東海大學生命科學系
碩士論文

指導教授：林仲平 博士

Dr. Chung-Ping Lin

島嶼型豆娘的種化模式

Speciation with Gene Flow in Island Damselflies

研究生：李逸鴻

Yat-Hung Lee

中華民國一百年七月

東海大學生命科學系碩士論文

島嶼型豆娘的種化模式

Speciation with Gene Flow in Island Damselflies

研究生：李逸鴻

Yat-Hung Lee

指導教授：林仲平 博士

Dr. Chung-Ping Lin

中華民國一百年七月

東海大學生命科學系
碩士論文學位考試審定書

生命科學系碩士班研究生 李逸鴻 君所撰寫之論文

（中文）

島嶼型豆娘的種化模式

（英文）

Speciation with Gene Flow in Island Damselflies

經本委員會審定通過，特此證明。

學位考試委員會　召集人 李壽先 （簽名）

委　員 林仲平

王弘毅

林思民

中華民國 100 年 7 月 22 日

致謝

能完成這篇論文，首要感謝林仲平老師在實驗研究與論文撰寫上的細心指導。以及特地前來幫我口試的各個委員，師大的李壽先老師與林思民老師、台大的王弘毅老師，謝謝您們的批評與建議。在野外的部分，感謝 Hidetoshi Ota 老師在我去石垣島時，熱情協助我的採集工作。也感謝所有在國外提供我們豆娘標本的採集者。在實驗室的分析部分，謝謝本系的邱偉欣老師與台大的王弘毅老師在數據分析上提供的協助，另外特別要感謝本校資工系的楊朝棟老師，楊老師非常大方的提供多台高速電腦，讓我得以分析重要的數據，沒有他的協助，這些分析絕對無法完成。

回顧這段碩士生涯，受到許多實驗室成員與朋友的照顧。謝謝黃仁磐學長，堪稱我們實驗室的智庫，只要分析軟體上有甚麼問題，問他就對了。謝謝紹彰學長、鎮盤學長、若凡學姊、明玉學姊、薇芸學姊提供我在實驗上的協助，還會陪我漫天閒聊，消憂解愁。謝謝許多生科系的同學與學弟妹，小關、宏年、小鑫、阿山、光佑、柏翰、政修、怡明、婉萍、侯玨、坤彰、泰申、玉婷、詩婷、廣瑜、小ａ、瑋良，這段時間我們在學業上互相鼓勵支持，在生活上一起同樂同遊，在你們身上我學到了很多課本沒有的東西，感謝你們的陪伴。另外遠在師大的書萍，與我一同解決研究上的難題，讓我發現了很多實驗分析上的盲點，謝謝你。

感謝佩瑜，一個鬼靈精怪的可愛女生，這段時間，你辛苦了，不管在課業還是生活上，沒有你的協助與幫忙，我一個人無法完成，你總是能在我最失意時幫我加油打氣，在我最需要人陪伴時，跟我聊天解悶，陪我哭陪我笑，陪我發瘋（這

你懂），這段時光……有你真好。最後要感謝我的家人，爸爸、媽媽、哥哥、外公、外婆、阿芳阿姨，你們總是無怨無悔的支持我，幫助我，照顧我，再多的感激也無法回報您們對我的愛。我也愛您們。接下來的我，即將邁入人生另一個階段，還有很多挑戰，很多驚喜等我去發現，希望能繼續帶著你們的祝福，一路過關斬將!!!

Contents

Abstract (in Chinese)

地理隔離在過去一直被認為是推動種化的主要力量。異域種化模式強調地理屏障和阻斷基因交流對於族群分化的重要性。在異域種化的模式下，基因交流被認為是種化的阻礙。對於過去有多少基因交流發生在分化中的野生族群或物種目前仍然不明。本研究我們利用兩個粒線體和十個核基因來估算兩對姊妹種幽螽屬豆娘—*E. formosa* + *E. yayeyamana* 以及 *E. decorata* + *E. ornata*—在過去種化進行時的基因交流程度。根據 *cox2* 和 *arr* 基因所重建的親緣關係顯示 *E. formosa* + *E. yayeyamana* 以及 *E. decorata* + *E. ornata* 為支持度高的姐妹種。在 *E. formosa* + *E. yayeyamana* 和 *E. decorata* + *E. ornata* 這兩對的多基因座分析結果皆棄卻狹義的地理隔離模式。在 *E. formosa* 和 *E. yayeyamana* 之間，我們發現中等至大量的雙向基因交流，但在 *E. decorata* 和 *E. ornata* 之間沒有太多的基因交流證據。*E. decorata* 和 *E. ornata* 的分化時間估計約為 0.511 個百萬年前，這個時間晚於 *E. formosa* 和 *E. yayeyamana* 的分化時間 (1.145 個百萬年前)。分析結果指出有基因交流的種化模式，最能解釋在 *E. formosa* 和 *E. yayeyamana* 族群內所觀察到的 DNA 序列變異，而沒有基因交流的異域種化模式則較適合描述 *E. decorata* 和 *E. ornata* 的分化。

關鍵字：異域種化、島嶼特有種、基因交流的種化模式、幽螽屬、幽螽科、蜻蛉目

Abstract

Geographic isolation has been proposed as a major force in speciation. Allopatric mode of speciation emphasizes the prominent role of physical barriers and restriction of gene flow on population divergence. Under allopatric model, gene flow is considered as an impediment for speciation. The question of how much historical gene flow has occurred in diverged natural populations and species is largely unknown. In this study, we investigated the level of historical gene flow during the speciating process of two sibling species pairs of *Euphaea* damselflies, *E. formosa* + *E. yayeyamana* and *E. decorata* + *E. ornata*, using two mitochondrial and ten unclear loci. The reconstructed species phylogeny based on *cox2* and *arr* genes indicated that *E. formosa* + *E. yayeyamana*, and *E. decorata* + *E. ornata*, are both valid sister species pairs. The results of multilocus analyses rejected the strict isolation model in *E. formosa* and *E. yayeyamana*, and *E. decorata* and *E. ornata.* Moderate to large two directional gene flows were detected between *E. formosa* and *E. yayeyamana*, but there is little evidence of gene flow between *E. decorata* and *E. ornata.* The divergence time of *E. decorata* and *E. ornata* was estimated at approximately 0.511 Mya, which was more recent than the split of *E. formosa* and *E. yayeyamana* (1.145 Mya). We concluded that the model of speciation with gene flow best describe the observed sequence variation in *E. formosa* and *E. yayeyamana*, whereas the model of

allopatric speciation without gene flow is more appropriate for *E. decorata* and *E. ornata*.

Key words: Allopatric speciation, island endemics, isolation with migration, *Euphaea*, Euphaeidae, Odonata

Introduction

Allopatric model has been proposed as a dominant mechanism of speciation in natural populations (Mayr 1963; Coyne & Orr 2004). It hypothesized that the emergence of physical barriers completely disrupt the gene flow between populations, the separated populations start accumulating genetic variation following initial geographical isolation, and eventually the diverging population become different species (Coyne & Orr 2004). The allopatric mode of speciation was based on the "biological species concept" (Mayr 1963), which in essence considered the whole genome of an individual as the unit of speciation. Wu (2001) proposed the "porous model" of speciation which treats individual genes and the portion of the genome as the unit of speciation. This genic view of speciation allows a portion of genes and the genome continually exchange between diverging populations (Osada & Wu 2005). But for genes that are differentially adapted and targeted by selection, they are not exchangeable due to the detrimental effect in reproduction (Wu & Ting 2004). Until such fixed genomic regions expand gradually by linkage or hitchhiking, two diverging populations then proceed to the final stage of becoming distinct species (Wu & Ting 2004). Under this genic view of speciation, studying mechanisms of speciation in natural populations focused on estimating the amount of historical gene flows among divergent populations and species (Osada & Wu 2005; Turner *et al*. 2005; Niemiller *et*

al. 2008; Nadachowska & Babik 2009; Li *et al*. 2010; Yeung *et al*. 2010; Zheng & Ge 2010; Wachowiak *et al*. 2011).

Wakeley and Hey (1997) proposed the "isolation model" of speciation in which the fixed genetic differences were accumulated and the shared genetic variations were reduced by genetic drift through time during speciation. Under this model, genetic drift plays a major role in determining the level of shared and fixed genetic variation among diverging populations (Gavrilets 2003). This model can be used as a null hypothesis to assess the validity of a "strict" allopatric speciation model. Nielsen and Wakeley (2001) later modified this isolation model into a model of isolation with migration (IM) by adding the parameter of gene flow. Using the IM model, the coalescent process of multiple loci can be modeled in consideration of all possible genealogies to estimate the important population genetic parameters including effective population size, migration rate and divergence time (Hey & Nielsen 2007).

The role of gene flow in speciation remains contentious among evolutionary biologists (Coyne & Orr 2004). Recent empirical evidences for the model of speciation with gene flow were accumulating and have been found in natural populations of a few "non-model" vertebrates, including chimpanzees (*Pan troglodytes troglodytes*, *P. t. verus* and *P. paniscus*, Won & Hey 2005), gorillas (*Gorilla gorilla* and *G. beringei*, Thalmann *et a*l. 2007), lizards (*Sceloporus magister***,**

Leaché & Mulcahy 2007), salamanders (*Gyrinophilus palleucus palleucus*, *G. p. necturoides* and *G. gulolineatus*, Niemiller *et al*. 2008; *Lissotriton vulgaris kosswigi* and *L. v. vulgaris*, Nadachowska & Babik 2009), threespine stickleback fish (*Gasterosteus aculeatus* species group, Berner *et al*. 2009) and birds (*Anas strepera* and *A. falcate*, Peters *et al*. 2007; *Manacus aurantiacus*, *M. candei*, *M. vitellinus* and *M. manacus*, Brumfield *et al*. 2008; *Leucodioptron taewanus* and *L. canorum canorum*, Li *et al*. 2010; *Platalea regia* and *P. minor*, Yeung *et al*. 2010). These studies demonstrated detectable historical gene flow in the diverging process of natural populations and species. Studies of insects such as fruit flies (*Drosophila pseudoobscura* and *D. persimilis*, Hey & Nielsen 2004), butterflies (*Heliconius cydno* and *H. melpomene*, Bull *et al*. 2006; *Heliconius erato*, *H. e. chestertonii* and *H. himera*, Arias *et al*. 2008) and grasshoppers (*Melanoplus oregonensis* and *M. montanus*, Carsten & Knowles 2007) revealed substantial amount of migration among diverging species that support the model of speciation with gene flow. In a genome-wide study of the M and S forms of the malaria mosquito, *Anopheles gambiae*, the sympatric populations share 97% of genetic variation but maintain differences at the genomic region of low rates of crossing over near the centromeres (speciation island or continents) (Turner *et al*. 2005; Lawniczak *et al*. 2010; Michel *et al*. 2010; White *et al*. 2010).

The gossamer-wing damselfly genus, *Euphaea*, contains thirty species distributing widely across Southeast Asia. The larvae of *Euphaea* damselflies inhabit the substrate of forest streams with strong current (Hayashi 1990). The adult male display aggressive territorial behavior in defending female's ovipositional sites of submerged rocks and vegetation (Dudgeon 1989; Huang & Lin 2011). *Euphaea formosa*, *E. yayeyamana*, *E. decorata* and *E. ornata* are four morphologically similar species in which the males have black bands on the hindwings and transparent forwings (Fig. 2). *E. formosa* and *E. yayeyamana* are endemic to Taiwan, Yaeyama islands, respectively, whereas *E. decorata* is widely distributed in southeast Asian mainland and *E. ornata* was only found in Hainan Island (Fig. 1). *E. formosa* and *E. yayeyamana* are estimated to have begun diverging approximately 2.6 Mya during the late Pliocene (Huang & Lin 2011). Although there is no information on the degree of reproductive isolation between species of these two pairs, the current disjunct distributions of these species separated by the oceans suggest low level of hybridization. These oceanic barriers have been considered as efficient isolation mechanisms for extant populations of these damselflies (Hayashi 1990; Huang & Lin 2011). However, current geographic range of a species often does not reflect its historical distribution, because the oceanic barrier may not consistently present throughout species's history (Losos & Glor 2003). Taiwan, Hainan and the adjacent

continental islands of Southeast Asia have experienced sea level changes during the recent glacial cycles (Voris 2000). The land bridges between the Asian continent and these islands were repeatedly formed and submerged during the glaciated periods (Kimura 2000; Voris 2000). Taiwan, Ryukyu Arc and the coastal regions including Hainan of Southeastern China merged together when the sea level was low. These glacial land bridges allowed intermittent secondary contacts between geographically isolated populations over the past few million years (Xu *et al.* 2009; Huang & Lin 2011). Therefore, a model of speciation with gene flow is more appropriate than a strict geographical isolation model to describe the diverging process of these endemic damselflies on continental islands.

