate was warm and moist in most of Southeast Asia (Zachos et al. 2001). Thus, the climate condition and forest biomes in this period could have provided habitats for the origin, survival, and radiation differentiation of Arisaema, as it likes to grow in the understory, along forest margins, and in light gaps in forest.

Murata et al. (1993), Wada & Uemura (1994), Vogel & Martens (2000), Renner et al. (2004), Thomason (2020) noted that fruits of Arisaema are few-seed berries with bright colors (orange, yellow and red) and represent striking adaptations for birds and small rodents dispersal. Seeds were generally resistant to separate digestive processes in birds and small rodents. Therefore, seeds can be dispersed far to other areas (Kleyheeg et al. 2018). Jetz et al. (2012) inferred that birds underwent a major increase in diversification rate since ca. 50 Ma. Therefore, it is likely that the dispersal and diversification of Arisaema was closely related to birds. Additionally, the islands of Southeast Asia appeared during the late Eocene to early Oligocene (Holloway & Hall 1998; Meimberg et al. 2001; Morley 2003; Clayton et al. 2009; Hall 2012; Zahirovic et al. 2014). Birds with the ability to travel at least several hundred kilometers may have spread the seeds of Arisaema widely.

In addition, Arisaema species have scents that attract mammals and especially rodents. Some rodent species act as pollinators and seed dispersers for Arisaema (Gusman & Gusman 2006), while Fabre et al. (2012) suggested that many extant families of rodent originated during the Paleogene before the end of the Oligocene (ca. 31 Ma). Oliver et al. (2017) suggested that Ctenodactylidae (Rodentia) from Mongolia has an initial burst in diversification in the early Oligocene. To settle in the evolving forest, birds and rodents of SE Asia supported the spread and diversification of Arisaema in Southeast Asia (Fig. 4), producing many endemic species in Southeast Asia, and generated high diversity in Asia, especially NJK region as suggested by Govaerts et al. (2002).

According to results from BSM (Tab. S3), within-area speciation is the most frequent type of biogeographical event in Arisaema. Within-area speciation events appear preponderant in most of the large clades of Arisaema that are endemic to single areas as defined in our study. Additionally, we found that dispersal without “range contractions” was the principal driver of range evolution which occurred about nine times more often than vicariance (Tab. S3). Dupin et al. (2017) suggested that dispersal events seem to be the most common factor for the distribution of plants, whereas, vicariant events are important factors for range evolution in many animals. Thus dispersal events between both proximal and distant regions appear to have been frequent in the historical biogeography of Arisaema and may have been facilitated by the presence of forests.

On morphology, Hara (1971) documented that a primitive type is a narrow cylindrical appendix character. This character occurs mainly in the sections Anomala and Fimbriatum. Both the two sections are distributed mainly in lowland SE Asia and Southern China. Also following Hara (1971), this character developed two evolution trends. One trend is appendix becoming more stout and clavate or even capitates forms at the upper portion and truncate at the base with a stipe. Those characters of the appendix are found in the sections Pistillata, Sinarisaema, Franchetiana, Decipientia, Nepenthoidea, and Dochafa. It means that Arisaema species possibly dispersed from SE Asia following two directions. One direction is going up to NJK region and to America and the other is from SE Asia up to high mountains of South-East Asia, Himalayas and Africa. The other trend is the appendix being elongated and often curved, flagellate tapering to a thin thread. This style of appendix was found in the sections Flagellarisae, Tenuipistillata, Tortuosa, and Arisaema. All of them are distributed in the Himalayas. The evolutionary trends
above can be developed parallel in different regions and different groups.

