Title: Molecular phylogeny and macroevolution of Chaitophorinae aphids (Insecta: Hemiptera: Aphididae)

Author: Tong-Yi Liu, Jing Chen, Xi-Chao Zhu, Cai-Ling Li, Li-Yun Jiang, Karina Wieczorek, Ge-Xia Qiao

Citation style: Liu Tong-Yi, Chen Jing, Zhu Xi-Chao, Li Cai-Ling, Jiang Li-Yun, Wieczorek Karina, Qiao Ge-Xia. (2021). Molecular phylogeny and macroevolution of Chaitophorinae aphids (Insecta: Hemiptera: Aphididae). "Systematic Entomology" (2021), iss. 0, s. 1-12. DOI: 10.1111/syen.12531

UNIWERSYTET ŚLĄSKI W KATOWICACH — Biblioteka Uniwersytetu Śląskiego — Ministerstwo Nauki i Szkolnictwa Wyższego

ORIGINAL ARTICLE

Molecular phylogeny and macroevolution of Chaitophorinae aphids (Insecta: Hemiptera: Aphididae)

Tong-Yi Liu[1,2] [ID] | Jing Chen[1] | Xi-Chao Zhu[1] | Cai-Ling Li[1] | Li-Yun Jiang[1] | Karina Wieczorek[3] | Ge-Xia Qiao[1,2] [ID]

[1]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

[2]College of Life Science, University of Chinese Academy of Sciences, Beijing, China

[3]Department of Zoology, Faculty of Biology and Environmental Protection, University of Silesia, Katowice, Poland

Correspondence
Ge-Xia Qiao, Institute of Zoology, Chinese Academy of Sciences, No. 1, Beichen West Road, Chaoyang District, Beijing 100101, China.
Email: qiaogx@ioz.ac.cn

Abstract

Chaitophorinae is a predominantly Northern Hemisphere aphid subfamily characterized by numerous setae on the body. Two constituent tribes are associated with different host plants, with Chaitophorini feeding on deciduous trees and shrubs and Siphini colonizing grasses. Based on data from multiple genes (COI, COII, Cytb and EF-1α), geographical distribution and host association, this study investigated the phylogeny and macroevolution of Chaitophorinae using phylogenetic reconstruction, molecular dating, model-based ancestral area and character estimations and diversification rate calculation. Our results support the monophyly of Chaitophorinae and two tribes, indicate that *Sipha* and the two largest genera *Chaitophorus* and *Periphyllus* are not monophyletic, and suggest a need for a change in the taxonomic status of *Lambersaphis*, which was nested within *Chaitophorus* in the phylogenetic tree. We recovered an origin of Chaitophorinae on *Acer* plants from eastern Asia during the Late Cretaceous to early Palaeocene, followed by multiple dispersals into other areas that were responsible for its contemporary distribution. The origins of Siphini and *Chaitophorus* + *Lambersaphis* coincided with colonizations of novel host plants. An increase in diversification rate occurred within *Chaitophorus* in the Miocene and was associated with range expansion and switching onto new host plants, highlighting the roles of dispersal and host shift in aphid diversification.

KEYWORDS
diversification rate, geographical dispersal, historical biogeography, host shift

INTRODUCTION

The aphid subfamily Chaitophorinae (Hemiptera: Aphididae) comprises nearly 200 described species and subspecies within 12 genera and 2 tribes (Favret, 2020) and is predominantly distributed in the Holarctic region (Blackman & Eastop, 2020; Wieczorek, 2010). Chaitophorinae aphids are characterized by numerous and conspicuous body setae. They are monoecious and holocyclic, whereas some species (e.g., *Sipha (Sipha) flava* (Forbes) and *Sipha (Rungsia) maydis* Passerini) can live parthenogenetically all year round in regions with mild winters (Blackman & Eastop, 2020; Wieczorek, 2010). Most species are monophagous or oligophagous, and this high host specificity is shown in different host-association patterns within the two tribes. Species of Chaitophorini feed on deciduous trees and shrubs, including *Populus* L. (Salicaceae), *Salix* L. (Salicaceae), *Acer* L. (Aceraceae), *Aesculus* L. (Hippocastanaceae)

Tong-Yi Liu and Jing Chen contributed equally to this study.

and *Koelreuteria* Laxm. (Sapindaceae), while Siphini is associated with herbaceous monocotyledonous plants, such as Poaceae, Cyperaceae, Juncaceae and Typhaceae (Blackman & Eastop, 2020; Wieczorek, 2010). Chaitophorini contains the bulk of species diversity in the subfamily (over 170 species and subspecies in seven genera), including the two largest genera *Chaitophorus* Koch (109 species and subspecies) and *Periphyllus* van der Hoeven (50 species and subspecies) (Favret, 2020). *Chaitophorus* is widely distributed in Eurasia and North America; *Chaitogenophorus* Zhang, Qiao & Chen, *Lambersaphis* Narzikulov, *Trichaitophorus* Takahashi and *Yamatochaitophorus* Higuchi occur in Asia; *Pseudopterocomma* MacGillivray is restricted to North America; *Periphyllus* and species belonging to Siphini are predominantly distributed in Eurasia (Ghosh, 1980; Qiao et al., 2003; Richards, 1972; Wieczorek, 2010; Zhang et al., 1999). Chaitophorinae is an economically important aphid group. Some species are agricultural, horticultural and forestry pests and can transmit plant viruses (Wieczorek, 2010).

In the past few decades, limited research has been directed to the phylogenetic relationships of Chaitophorinae. In high-level phylogenetic studies of Aphididae, very few chaitophorine representatives were included, and the tribal monophyly and inner relationships were not discussed (J. Chen et al., 2017; Nováková et al., 2013; Ortiz-Rivas et al., 2004; Ortiz-Rivas & Martínez-Torres, 2010; Papasotiropoulos et al., 2013; Rebijith et al., 2017; von Dohlen & Moran, 2000). Based on 40 morphological and 2 ecological characters and by cladistic analysis, Qiao (1996) estimated the phylogeny of Chaitophorinae but failed to retrieve a monophyletic Siphini. The monophyly of Siphini was later confirmed in the cladistic analysis conducted by Wieczorek (2010). Wieczorek and Kajtoch (2011) then reconstructed a total-evidence phylogeny of Siphini using four genes (COI-tRNA-Leu-COII, ND1, EF-1α and 18S), 29 morphological and 2 ecological characters and discussed its inner relationships. Recently, Wieczorek et al. (2017) completed a relatively detailed phylogenetic study of Chaitophorinae, which tested the monophyly of this subfamily and 2 constituent tribes and investigated generic relationships using 2 molecular (COI and EF-1α), 83 morphological and 8 biological characters. In their study, the monophyly of Chaitophorinae and Siphini was supported, but Chaitophorini was not recovered as monophyletic. Previous phylogenetic studies have focused mostly on Siphini, and the inner relationships of Chaitophorini remain mostly unexplored. The insufficient sampling of Chaitophorini, especially considering its great species diversity, makes it necessary to examine previous conclusions based on a broader range of taxa.

Fossil records of the Chaitophorinae aphids are too sparse and young to provide valuable information on the origin of this group. Only one fossil of the extant subspecies *Chaitophorus salijaponicus niger* Mordvillko is known from Greenland and dates to 2.5 Ma (Heie, 1995). Aphid origin and diversification are closely linked with their host plants (Heie, 1987, 1996; von Dohlen & Moran, 2000). The same is true for the chaitophorine aphids. Their ancestor is thought to be monoecious on angiosperms (Moran, 1992). Shingleton and Stern (2003) and Wieczorek et al. (2017) investigated the history of host use for *Chaitophorus*, which was regarded as the ancient lineage

within Chaitophorinae by taxonomists (Chakrabarti & Mandal, 1986; Szelegiewicz, 1961). They both supported the hypothesis that *Chaitophorus* was ancestrally associated with *Populus* and that several independent shifts from *Populus* to *Salix* occurred during its evolution. Qiao (1996) speculated that the Chaitophorinae aphids might have originated from a common ancestor feeding on woody plants older than Salicaceae and that the diversification of Siphini was accompanied by host switches to Poaceae plants. Based on analysis of morphological, anatomical and palaeobotanical data, Wieczorek (2010) proposed a similar but more specific hypothesis that Chaitophorinae had arisen on woody plants during the Late Cretaceous to Palaeogene and that a complete host shift from woody to herbaceous plants had taken place for Siphini in the Miocene, which coincided with the origin of this tribe from Eurasia. Wieczorek and Kajtoch (2011) again repeated the hypothesis that the Siphini originated on the steppes of Eurasia in the Miocene.