The aims of this study are to estimate the amount of historical gene flow and to identify the geographic mode of speciation between two pairs of *Euphaea* sibling damselflies. We used two mitochondrial and ten nuclear loci to test the null hypothesis of a strict allopatric speciation in two pairs of sister species. The detection of substantial amount of historical gene flow during the speciation process of these four damselfly species was used as evidence in supporting the model of speciation with gene flow.

Materials and methods

Sampling

Damselfly specimens of the four *Euphaea* species were collected from five localities in Taiwan, three in Yaeyama island, two in Hainan, three in Hong Kong and two in Vietnam (Table 1). Representative specimens for each species were collected from at least two localities within these islands. The specimens of additional eight *Euphaea* species were obtained from various localities of Southeast Asia and used as outgroups in phylogenetic analyses (Table 1). All collected specimens were immersed in 95% EtOH immediately after capturing and later stored at –80°C until DNA extraction.

DNA extraction and sequencing

The muscle tissues from the thorax or hind femur of the specimen were dissected and used for extracting genomic DNA with MasterPure™ Complete DNA and RNA Purification Kit (EPICENTRE, Wisconsin, USA). Two mitochondrial (*cox2* and *nad5*), the ribosomal internal transcribed spacer (*ITS*), eight nuclear protein-coding including actin (*act*), arrestin 2 (*arr*), abnormal wing discs 2 (*awd2*), elongation factor 1alpha (*EF1α*), ferritin (*fer*), myosin light chain (*mlc*), long wavelength opsin 1 (*lop1*) and succinate dehydrogenase B (*sdhB*), and one anonymous (*anon*) genes were amplified using newly designed primers based on whole genome sequencings of *E. formosa* and

E. yayeyamana (Lee & Lin unpublished) and other published primers (Weekers *et al.* 2001; Jordan *et al.* 2003; Huang & Lin 2011) (Table 2). The polymerase chain reaction (PCR) in reaction volumes of 50μl contained 100–300 ng of template DNA, 0.4μM of forward and reverse primers, 0.2mM dNTP, 10mM Tris-HCl, pH 9.0, 50mM KCl, 1.5mM $MgCl_2$ and 0.04U Pro Taq polymerase (Protech Technology, Taiwan). After 94°C initial denaturation for three minutes, the PCR reaction was set to repeat 35 cycles of denaturation at 94°C for one minute, followed by an annealing step for one minute at the optimal annealing temperature (Tm) (Table 2), and then a DNA extension step at 72°C for one minute. Final extension step was set to 72°C for 10 minutes. Most PCR products contained multiple DNA bands on the gel electrophoresis except for *cox2*. The target PCR products with desired length were cut off from the agarose gel and then extracted. Some of PCR products (*arr2*, *fer*, *mlc* and *lop1*) were cloned into DH-5α competent cells (Protech Technology, Taiwan) using TOPO TA Cloning Kit (Topoisomerase I mediated TA cloning) (Invitrogen, Carlsbad, USA). We randomly chose at least three positive clones for DNA sequencing from both directions on an ABI PRISM^{TM} 377 automatic sequencer (Perkin Elmer, USA) at the Mission Biotech, Taiwan. DNA sequences were manually edited in EditSeq, aligned using Clustal W algorithm of MegAlign (DNASTAR Lasergene 7 software, Madison, USA), and searched and annotated in the GeneBank (NCBI,

http://www.ncbi.nlm.nih.gov) using Basic Local Alignment Search Tool (BLAST). For DNA sequences with heterozygosity, we used the PHASE algorithm (Stephens *et al.* 2001; Stephens & Donelly 2003) implemented in DnaSP (v. 5, Librado & Rozas 2009) to reconstruct haplotypes from population genotype data using Maximum Likelihood (ML) method. After discarding 1000 steps as burn-ins, we set the MCMC iterations to 100,000 and thinned every 100 intervals to reconstruct unphased data set of each heterozygous locus. The DNA sequences used in this study were deposited in GenBank (#~#). The phylogenetic trees, gene trees and the associated sequence alignments can be found at TreeBASE (#).

Phylogenetic reconstruction

Maximum Parsimony (MP) and Bayesian Inference (BI) were used to reconstruct the species phylogeny of the *Euphaea* based on a combined *cox2* and *arr* data set. For MP analyses, we used parsimony ratchet algorithm (Nixon 1999) implemented in PAUPRat (v. 1, Sikes & Lewis 2001) and run the tree search in PAUP* (v. 4.0b10, Swofford 2002) to obtain the most parsimonious trees. 20 independent analyses with the default setting of 200 replicates were run. Parsimony branch supports were calculated in PAUP* using the tree bisection and reconnection (TBR) branch swapping and 1000 bootstrap replicates, each containing 100 iterations of random

sequence addition. For BI analyses, we used MrBayes (v. 3.12, Huelsenbeck & Ronquist 2001) to calculate the posterior probability of trees. The substitution models and prior values of the analyses were estimated using Bayesian Information Criterion (BIC) in jModeltest (v. 0.1.1, Posada 2008). Two independent Bayesian analyses, each containing four independent chains, were run. The MCMC chains were run for 2×10^8 generations with trees being sampled for every 2×10^4 generations. The MCMC processes were run until the average standard deviation of split frequencies falling below 0.01 for the convergence of two separate runs. The 2500 trees sampled before reaching stationary plateau were discarded as burn-ins. The Bayesian posterior probability (BPP) of the tree branches was calculated by constructing a 50% majority consensus tree in PAUP*. A species phylogeny was also estimated from the jointed posterior probability of *cox2* and *arr* gene trees using *BEAST (v. 1.6, Heled & Drummond 2010). The best fitted model of substitution was chosen using jModeltest and the two genes were set to be unlinked. A random local clock model was used for the analyses. The species tree prior was set to the default of Yule process. Two independent runs of 1×10^8 generations were performed using four Markov chains for each run with trees being sampled for every 1×10^4 generations. The prior values and model settings given in BEAUTi (v. 1.6, Heled & Drummond 2010) were used to generate XML files for *BEAST analyses. After discarding the first 1×10^7

generations (10%) as burn-ins, the Tracer (v. 1.5, Drummond & Rambaut 2007) was used to determine the convergence of runs where the ESS (effective sample size) of parameters were higher than 100. The resulting trees were summarized using TreeAnnotator (v. 1.6, Drummond & Rambaut 2007). The individual gene trees of the twelve loci were reconstructed using the haplotype data sets in MrBayes with substitution models and prior values estimated by jModeltest. Two independent analyses, each containing four independent chains, were run for 1×10^7 generation with sampling frequency of 1×10^3 generation. The first 2500 trees (25%) were discarded as burn-ins.

Summary statistics

We calculated summary statistics of the sampled populations, including polymorphic site (*S*), haplotype diversity (H_d, Nei 1987) and nucleotide diversity (π, the average number of nucleotide differences per site between two sequences, Nei 1987) using DnaSP. The net genetic distance per site per locus between the species were calculted under the Tamura-Nei substitution model (Tamura & Nei 1993) in MEGA 5 (Tamura *et al*. 2011) to measure the level of interspecific differentiation within the sister pairs. The population migration rates of each locus were calculated using the method of Hudson *et al*. (1992) implemented in DnaSP. For autosomal loci, the minimum

number of recombination events (*Rm*) were estimated based on the four-gamete test (Hudson & Kaplan 1985). We used IMgc (Woerner *et al.* 2007) to extract non-recombining blocks for loci with detectable recombination events.

Testing the isolation model

The isolation model in WH program was used to test the fit of observed sequence data to a strict allopatric speciation model (Wakeley & Hey 1997; Wang *et al*. 1997). We calculated four types of polymorphism (S_{x1} and S_{x2}: exclusive polymorphisms in species 1 and species 2, respectively; S_s: shared polymorphisms between species; S_f: fixed differences between species) and estimated population recombination rate ($4Nc$, where N is effective population size and c is recombination rate per generation, Hey & Wakeley 1997) in SITES as the input file for WH. Three parameters were estimated in WH, including effective population size in mutational units (θ) of descendent and ancestral species, and the divergence time (T) in unit of $2N_1$ generation (N_1 is the effective population size in demographic unit of species 1) (Wang *et al.* 1997). The WWH (which measured the variance of shared and fixed polymorphism among loci, Wang *et al*. 1997) and the chi-square test (which measured the deviation of exclusive, shared and fixed polymorphism between the observed and simulated values, Kliman *et al*. 2000) were used to evaluate the level of model fitting. The expected statistics

were obtained by the coalescent simulation of 1×10^4 generations in WH. The significance of the deviation between observed and simulated values was also tested in the linkage disequilibrium (LD) test (Machado *et al*. 2002), in which the value of DSS (the average of LD among all site pairs of shared polymorphisms) minus DSX (the average of LD among all site pairs in which one site is shared and the other is exclusive) was calculated. If the gene flow occurred, the DSS should be positive and DSX should be negative. Thus, the value of DSS minus DSX will be larger than the value under the scenario of no gene flow (Machado *et al*. 2002). To account for the influence of heterogeneity in allele frequency, we used *D* (The basic gametic measurement of LD, *D*, divided by the maximum possible value of *D*, D_{max}) in the LD test (Machado *et al*. 2002).

Testing the assumptions of the isolation with migration model

Stochastic process such as incomplete lineage sorting may result in different topologies between the gene tree and species tree even without gene flow (Edwards & Beerli 2000; Nichols 2001). Therefore, testing the fit of the isolation model is only the first step to examine the mechanism of speciation process. We further examined the level of historical gene flow between sister *Euphaea* species using the IM model. The IM model assumed that diverging populations continuously exchange gene over the

process of speciation. We used IMa2 program (Hey & Nielsen 2007; Hey 2010) to examine the fit between the observed data to the isolation with migration model. The speciation model in IMa2 has four assumptions (Pinho & Hey 2010): 1) The history of sampled populations (or species) can be fitted by an IM model reasonably, 2) No selection force act on the sampled loci, 3) No recombination within sampled loci and free recombination among them, and 4) No mis-application of the mutation model (i.e., infinite site or HKY model for non-microsatellite sequences) for sampled loci (Pinho & Hey 2010). We specifically tested the assumption 2, 3 and 4 using various statistical analyses. Tajima's *D* (Tajima 1989a) and Fay and Wu's *H* (Fay & Wu 2000) for each locus were calculated in DnaSP to test the neutrality of sampled loci. The parameter estimates from these two neutrality tests were used to perform 1000 coalescent simulations to test if the results significantly deviated from the neutral expectation by simultaneously considering the estimates of population recombination rate, *R* (equal to 4*Nr*, where *N* is population size and *r* is the recombination rate per gene, Hudson 1987). The multilocus HKA test (Hudson *et al*. 1987) was conducted to test the deviation from neutrality by comparing the distribution between observed and expected values using 10000 coalescent simulations in HKA program (http://genfaculty.rutgers.edu/hey/software#HKA). The summary statistics for HKA program were calculated in SITES

(http://genfaculty.rutgers.edu/hey/software#SITES). The clustering analyses implemented in STRUCTURE (v. 2.3.3, Pritchard *et al.* 2000) were used to examine the assumption of no population structure of the sampled populations. The data set contained haplotype assignment for each locus with each individual treated as having a single allele. The heterozygous sequences were unphased in DnaSP and randomly chosen for the analysis. An admixture model was used to estimate which parts of the genome of an individual coming from different populations (Hubisz *et al*. 2009). An independent model was used to calculate allele frequencies in which the allele frequency of each population is independently drawn from the prior distribution of the allele frequencies (Pritchard *et al*. 2000). The number of populations (*K*) in each species was set to range from one to six. Twenty independent MCMC runs of 1×10^6 steps were performed for each number of *K* with the first 1×10^5 steps discarded as burn-ins. The log likelihood values were monitored during the runs to ensure the convergence of MCMC process. For nucleotide substitution models, we evaluated the assumption of the Infinite Sites (IS) model that there is no more than one variant at a polymorphic site (Kimura 1969). We applied the IS model to *ITS* and *awd2* in *E. formosa* and *E. yayeyamana* and to *act* and *sdhB* in *E. decorata* and *E. ornata*. For the remaining loci violating the assumption of the IS model, we applied the HKY model (Hasegawa *et al.* 1985) to them in the IMa2 analyses.