Migration to Himalayas and Africa from Southeast Asia

In the middle of the tree, the ancestral area reconstruction inferred that Arisaema migrated from Southeast Asia to Himalayas several times (Fig. 4). In addition, we found that Arisaema migrated from Southeast Asia to Himalayas and with no recognized reversed migration from Himalayas even though the two areas are very close. The Indian subcontinent began rifting from Australia-Antarctica ca. 136 Ma (Gibbons et al. 2013), and then collided with Eurasia at 43-50 Ma (McLoughlin 2001; Morley 2003; Zahirovic et al. 2014). Our results indicated that all the Himalayas samples diverged after 21.28 Ma: A. jacquemontii ca. 21.28 Ma (95 % HPD: 13.88, 28.91 Ma), A. album ca. 8.06 Ma (95 % HPD: 4.22, 12.83 Ma), A. sayadricum and A. murrayi ca. 1.17 Ma (95 % HPD: 0.06, 3.68 Ma), A. auriculatum and A. shimienense ca. 0.67 Ma (95 % HPD: 0, 3.51 Ma), A. exappendiculatum ca. 3.07 Ma (95 % HPD: 0.78, 6.31 Ma), A. erubescens ca. 1.13 Ma (95 % HPD: 0.03, 3.20 Ma), A. ciliatum ca. 2.76 Ma, much later than the collision between India and Eurasia. Moreover, Nauheimer et al. (2012) indicated that all genera of Pistia clade in Aroideae originated in Eurasia (excluding India) except Pistia and Protarum. Thus, the Himalayas members of Arisaema could only be derive from Southeast Asian ancestor by bird and small rodent dispersal (Fig. 4), and the uplift of the Himalayas may have hindered the reverse migration to Asia by the Himalayan members, thus forming more species endemic in that area. For the case of A. album, the endemic species from NE India, our molecular results strongly supported its placement within Attenuata section. During the colonization of Arisaema in Southeast Asia, A. album likely originated by within-area speciation events at 8 Ma after the collision between India and Eurasia when the two areas were continuous and the colonization took place near India.

Additionally, our ancestral area reconstruction analyses indicated that Arisaema migrated from Southeast Asia to Africa only once at node 6 ca. 8.83 Ma (95 % HPD: 2.52, 15.51 Ma) during the late Miocene. The Asia–Africa disjunction could be explained by three major competing hypotheses: (1) migration via the Eocene boreotropical forests (Davis et al. 2002; Weeks et al. 2005; Muellner et al. 2006); (2) overland migration via land bridges (Zhou et al. 2012); and (3) recent transoceanic long-distance dispersal (Renner 2004; Yuan et al. 2005; Clayton et al. 2009). The first hypothesis relied on the existence of extensive boreotropical forests from Europe across South Asia associated with the Eocene thermal maximum ca. 50–52 Ma (Zachos et al. 2001; 2008). During the optimum climatic period, most regions of northern Africa, including areas now occupied by the Sahara, were dominated by rainforests (Raven & Axelrod 1974; Axelrod & Raven 1978). These tropical forests could have extended much further north than extant forests, enabling floristic exchanges between Africa and Asia in both directions (Wolfe 1975; Morley 2000). This hypothesis has been used to explain the disjunction between Asia and Africa in some angiosperm families (Davis et al. 2002, 2005; Weeks et al. 2005; Lu et al. 2013; Liu et al. 2018; Zhou et al. 2019). However, the time of this expansion of boreotropical forests did not coincide with the Asian–African divergence time of Arisaema during the late Miocene. The second hypothesis “overland migration via land bridges” indicated that in the early Miocene ca. 20 Ma, the collision between the Africa continent with the Eurasia continent and the tentative closure of the Tethys sea enabled land connections and land bridge formation between Africa and Southwest Asia. These geological events, coinciding with a warm phase in the middle Miocene climatic optimum, facilitated extensive tropical biotic exchange between Asia and Africa across the Arabian Peninsula. However, the divergence time of Arisaema is much later than the time of the hypothesis. The last hypothesis “recent transoceanic long-distance dispersal” has been used to explain the disjunction between Asia and Africa in several genera of angiosperms: Adansonia (Malvaceae) (Baum et al. 1998); Osbeckia (Melastomataceae) (Renner & Meyer 2001; Renner 2004); Exacum (Gentianaceae) (Yuan et al. 2005); Eurycoma, Brucea and Soulamea (Simaroubaceae) (Clayton et al. 2009); Loranthaceae (Liu et al. 2018). This hypothesis may be more feasible for Arisaema, with a long-distance dispersal from Southeast Asia to Africa at ca. 8.83 Ma (95 % HPD: 2.52, 15.51 Ma) in the late Miocene (node 6 in Tab. 2 and Fig. 4). The taxa involved in these migrations possibly evolved into the Asian endemic species, African endemic species and species shared by both areas.