Historical biogeographic studies are important to understand the aphid diversification and historical processes that have given rise to contemporary distribution. However, few studies have explored the biogeographical scenarios of Chaitophorinae, especially in a statistically rigorous framework. The lack of timing data and statistical tests for ancestral areas makes previous inferences (Wieczorek, 2010; Wieczorek & Kajtoch, 2011) unreliable. In addition, considering the importance of host plants to aphid evolution and the strong host specificity of Chaitophorinae species, it is necessary to ascertain the history of host association and evaluate its impacts on Chaitophorinae diversification.

Here, we present the most extensive phylogenetic study of Chaitophorinae to date based on one nuclear gene and three mitochondrial genes and explore its macroevolutionary history. We estimated divergence times, reconstructed ancestral areas and host associations and calculated diversification rates. The aims of this study were to obtain a robust and detailed phylogeny of Chaitophorinae, to establish a temporal framework and historical scenario for the evolution of Chaitophorinae aphids, and to determine how the current distribution pattern formed and how biogeographic and host-association changes influenced the diversification of Chaitophorinae.

MATERIALS AND METHODS

Taxon sampling and molecular data

A total of 52 species/subspecies from 10 of the 12 described genera of Chaitophorinae were included in this study. Forty-one ingroup species/subspecies and five genera were sampled from Chaitophorini. The ingroup sampling covered all major biogeographic regions of Chaitophorinae aphids. Twelve species from Adelgidae, Phylloxeridae, Aphidinae, Calaphidinae and Greenideinae were selected as outgroups based on previous phylogenetic studies (Nováková et al., 2013; Ortiz-Rivas & Martínez-Torres, 2010; von Dohlen & Moran, 2000).

Three mitochondrial markers (COI, COII and Cytb) and one nuclear marker (EF-1α) were used for phylogenetic inference.

Sequences retrieved from GenBank were combined with 47 novel sequences obtained from this study. The detailed procedures of DNA extraction, PCR, and sequencing followed Liu et al. (2015). Sequences were assembled using SeqMan II (DNAStar, Madison, WI, U.S.A.) and verified for protein-coding frame shifts with EditSeq (DNAStar). Introns of EF-1α sequences were identified and removed before further analysis. Novel sequences have been deposited in GenBank. All slide-mounded voucher specimens and ethanol-preserved samples were deposited in the National Zoological Museum of China, Institute of Zoology, Chinese Academy of Sciences, Beijing, China. Voucher information and GenBank accession numbers are summarized in Table S1.

Phylogenetic analyses

Multiple alignments were carried out with the MAFFT online server (Katoh et al., 2019; Kuraku et al., 2013), followed by manual adjustments in MEGA X (Kumar et al., 2018). The characteristics for each gene, such as the numbers of parsimony-informative and variable sites, were measured using DnaSP v5 (Librado & Rozas, 2009) and are summarized in Table S2. Then, we concatenated individual gene datasets using SequenceMatrix 1.8 (Vaidya et al., 2011), producing a final matrix of 2422 bp.

Phylogenetic trees were estimated using maximum-likelihood (ML) approach and Bayesian inference (BI). The combined dataset was partitioned by codon position. We used PartitionFinder v2.1.1 (Lanfear et al., 2017) to simultaneously assess the optimal partitioning strategy and substitution model with the following settings: branchlengths = linked (higher likelihood than the unlinked option), models = all, model_selection = bic (Bayesian information criterion) and search = greedy. The resulting partitioning scheme and models are summarized in Table S3. BI analysis was performed in MrBayes 3.2.6 (Ronquist et al., 2012) under default priors, with each partition unlinked for parameter estimations. Two concurrent runs, each with four chains, were conducted for 20 million generations. The chains were sampled every 500 generations. Stationarity was assumed when the average standard deviation (SD) of split frequencies fell below 0.01 and the effective sample size (ESS) values of all parameters were greater than 200. The first 25% of trees were discarded as burn-in. ML analysis was inferred using RAxML v8.2.10 (Stamatakis, 2014), with the GTRGAMMA model for each partition and 1000 rapid bootstrapping replicates.

Four morphologically defined genera of Chaitophorinae (i.e., *Chaitophorus*, *Periphyllus*, *Sipha* Passerini and *Trichaitophorus*) were not retrieved as monophyletic in the above phylogenetic analyses. We thereby performed approximately unbiased (AU) tests (Shimodaira, 2002) to assess their monophyly. We first constrained the monophyly of a specific genus and inferred the ML tree using RAxML. The site-wise log-likelihoods were calculated for each topology by TREE-PUZZLE 5.3 (Schmidt et al., 2002). *p*-Values of AU tests were then calculated with CONSEL v.0.1j (Shimodaira & Hasegawa, 2001) to assess the level of statistical support for the constrained and unconstrained ML trees.

Molecular dating

Divergence times were estimated with BEAST v2.6.6 (Bouckaert et al., 2019) using the combined molecular dataset. The partitions and substitution models were employed based on the result of PartitionFinder analysis (Table S3). An uncorrelated lognormal relaxed clock model with a birth-death tree prior was applied. The best tree resulting from RAxML analysis was scaled by the calibration ages stated below using the *chronos* function in the R package APE 5.2 (Paradis et al., 2004). This tree was utilized as the starting tree, but the topology was not fixed. Doing this avoided the failure in the initialization phase of BEAST analysis, which might arise from the conflict between the randomly calculated starting tree and node age prior. Four independent analyses were run, each of 500 million generations with sampling every 50,000 generations. Convergence of the chains was monitored using Tracer v1.7.1 (Rambaut et al., 2018), with target ESS values greater than 200 for all parameters. Runs were combined and resampled with LogCombiner v2.6.6 (Bouckaert et al., 2014), with 25% of samples in each run were discarded as burn-in, resulting in approximately 10,000 samples in the posterior distribution. TreeAnnotator v2.6.6 (Bouckaert et al., 2014) was used to generate the maximum clade credibility (MCC) tree, with the mean node age estimates and 95% high posterior density (HPD) intervals.

Four fossils calibrations were selected following the recommendations of Parham et al. (2012). The fossil record of the extant subspecies *C. salijaponicus niger* Mordvillko from the Kap København Formation in Greenland has been dated back to 2.5 Ma (Heie, 1995). We thereby set the divergence time between *C. salijaponicus niger* and its sister *Chaitophorus salicti* to a minimum age of 2.5 Ma using a lognormal prior with an offset of 2.5 Ma and an SD of 1. Two other fossils belong to Aphidinae. The oldest fossil record of Aphidinae is the fossils †*Aphidocallis caudata* Kononova from Taimyr amber dated to the Santonian, Late Cretaceous (85.8–83.5 Ma) (Heie & Wegierek, 2011; Kononova, 1978). It can be placed in Aphidinae based on characters of the cauda, but not within any particular group (Kononova, 1978). Therefore, we placed this fossil at the stem of Aphidinae with a minimum age of 83.5 Ma (offset 83.5; SD 1.5). The second is †*Halajaphis siphonosetae* Wegierek from Baltic amber in Poland and has been dated to late Priabonian, in the upper Eocene (Wegierek, 1996; Heie & Wegierek, 2011). The triangular cauda and cylindrical siphunculi place this fossil within Macrosiphini. We used it to calibrate the crown of the tribe Macrosiphini with a lognormal distribution bounded by a minimum age of 33.9 Ma (offset 33.9; mean = 1.0; SD 1.3). For the age of the common ancestor of the Aphididae, Adelgidae and Phylloxeridae (i.e., the root), a maximum age constraint of 165 Ma was utilized to avoid unrealistic divergence time estimates using a uniform prior distribution. The maximum age constraint was based on the oldest undoubted fossil †*Daoaphis magnalata* from the Middle Jurassic in Inner Mongolia, China (Huang et al., 2015), which belongs to the extinct family Oviparosiphidae thought to be sister to Aphididae (Havill et al., 2007; Heie, 1987; von Dohlen & Moran, 2000).