Testing the isolation with migration model

The six population genetic parameters were estimated in IMa2, including the effective population sizes of two descendants (θ_1 and θ_2) and one ancestral populations (θ_A), two directional migration rates (m_1 and m_2) and divergence time (t). The MCMC and load-genealogies mode were performed in IMa2 (Hey & Nielsen 2007). The MCMC mode applied MCMC simulations to sample the probable genealogies. The load-genealogies mode used the genealogies saved in MCMC mode to estimate the jointed posterior densities of the parameters and to perform the likelihood-ratio tests of nested speciation models (Hey & Nielsen 2007). Three short MCMC runs were carried out in MCMC mode to obtain initial estimates of suitable ranges of prior settings, followed by three independent long runs to estimate the parameters. To obtain a better mixing property, we applied geometric heating scheme under the Metropolis-couplings criterion (Geyer 1991). Twenty Metropolis-coupled chains were run simultaneously as suggested in the IMa2 manual for the medium sized data set (< 15 loci). The β term (the degree of heat and the swapping rate between chains) of the geometric heating scheme was adjusted to be between 0.890~0.997 to obtain a moderate heating but high swapping rate for the multiple chains (*h1*=0.96, *h2*=0.9, *h1* specifies the degree of non-linearity and *h2* defines the

lowest value of β). The long runs were restarted when the swapping rate between chain 0 and 1 was lower than 5%. Before the MCMC chains of long runs reaching the plateau, all samples were considered as burn-ins. The duration of the long runs were set to be continuous and the runs were monitored for the values of effective sample size (ESS), the trend-line plot of the splitting time parameter and the swapping rate between successive MCMC chains to ensure proper mixing and convergence. After the completion of long runs, a total of 3×10^5 genealogies were saved and used to calculate the parameter values and the likelihood ratio tests of nested models in the load-genealogies mode.

Estimating substitution rate

A revised universal molecular clock of *cox1* for insects (1.77×10^{-8} per site per year, Papadopoulou *et al*. 2010) and a synonymous substitution rate of *arr* in *Drosophila* (1.2×10^{-8} per site per year, Moriyama & Gojobori 1992) were used to calculate the population size and divergence time in demographic units for mitochondrial and nuclear genes, respectively. The substitution rates of remaining genes were calculated using the relative ratios of net genetic distances between *cox1* and the two mitochondrial genes, and between *arr* and nine nuclear genes. The substitution rates were calculated separately for the two species pairs. The geometric

means of substitution rates from all loci in the two sister pairs were calculated separately and used to convert the population genetic parameters ($\theta = 4N\mu$ and $t = T\mu$) into demographic units (N, number of individuals and T, years) (Hey & Nielsen 2004). The four *Euphaea* damselflies were assumed to have one to two generations per year based on the available field observations (Dudgeon 1989; Hayashi 1990; Huang & Lin 2011). Here, we use one generation per year of these four damselflies to convert parameters into demographic units.

Results

Species phylogeny and gene trees

A sequence alignment of 2124 bp (*cox2*, 705 bp; *arr*, 1419 bp) was obtained for eleven *Euphaea* species. There were 170 and 153 parsimoniously informative sites in *cox2* and *arr*, respectively. A total of 4020 equally parsimonious trees (1390 steps) were found. TPM2uf + Γ (*cox2*, –lnL = 4923.9660) and HKY + Γ (*arr*, –lnL = 10519.3720) were selected as the best-fitted models of sequence evolution for BI. The MP and Bayesian trees recovered the sister relationships among the four *Euphaea* species (Fig. 2). The two sister groups, *E. formosa* + *E. yayeyamana* and *E. decorata* + *E. ornata*, were found in all analyses with high branch supports (Fig. 2). The estimated divergence time from *BEAST analyses was 0.30 Mya for *E. formosa* vs. *E.*

yayeyamana and 0.09 Mya for *E. decorata* vs. *E. ornata* with a wide 95% credible intervals (Fig. 2). For individual gene trees, *cox2*, *nad5*, *ITS*, *arr*, *fer* and *sdhB* indicated reciprocal monophyly of the four *Euphaea* species. Topology of the remaining gene tree (*act*, *awd2*, *EF1a*, *mlc*, *lop1* and *anon*) are not monophy suggesting incomplete lineage sorting or historical gene flow (Fig. 3).

Summary statistics and model assumptions

A total of 14 to 31 sequences per species were amplified for the twelve loci (Table 3). The overall haplotype diversities (H_d) were similar among the four species, whereas the overall nucleotide diversity (π) was higher in *E. formosa* than the other three species. The net genetic distance between *E. formosa* and *E. yayeyamana* (0.014 $\pm$ 0.017) was approximately two times greater than that between *E. decorata* and *E. ornata* (0.007 $\pm$ 0.006) (Table 4). For mitochondrial loci, the net genetic distance between *E. formosa* and *E. yayeyamana* (*cox2*, 0.047; *nad5*, 0.051) were approximately five to nine times higher than that of between *E. decorata* and *E. ornata* (*cox2*, 0.005; *nad5*, 0.011) (Table 4). The population migration rates indicated different levels of gene flow among individual loci and ranged from 0.03 to 3.86 in *E. formosa* vs. *E. yayeyamana* and from 0 to 2.37 in *E. decorata* vs. *E. ornata* (Table 4). Apparently large number of recombination events (> 5) were detected in *mlc*, *lop1*

and *anon* of *E. formosa*, *arr* of *E. yayeyamana*, *arr*, *fer* and *anon* of *E. decorata*, and *arr* and *lop1* of *E. ornata*. We then extracted maximum non-recombined regions (7780 bp in *E. formosa* and *E. yayeyamana*, 7921 bp in *E. decorata* and *E. ornata*) from all loci and used them in additional IMa2 analyses. The mutlilocus HKA tests demonstrated no significant departure from neutral expectation for all twelve loci (*E. formosa* and *E. yayeyamana*, sum of deviations = 25.118, df = 22, $p = 0.291$; *E. decorata* and *E. ornata*, sum of deviations = 11.294, df = 22, $p = 0.971$). The mutlilocus Tajima's *D*s were not significantly different from neutral expectation (*E. formosa*, D = –0.462, $p = 0.097$; *E. yayeyamana*, $D = 0.097$, $p = 0.660$; *E. decorata*, $D = 0.006$, $p = 0.640$; *E. ornata*, $D = –0.451$, $p = 0.077$). Tajima's *D* tests of the single locus demonstrated that most loci had negative values of *D*, but they were not significantly different from zero except for *nad5* in *E. formosa* ($D = –2.192$, $p < 0.001$) and *E. ornata* ($D = –1.839$, $p = 0.006$), and *ITS* in *E. ornata* ($D = –1.999$, $p = 0.001$) (Table 3). Fay and Wu's *H* tests indicated that significant departure from neutrality was found in *nad5* of *E. formosa* and *E. yayeyamana*, *fer* of *E. yayeyamana*, *mlc* of *E. formosa*, *E. yayeyamana* and *E. decorata*, and *anon* of *E. ornata* (Table 3). When all intron regions within *mlc* were exclude, Fay and Wu's *H* showed no significant departure from neutrality in four species. Significantly negative values of Tajima's *D* were caused by positive selection, purifying selection or recent population expansion

(Tajima 1989b). Fay and Wu's *H* compares the intermediate and high-frequency variants, which is less sensitive to demographic fluctuation such as rapid population growth and can be used to distinguish positive selection from purifying selection which is expected to have similar abundance between intermediate and high-frequency variants (Fu 1997). Therefore, we performed additional WWH test, chi-square test, LD test and IMa2 analyses with the exclution of *nad5*, *fer* and *mlc* of *E. formosa* and *E. yayeyamana*, and *mlc* and *anon* of *E. decorata* and *E. ornata*, which may have been under recent positive selection as indicated by Fay and Wu's *H* (6645 bp in *E. formosa* and *E. yayeyamana*; 8396 bp in *E. decorata* and *E. ornata*). Additional data sets without recombined blocks and non-neutral loci were analyzed in sister pairs by IMa2. For testing the existence of population structure, the highest probability of the model was found when *K* was equal to one (Table 5). The probabilities of grouping individuals in each populations were symmetrically equal to 1/*K* for all individuals. These results indicated no population structure within samples of the four *Euphaea* species.

Testing the isolation model

The different loci appear to have different evolutionary histories as indicated by the variation in the number of shared polymorphisms and fixed differences (Table 6).

Nad5, *awd2*, *EFlα*, *mlc*, *lop1* and *anon* have no fixed differences, but *mlc* and *lop1* do reveal a large number of shared polymorphism between *E. formosa* and *E. yayeyamana*. In contrast, *cox2* and *ITS* have large numbers of fixed differences (> 13) but very little shared polymorphism between *E. formosa* and *E. yayeyamana*. Between *E. decorata* and *E. ornata*, most loci have none or small number of fixed differences except for *nad5*, *ITS* and *arr*. Only *arr*, *awd2*, *fer* and *anon* have a larger number (> 11) of shared polymorphism between *E. decorata* and *E. ornata.* The population size parameter ($\theta = 4N\mu$) estimated by WH showed that the size of extant *E. formosa* (123.0) and *E. decorata* (81.5) were larger than their sister species, *E. yayeyamana* (52.2) and *E. ornata* (68.9), respectively (Table 7). The estimated population size of ancestral species (θ_A) in both sister pairs (292.1 for *E. formosa* + *E. yayeyamana*, 189.5 for *E. decorata* + *E. ornata*) were larger than that of their descendent species (Table 7). The estimated divergence time of *E. formosa* and *E. yayeyamana* (0.20) were more recent than the split of *E. decorata* and *E. ornata* (0.35). The chi-square statistics between the observed and expected values from the isolation model were significantly large (*E. formosa* and *E. yayeyamana*, $\chi^2 = 711.11$, $P_{\chi 2} = 0.007$; *E. decorata* and *E. ornata*, $\chi^2 = 365.27$, $P_{\chi 2} = 0.005$) (Table 7), indicating that the observed sequence data of the two sister pairs did not fit well with the isolation model. However, the isolation model was not rejected (*E. formosa* and *E. yayeyamana*,

WWH = 42, P_{WWH} = 0.677; *E. decorata* and *E. ornata*, WWH = 32, P_{WWH} = 0.630) in the WWH tests. When non-neutral loci were excluded, the isolation model for both sister pairs was rejected using chi-square statistics but not in the WWH tests (Table 7). In LD tests, significantly positive mean values were found in *E. formosa* and *E. decorata*, indicating that one directional gene flow has occurred into these two species (Table 8). For individual locus, *fer* and *mlc* in *E. formosa* and *nad5* in *E. yayeyamana*, and *arr* and *awd2* in *E. decorata* were found to be significantly deviated from the expected values under the isolation model.

Parameter estimation and model selection of isolation with migration

Using the full data set, multiple runs of simulation for the two *Euphaea* species pairs resulted in unambiguous marginal density distributions of parameters values for population size (θ), migration (m) and divergence time (t) (Fig. 4). The ESS of splitting time parameters in different runs ranged from 481 to 3783 in *E. formosa* vs. *E. yayeyamana* and from 255 to 2555 in *E. decorata* vs. *E. ornata*. The parameters scaled by neutral mutation rate (θ, m and t) were converted into demographic units using the geometric means of mutation rates for each species pairs (3.58×10^{-6} per locus per year for *E. formosa* vs. *E. yayeyamana* ; 3.32×10^{-6} per locus per year for *E. decorata* vs. *E. ornata*). The estimated population size (θ) was 0.711 million (HPD:

0.561–0.903) for *E. formosa* and 0.298 million (HPD: 0.224–0.393) for *E. yayeyamana* (Table 9). The estimated population size of ancestral species for *E. formosa* and *E. yayeyamana* was 0.701 million, with zero included in the lowest 95% HPD interval (HPD: 0–2.408) due to non-convergent estimation (Fig. 4A). The marginal density curve of divergence time (t) for *E. formosa* and *E. yayeyamana* had a sharp peak at 4.095 (Fig. 4B), which was equal to an estimation of 1.145 Mya (HPD: 0.642–8.334) (Table 9). This divergence time estimation was not converged at the highest 95% HPD interval as indicated by a long right tail of the distribution (Fig. 4B). There was significant two directional gene flow (m) from *E. formosa* to *E. yayeyamana* ($m_{fy} = 0.132$, $\chi^2 = 38.192$, $p < 0.001$) and from *E. yayeyamana* to *E. formosa* ($m_{yf} = 0.0367$, $\chi^2 = 30.821$, $p < 0.001$) (Fig. 4C) (Table 9). The population migration rate per generation from *E. formosa* to *E. yayeyamana* ($2N_f M_{fy}$) was close to 0.6749, whereas the reverse was estimated to be 0.0785 ($2N_y M_{yf}$) (Table 9). The estimated number of migration events from *E. formosa* to *E. yayeyamana* for individual loci ranged from one to five, and the reverse ranged from zero to two (Table 10). The mean time of migration events varied among loci and occurred ranging from 0.02 to 4.04 for migrating from *E. formosa* to *E. yayeyamana* and from 0.32 to 1.85 for migrating from *E. yayeyamana* to *E. formosa* (Table 10). Only one migration event of *anon* was detected from *E.decorata* to *E. ornata*. The likelihood

ratio tests of the nested models rejected all models with zero migration rates, equal rates of m_{fy} and m_{yf}, and equal population sizes for *E. formosa*, *E. yayeyamana* and ancestral species (Table 11). The model of equal population sizes for *E. formosa* and ancestral species was not rejected ($p = 0.189$).

