Electronic supplementary material for Luo, S-X., L-J. Zhang, S. Yuan, Z-H. Ma, D-X. Zhang, and S. S. Renner. 2017. The largest early-diverging angiosperm family is mostly pollinated by ovipositing insects and so are most surviving lineages of early angiosperms. Proc. Roy. Soc. B, accepted 16 Nov. 2017

**Materials and methods**

1. Study species, study sites, and observation periods

Flower phenology and morphology were monitored every few weeks by local assistants who visited marked plants and took photos of the floral (or fruiting) stages, which they sent the first author via email, so he could plan his visits to coincide with maximal flowering. To determine pollen viability and the period of stigma receptivity, we used 3-(4 5-dimethylthiazol-2-yl)-2 5-diphenyltetrazolium bromide to test for the presence of dehydrogenase on pollen grains and stigma papillae (Rodriguez-Riaño and Dafni 2000).

1. Behavior, morphology, and life cycle of resin gall midges

Nocturnal and diurnal observations of flowering visitors were made for each species, covering the entire flowering period. Midge visiting and ovipositing behavior were studied by direct observation and by filming. Midges that had visited flowers were examined for pollen loads under a stereoscope as well as under SEM, focusing particularly on the abdomens and distal cerci (the ovipositor cannot easily be seen, but the cerci are conspicuous). We also used the videos to analyze details of ovipositing behavior. Numbers of eggs and larvae in flowers were counted, and egg hatching and larval development were monitored and recorded. Previous work shows that flowers of at least four species of *Kadsura* produce resin in floral tissues wounded by midges, the larva of which then feed on the resin (Luo et al., 2017). We therefore checked for the presence of resin in all studied species and monitored feeding behavior of midge larvae. We also checked the precise ovipositioning sites of the midges in flowers of different morphologies (bisexual or unisexual).

For three species that turned out to be beetle pollinated, we monitored the beetles’ visiting behavior and documented whether they carried pollen. Beetles were stored in a jar with silica gel, while midges collected for identification were preserved in 95% ethanol. Voucher specimens are in the first author’s collection.

(c) Population- and community-level investigations

To test if gall midges might be species-specific pollinators, we sequenced multiple midges per study species and site as well as midges from two or three populations of a few widespread species that we studied at multiple sites (as shown in Table S2: *Illicium lanceolatum, I. micranthum* [I. verum](http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Info&id=124778&lvl=3&lin=f&keep=1&srchmode=1&unlock), Kadsura coccinea, K. longipedunculata, [Schisandra sphenanthera](http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Info&id=13674&lvl=3&lin=f&keep=1&srchmode=1&unlock)). We also collected midges from species co-occurring at the same location. This was done at Emeishan in the west of China where we studied S. henryi, S. micranthum, S. rubriflora, and S. sphenanthera; at Hengshan in the centre of China where we studied Schisandra spec. Wang 193, S. lanceolatum, and [S. sphenanthera](http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Info&id=13674&lvl=3&lin=f&keep=1&srchmode=1&unlock); at Tianjinshan in the south of China where we studied Illicium spec. Cui 184, llicium jiadifengpi, and Kadsura coccinea, and at Daweishan in the Southwest of China where we studied Illicium majus and Schisandra spec. Luo 761.

1. Taxon sampling, DNA barcoding of midges, and co-phylogenetic analyses

For the plant phylogeny, we enlarged the matrix from Luo et al. (2017), which comprises the nuclear ribosomal DNA region ITS1-5.8S-ITS2 and two plastid regions (trnL intron, trnL-trnF spacer, and trnS-trnG), which in combination have proven useful in earlier studies of Schisandraceae (Liu et al., 2006; Morris et al., 2007; Luo et al., 2010, 2017). Primers and PCR conditions were as in Morris et al. (2007). Newly generated sequences were submitted to GenBank, and all sequences, along with voucher information (also for all sequences from earlier studies), are shown in Table S3. Sequence alignment was performed in MAFFT v. 7 in the online server (<http://mafft.cbrc.jp/alignment/> Katoh & Standley, 2013) under standard parameters except for the ITS region which was aligned under Q-INS-i optimization, which takes rRNA secondary structure into consideration. The final matrix consisted of 2416 base pairs sequenced for 46 of the ca. 90 species of Schisandraceae. We rooted our trees between Illicium and Kadsura/Schisandra, a sister group relationship found in previous studies (e.g., Liu et al., 2006; Morris et al., 2007). Maximum likelihood analyses of the plant matrix (as well as the midge matrix; below) relied on RAxML v8.0 (Stamatakis et al., 2008), using the GTR + G model of substitution with 100 replicate heuristic searches.