Ancestral area estimation

We performed ancestral area estimation on the BEAST MCC tree with outgroups removed using the R package BioGeoBEARS v1.1.1 (Matzke, 2018). Threetime slices were defined (80–66 Ma, 66–30 Ma, 30–0 Ma), and dispersal rates were assigned for each time slice from 0.10 for well-separated areas to 1.00 for well-connected areas (Table S4). We applied and compared six biogeographical models implemented in BioGeoBEARS: (i) the dispersal-extinction-cladogenesis (DEC) model (Ree et al., 2005); (ii) DEC + J; (iii) DIVALIKE, a likelihood version of dispersal-vicariance analysis (DIVA) (Ronquist, 1997); (iv) DIVALIKE + J; (v) BAYAREALIKE, a likelihood version of the Bayesian model implemented in BayArea (Landis et al., 2013); and (vi) BAYAREALIKE + J. The '+ J' models allow for founder-event speciation (Matzke, 2014). Model fit was assessed using the Akaike information criterion (AIC) and Akaike weights (Wagenmakers & Farrell, 2004).

Distribution data of Chaitophorinae species were recovered from the online datasets of Global Biodiversity Information Facility (www.gbif.org), Barcode of Life Data System v4 (BOLD; www.boldsystems.org) and Aphids on the World's Plants: An Online Identification and Information Guide (www.aphidsonworldsplants.info), and literature (Ghosh, 1980; Higuchi, 1972; Pintera, 1987; Richards, 1972; Wieczorek, 2010; Zhang et al., 1999). Invasion records were not included. Five geographical regions were defined: Nearctic (NA), West Palaearctic (WP), East Palaearctic (EP), Neotropical (NT) and Oriental (OL). The regions were taken from Terrestrial Ecoregions of the World (Olson et al., 2001), with the Palaearctic region divided into western and eastern areas. The distribution matrix coding the presence/absence of each species was produced using SpeciesGeoCoder v1.1.0 (Töpel et al., 2017).

Ancestral host reconstruction

Ancestral state reconstruction was conducted to infer the evolutionary history of host association within Chaitophorinae using the Multistate submodel and Markov chain Monte Carlo (MCMC) in BayesTraits (Meade & Pagel, 2017). Host plant information was obtained from the online dataset aphids on the world's plants: An Online Identification and Information Guide (www.aphidsonworldsplants.info) and literature (Wieczorek, 2010; Zhang & Zhong, 1982). Each sampled species of Chaitophorinae was encoded as follows: (0) *Salix*, (1) *Populus*, (2) *Acer*, (3) *Aesculus*, (4) *Koelreuteria*, (5) Juncaceae, (6) Poaceae and (7) Cyperaceae. Ancestral states were reconstructed for all nodes. Reverse jump MCMC analyses (Pagel & Meade, 2006) were performed based on the BEAST MCC tree with outgroup taxa pruned. A hyper exponential prior seeded from a uniform distribution on the interval 0 to 2 was used. Two analyses were run for 101 million iterations, sampling every 1000 iterations after a burn-in of 1 million iterations. Thereafter, we summarized the posterior probabilities of host association as the mean of the remaining sampling. Given the similar results of the two runs, we report only one of them here.

Diversification analyses

Semi-logarithmic lineage through time plots (LTT) were constructed using the R package APE 5.2 (Paradis et al., 2004). One thousand randomly sampled post-burn-in trees from the BEAST analysis were used to generate the confidence intervals. We used BAMM v2.5.0 (Rabosky, 2014) to estimate the variation in diversification rates over time and among lineages based on the BEAST MCC tree with outgroups pruned. The priors were determined using the *setBAMMpriors* command in the R package BAMMtools v2.1.6 (Rabosky, Grundler, et al., 2014). To account for incomplete taxon sampling, the sampling fraction of each genus was provided. Two runs of 10 million generations with a sampling frequency of 1000 were conducted. The convergence was checked by plotting the log-likelihood trace and estimating the ESS values with the R package CODA v0.19-2 (Plummer et al., 2006). The first 10% of the results were discarded as burn-in. To test the sensitivity of the posterior estimates to the prior (Moore et al., 2016), we set different 'expected number of rate shift' priors (0.1, 1.0 and 10). Three priors yielded similar results. Post-run analyses and data visualizations were performed using the following methods in BAMMtools: (i) plotting the single best shift configuration with the highest posterior probability; (ii) macroevolutionary cohort analysis summarizing the macroevolutionary rate heterogeneity on the phylogenetic tree (Rabosky, Donnellan, et al., 2014) and (iii) calculating the marginal (branch-specific) shift probability for each branch.

RESULTS

Phylogenetic reconstructions

The phylogenetic analyses under ML and BI approaches yielded largely consistent ingroup topology (Figure 1). The subfamily Chaitophorinae was retrieved as monophyletic with strong support (bootstrap support, BS = 99%; posterior probability, PP = 1). Its two constituent tribes, Siphini and Chaitophorini, were also monophyletic. Within the clade of Siphini (BS = 100%, PP = 1), *S. (S.) flava* split off earliest from other taxa. *Laingia psammae* was then placed as sister to the remaining species, which were clustered into two clades. One clade (BS = 93%, PP = 1) consisted of *Sipha (Rungsia) burakowskii*, *Atheroides* Haliday and *Chaetosiphella* Hille Ris Lambers, with the latter two monophyletic genera forming sister taxa in a well-supported clade (BS = 99%, PP = 1); the other clade (BS = 95%, PP = 1) included *Caricosipha paniculatae* and the remaining species of *Sipha (Rungsia)*. The genus *Sipha* and its subgenus *Rungsia* Mimeur were both polyphyletic. The tribe Chaitophorini was recovered as monophyletic with good support (BS = 71%, PP = 0.95). The earliest branching lineage (BS = 98%, PP = 1) comprised species of *Periphyllus*, which were separated into two clusters. The rest of the Chaitophorini representatives were split into two well-supported clades (BS = 100% and 99%, PP = 1). One clade included the monophyletic *Yamatochaitophorus*, the polyphyletic *Trichaitophorus* and the remaining *Periphyllus* species. All representatives of the genus *Chaitophorus* formed the other clade,
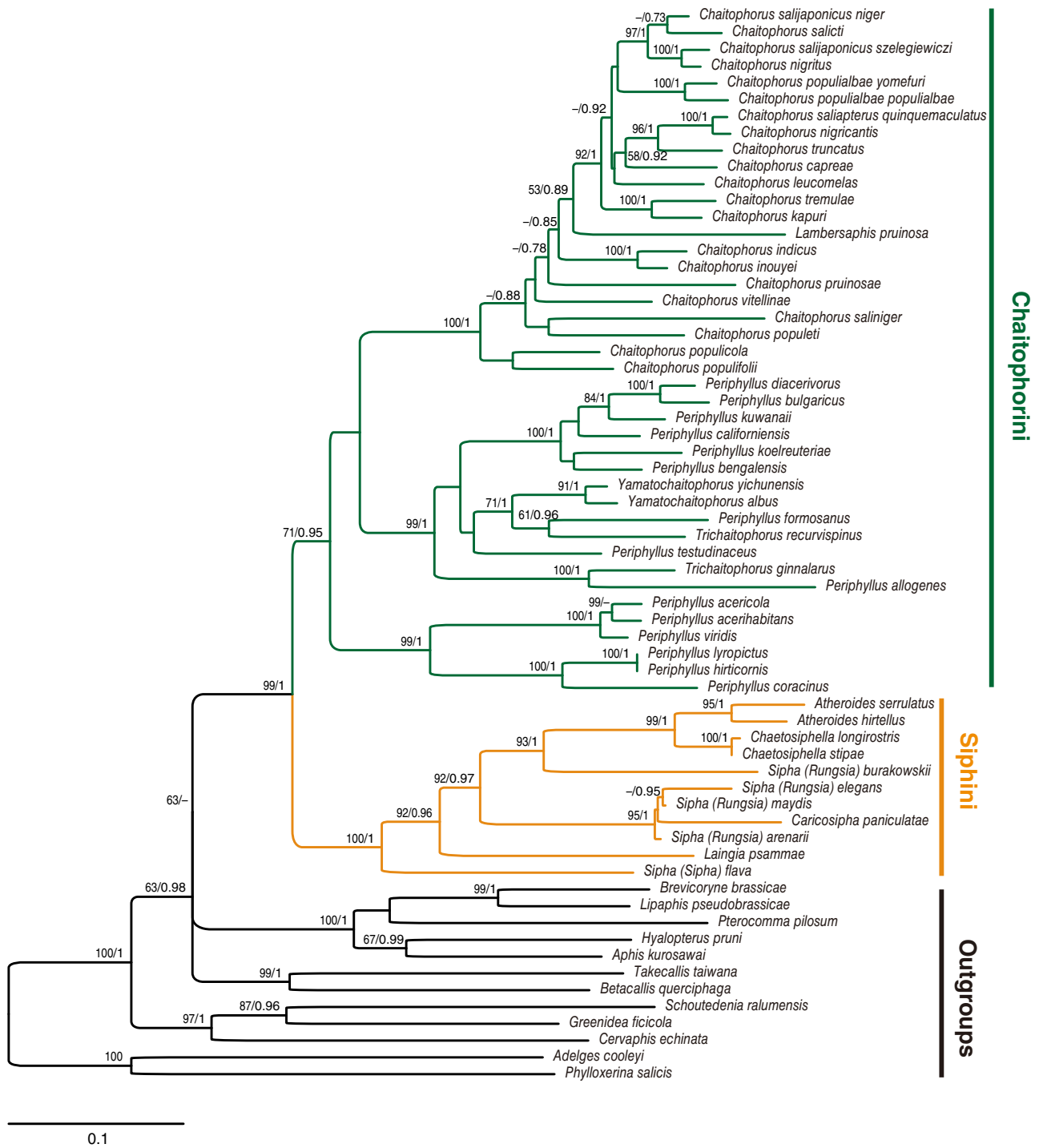
FIGURE 1 Maximum-likelihood phylogeny of Chaitophorinae. Node values indicate maximum-likelihood bootstrap supports (>50%) and Bayesian inference posterior probabilities (>0.7)