The estimated population size (θ) of *E. decorata* (0.489 million, HPD: 0.352–0.661) and *E. ornata* (0.521 million, HPD: 0.374–0.698) were similar in model of isolation with migration (Table 9). The population size of ancestral species for *E.decorata* and *E. ornata* (0.397 million, HPD: 0.237–0.625) was smaller than that of its descendants (Table 9). The estimated divergence time (t) for *E. decorata* and *E. ornata* was 0.511 Mya (HPD: 0.357–0.726) (Table 9), which was much more recent than the split between *E. formosa* and *E. yayeyamana* of approximately 1.145 Mya. The estimated migration rate (m) between *E. decorata* and *E. ornata* during speciation was much lower than 0.001 (Fig. 4F) (Table 9). The likelihood ratio tests in isolation with migration model did not reject the models with zero migration and equal population sizes for *E. decorata* and *E. ornata* (Table 11).

With the data set of no recombined blocks, the estimated population sizes (θ) were all smaller than the full data set (Fig. 5A, D) (Table 9). The estimated divergence times (t) were more recent for *E. decorata* vs. *E. ornata* (Fig. 5E). The divergence time (t) of *E. formosa* and *E. yayeyamana* was estimated poorly in this

data set (Fig. 5B). The estimated migration parameters (*m*) indicated similar pattern of gene flow with higher values (Fig. 5C, F) (Table 9). Using data set with non-neutral loci excluded, the estimated population sizes (θ) in *E. formosa* and *E. yayeyamana* were smaller than that in full data set (Fig. 6A, D) (Table 9). The estimated divergence time (*t*) for *E. decorata* vs. *E. ornata* was comparable to the full data set (Fig. 6E), but the estimation of divergence time failed to converge for *E. formosa* vs. *E. yayeyamana* (Fig. 6B). The estimated migration rate (*m*) had similar pattern as full data set, but the migration rate from *E. yayeyamana* to *E. formosa* (m_{yf}) included zero estimation in the lowest 95% HPD interval (Fig. 6C). When the data set without recombined blocks and non-neutral loci was analyzed, the estimated population sizes (θ) in *E. yayeyamana* and *E. formosa* were smaller than that of data set with non-neutral loci excluded (Fig. 7A). The estimated population sizes (θ) in *E. decorata* and *E. ornata* were similar to the data set without recombined blocks (Fig. 7C). The estimated divergence time (*t*) in *E. yayeyamana* vs. *E. formosa* was failed to converge (Fig. 7B). The estimated migration rate (*m*) was similar to data set without non-neutral loci (Fig. 7C, F). The results of nested model tests for the reduced data sets were similar to the full data set, except the model of equal population size between *E. formosa* and ancestral species were rejected (Table 11).

Discussion

In this study, substantial two directional gene flow was detected during the speciation process of *E. formosa* and *E. yayeyamana*. The level and timing of gene flow between *E. formosa* and *E. yayeyamana* varies among loci (Table 4 and 10). The results support the "porous model" of speciation where some portions of genome involving in reproductive isolation cease gene flow in the early stage of speciation while the rest of the genome exchange gene continually (Wu 2001, Wu & Ting 2004). Detection of historical gene flow during the diverging process between *E. formosa* and *E. yayeyamana* clearly reject the scenario of a strict allopatric model and in favor of an isolation with migration model. In the strict allopatric speciation, reproductive isolation is a byproduct resulted from continuous accumulation of variation in isolated population after the initial geographical isolation. Therefore, mutation and random genetic drift play more important roles in driving population divergence and speciation (Gavrilets 2003). In an isolation with migration model of speciation, divergent selection or sexual selection are more important in initiating and maintaining the divergence populations.

The assumptions of no recombination within locus and selective neutrality for IM model were violated in a few loci. The simulation studies indicated that the violation of model assumptions in isolation with migration can cause bias in

parameter estimation (Becquet & Molly 2009; Straburg & Rieseberg 2010). The results of simulation suggested that the parameters of population size (θ) and divergence time (t) are more sensitive to the violation of model assumptions than that of migration rate (m) (Straburg & Rieseberg 2010). The estimated population sizes of *Euphaea* species are smaller using data set containing only non-recombined regions, which may result from decreased level of polymorphisms. The population size of ancestral species and the divergence time in *E. formosa* and *E. yayeyamana* (but not in *E. decorata* and *E. ornata*) were poorly estimated, which is likely the result of insufficient information in the sampled loci (Becquet & Przeworski 2007; Morgan *et al*. 2010; Wachowiak *et al*. 2011). The simulation studies found that the estimated migration rate in IM were robust even when the sampled loci underwent intralocus recombination or selective sweeps (Straburg & Rieseberg 2010). The false positive of detecting gene flow may occurred when focal species had gene flow with the third unsampled species (Straburg & Rieseberg 2010), or when species diverged allopatrically followed by secondary contact (Becquet & Przeworski 2009). These two scenarios may occur in hybridizing species or currently sympatric species (Straburg & Rieseberg 2010). In *E. formosa* and *E. yayeyamana*, the above two scenarios can be ruled out because the two damselflies are distinct in aspect of morphology, genetics, and phylogeny which are distantly related to their closest

Euphaea species. The two species are currently not sympatric and have disjunct geographical distribution separated by the oceanic barrier (Hayashi 1990; Huang & Lin 2011; Lee & Lin, unpublished).

In *E. decorata* and *E. ornata*, no detectable historical gene flow was found in sampled sequences. The past gene flow between *E. decorata* and *E. ornata* was either too small to be detected or occurred only in the early stage of speciation. The isolation with migration model in IMa2 estimated the "average" gene flow during the speciation process (Hey & Nielsen 2004, 2007) in which the exact timing of gene flow cannot be estimated (Straburg & Rieseberg 2011). The differences in rates of gene flow over time can cause estimation of zero gene flow in IMa2, and the resulting timing of gene flow cannot be used to distinguish between scenarios of a strict allopatric divergence and an isolation with migration model (Becquet & Przeworski 2007).

The results of *E. yayeyamana* having a much smaller population size (0.298 million) than *E. formosa* (0.701 million) and their ancestral species (0.711 million), indicates that *E. yayeyamana* experienced a population reduction or bottleneck after colonizing Ishigaki and Iriomote islands. The current size of Taiwan is over 100 times larger than Ishigaki and Iriomote (Taiwan: 36,000 km^2; Ishgaki: 222.6 km^2; Iriomote: 289 km^2). However, the estimated population size of extant *E. formosa* was only 2.3

times larger than *E. yayeyamana*. The current population size of *E. formosa* appears to be underestimated and likely due to limited geographical sampling. The estimated divergence time of *E. formosa* and *E. yayeyamana* was approximately 1.15 Mya (0.64–8.33) during the early Pleistocene and was approximately two fold more recent than the previous estimation (2.6 Mya, Huang & Lin 2010). The difference in divergence time estimation may result from sample size, number of genetic makers or substitution rates.

The existence of two directional migration during diverging process of *E. formosa* ($2N_f M_{fy} = 0.675$) and *E. yayeyamana* ($2N_y M_{yf} = 0.079$) indicates that speciation in these two *Euphaea* damselfies can occur in spite of substantial gene flow. Most studies of gene flow among natural populations had migration estimated to range from 0 to 0.1 migrants per generation (Pinho & Hey 2010). The level of estimated historical gene flow was considered to be moderate from *E. yayeyamana* to *E. formosa* and large from *E. formosa* to *E. yayeyamana*. In model of speciation with gene flow, the force of divergent selection can overwhelm the homogenizing effect of gene flow (Gavrilets 2003). An earlier study of morphological differentiation of *E. formosa* and *E. yayeyamana* found that these sibling species exhibit distinct wing shapes which are likely subject to natural or sexual selection (Lee & Lin, unpublished). Ecological factors can have profound effects in speciation by increasing

the rate of divergence of locally adapted populations (Gavrilets 2003). Taiwan, Ishigaki and Iriomote island are different not only in size but also in stream ecology. The apparent difference in ecology between Taiwan and Ishigaki and Iriomote determine the abundance and availability of larvae prey for *Euphaea* species in the stream habitats (Hayashi 1990). The abundance of larvae prey may plays a major factor in *E. formosa* and *E. yayeyamana* and finally result to ecological speciation.

Acknowledgements

We thank Jen-Pan Huang and Hidetoshi Ota for help in collecting samples. We also thank Wei-Hsin Chiu, Hurng-Yi Wang and Chao-Tung Yang for the access to their high-performance computing systems.

References

Arias CF, Muñoz AG, Jiggins CD, Mavárez J, Bermingham E, Linares M (2008) A hybrid zone provides evidence for incipient ecological speciation in *Heliconius* butterflies. *Molecular Ecology*, **17**, 4699–4712.

Becquet C, Przeworski M (2007) A new approach to estimate parameters of speciation models with application to apes. *Genome Research*, **17**, 1505–1519.

Becquet C, Przeworski M (2009) Learning about modes of speciation by computational approaches. *Evolution*, **63**, 2547–2562.

Berner D, Grandchamp A, Hendry A (2009) Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution*, **63**, 1740–53.

Brumfield RT, Liu L, Lum DE, Edwards SV (2008) Comparison of species tree methods for reconstructing the phylogeny of bearded manakins (Aves: Pipridae, *Manacus*) from multilocus sequence data. *Systematic Biology*, **57**, 719–731.

Bull V, Beltrán M, Jiggins CD, McMillan WO, Bermingham E, Mallet J (2006) Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Biology*, **4**, 11.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates Inc, Sunderland, Massachusetts.

Drummond AJ, Rambaut A (2007) BEAST: bayesian evolutionary analysis by 629 sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Dudgeon D (1989) Life cycle, production, microdistribution and diet of the damselfly *Euphaea decorata* (Odonata: Euphaeidae) in a Hong Kong forest stream. *Journal of Zoology*, **217**, 57–72.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence times in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics*, **155**, 1405–1413.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Gavrilets S (2003) Perspective: models of speciation: what have we learned in 40 years? *Evolution*, **57**, 2197–2215.

Geyer CJ (1991) Markov chain Monte Carlo maximum likelihood. In: *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface* (eds Keramidas EM), pp. 156–163. Interface Foundation of North America, Seattle, Washington.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a

molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hayashi F (1990) Convergence of insular dwarfism in damselflies (*Euphaea*) and dobsonflies (*Protohermes*). *Freshwater Biology*, **23**, 219–231.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.

Hey J (2010) The divergence of chimpanzee species and subspecies as revealed in multi-population isolation-with-migration analyses. *Molecular Biology and Evolution*, **27**, 921–933.

Hey J, Wakeley J (1997) A coalescent estimator of the population recombination rate. *Genetics*, **145**, 833–846.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 2785–2790.

Huang JP, Lin CP (2011) Lineage-specific late Pleistocene expansion of an endemic subtropical gossamer-wing damselfly, *Euphaea formosa*, in Taiwan. *BMC*

Evolutionary Biology, **11**, 94.

Hubisz M, Falush D, Stephens M, Pritchard J (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Hudson RR (1987). Estimating the recombination parameter of a finite population model without selection. *Genetical Research*, **50**, 245–250.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111,** 147–164.

Hudson RR, Kreitman M, Aguade M (1987). A test of neutral molecular evolution based on nucleotide data. Genetics, **116**, 153–159.

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics*, **132**, 583–589.

Huelsenbeck JP*,* Ronquist F (2001) MrBayes v. 3.12: Bayesian inference of phylogeny. *Bioinformatics,* **17***,* 754–755.