For the midge phylogeny, we enlarged the matrix of Luo et al. (2017), following the same methods of DNA isolation, sequencing, and alignment. Table S4 provides specimen information and GenBank accession numbers for all newly sequenced midges. We combined our sequences with additional cytochrome oxidase subunit I gene (COI) sequences of super-tribe Cecidomyiidi from GenBank, most importantly those of Yukawa et al. (2009, 2011). This supertribe has 2395 species in 11 tribes of which we included representatives of Aphidoletini, Asphondyliini, Cecidomyiini, Lestodiplosini, Lopesiini, with three species of Lasiopteridi (genus Asteromyia) used as the outgroup based on (Joy, 2013).

1. Molecular clock dating

Molecular clock dating relied on BEAST 1.8.4 (Drummond et al., 2012) under an uncorrelated lognormal (UCLN) clock model and the GTR+G model of substitution, using Yule tree priors, with Markov chain Monte Carlo (MCMC) chain length of 10 to 25 million trees, sampling every 1000th or 5000th generation with chain length depending on convergence as determined by examining the log files in Tracer v. 1.6 (<http://beast.bio.ed.ac.uk/>) after removal of an initial burn-in proportion of 10% of the trees (1000 of 10000 trees). We made sure that all effective sample size values were well above 300 as recommended in the BEAST manual. To calibrate the Schisandraceae phylogeny, we used three Schisandraceae fossils as justified in detail in Luo et al. (2017). First, Schisandra oregonensis seed fossils from the Middle Eocene Clarno Formation of Oregon dated to 44 million years (my; Manchester, 1994), which provide a minimum age for the split between the S. glabra clade (S. glabra being the only Schisandra in North America) and the remaining Schisandra species. For this constraint, we used a normal prior distribution with a range of 42-46 my. Second, leaves of Illicium spec. 1 and 2 from the Middle Eocene Geiseltal flora near Halle, Germany, resemble the modern East Asian species I. simonsii, I. lanceolatum, I. henryi, I. dunnianum, I. arborescens, and I. verum (Oh et al., 2003) and provide a minimum age for the Illicium crown group of 45 my. For this constraint, we used a normal prior distribution with a range of 43-47 my. Thirdly, *Illicium avitum*, known from seven immature fruiting axes from the Brandon Lignite of Vermont dated to the Early Miocene (16.4– 23.8 my; Tiffney & Barghoorn, 1979), provides a minimum age for the New World *Illicium* clade. For this constraint, we used a lognormal prior distribution with an offset at 16.4 my and a range back to 23.5 my.

To calibrate the midge phylogeny, we removed zero-length branches (from population-level sampling of midges) and then used the UCLN clock model with a COI substitution rate of 0.01% per million years as in other Cecidomyiidae dating studies (Stireman et al., 2010). The rate was assigned a normal distribution, with mean 0.01 and SD of 1.0. Using these settings, we obtained an age of 3.94 myfor the *Asteromyia laeviana* divergence, while Stireman et al. (2010) obtained an age of 3.76 my for the same node. Also, the standard deviations of the rate estimates (ucld.stdev) were relatively low (e.g., 0.5), suggesting that substitution accumulation in this data set is relatively clock like.

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7 Mol. Bio. Evol. 29: 1969–1973.

Joy, J. B. 2013. Symbiosis catalyses niche expansion and diversification. *Proc Roy. Soc. B* 280: 20122820. <http://dx.doi.org/10.1098/rspb.2012.2820>.

Katoh, K., and Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30: 772780.

Kawakita, A, and Kato M, 2006. Assessment of the diversity and species specificity of the mutualistic association between Epicephala moths and Glochidion trees Mol. Ecol. 15: 3567–3581.

Manchester SR. 1994. Fruits and seeds of the Middle Eocene Nut Beds flora, Clarno Formation, Oregon. *Palaeontograph. Am*. 58: 1–205.

Morris, A. B., Bell, C. D., Clayton, J. W., Judd, W., Soltis, D. E. and Soltis, P. S. 2007. Phylogeny and divergence time estimation in Illicium with implications for New World biogeography. Syst. Bot. 32: 236–249.

Oh, I-C., Denk, T., and Friis, E. M. 2003. Evolution of Illicium (Illiciaceae): Mapping morphological characters on the molecular tree. Plant Syst. Evol. 240: 175–209.

Rodriguez-Riaño T, Dafni A. 2000. A new procedure to assess pollen viability. *Sexual Plant Reproduction* 12: 241–244.

Stamatakis, A., Hoover, P., and Rougemont, J. 2008. A rapid bootstrap algorithm for the RAxML web-servers. Syst. Bio. 75: 758–771.

Stireman, J. O., Devlin, H., Carr, T. G., and Abbot, P. 2010. Evolutionary diversification of the gall midge genus Asteromyia (Cecidomyiidae) in a multitrophic ecological context. Mol. Phyl. Evol. 54. 194–210.

Tiffney, B. H. and E. S. Barghoorn. 1979. Flora of the Brandon Lignite. IV. Illiciaceae. Am. J. Bot. 66: 321–329.

Wikström, N., Savolainen, V., and Chase, M. W. 2001. Evolution of the angiosperms: calibrating the family tree. *Proc. Roy. Soc. B* 268: 2211–2220.