with *Lambersaphis pruinosa* being nested within it, thereby making *Chaitophorus* paraphyletic. Two subspecies, *Chaitophorus populialbae populialbae* and *Chaitophorus populialbae yomefuri*, formed a single clade (BS = 100%, PP = 1), whereas *C. salijaponicus niger* did not cluster with *Chaitophorus salijaponicus szelegiewiczi*.

In the AU tests of specific phylogenetic hypotheses (Table S5), the monophyly of genera *Periphyllus*, *Sipha* and *Trichaitophorus* were all significantly rejected ($p < 0.05$), which supported the phylogenetic

topologies shown in Figure 1. For *Chaitophorus*, the topology constrained to monophyly could not be rejected ($p = 0.393$).

Divergence times

Divergence time estimates for Chaitophorinae are shown in Figure S1. Chaitophorinae split from its sister lineage about 85 Ma

(95% HPD: 88–84 Ma). The Chaitophorinae crown was estimated to have arisen at 69 Ma (95% HPD: 75–61 Ma). The crown ages of Siphini and Chaitophorini were inferred to be 54 Ma (95% HPD: 64–47 Ma) and 60 Ma (95% HPD: 67–52 Ma), respectively. Within Chaitophorini, the major clades (nodes 6, 9 and 11 in Figure 2) were dated between 42 and 33 Ma.

Historical biogeography

The statistical results of six biogeographic models implemented in BioGeoBEARS are summarized in Table 1. Comparative statistics show that the BAYAREALIKE model received the lowest AIC score, and therefore, best fitted our data (log-likelihood = −105.083, AIC = 214.2, AIC weight = 0.64; Table 1), suggesting that the biogeographical patterns of Chaitophorinae aphids may have been mainly shaped by dispersal and sympatric cladogenesis events (Matzke, 2013).

According to the preferred BAYAREALIKE model, the ancestors of Chaitophorinae and its two constituent tribes Siphini and Chaitophorini most likely inhabited the EP, which was also inferred as the ancestral area for most internal nodes (Figures 2a and S2). Dispersal was estimated to be dominant in the biogeographic history of Chaitophorinae. During the late Oligocene to early Miocene (26–16 Ma, Figure S1), the ancestors of nodes 7 and 12 (Figure 2a) colonized the WP and NA, respectively, from the EP by dispersal; the ancestor of node 10 expanded its range into the OL region, and the ancestors of nodes 1, 13 and 4 expanded into the WP. More recently (c. 10 Ma, Figure S1), one range expansion into the OL region was found in the ancestor of node 15, the recent ancestor of which inhabited the whole Palaearctic region. Dispersal events have occurred more frequently in terminal taxa, mainly resulting in expansion of geographical ranges. Range contractions were also observed in several species (e.g., *C. paniculatae* and *Trichaitophorus recurvispinus*). Through dispersal, *S. (S.) flava* and *Periphyllus bengalensis* colonized the
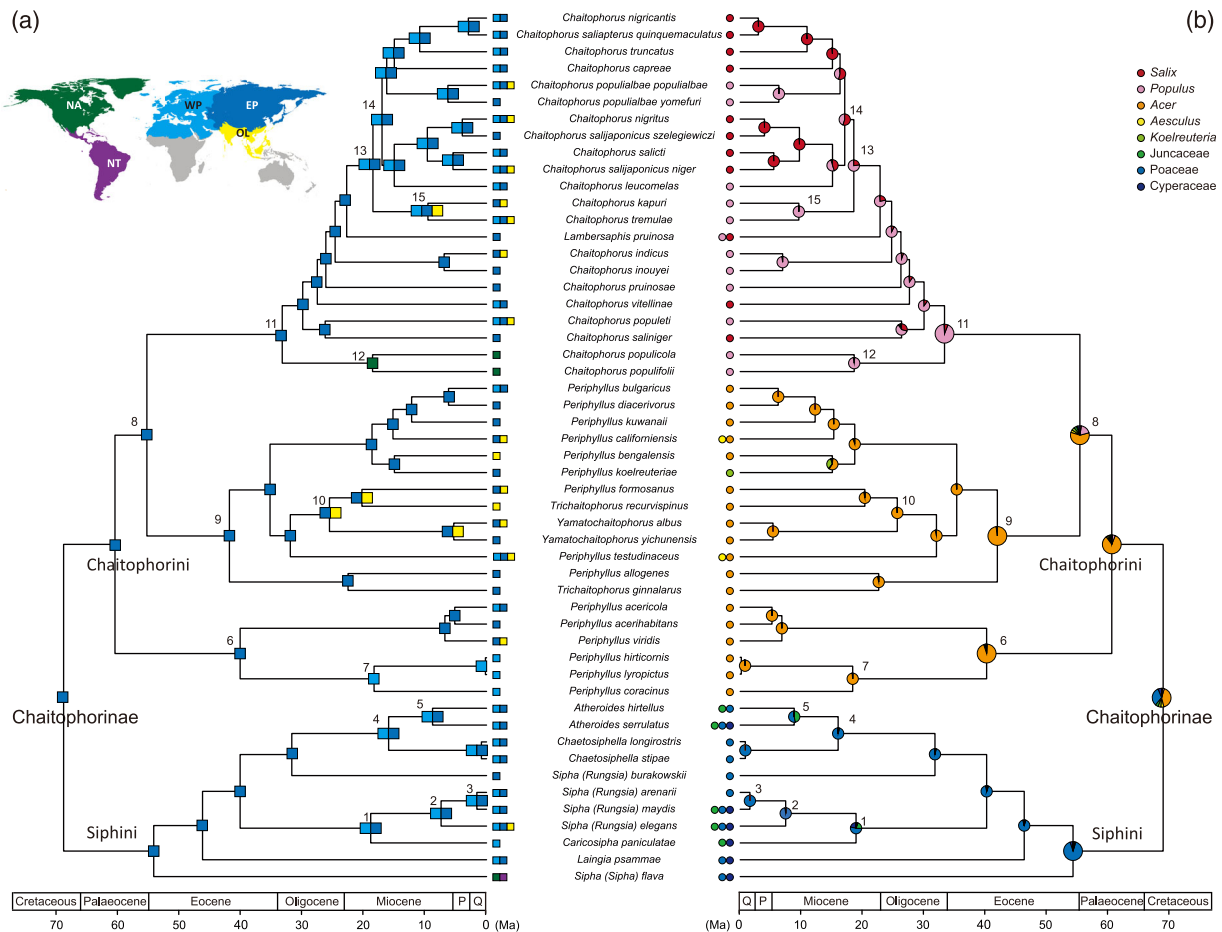

FIGURE 2 (a) Ancestral area estimations of Chaitophorinae based on the BAYAREALIKE model. The coloured box at each node represents the ancestral area with the highest relative probability. A single box indicates a single geographical area; combined boxes indicate a distribution including two or more areas. (b) Bayesian reconstructions for host association within Chaitophorinae. The pie chart at each node shows the posterior probabilities for character states. EP, East Palaearctic; NA, Nearctic; NT, Neotropical; OL, Oriental; P, Pliocene; Q, Quaternary; WP, West Palaearctic

TABLE 1 Parameter estimates of different models and model comparison in BioGeoBEARS.