In addition, our results of the divergence between African A. schimperianum and Himalayan A. costatum at 8.83 Ma (95 % HPD: 2.52, 15.51 Ma) (node 6, Figs. 3, 4) is younger than Renner et al. (2004) (19 to 31 Ma). Renner et al. (2004) noted that their age estimates between Africa and Himalayas should be improved in future study. Thus, our result is more reliable than Renner et al. (2004) with the advances of fossil constraint, analysis methods and intensive sampling.

Dispersal from Southeast Asia to North China, Japan and the Korean Peninsula and vice versa

Our ancestral area reconstruction inferred that Arisaema dispersal from Southeast Asia to NJK region and vice versa multiple times since the early Miocene at node 9 ca. 21.27 Ma (Fig. 4). The global climate has been warm since the late Oligocene and early Miocene before it became optimum in the Mid Miocene; after that period, the climate started cooling (Zachos et al. 2001). The forests have been present since the early Cenozoic in the Northern Hemisphere. Moreover, the two areas are very close, and always contiguous. Thus we hypothesize that Southeast
Asian Arisaema was dispersed to the NJK region during a warm climate and via moist forest areas.

Based on our observation in the field, Arisaema species in subtropical and tropical areas are evergreen leaves, whereas members in temperature or mountainous forests at high elevation have seasonal leaves. Our results recognized that in the NJK region, Arisaema diverged during periods of global cooling of Miocene and Pliocene (Fig. 4). In fact, based on our observations in the field, Arisaema in the NJK region or in mountainous forests at high elevation always has seasonal leaves. This would be a mechanism of adaptation to living conditions when dispersal from tropical region to temperate region or mountainous forests at high elevation regions. In addition, the mechanism of adaptation to living conditions can be made the diversification of Arisaema in the NJK region (Figs. 4, S1).

Dispersal to America from North China, Japan and the Korean Peninsula during Miocene

The biogeographical analyses show that the colonization to America of Arisaema has occurred multiple times (Fig. 4). The disjunction of Arisaema between the NJK region and America can be explained by the hypothesis of the Bering land bridge. The Bering land bridge was mostly exposed during the Neogene for trans-Pacific floristic changes until about 3.5 Ma (Wen 1999; Ickert-Bond et al. 2009). Our ancestral area reconstruction indicated that Arisaema dispersal to America from NJK region ca. 18 Ma and 5 Ma (Figs. 3, 4). The sharp global temperature was decreased in the Paleogene (Zachos et al. 2001) and subsequent events, such as continental-wide glaciations, drying of the mid-continental region of North America and orogens in western North America and the western part of eastern Asia. However, global climate optimum in the mid Miocene (ca. 17-15 Ma) could support the dispersal of Arisaema. Thus we hypothesized that Arisaema was dispersed to America from the NJK region by bird via the Bering land bridge during the Miocene. These results highlight that active migrations during favorable periods may have resulted in the widespread distributions of many lineages. Our results also indicate the importance of the Miocene in the Northern Hemisphere intercontinental disjunction of temperate and tropical lineages.

Conclusions

This study represents a comprehensive phylogenetic and biogeographic reconstruction of the genus Arisaema, providing an understanding of the evolutionary history of the genus through time and space. Arisaema originated in Southeast Asia at ca. 31 Ma during the Oligocene with the presence and development of the tropical rainforest and suitable climatic conditions. Important events in the historical biogeography of Arisaema occurred from Oligocene to Miocene, with Southeast Asia as the center of origin and diversification of the genus in North China, Japan and the Korean Peninsula. Finally, this study demonstrates the usefulness of using large secondary molecular datasets and biogeography to determine the evolutionary origin and diversification of plant taxa that occur in Asia.

Acknowledgments

This research is funded by Graduate University of Science and Technology under grant number GUST.STS.DT2020-ST01. We also thank Dr. Lars Nauheimer of Australian Tropical Herbarium, Australia, and Dr. Peter C. Boyce from Ludwig-Maximilians-Universität München, Germany for their help in initiating this study and critical reading of the manuscript.

References


Reconstruction of the evolutionary biogeography reveals the origin and diversification of Arisaema (Araceae)


Perspectives in Plant Ecology, Evolution and Systematics 57: 5-20.