Jordan S, Simon C, Polhemus D (2003) Molecular systematics and adaptive radiation of Hawaii's endemic damselfly genus *Megalagrion* (Odonata: Coenagrionidae). *Systematic Biology*, **52**, 89–109.

Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, **61**, 893–903.

Kimura M (2000) Paleogeography of the Ryulryu Islands. *Tropics*, **10**, 5–24.

Kliman RM, Andolfatto P, Coyne JA, Depaulis F, Kreitman M, Berry AJ, McCarter J, Wakeley J, Hey J (2000) The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics*, **156**, 1913–1931.

Lawniczak MKN, Emrich SJ, Holloway AK, Regier AP, Olson M, White B, Redmond S, Fulton L, Appelbaum E, Godfrey J, Farmer C, Chinwalla A, Yang SP, Minx P, Nelson J, Kyung K, Walenz BP, Garcia-Hernandez E, Aguiar M, Viswanathan LD, Rogers YH, Strausberg RL, Saski CA, Lawson D, Collins FH, Kafatos FC, Christophides GK, Clifton SW, Kirkness EF, Besansky NJ (2010) Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science*, **330**, 512–514.

Leaché A, Mulcahy D (2007) Phylogeny, divergence times and species limits of spiny lizards (*Sceloporus magister* species group) in western North American deserts and Baja California. *Molecular Ecology*, **16**, 5216–33.

Lee YH, Lin CP. Primers for amplification of nuclear introns in four black-banded gossamer-wing *Euphaea* damselflies of East Asia. Unpublished.

Lee YH, Lin CP. Morphometric and genetic differentiation of two sibling gossamer-wing damselflies, *Euphaea formosa* and *E. yayeyamana* (Insecta, Odonata, Euphaeidae), and adaptive trait divergence in subtropical East asian

islands. Unpublished.

Li JW, Yeung CKL, Tsai PW, Lin RC, Yeh CF, Yao CT, Han L, Hung LM, Ding P, Wang Q, Li SH (2010) Rejecting strictly allopatric speciation on a continental island: prolonged postdivergence gene flow between Taiwan (*Leucodioptron taewanus*, Passeriformes Timaliidae) and Chinese (*L. canorum canorum*) hwameis. *Molecular Ecology*, **19**, 494–507.

Librado P, Rozas J (2009) DnaSP v. 5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Losos JB, Glor RE (2003) Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, **18**, 220–227.

Machado CA, Kliman RM, Markert JA, Hey J (2002) Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Molecular Biology and Evolution*, **19**, 472–488.

Mayr E (1963) *Animal Species and Evolution*. Harvard University Press, Cambridge, Massachusetts.

Michel AP, Sim S, Powell THQ, Taylor MS, Nosil P, Feder JL (2010) Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 9724–9729.

Morgan K, Linton YM, Somboon P, Saikia P, Dev V, Socheat D, Walton C (2010) Inter-specific gene flow dynamics during the Pleistocene-dated speciation of forest-dependent mosquitoes in Southeast Asia. *Molecular Ecology*, **19**, 2269–2285.

Moriyama EN, Gojobori T (1992) Rates of synonymous substitution and base composition of nuclear genes in *Drosophila*. *Genetics*, **130**, 855–864.

Nadachowska K, Babik W (2009) Divergence in the face of gene flow: the case of two newts (Amphibia: Salamandridae). *Molecular Biology and Evolution*, **26**, 829–841.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nichols R (2001) Gene trees and species trees are not the same. *Trends in Ecology and Evolution*, **16**, 358–364.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov Chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Niemiller ML, Fitzpatrick BM, Miller BT (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Molecular Ecology*, **17**, 2258–2275.

Nixon KC (1999) The parsimony ratchet, a new method for rapid parsimony analysis.

Cladistics, **15**, 407–414

Osada N, Wu CI (2005) Inferring the mode of speciation from genomic data: a study of the great apes. *Genetics*, **169**, 259–264.

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-aegean trench calibration. *Molecular Biology and Evolution*, **27**, 1659–1672.

Peters JL, Zhuravlev Y, Fefelov I, Logie A, Omland KE (2007) Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (anas spp.). *Evolution*, **61**, 1992–2006.

Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 215–230.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Sikes DS, Lewis PO (2001) PAUPRat: PAUP* implementation of the parsimony ratchet. Distributed by the authors. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, USA.

Stephens M, Donelly P (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

Stephens M, Smith N, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics,* **68**, 978–989.

Strasburg JL, Rieseberg LH (2010) How robust are "isolation with migration" analyses to violations of the IM model? A simulation study. *Molecular Biology and Evolution*, **27**, 297–310.

Strasburg JL, Rieseberg LH (2011) Interpreting the estimated timing of migration events between hybridizing species. *Molecular Ecology*, **20**, 2353–2366.

Swofford DL (2002) *PAUP* v. 4.0b10: Phylogenetic Analysis Using Parsimony (*and Other Methods).* Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1989a) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tajima F (1989b) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597–601.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular*

Biology and Evolution, **10**, 512–526.

Tamura K, Peterson D, Peterson N Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* (In Press).

Thalmann O, Fischer A, Lankester F, Paabo S, Vigilant L (2007) The complex evolutionary history of gorillas: insights from genomic data. *Molecular Biology and Evolution*, **24**, 146–58.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, **3**, e285.

Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**, 1153–1167.

Wachowiak W, Palmé AE, Savolainen O (2011) Speciation history of three closely related pines *Pinus mugo* (T.), *P. uliginosa* (N.) and *P. sylvestris* (L.). *Molecular Ecology*, **20**, 1729–1743.

Wakeley J, Hey J (1997) Estimating ancestral population parameters. *Genetics*, **145**, 847–855.

Wang RL, Wakeley J, Hey J (1997) Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics*, **147**, 1091–1106.

Weekers PHH, De Jonckheere JF, Dumont HJ (2001) Phylogenetic relationships inferred from ribosomal ITS sequences and biogeographic patterns in representatives of the genus *Calopteryx* (Insecta: Odonata) of the west Mediterranean and adjacent west European zone. *Molecular Phylogenetics and Evolution*, **20**, 89–99.

White BJ, Cheng C, Simard F, Costantini C, Besansky NJ (2010) Genetic association of physically unlinked islands of genomic divergence in incipient species of *Anopheles gambiae*. *Molecular Ecology*, **19**, 925–939.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Won YJ, Hey J (2005) Divergence population genetics of chimpanzees. *Molecular Biology and Evolution*, **22**, 297–307.

Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Wu CI, Ting CT (2004) Genes and speciation. *Nature Reviews Genetics*, **5**, 114–122.

Xu J, Chan TY, Tsang LM, Chu KH (2009) Phylogeography of the mitten crab *Eriocheir sensu stricto* in East Asia: Pleistocene isolation, population expansion and secondary contact. *Molecular Phylogenetics and Evolution*, **52**, 45–56.

Yeung CKL, Tsai PW, Chesser RT, Lin RC, Yao CT, Tian XH, Li SH (2011) Testing

founder rffect speciation: divergence population genetics of the spoonbills *Platalea regia* and *P. minor* (Threskiornithidae, Aves). *Molecular Biology and Evolution*, **28**, 473–482.

Zheng XM, Ge S (2010) Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology*, **19**, 2439–2454.

Table 1 Localities and collecting information of the *Euphaea* specimens.

Species	Country	Localities	Collector	Date	Coordinates	
Bayadera brevicauda (Fraser, 1873)	Taiwan	Fushan, Yilan	YHL	2008.7	N 24° 45' 28.0"	E 121° 37' 23.2"
Euphaea amphicyana (Ris, 1930)	Philippines	Ebol Fall, Mindanao	CPL	2009.5	N 07° 29' 10.1"	E 125° 14' 02.4"
Euphaea cora (Ris, 1930)	Philippines	Sian, Mindanao	CPL	2009.5	N 08° 14' 38.3"	E 126° 16' 01.8"
Euphaea decorata (Selys, 1853)	Hong Kong	Tai Po Kau, Tai Po	YHL	2008.6	N 22° 25' 34.2"	E 114° 10' 52.6"
		Shing Mun Country Park, Tsuen Wan	YHL	2008.6	N 22° 23' 26.2"	E 114° 08' 41.4"
		Wu Kau Tang	YHL	2008.6	N 22° 30' 20.4"	E 114° 14' 30.5"
	Vietnam	Tam Dao National Park, Vinh Phuc	CPL	2008.6	N 21° 27' 21.9"	E 105° 38' 41.5"
		Me Linh Biodiversity Station, Vinh Phuc	CPL	2008.6	N 21° 22' 19.8"	E 105° 42' 53.1"
Euphaea formosa (Selys, 1869)	Taiwan	Tsengwen	JPH	2005.9	N 23° 15' 50.0"	E 120° 35' 30.0"
		Ta-an	JPH	2005.9	N 24° 17' 30.0"	E 120° 52' 30.0"
		Lienhuachi	JPH	2005.6	N 23° 55' 00.0"	E 120° 52' 30.0"
		Touchien	JPH	2005.7	N 24° 43' 10.0"	E 121° 15' 05.0"
		Shimenkeng	JPH	2005.7	N 25° 17' 30.0"	E 121° 34' 20.0"
Euphaea masoni (Selys, 1879)	Vietnam	Me Linh Biodiversity Station, Vinh Phuc	CPL	2008.6	N 21° 22' 19.8"	E 105° 42' 53.1"
Euphaea ornata (Campion, 1924)	China	Jianfengling, Hainan	CPL	2008.4	N 18° 45' 21.8"	E 108° 53' 18.7"
		Mt. Diaoluo, Hainan	CPL	2008.4	N 18° 43' 30.4"	E 109° 52' 07.4"
Euphaea querini (Rambur, 1842)	Vietnam	Me Linh Biodiversity Station, Vinh Phuc	CPL	2008.6	N 21° 22' 19.8"	E 105° 42' 53.1"
Euphaea refulgens (Selys, 1853)	Philippines	Mt. Isarog, Luzon	CPL	2006.2	N 13° 39' 47.5"	E 123° 19' 56.6"
Euphaea subcostalis (Selys, 1873)	Indonesia	Gurah, Sumatra	CPL	2005.9	N 03° 40' 55.0"	E 97° 39' 13.0"
Euphaea yayeyamana (Oguma, 1913)	Japan	Hoshino, Ishigaki	YHL	2008.7	N 24° 25' 59.7"	E124° 14' 26.9"
		Fuanasogawa, Ishigaki	YHL	2008.7	N 24° 26' 20.4"	E124° 14' 56.1"
		Shiiragawa, Iriomote	YHL	2008.7	N 24° 18' 49.9"	E123° 54' 22.7"

CPL, Chung-Ping Lin; JPH, Jen-Pan Huang; YHL, Yat-Hung Lee.

Table 2 List of primers for amplifying mitochondrial and nuclear loci used in this study.

Locus	Tm (°C)	Exon	Intron	Primer	Sequence	Reference
cox2	54	1	0	C2-J-3102	5'-AAATGGCAACATGAGCACAAYT-3'	Jordan *et al.* 2003
				E-C2-N-3740	5'-TCATCTAGTGAGGCTTCA-3'	Huang & Lin 2011
nad5	54	1	0	ND5-143F	5'-CCATGATCAAATCTCTTAACTAA-3'	This study
				ND5-1133R	5'-TGCTGCTATRACYAARAGTGC-3'	This study
ITS	52	-	-	18SF	5'-TAGAGGAAGTAAAAGTCG-3'	Weekers *et al.* 2001
				28SR	5'-GCTTAAATTCAGCGG-3'	Weekers *et al.* 2001
act	56~58	1	0	Act-115F	5'-TGCTATGTSGCYCTKGACTTC-3'	Lee & Lin unpublished
				Act-539R	5'-ACTCGTCGTAYTCCTGYTTSGA-3'	Lee & Lin unpublished
arr	54~56	3	2	Arr-805F	5'-GARATCTACTACCAYGGYGAGAA-3'	Lee & Lin unpublished
				Arr-1087R	5'-ACCARRGTGGARGABGCMARGTT-3'	Lee & Lin unpublished
awd2	53~55	2	1	Awd-195F	5'-GGAGGGACTCAATTCTGTCAAA-3'	Lee & Lin unpublished
				Awd-329R	5'-AATTTCCTTCTGAGCTGACTCAA-3'	Lee & Lin unpublished
EF1α	54	4	3	EF-1-F-2361	5'-YGGMCACAGRGATTTCATCAA-3'	Jordan *et al.* 2003
				EF1-R-3093	5'-CCAGGRTGGTTRAGCACRATGA-3'	Jordan *et al.* 2003
fer	53~55	2	1	Fer-29F	5'-CAATSGCAWACTAYTTYGAYCAAG-3'	Lee & Lin unpublished
				Fer-344R	5'-TCYTTRATGGCCTCRACYTGYT-3'	Lee & Lin unpublished
mlc	53	3	2	mlc-44F	5'-CCAATCAACTTCACCCAACTGCT-3'	Lee & Lin unpublished
				mlc-182R	5'-GAACTTGTCTCCCCAGGTCATAAG-3'	Lee & Lin unpublished
lop1	54~56	2	1	E-2409-opsin-F-417	5'-CCTTTGGCACGGAATCTTAG-3'	Lee & Lin unpublished
				E-2409-opsin-R-710	5'-CATTGTATCTGTCCATGGCG-3'	Lee & Lin unpublished
sdhB	54~56	3	2	SdhB-16F	5'-TAYCGATGGAAYCCRGAHAAGSC-3'	Lee & Lin unpublished
				SdhB-268R	5'-AYNACRTACATRTGBGGYARHGG-3'	Lee & Lin unpublished
anon (*Anonymous*)	53	-	-	mlc-44F	5'-CCAATCAACTTCACCCAACTGCT-3'	Lee & Lin unpublished
				mlc-182R	5'-GAACTTGTCTCCCCAGGTCATAAG-3'	Lee & Lin unpublished

Table 3 Summary statistics of sampled populations in four *Euphaea* species.