Model	LnL	Number of parameters	d	e	j	AIC	AICwt
DEC	−115.928	2	0.021	0.002	0	235.86	0.73
DEC + J	−115.928	3	0.021	0.002	1×10^{-5}	237.9	0.27
DIVALIKE	−124.763	2	0.023	0.001	0	253.53	0.73
DIVALIKE + J	−124.763	3	0.023	0.001	1×10^{-5}	255.5	0.27
BAYAREALIKE	−105.083	2	0.014	0.009	0	214.2	0.64
BAYAREALIKE + J	−104.640	3	0.013	0.007	0.014	215.3	0.36

Abbreviations: AIC, Akaike information criterion; AICwt, AIC weight; d, rate of dispersal; e, rate of extinction; j, relative probability of founder-event speciation; LnL, log-likelihood.
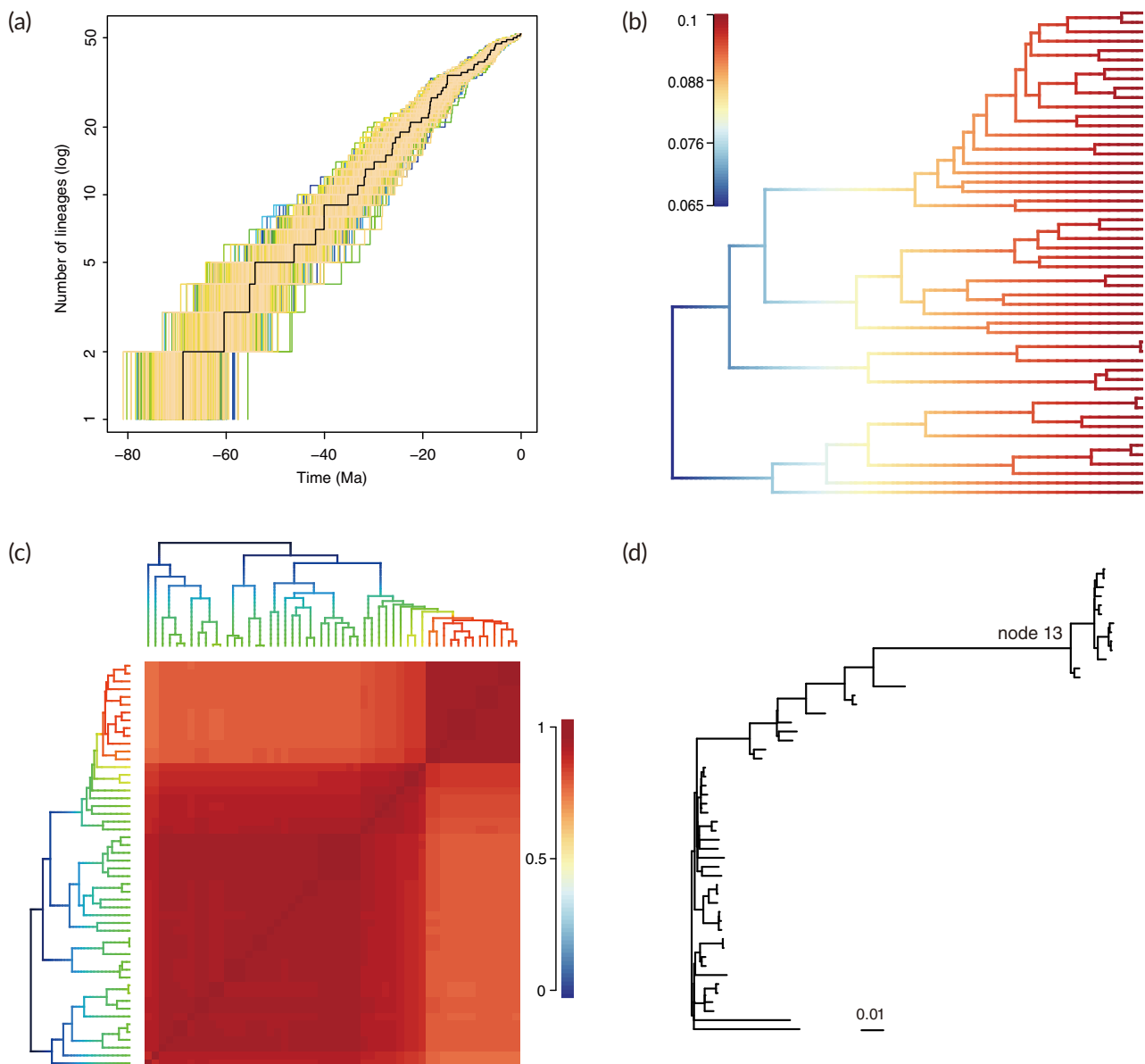
FIGURE 3 (a) Semi-logarithmic plots of lineage through time. The black line represents the maximum clade credibility tree, and the coloured lines show the age uncertainty of 1000 randomly sampled post-burn-in trees from the BEAST analysis. (b) The best shift configuration from the BAMM analysis indicating the diversification rate across lineages. Warmer colour indicates a faster rate. (c) Macroevolutionary cohort matrix illustrating the pairwise probability of any two species sharing a common macroevolutionary rate dynamic. Warmer matrix colour represents a higher probability. (d) Chaitophorinae phylogeny with the branch lengths scaled by branch-specific probability that they contain a rate shift event

Americas and the OL region, respectively, which were not part of the distribution areas of their respective recent ancestors.

Ancestral host reconstruction

The ancestral state reconstructions for host association within Chaitophorinae are shown in Figure 2b. The ancestor of Chaitophorinae most likely fed on *Acer* plants (PP = 0.43). Poaceae and *Acer* were strongly suggested as ancestral host plants for Siphini (PP = 0.91) and Chaitophorini (PP = 0.84), respectively. Within the tribe Siphini, Poaceae was favoured for most internal nodes, while the ancestral state for node 5 was equivocal. Host expansion onto Cyperaceae occurred once each in *S. (S.) flava* and *L. psammae*. At least one expansion to Juncaceae and to Cyperaceae likely took place in both *S. (S.) maydis* and *S. (S.) elegans*. In addition, at least one host expansion onto Juncaceae in the ancestor of *Atheroides* (node 5) and one expansion onto Cyperaceae in *A. serrulatus* were identified. Within the tribe Chaitophorini, all internal nodes were estimated to be associated with *Acer* except the clade comprising *Chaitophorus* and *Lambersaphis*, which the ancestor (node 11) colonized *Populus* plants. For the sampled Chaitophorini representatives, at least eight host-association transitions may have occurred: one host shift to *Koelreuteria* in *Periphyllus koelreuteriae*; at least one shift to *Populus* in the ancestor of node 11; at least three shifts to *Salix* with one in each of *Chaitophorus saliniger* and *Chaitophorus vitellinae* and at least one in the ancestor of node 14; two host expansions onto *Aesculus* with one in each of *Periphyllus californiensis* and *Periphyllus testudinaceus*; and one expansion onto *Salix* in *L. pruinosa*.

Diversification analyses

The semi-logarithmic LTT plots were approximately linear, suggesting relatively constant speciation and extinction rates (Figure 3a). This is consistent with the BAMM analysis, which showed a slight increase in diversification rate over time (Figure 3b). Nevertheless, the macroevolutionary cohort analysis (Figure 3c) revealed two macroevolutionary dynamics across Chaitophorinae: the sister clade of *L. pruinosa* (node 13 in Figure 2) that underwent a major increase in diversification rate and the remaining Chaitophorinae species. Node 13 also showed the highest branch-specific rate shift probability (Figure 3d).

DISCUSSION

Phylogenetic relationships

The monophyly of Chaitophorinae was strongly supported in the present study, in agreement with the molecular phylogeny of Wieczorek et al. (2017). The tribe Siphini was recovered as monophyletic, which is consistent with previous studies (Wieczorek, 2010; Wieczorek et al., 2017; Wieczorek & Kajtoch, 2011). Species of Siphini are

characterized by 5- or rarely 4-segmented antennae, siphunculi pore-like or slightly elevated without reticulation, apterous males and feeding on herbaceous monocotyledonous plants. The genus *Sipha* and its subgenus *Rungsia* were not monophyletic. The monophyly of *Sipha* was also not supported in the morphological and total-evidence phylogenies obtained in Wieczorek and Kajtoch (2011). In the morphological cladistic analysis of Siphini (Wieczorek, 2010), although a monophyletic *Sipha* was retrieved, Wieczorek pointed out that its monophyly was only supported by a single homoplastic synapomorphy, empodial setae pointed, which is also common in other genera. Therefore, the genus *Sipha* should be reviewed thoroughly. *Atheroides* and *Chaetosiphella* were well-supported sister groups. Several morphological characteristics are shared by these two genera, such as dorsal setae of the body with variable shapes of apices, siphunculi pore-like, cauda broadly rounded and 8-shaped pseudosensoria on hind tibiae in oviparous females (Wieczorek, 2010).

The tribe Chaitophorini was retrieved as monophyletic in our phylogenetic analyses. This tribe is characterized by 6-segmented antennae, stump-shaped siphunculi with reticulation and feeding on deciduous trees and shrubs. The genus *Periphyllus* was polyphyletic, with some species being placed as the sister clade to all remaining Chaitophorini representatives while others were forming a separate clade wherein species of *Yamatochaitophorus* and *Trichaitophorus* were nested. *Periphyllus*, *Yamatochaitophorus* and *Trichaitophorus* all live on maples and are morphologically similar in alate morphs, especially in the setal pattern of embryos (Chakrabarti & Mandal, 1986; Quednau & Chakrabarti, 1976). The most species-rich genus, *Chaitophorus*, was paraphyletic, with the monotypic genus *Lambersaphis* nested within it. Both *Chaitophorus* and *Lambersaphis* inhabit poplars and willows. *Lambersaphis* occurs in central Asia and is distinguished by very short processus terminalis, short and crater-like siphunculi without reticulation, semicircular cauda, short and sparce dorsal setae of body and fuscous-bordered wing veins (Pintera, 1987; Qiao et al., 2003). However, the latter three characteristics are also shared by some *Chaitophorus* species (Pintera, 1987). Therefore, our phylogenetic inferences suggest that the taxonomic validity of *Lambersaphis* needs re-assessment.

Origins of Chaitophorinae and Siphini

Previous studies speculated that Chaitophorinae might have originated on woody plants (Qiao, 1996; Wieczorek, 2010). In the present study, the Chaitophorinae crown dated to 69 Ma. Reconstruction analyses indicated that the common ancestor of Chaitophorinae aphids occupied the EP and fed on *Acer* plants (Figure 2). Fossil leaves and samaras of *Acer* are abundant throughout the Northern Hemisphere, with many found in Palaeocene strata (Wolfe & Tanai, 1987; Zhu & Manchester, 2020). To date, the earliest maple fossil record, *Acer amboyense* Newberry, dates to the Late Cretaceous (c. 100–94 Ma) (Jin, 2009; Newberry, 1895). A recent maple biogeographic study employing phylogenomic data inferred that the *Acer* ancestor had arisen by the early Palaeocene from eastern Asia, which was also the

modern centre of *Acer* diversity (Li et al., 2019). Although their estimated crown age of *Acer* is younger than the oldest fossil record, the hypothesis of an Asian origin of *Acer* is consistent with the perspectives of some botanists (Pojàrkova, 1933; Xu, 1998). Therefore, we hypothesize that Chaitophorinae may have originated from eastern Asia during the Late Cretaceous to early Palaeocene, coinciding with the occurrence of their host plants *Acer*.

The tribe Siphini was estimated to originate in the early Eocene after the origin of Chaitophorini in the Palaeocene (Figure S1), supporting the viewpoint of Wieczorek (2010) that Siphini is a young clade within Chaitophorinae. Wieczorek (2010) and Wieczorek and Kajtoch (2011) proposed that the Siphini might have originated from Eurasia in the Miocene, accompanied by a host shift to grasses. The hypothesis concluded from our reconstructions (Figure 2) is roughly consistent with previous inference, but a much earlier origin for Siphini was estimated here. The common ancestor of Siphini may have switched from *Acer* to Poaceae plants and then arisen in Asia by the early Eocene. The origin and early divergences of Poaceae are still uncertain. It is now generally accepted that Poaceae originated from Gondwana (South America or Africa) (Bouchenak-Khelladi et al., 2010; Bremer, 2002). Wu et al. (2018) reported the earliest known Poaceae fossils from the late Early Cretaceous (*c.* 113–101 Ma) of China and speculated, based on fossil records that the deep-diverging lineages of Poaceae could have quickly achieved a worldwide distribution (excl. North America) during *c.* 129–125 Ma. Therefore, despite the absence of fossil records (Strömberg, 2011), it is reasonable to assume a distribution of grasses in Asia during the early Eocene, which provided a necessary condition for the origin of Siphini aphids. In addition, the latest Cretaceous Poaceae fossils were found on the Indian subcontinent (Prasad et al., 2005). It is possible that the Indian grasses had migrated into Asia following the India-Asia collision at 55–50 Ma (Clyde et al., 2003; Strömberg, 2011), which may also have given the Siphini ancestor an opportunity to colonize new hosts.

Biogeographical and host-association history and diversification

During the late Palaeocene to early Oligocene, early divergences within Chaitophorini and Siphini took place in eastern Asia where their ancestors lived. Deep-branching lineages of Chaitophorini continued to colonize the original host maples until approximately 33 Ma when a host shift to poplars occurred (node 11 in Figure 2b). According to fossil records (Hsü, 1983), *Populus* was distributed in eastern Asia during that time, making such complete host-plant switching associated with the emergence of MRCA of *Chaitophorus* and *Lambersaphis* possible. By the early Miocene, *Chaitophorus* dispersed into North America directly through the Beringia Land Bridge (Sanmartín et al., 2001) and still lived on poplars that were then present in western North America (Collinson, 1992). Soon after (*c.* 18 Ma), a clade of poplar-feeding *Chaitophorus* (node 13 in Figure 2) expanded into Europe, coupled with an acceleration in diversification

rate (Figures 3c,d). The Turgai Strait that separated the Palaearctic dried up at 30 Ma (Sanmartín et al., 2001), and *Populus* were widely distributed in central Europe during the early Miocene (Collinson, 1992), allowing the range expansion of *Chaitophorus* from Asia into Europe, which subsequently led to an increase in Eurasian distributions in the Miocene. This westward expansion coincided with a diversification rate shift within *Chaitophorus*, suggesting that geographical dispersal might be an important factor behind aphid diversification. Migration into new geographical areas could expose dispersing species to novel biotic or abiotic conditions and provide new ecological opportunities. Dispersal followed by geographical isolation or exploitation of new resources (e.g., colonizing novel host plants) may lead to speciation. Within this westward-dispersing clade of *Chaitophorus*, the host shifted from poplars to willows (node 14 in Figure 2b) occurred during the middle Miocene when *Salix* was available in Eurasia (Collinson, 1992; Hui et al., 2011). Previous studies have highlighted the importance of host switching in aphid diversification (R. Chen et al., 2016; Favret & Voegtlin, 2004; Liu et al., 2015; Moran et al., 1999). Therefore, dispersal and host shift may have been responsible for the increased diversification of *Chaitophorus* in the Miocene. During the late Oligocene–Miocene interval, some Chaitophorini and Siphini lineages (nodes 10, 7, 4 and 1 in Figure 2) that lived on their respective ancestral host plants successively dispersed from the EP into other areas (i.e., the OL region and WP). In addition, some extant species have undergone geographical dispersal or host expansion or shift since they diverged, leading to current patterns of distribution and host association.