Locus	Species	Coding	Non-coding	N	*S*	*Hd*	π	*Rm*	*H*	*D*
cox2	*formosa*	500	0	19	42	0.947	0.0289	-	2.111	0.497
	yayeyamana	500	0	16	8	0.850	0.0032	-	–2.750	–1.225
	decorata	606	0	17	9	0.647	0.0053	-	0.191	0.756
	ornata	606	0	16	12	0.908	0.0035	-	1.800	–1.602
nad5	*formosa*	912	0	16	62	0.908	0.0101	-	–85.983**	–2.192**
	yayeyamana	912	0	16	16	0.925	0.0050	-	–6.833*	–0.202
	decorata	912	0	16	14	0.742	0.0058	-	–1.767	0.950
	ornata	912	0	15	17	0.943	0.0033	-	–0.343	–1.839*
ITS	*formosa*	666	0	19	15	0.942	0.0041	0	0.906	–1.385
	yayeyamana	666	0	31	6	0.596	0.0015	0	–0.024	–1.031
	decorata	680	0	25	24	0.970	0.0073	1	–1.583	–0.881
	ornata	680	0	29	11	0.727	0.0016	0	–0.924	–1.999*
act	*formosa*	445	0	19	3	0.591	0.0015	0	0.444	–0.607
	yayeyamana	445	0	18	1	0.503	0.0011	0	0.183	1.378
	decorata	445	0	18	1	0.503	0.0011	0	0.183	1.378
	ornata	445	0	24	1	0.489	0.0011	0	–0.326	1.391
arr	*formosa*	305	1029	15	25	0.838	0.0056	0	–2.590	–0.151
	yayeyamana	305	1029	19	51	0.883	0.0108	7	6.368	–0.154
	decorata	305	1016	21	64	0.952	0.0125	13	1.752	–0.363
	ornata	305	1016	20	49	0.947	0.0098	5	6.358	–0.329
awd2	*formosa*	154	212	19	5	0.789	0.0039	0	0.561	–0.084
	yayeyamana	154	212	17	16	0.809	0.0115	0	2.919	–0.695
	decorata	154	212	21	14	0.805	0.0095	1	1.214	–0.421
	ornata	154	212	19	12	0.813	0.0083	0	2.041	–0.705
EF1α	*formosa*	785	243	18	11	0.784	0.0018	0	0.601	–1.561
	yayeyamana	785	243	19	8	0.889	0.0020	2	1.076	–0.292
	decorata	782	243	21	10	0.814	0.0019	1	1.505	–1.027
	ornata	782	243	20	9	0.637	0.0014	0	–0.632	–1.441
fer	*formosa*	337	876	24	43	0.884	0.0059	3	–2.377	–1.576
	yayeyamana	337	876	25	18	0.900	0.0032	3	–9.930**	–0.705
	decorata	337	878	20	84	0.974	0.0120	7	–8.389	–1.624
	ornata	337	878	14	38	0.945	0.0139	2	2.022	1.699
mlc	*formosa*	193	696	22	53	0.952	0.0163	6	–9.333*	–0.313
	yayeyamana	193	696	20	40	0.953	0.0133	0	–12.294*	–0.412
	decorata	193	700	19	26	0.883	0.0085	1	–9.994*	–0.337
	ornata	193	700	24	53	0.975	0.0176	1	0.964	0.115
lop1	*formosa*	311	1070	18	85	0.922	0.0248	5	4.837	1.375
	yayeyamana	311	1070	16	38	0.875	0.0131	0	0.717	2.392*
	decorata	311	1057	23	47	0.960	0.0072	0	6.917	–0.920
	ornata	311	1057	21	52	0.967	0.0083	5	7.957	–0.890

sdhB	*formosa*	275	200	21	14	0.914	0.0080	1	–0.771	–0.344
	yayeyamana	275	200	17	6	0.750	0.0056	1	1.581	0.962
	decorata	258	200	14	3	0.495	0.0032	0	–1.187	1.753
	ornata	258	200	21	1	0.429	0.0009	0	0.257	0.959
anon	*formosa*	0	450	17	33	0.919	0.0228	6	0.816	–0.090
	yayeyamana	0	450	17	7	0.875	0.0048	1	–2.265	0.079
	decorata	0	459	20	35	0.953	0.0250	5	1.842	0.383
	ornata	0	459	24	32	0.902	0.0131	3	–10.601*	–1.121
Total / Mean	*formosa*	4494	5165	227	32.6	0.866	0.011	21	-	0.021
	yayeyamana	4494	5165	231	17.9	0.817	0.006	14	-	0.525*
	decorata	4580	5168	235	27.6	0.808	0.008	29	-	0.595**
	ornata	4580	5168	247	23.9	0.807	0.007	16	-	–0.116

N, number of sequences; *S*, number of polymorphic sites; H_d, haplotype diversity; π, nucleotide diversity (Nei 1987); *Rm*, minimum number of recombination events; *H*, Fay and Wu' s *H* (Fay & Wu 2000); *D*, Tajima's *D* test (Tajima 1989a); *, $p < 0.05$; **, $p < 0.01$.

Table 4 Estimates of net genetic divergence per site and population migration rates.

	Net genetic divergence per site [a]		Population migration rate (*Nm*) [b]	
Locus	*E. formosa* vs. *E. yayeyamana*	*E. decorata* vs. *E. ornata*	*E. formosa* vs. *E. yayeyamana*	*E. decorata* vs. *E. ornata*
cox2	0.0470	0.0050	0.18	0.55
nad5	0.0510	0.0110	0.08	0.21
ITS	0.0210	0.0170	0.03	0.05
act	0.0001	0.0020	0.15	0.00
arr	0.0120	0.0160	0.18	0.18
awd2	0.0010	0.0010	3.86	2.37
EF1α	0.0010	0.0010	0.85	0.38
fer	0.0070	0.0070	0.16	0.45
mlc	0.0090	0.0140	0.47	0.25
lop1	0.0060	0.0020	0.86	1.22
sdhB	0.0060	0.0030	0.26	0.19
anon	0.0120	0.0060	0.29	0.81
Mean	0.0144	0.0071	0.61	0.56
SD	0.0172	0.0059	1.06	0.67

[a], Calculated using Tamura-Nei model (Tamura & Nei 1993); [b], Estimated using the method of Hudson *et al*. (1992).

Table 5 The mean value of log likelihood of each *K* calculated in STRUCTURE.

K	*E. formosa*	*E. yayeyamana*	*E. decorata*	*E. ornata*
1	–373.505	–318.210	–336.305	–311.090
2	–382.715	–320.590	–338.440	–311.510
3	–388.175	–324.675	–342.180	–312.675
4	–384.200	–323.315	–341.360	–314.395
5	–384.050	–320.215	–340.400	–314.520
6	–381.220	–321.485	–337.945	–314.370

Table 6 The four types of observed and simulated polymorphic sites calculated in SITES.

	E. formosa vs. E. yayeyamana				*E. decorata* vs. *E. ornata*			
	Observed / Simulated				Observed / Simulated			
Locus	S_{xf}	S_{xy}	S_s	S_f	S_{xd}	S_{xo}	S_s	S_f
cox2	40 / 35.15	6 / 12.50	2 / 0.24	18 / 18.12	9 / 8.09	12 / 6.62	0 / 0.03	1 / 7.27
nad5	54 / 36.39	8 / 13.56	8 / 0.25	0 / 19.80	12 / 12.78	15 / 10.45	2 / 0.04	6 / 11.73
ITS	16 / 18.44	6 / 9.34	0 / 6.59	13 / 0.63	24 / 20.45	9 / 18.18	2 / 6.18	11 / 1.19
act	2 / 2.21	0 / 0.97	1 / 0.74	1 / 0.08	0 / 0.43	0 / 0.40	1 / 0.14	0 / 0.03
arr	17 / 42.55	43 / 20.72	8 / 14.92	12 / 1.81	49 / 52.06	35 / 43.95	15 / 15.56	16 / 3.44
awd2	1 / 9.43	12 / 4.08	4 / 3.14	0 / 0.36	3 / 6.83	1 / 5.68	11 / 2.03	0 / 0.46
EF1α	10 / 9.81	7 / 4.47	1 / 3.35	0 / 0.37	10 / 8.60	9 / 7.26	0 / 2.57	0 / 0.57
fer	35 / 30.31	10 / 13.52	8 / 10.23	2 / 0.95	69 / 50.88	23 / 38.96	15 / 14.50	1 / 3.66
mlc	35 / 43.91	22 / 18.98	22 / 14.62	0 / 1.49	19 / 33.57	51 / 30.75	7 / 10.42	0 / 2.26
lop1	61 / 54.85	14 / 23.77	24 / 18.22	0 / 2.16	44 / 43.77	49 / 36.51	3 / 12.96	0 / 2.76
sdhB	10 / 9.54	2 / 4.00	4 / 3.12	1 / 0.34	3 / 1.68	1 / 1.64	0 / 0.54	0 / 0.14
anon	34 / 22.42	6 / 10.10	1 / 7.59	0 / 0.90	20 / 22.85	16 / 20.61	16 / 7.04	0 / 1.50
Total	315	136	83	47	262	221	72	35

S_{xf}, S_{xy}, S_{xd} and S_{xo}, exclusive polymorphisms in *E. formosa*, *E. yayeyamana*, *E. decorata* and *E. ornata*, respectively; S_s, shared polymorphisms between species; S_f, fixed differences between species.

Table 7 Estimated parameters and statistics of fitting the isolation model.

E. formosa vs. *E. yayeyamana*								
	θ_f	θ_y	θ_A	T	χ^2	$P_{\chi 2}$	WWH	P_{WWH}
Full data	123.0	52.2	292.1	0.20	711.11	0.007**	42	0.677
	37.7–304.7	18.3–102.6	116.4–521.6	0.08–0.33				
Non-neutral loci excluded	52.1	27.6	205.1	0.30	275.05	0.029*	42	0.594
	0.002–148.7	0.001–67.3	67.65–406.9	0.13–0.57				
E. decorata vs. *E. ornata*								
	θ_d	θ_o	θ_A	T	χ^2	$P_{\chi 2}$	WWH	P_{WWH}
Full data	81.5	68.9	189.5	0.35	365.27	0.005**	32	0.630
	42.5–136.6	37.0–114.6	95.9–307.9	0.21–0.54				
Non-neutral loci excluded	72.1	50.6	159.1	0.34	360.02	0.005**	31	0.562
	37.0–119.4	27.4–83.9	71.0–262.8	0.20–0.54				

θ_f, θ_y, θ_d and θ_o, estimated effective population size in mutational units of *E. formosa*, *E. yayeyamana*, *E. decorata* and *E. ornata*, respectively; θ_A, estimated effective population size in mutational units of ancestral species; T, the estimated divergence time in unit of $2N_1$ generation (N_1 is the effective population size in demographic unit of species 1) (Wang *et al.* 1997); *, $p < 0.05$; **, $p < 0.01$. Values below the parameter estimates are the 95% confidence intervals.

Table 8 The statistics of LD tests.