CONCLUSIONS

Based on broad sampling and multiple genes, we provide valuable insights into the phylogeny, biogeography and diversification of the aphid subfamily Chaitophorinae. We infer that the Chaitophorinae originated from eastern Asia, along with the emergence of *Acer* during the Late Cretaceous to early Palaeocene. The origins of Siphini and the common ancestor of *Chaitophorus* and *Lambersaphis* in their ancestral areas were both associated with changes in host use, highlighting the contribution of host shift to the divergence of major lineages within Chaitophorinae. Dispersal was estimated to be the dominant process in the history of Chaitophorinae. A dispersal from Asia into Europe and host shifts from poplars to willows may have triggered the increased diversification of *Chaitophorus* during the Miocene. We thereby provide a good example of an aphid macroevolution study in which histories of biogeography and host association were inferred simultaneously and the effects of these two processes on aphid diversification were evaluated. In the future, global biogeography of more monophyletic aphid groups with a good taxonomic basis needs to be investigated from a historical point of view. Taking geological events, climate changes and palaeobotanical evidence into account is undoubtedly important to explore the origin and diversification of aphids.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov, reference numbers MW356901-MW356908, MW353804-MW353842.

ORCID

Tong-Yi Liu 🆔 https://orcid.org/0000-0002-6831-0537
Ge-Xia Qiao 🆔 https://orcid.org/0000-0002-7300-6812

REFERENCES

Blackman, R.L. & Eastop, V.F. (2020) Aphids on the world's plants: an online identification and information guide [WWW Document]. Available at: http://www.aphidsonworldsplants.info [Accessed 11th September 2020].

Bouchenak-Khelladi, Y., Verboom, G.A., Savolainen, V. & Hodkinson, T.R. (2010) Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Botanical Journal of the Linnean Society*, 162, 543–557.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D. et al. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10, e1003537.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A. et al. (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15, e1006650.

Bremer, K. (2002) Gondwanan evolution of the grass alliance of families (Poales). *Evolution*, 56, 1374–1387.

Chakrabarti, S. & Mandal, A.K. (1986) An account of *Trichaitophorus* and *Yamatochaitophorus* (Homoptera: Aphididae) with a note on their phylogeny. *Zoological Journal of the Linnean Society*, 88, 329–338.

Chen, R., Favret, C., Jiang, L., Wang, Z. & Qiao, G. (2016) An aphid lineage maintains a bark-feeding niche while switching to and diversifying on conifers. *Cladistics*, 32, 555–572.

Chen, J., Wang, Y., Jiang, L. & Qiao, G. (2017) Mitochondrial genome sequences effectively reveal deep branching events in aphids (Insecta: Hemiptera: Aphididae). *Zoologica Scripta*, 46, 706–717.

Clyde, W.C., Khan, I.H. & Gingerich, P.D. (2003) Stratigraphic response and mammalian dispersal during initial India-Asia collision: Evidence from the Ghazij Formation, Balochistan, Pakistan. *Geology*, 31, 1097–1100.

Collinson, M.E. (1992) The early fossil history of Salicaceae: a brief review. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, 98, 155–167.

von Dohlen, C.D. & Moran, N.A. (2000) Molecular data support a rapid radiation of aphids in the Cretaceous and multiple origins of host alternation. *Biological Journal of the Linnean Society*, 71, 689–717.

Favret, C. (2020) Aphid species file. Version 5.0/5.0 [WWW Document]. Available at: http://Aphid.SpeciesFile.org [Accessed 11th September 2020].

Favret, C. & Voegtlin, D.J. (2004) Speciation by host-switching in pinyon *Cinara* (Insecta: Hemiptera: Aphididae). *Molecular Phylogenetics and Evolution*, 32, 139–151.

Ghosh, A.K. (1980) *The Fauna of India and the adjacent countries. Homoptera, Aphidoidea, part I, Chaitophorinae*. Chennai: Amra Press.

Havill, N.P., Foottit, R.G. & von Dohlen, C.D. (2007) Evolution of host specialization in the Adelgidae (Insecta: Hemiptera) inferred from molecular phylogenetics. *Molecular Phylogenetics and Evolution*, 44, 357–370.

Heie, O.E. (1987) Palaeontology and phylogeny. In: Minks, A.K. & Harrewijn, P. (Eds.) *Aphids: their biology, natural enemies, and control*, Vol. 2A. Amsterdam: Elsevier, pp. 367–391.

Heie, O.E. (1995) An aphid from the Plio-Pleistocene Kobenhavn Formation, North Greenland (Homoptera, Aphidoidea). *Entomologiske Meddelelser*, 63, 17–18.

Heie, O.E. (1996) The evolutionary history of aphids and a hypothesis on the coevolution of aphids and plants. *Bollettino di Zoologia Agraria e di Bachicoltura, Series II*, 28, 149–155.

Heie, O.E. & Wegierek, P. (2011) A list of fossil aphids (Hemiptera, Sternorrhyncha, Aphidomorpha). *Monographs of the Upper Silesian Museum*, 6, 1–82.

Higuchi, H. (1972) A taxonomic study of the subfamily Callipterinae in Japan (Homoptera: Aphididae). *Insecta Matsumurana*, 35, 19–126.

Hsü, J. (1983) Late Cretaceous and Cenozoic vegetation in China, emphasizing their connections with North America. *Annals of the Missouri Botanical Garden*, 70, 490–508.

Hui, Z., Li, J., Xu, Q., Song, C., Zhang, J., Wu, F. et al. (2011) Miocene vegetation and climatic changes reconstructed from a sporopollen record of the Tianshui Basin, NE Tibetan Plateau. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 308, 373–382.

Huang, D., Wegierek, P., ŻYŁA, D. & Nel, A. (2015) The oldest aphid of the family Oviparosiphidae (Hemiptera: Aphidoidea) from the Middle Jurassic of China. *European Journal of Entomology*, 112, 187–192.

Jin, J. (2009) Two Eocene fossil fruits from the Changchang Basin of Hainan Island, China. *Review of Palaeobotany and Palynology*, 153, 150–152.

Katoh, K., Rozewicki, J. & Yamada, K.D. (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, 20, 1160–1166.

Kononova, E.L. (1978) New aphid species (Homoptera, Aphidinea) from Upper Cretaceous deposits of the Taymyr. *Entomological Review*, 56, 72–80.

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547–1549.

Kuraku, S., Zmasek, C.M., Nishimura, O. & Katoh, K. (2013) aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Research*, 41, W22–W28.

Landis, M.J., Matzke, N.J., Moore, B.R. & Huelsenbeck, J.P. (2013) Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, 62, 789–804.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773.

Li, J., Stukel, M., Bussies, P., Skinner, K., Lemmon, A.R., Lemmon, E.M. et al. (2019) Maple phylogeny and biogeography inferred from phylogenomic data. *Journal of Systematics and Evolution*, 57, 594–606.

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

Liu, Q.H., Chen, J., Huang, X.L., Jiang, L.Y. & Qiao, G.X. (2015) Ancient association with Fagaceae in the aphid tribe Greenideini (Hemiptera: Aphididae: Greenideinae). *Systematic Entomology*, 40, 230–241.

Matzke, N.J. (2013) Probabilistic historical biogeography: new models for founderevent speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, 5, 242–248.

Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in Island clades. *Systematic Biology*, 63, 951–970.

Matzke, N.J. (2018) *BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis with R scripts*. Version 1.1.1 [WWW Document]. Available at: https://github.com/nmatzke/BioGeoBEARS [Accessed 11th December 2019].

Meade, A. & Pagel, M. (2017) *BayesTraits V3.0.1* [WWW Document]. Available at: http://www.evolution.rdg.ac.uk/ [Accessed 11th December 2018].

Moore, B.R., Höhna, S., May, M.R., Rannala, B. & Huelsenbeck, J.P. (2016) Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 9569–9574.

Moran, N.A. (1992) The evolution of aphid life cycles. *Annual Review of Entomology*, 37, 321–348.

Moran, N.A., Kaplan, M.E., Gelsey, M.J., Murphy, T.G. & Scholes, E.A. (1999) Phylogenetics and evolution of the aphid genus *Uroleucon* based on mitochondrial and nuclear DNA sequences. *Systematic Entomology*, 24, 85–93.

Newberry, J.S. (1895) The flora of the Amboy Clays. *Monograph of the United States Geological Survey*, 26, 1–260.

Nováková, E., Hypša, V., Klein, J., Foottit, R.G., von Dohlen, C.D. & Moran, N.A. (2013) Reconstructing the phylogeny of aphids (Hemiptera: Aphididae) using DNA of the obligate symbiont *Buchnera aphidicola*. *Molecular Phylogenetics and Evolution*, 68, 42–54.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C. et al. (2001) Terrestrial ecoregions of the world: a new map of life on earth. *Bioscience*, 51, 933–938.

Ortiz-Rivas, B. & Martínez-Torres, D. (2010) Combination of molecular data support the existence of three main lineages in the phylogeny of aphids (Hemiptera: Aphididae) and the basal position of the subfamily Lachninae. *Molecular Phylogenetics and Evolution*, 55, 305–317.

Ortiz-Rivas, B., Moya, A. & Martínez-Torres, D. (2004) Molecular systematics of aphids (Homoptera: Aphididae): new insights from the long-wavelength opsin gene. *Molecular Phylogenetics and Evolution*, 30, 24–37.

Pagel, M. & Meade, A. (2006) Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist*, 167, 808–825.

Papasotiropoulos, V., Tsiamis, G., Papaioannou, C., Ioannidis, P., Klossa-Kilia, E., Papapanagiotou, A. et al. (2013) A molecular phylogenetic study of aphids (Hemiptera: Aphididae) based on mitochondrial DNA sequence analysis. *Journal of Biological Research (Thessaloniki)*, 20, 195–207.

Paradis, E., Claude, J. & Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289–290.

Parham, J.F., Donoghue, P.C.J., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P.A. et al. (2012) Best practices for justifying fossil calibrations. *Systematic Biology*, 61, 346–359.

Pintera, A. (1987) Taxonomic revision of the species of genus *Chaitophorus* Koch in Palaearctis (Homoptera: Aphidoidea). *Deutsche Entomologische Zeitschrift*, 34, 219–340.

Plummer, M., Best, N., Cowles, K. & Vines, K. (2006) CODA: convergence diagnosis and output analysis for MCMC. *R News*, 6, 7–11.

Pojàrkova, A.I. (1933) Botanico-geographical survey of the maples of the USSR in connection with the history of the whole genus *Acer* L. *Acta Instituti Botanici Academiae Scientiarum USSR, Series*, 1(1), 225–374.

Prasad, V., Strömberg, C.A.E., Alimohammadian, H. & Sahni, A. (2005) Dinosaur coprolites and the early evolution of grasses and grazers. *Science*, 310, 1177–1180.

Qiao, G.X. (1996) *Studies on the phylogeny of Hormaphididae, Thelaxidae, Chaitophoridae and Greenideidae etc. 7 families*. PhD Thesis, Chinese Academy of Sciences, Beijing.

Qiao, G.X., Jiang, L.Y. & Zhang, G.X. (2003) *Lambersaphis* Narzikulov, a newly recorded genus in China (Homoptera, Aphididae). *Acta Zootaxonomica Sinica*, 28, 499–501.

Quednau, F.W. & Chakrabarti, S. (1976) A new genus and three new species of aphids (Homoptera: Aphididae) from north-west India. *The Canadian Entomologist*, 108, 461–468.

Rabosky, D.L. (2014) Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One*, 9, e89543.

Rabosky, D.L., Donnellan, S.C., Grundler, M. & Lovette, I.J. (2014) Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. *Systematic Biology*, 63, 610–627.

Rabosky, D.L., Grundler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W. et al. (2014) BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 5, 701–707.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67, 901–904.

Rebijith, K.B., Asokan, R., Hande, H.R., Joshi, S., Surveswaran, S., Ramamurthy, V.V. et al. (2017) Reconstructing the macroevolutionary patterns of aphids (Hemiptera: Aphididae) using nuclear and mitochondrial DNA sequences. *Biological Journal of the Linnean Society*, 121, 796–814.

Ree, R.H., Moore, B.R., Webb, C.O. & Donoghue, M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, 59, 2299–2311.

Richards, W.R. (1972) The Chaitophorinae of Canada (Homoptera: Aphididae). *The Memoirs of the Entomological Society of Canada*, 104, 1–109.

Ronquist, F. (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, 46, 195–203.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S. et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539–542.

Sanmartín, I., Enghoff, H. & Ronquist, F. (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society*, 73, 345–390.

Schmidt, H.A., Strimmer, K., Vingron, M. & von Haeseler, A. (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, 18, 502–504.

Shimodaira, H. (2002) An approximately unbiased test of phylogenetic tree selection. *Systematic Biology*, 51, 492–508.

Shimodaira, H. & Hasegawa, M. (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, 17, 1246–1247.

Shingleton, A.W. & Stern, D.L. (2003) Molecular phylogenetic evidence for multiple gains or losses of ant mutualism within the aphid genus *Chaitophorus*. *Molecular Phylogenetics and Evolution*, 26, 26–35.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.

Strömberg, C.A.E. (2011) Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences*, 39, 517–544.

Szelegiewicz, H. (1961) Die polnischen Arten der Gattung *Chaitophorus* Koch s. lat. (Homoptera, Aphididae). *Annales Zoologici*, 19, 229–352.

Töpel, M., Zizka, A., Calió, M.F., Scharn, R., Silvestro, D. & Antonelli, A. (2017) SpeciesGeoCoder: fast categorization of species occurrences for analyses of biodiversity, biogeography, ecology, and evolution. *Systematic Biology*, 66, 145–151.

Vaidya, G., Lohman, D.J. & Meier, R. (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, 27, 171–180.

Wagenmakers, E.-J. & Farrell, S. (2004) AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11, 192–196.

Wegierek, P. (1996) Aphid species (Homoptera: Aphidinea) from the collection of the Baltic amber in the Museum of the Earth, Polish Academy of Sciences, Warsaw (Part two). *Prace Muzeum Ziemi*, 44, 25–36.

Wieczorek, K. (2010) *A monograph of Siphini Mordvilko, 1928 (Hemiptera, Aphidoidea: Chaitophorinae)*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Wieczorek, K. & Kajtoch, Ł. (2011) Relationships within Siphini (Hemiptera, Aphidoidea: Chaitophorinae) in light of molecular and morphological research. *Systematic Entomology*, 36, 164–174.

Wieczorek, K., Lachowska-Cierlik, D., Kajtoch, Ł. & Kanturski, M. (2017) The relationships within the Chaitophorinae and Drepanosiphinae (Hemiptera, Aphididae) inferred from molecular-based phylogeny and comprehensive morphological data. *PLoS One*, 12, e0173608.

Wolfe, J.A. & Tanai, T. (1987) Systematics, phylogeny, and distribution of *acer* (maples) in the Cenozoic of western North America. *Journal of the Faculty of Science of Hokkaido University, Series IV*, 22, 1–246.

Wu, Y., You, H.L. & Li, X.Q. (2018) Dinosaur-associated Poaceae epidermis and phytoliths from the early cretaceous of China. *National Science Review*, 5, 721–727.

Xu, T.Z. (1998) The systematic evolution and distribution of the genus *Acer*. *Acta Botanica Yunnanica*, 20, 383–393.

Zhang, G. & Zhong, T. (1982) New genera and new species of Chinese Callaphididae and Chaitophoridae. *Acta Zootaxonomica Sinica*, 7, 67–77.

Zhang, G., Chen, X., Qiao, G., Zhong, T. & Li, J. (1999) *Fauna of agricultural and forestry aphids of Northwest China (Insecta: Homoptera: Aphidinea)*. Beijing: China Environmental Science Press.

Zhu, H. & Manchester, S.R. (2020) Red and silver maples in the Neogene of western North America: fossil leaves and samaras of *Acer* section *Rubra*. *International Journal of Plant Sciences*, 181, 542–556.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Table S1. Voucher information and GenBank accession numbers of aphid species used in this study. * indicates sequences newly generated in the present study.

Table S2. The information of sequence characters of each gene.

Table S3. Best-fit partition schemes and models of molecular evolution.

Table S4. Dispersal rates assigned for each time slice in BioGeoBEARS. NA, Nearctic; WP, West Palaearctic; EP, East Palaearctic; NT, Neotropical; OL, Oriental.

Table S5. Statistical testing of particular phylogenetic hypotheses. Obs, observed log-likelihood difference to the best topology; AU, approximately unbiased test. * indicates that the hypothesis received a *p*-value <0.05 and can be rejected.

Figure S1. Maximum clade credibility tree of Chaitophorinae resulting from the BEAST analysis. Red circles at nodes mark calibration points. Blue bars indicate 95% high posterior density intervals of the age estimates. Mean ages of some nodes are shown above the bars. P, Pliocene; Q, Quaternary.

Figure S2. Biogeographical scenario for Chaitophorinae based on the BAYAREALIKE model. The pie charts on each node represent the most likely areas. Letters in box represent the geographical regions (see key in Figure 2).