	E. formosa vs. *E. yayeyamana*						*E. decorata* vs. *E. ornata*					
	E. formosa			*E. yayeyamana*			*E. decorata*			*E. ornata*		
Locus	Observed	Simulated	*p*-value	Observed	Simulated	*p*-value	Observed	Simulated	*p*-value	Observed	Simulated	*p*-value
cox2	NA	0.065	NA	NA	0.107	NA	NA	NA	NA	NA	NA	NA
nad5	NA	0.025	NA	1.187	0.079	0.027*	NA	–0.150	NA	NA	–0.098	NA
ITS	NA	0.383	NA	NA	0.286	NA	NA	0.106	NA	NA	0.094	NA
act	NA	0.085	NA	NA	0.152	NA	NA	0.222	NA	NA	0.042	NA
arr	–0.017	0.059	0.749	0.032	0.033	0.419	0.328	0.044	0.037*	0.128	0.040	0.190
awd2	1.333	0.266	0.056	NA	0.18	NA	1.499	0.080	0.018*	NA	0.059	NA
EF1α	NA	0.280	NA	NA	0.188	NA	NA	0.204	NA	NA	0.187	NA
fer	1.196	0.081	0.001**	0.115	0.044	0.259	0.270	0.059	0.090	–0.031	0.046	0.697
mlc	0.629	0.123	0.032*	–0.010	0.085	0.620	–0.121	0.087	0.871	–0.044	0.081	0.705
lop1	0.541	0.145	0.058	0.058	0.105	0.439	NA	0.064	NA	NA	0.057	NA
sdhB	NA	0.065	NA	0.001	0.047	0.465	NA	0.021	NA	NA	0.075	NA
anon	NA	0.102	NA	NA	0.072	NA	0.611	0.094	0.06	–0.210	0.094	0.918
Mean	0.736	0.135	0.0005***	0.230	0.094	0.156	0.517	0.077	0.003*	–0.039	0.070	0.855
SD	0.544	0.258	0.780	0.471	0.277	0.155	0.608	0.218	0.023*	0.138	0.221	0.675

NA, not available due to small number of shared or exclusive polymorphisms; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Table 9 Estimated parameters of isolation with migration model in IMa2.

E. formosa vs. *E.yayeyamana*												
Full data	θ_f	θ_y	θ_A	m_{fy}	m_{yf}	t	N_f	N_y	N_A	$2N_f M_{fy}$	$2N_y M_{yf}$	T
MLE	10.180	4.270	10.030	0.1326	0.0367	4.095	0.711	0.298	0.701	0.6749	0.0785	1.145
HPD95Lo	8.025	3.210	0	0.0498	0.0092	2.295	0.561	0.224	0	0.1998	0.0147	0.642
HPD95Hi	12.930	5.630	29.310	0.2510	0.0935	29.81	0.903	0.393	2.048	1.6227	0.2631	8.334
Recombined blocks excluded												
MLE	6.625	3.290	9.117	0.2208	0.0386	29.02	0.463	0.230	0.637	0.7314	0.0634	8.113
HPD95Lo	5.125	2.430	0	0.0958	0.0059	1.845	0.358	0.170	0	0.2454	0.0071	0.516
HPD95Hi	8.525	4.410	32.220	0.3787	0.1210	29.98	0.596	0.308	2.252	1.6142	0.2668	8.381
Non-neutral loci excluded												
MLE	8.325	2.350	7.180	0.2831	0.0165	23.940	0.582	0.164	0.502	1.1784	0.0194	6.693
HPD95Lo	6.275	1.570	0	0.1313	0	5.140	0.439	0.110	0.000	0.4120	0	1.437
HPD95Hi	10.930	3.450	37.380	0.5330	0.0879	39.980	0.764	0.241	2.612	2.9128	0.1516	11.177
Recombined blocks and non-neutral loci excluded												
MLE	5.625	1.79	7.675	0.3965	0.01785	94.62	0.393	0.125	0.536	1.1152	0.0160	26.451
HPD95Lo	4.075	1.15	0	0.1975	0	13.26	0.285	0.080	0	0.4024	0	3.707
HPD95Hi	7.525	2.69	47.380	0.7535	0.1075	119.9	0.526	0.188	3.311	2.8350	0.1446	33.519
E. decorata vs. *E.ornata*												
Full data	θ_d	θ_o	θ_A	m_{do}	m_{od}	t	N_d	N_o	N_A	$2N_d M_{do}$	$2N_o M_{od}$	T
MLE	6.492	6.912	5.265	0.0009	0.0001	1.696	0.489	0.521	0.397	0.0029	0.0003	0.511
HPD95Lo	4.673	4.963	3.141	0	0	1.184	0.352	0.374	0.237	0	0	0.357
HPD95Hi	8.768	9.262	8.289	0.0735	0.0877	2.408	0.661	0.698	0.625	0.3222	0.4061	0.726
Recombined blocks excluded												
MLE	3.482	4.112	4.473	0.0068	0.0394	1.316	0.263	0.310	0.337	0.0118	0.0810	0.397
HPD95Lo	2.293	2.813	1.917	0	0	0.676	0.173	0.212	0.145	0	0	0.204
HPD95Hi	4.952	5.737	11.53	0.1477	0.2954	7.988	0.373	0.432	0.869	0.3657	0.8474	2.409
Non-neutral loci excluded												

MLE	6.213	5.737	4.077	0.0001	0.0001	1.472	0.468	0.432	0.307	0.0003	0.0003	0.444
HPD95Lo	4.218	3.837	2.313	0	0	1.004	0.318	0.289	0.174	0	0	0.303
HPD95Hi	8.697	8.063	6.885	0.0855	0.0885	2.039	0.656	0.608	0.519	0.3848	0.3447	0.615
Recombined blocks and non-neutral loci excluded												
MLE	3.015	2.925	3.795	0.0003	0.0003	0.227	0.221	0.286	0.0004	0.0004	0.335	0.227
HPD95Lo	1.845	1.785	1.485	0	0	0.139	0.135	0.112	0	0	0.190	0.139
HPD95Hi	4.755	4.665	27.610	0.3408	0.3172	0.358	0.352	2.081	0.8103	0.7399	3.616	0.358

MLE, maximum likelihood estimates; HPD95Lo, the 95 % lowest posterior density interval; HPD95Hi, the 95 % highest posterior density interval; θ_f , θ_y, θ_d and θ_o, estimated effective population size in mutational units of *E. formosa*, *E. yayeyamana*, *E. decorata* and *E. ornata*, respectively; θ_A, estimated effective population size in mutational units of ancestral species; m_{fy} and m_{yf}, estimated migration rate in mutational units from *E. formosa* to *E. yayeyamana* and the reverse way, respectively; m_{do} and m_{od}, estimated migration rate in mutational units from *E. decorata* to *E. ornata* and the reverse way, respectively; t, divergence time in mutational units; N_f, N_y, N_d and N_o, effective population size in demographic unit ($\times 10^6$ individuals) of *E. formosa*, *E. yayeyamana*, *E. decorata* and *E. ornata*, respectively; N_A, effective population size in demographic unit ($\times 10^6$ individuals) of ancestral species; $2N_f M_{fy}$ and $2N_y M_{yf}$, population migration rate per generation from *E. formosa* to *E. yayeyamana* and the reverse way, respectively; $2N_d M_{do}$ and $2N_o M_{od}$, population migration rate per generation from *E. decorata* to *E. ornata*; T, divergence time in demographic unit ($\times 10^6$ years).

Table 10 The number and mean time of migration events estimated in IMa2.

	E. formosa vs. *E. yayeyamana*								*E. decorata* vs. *E. ornata*							
	m_{fy}				m_{yf}				m_{do}				m_{od}			
	event[a]	probability	time[b]	probability	event[a]	probability	time[b]	probability	event[a]	probability	time[b]	probability	event[a]	probability	time[b]	probability
cox2	1	0.689	1.94	0.011	0	0.760	NA	NA	0	0.938	NA	NA	0	0.884	NA	NA
nad5	1	0.520	4.04	0.004	1	0.795	0.02	0.419	0	0.966	NA	NA	0	0.961	NA	NA
ITS	1	0.424	3.20	0.003	0	0.678	NA	NA	0	0.970	NA	NA	0	0.871	NA	NA
act	1	0.387	3.05	0.004	0	0.547	NA	NA	0	0.682	NA	NA	0	0.686	NA	NA
arr	5	0.258	0.26	0.076	0	0.462	NA	NA	0	0.969	NA	NA	0	0.818	NA	NA
awd2	5	0.257	0.02	0.402	0	0.352	NA	NA	0	0.616	NA	NA	0	0.521	NA	NA
EF1α	3	0.300	0.05	0.109	0	0.398	NA	NA	0	0.659	NA	NA	0	0.848	NA	NA
fer	1	0.527	2.66	0.009	0	0.471	NA	NA	0	0.924	NA	NA	0	0.807	NA	NA
mlc	3	0.434	0.08	0.129	1	0.400	0.62	0.042	0	0.834	NA	NA	0	0.950	NA	NA
lop1	1	0.269	0.56	0.011	2	0.325	0.32	0.081	0	0.772	NA	NA	0	0.835	NA	NA
sdhB	1	0.479	3.17	0.006	0	0.520	1.85	0.004	0	0.824	NA	NA	0	0.774	NA	NA
anon	1	0.428	1.91	0.010	2	0.269	1.55	0.018	1	0.500	0.20	0.010	0	0.520	NA	NA

[a], migration events observed over the simulation process in IMa2; [b], mean time values (scaled by neutral mutation rate, as with *t*) at which at least one migration event occurred; NA, not available due to zero migration event.

Table 11 Statistics of likelihood ratio tests of nested models selection in IMa2.

E. formosa vs. *E. yayeyamana*										*E. decorata* vs. *E. ornata*									
Model	df	-2LLR[a]	*p*	-2LLR[b]	*p*	-2LLR[c]	*p*	-2LLR[d]	*p*	Model	df	-2LLR[a]	*p*	-2LLR[b]	*p*	-2LLR[c]	*p*	-2LLR[d]	*p*
$m_{fy} = m_{yf}$	1	35.69	***	17.44	***	76.83	***	23.24	***	$m_{do} = m_{od}$	1	0.001	1.000	0.028	0.867	0.301	0.583	0.001	1.000
$m_{yf} = 0$	1[e]	110.10	***	92.02	***	19.67	***	9.20	***	$m_{od} = 0$	1[e]	0.001	0.500	0.028	0.433	0.682	0.205	0.001	0.500
$m_{fy} = 0$	1[e]	272.20	***	245.30	***	166.00	***	395.00	***	$m_{do} = 0$	1[e]	0.001	1.000	0.028	0.433	0.492	0.242	0.001	0.500
$\theta_f = \theta_y$	1	132.50	***	45.86	***	158.20	***	82.26	***	$\theta_d = \theta_o$	1	0.587	0.444	1.104	0.293	0.476	0.490	0.690	0.406
$\theta_f = \theta_y$, $m_{fy} = m_{yf}$	2	137.90	***	60.17	***	181.70	***	103.20	***	$\theta_d = \theta_o$, $m_{do} = m_{od}$	2	0.587	0.745	1.132	0.568	0.506	0.776	0.690	0.708
$\theta_f = \theta_y$, $m_{yf} = 0$	2[e]	245.20	***	132.00	***	185.20	***	112.10	***	$\theta_d = \theta_o$, $m_{od} = 0$	2[e]	0.587	0.373	1.285	0.263	0.691	0.354	0.690	0.354
$\theta_f = \theta_y$, $m_{fy} = 0$	2[e]	316.10	***	258.00	***	269.70	***	455.10	***	$\theta_d = \theta_o$, $m_{do} = 0$	2[e]	0.587	0.373	1.366	0.253	0.539	0.382	0.690	0.354
$\theta_f = \theta_A$	1	1.73	0.189	9.30	***	7.249	**	13.56	***	$\theta_d = \theta_A$	1	0.006	0.940	1.790	0.181	0.194	0.660	0.080	0.777
$\theta_f = \theta_A$, $m_{fy} = m_{yf}$	2	40.24	***	26.74	***	76.90	***	65.06	***	$\theta_d = \theta_A$, $m_{do} = m_{od}$	2	0.006	0.997	1.790	0.409	0.495	0.781	0.080	0.961
$\theta_f = \theta_A$, $m_{yf} = 0$	2[e]	111.80	***	118.50	***	26.92	***	13.56	***	$\theta_d = \theta_A$, $m_{od} = 0$	2[e]	0.006	0.499	1.364	0.253	0.682	0.356	0.080	0.481
$\theta_f = \theta_A$, $m_{fy} = 0$	2[e]	273.00	***	249.50	***	166.50	***	401.90	***	$\theta_d = \theta_A$, $m_{do} = 0$	2[e]	0.006	0.499	1.464	0.240	0.494	0.391	0.080	0.481
$\theta_y = \theta_A$	1	27.99	***	66.59	***	30.43	***	31.75	***	$\theta_o = \theta_A$	1	0.219	0.640	1.397	0.237	0.924	0.336	0.348	0.555
$\theta_y = \theta_A$, $m_{fy} = m_{yf}$	2	73.98	***	81.84	***	76.90	***	74.28	***	$\theta_o = \theta_A$, $m_{do} = m_{od}$	2	0.219	0.896	1.400	0.497	0.123	0.941	0.348	0.840
$\theta_y = \theta_A$, $m_{yf} = 0$	2[e]	138.10	***	154.40	***	56.64	***	51.65	***	$\theta_o = \theta_A$, $m_{od} = 0$	2[e]	0.219	0.448	1.370	0.252	0.683	0.355	0.348	0.420
$\theta_y = \theta_A$, $m_{fy} = 0$	2[e]	310.00	***	265.20	***	253.20	***	466.70	***	$\theta_o = \theta_A$, $m_{do} = 0$	2[e]	0.219	0.448	1.501	0.236	0.501	0.389	0.348	0.420
$\theta_f = \theta_y = \theta_A$	2	141.40	***	85.97	***	158.60	***	90.83	***	$\theta_d = \theta_o = \theta_A$	2	0.634	0.728	1.794	0.408	0.487	0.784	0.769	0.681
$\theta_f = \theta_y = \theta_A$, $m_{fy} = m_{yf}$	3	156.10	***	109.70	***	194.10	***	117.30	***	$\theta_d = \theta_o = \theta_A$, $m_{do} = m_{od}$	3	0.634	0.889	1.794	0.616	0.517	0.915	0.769	0.857
$\theta_f = \theta_y = \theta_A$, $m_{yf} = 0$	3[e]	251.50	***	177.50	***	185.30	***	135.80	***	$\theta_d = \theta_o = \theta_A$, $m_{od} = 0$	3[e]	0.634	0.445	1.371	0.356	0.691	0.438	0.769	0.429
$\theta_f = \theta_y = \theta_A$, $m_{fy} = 0$	3[e]	329.30	***	269.70	***	289.90	***	486.40	***	$\theta_d = \theta_o = \theta_A$, $m_{do} = 0$	3[e]	0.634	0.445	1.514	0.340	0.540	0.455	0.769	0.429

[a], log likelihood value and *p* value of nested model test using full data set; [b], log likelihood value and *p* value of nested model test data set without recombined blocks; [c], log likelihood value and *p* value of nested model test using data set without non-neutral loci; [d], log likelihood value and *p* value of nested model test using data set without recombined blocks and non-neutral loci; [e], distribution of test statistics of 2LLR is a mixture, the *p* values were divided by two (Hey & Nielsen 2007); **, $p < 0.01$; ***, $p < 0.001$.

Figure Legends

Figure 1. A map of East Asia shows the current geographical distribution of the four *Euphaea* species.

Figure 2. Species phylogeny and estimated divergence time of *Euphaea* damselflies based on *cox2* and *arr* genes (BS, Bayesian posterior probabilities of species tree; BPP, Bayesian posterior probabilities; PB, parsimony bootstrapping).

Figure 3. Gene trees of twelve loci of four *Euphaea* damselflies. The asterisk marks near by the nodes represented Bayesian posterior probabilities higher than 0.90.

Figure 4. The posterior marginal density of parameters scaled by mutation rate in IM model using full data set.

Figure 5. The posterior marginal density of parameters scaled by mutation rate in IM model using data set without recombined blocks.

Figure 6. The posterior marginal density of parameters scaled by mutation rate in IM model using data set without non-neutral loci.

Figure 7. The posterior marginal density of parameters scaled by mutation rate in IM model using data set without recombined blocks and non-neutral loci.

Figure 1


E. formosa
E. yayeyamana
E. decorata
E. ornata
China
Fujian
Taiwan Strait
Ishigaki
Iriomote
Guangxi
Guangdong
Hong Kong
Taiwan
Qiongzhou Strait
Hainan
Philippines
Vietnam
29º N
19º N
116º E
126º E
N
250 km

Figure 2

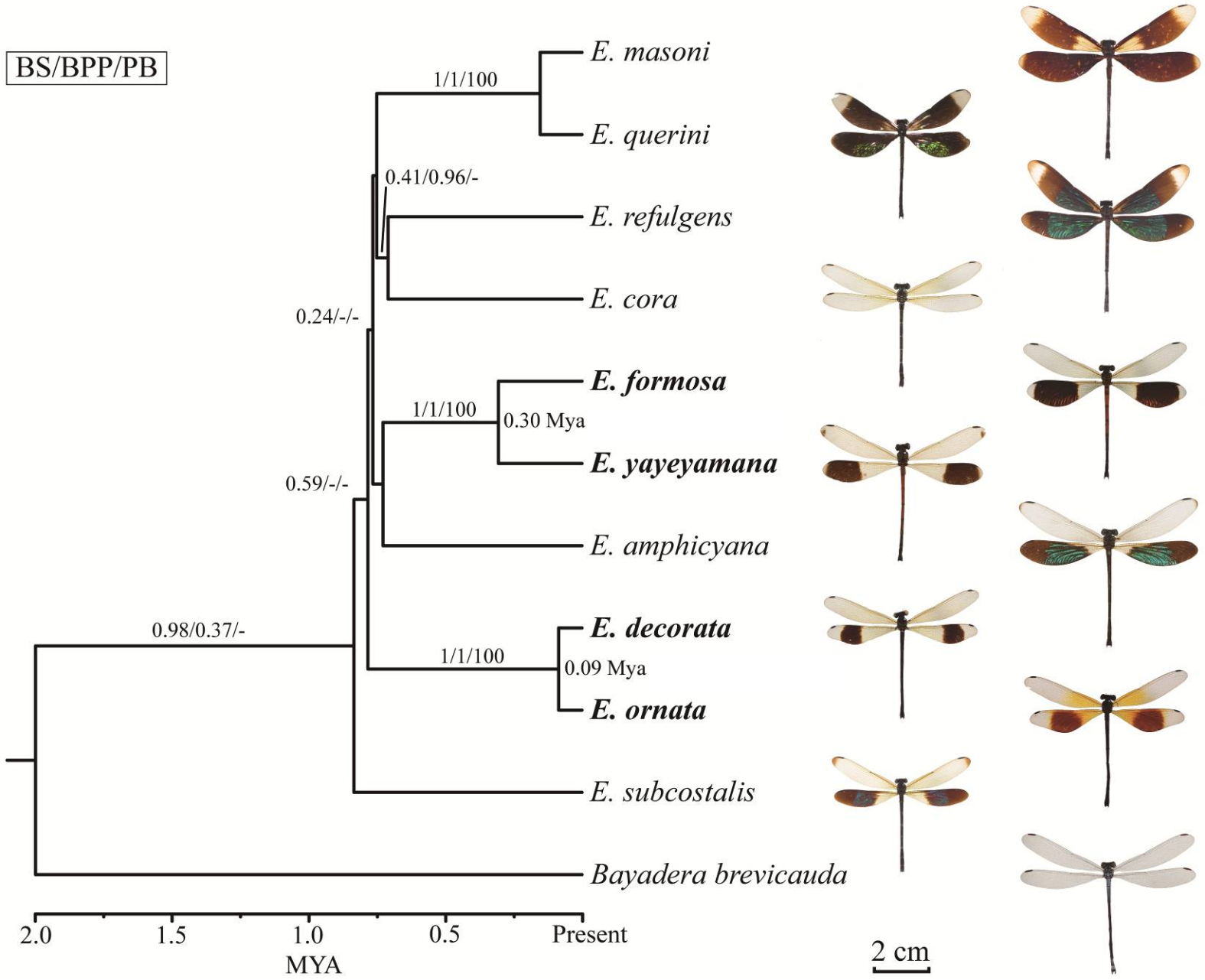

BS/BPP/PB
E. masoni
E. querini
E. refulgens
E. cora
E. formosa
E. yayeyamana
E. amphicyana
E. decorata
E. ornata
E. subcostalis
Bayadera brevicauda
1/1/100
0.41/0.96/-
0.24/-/-
1/1/100
0.30 Mya
0.59/-/-
0.98/0.37/-
1/1/100
0.09 Mya
2.0
1.5
1.0
0.5
Present
MYA
2 cm

Figure 3

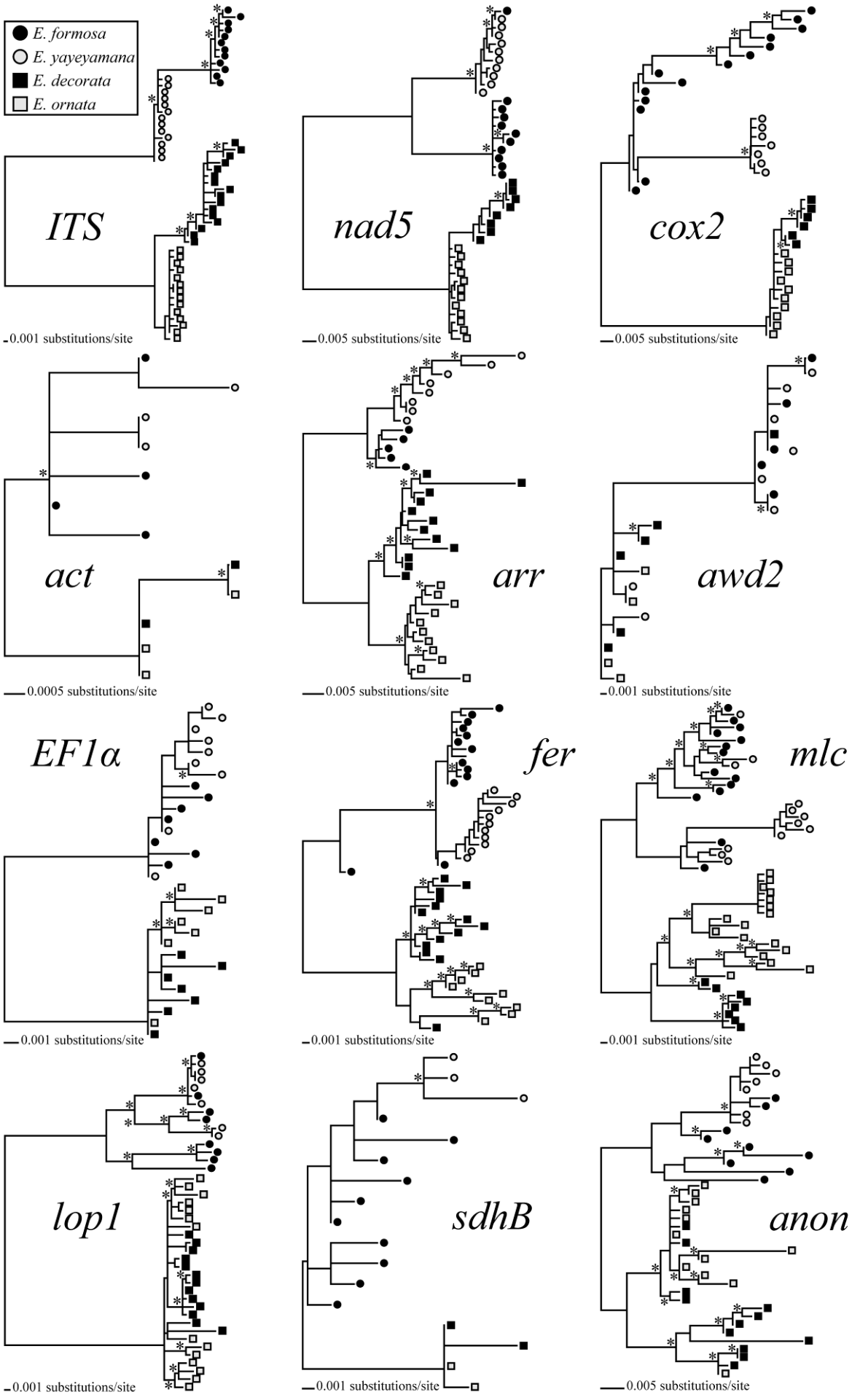

E. formosa
E. yayeyamana
E. decorata
E. ornata
ITS
0.001 substitutions/site
nad5
0.005 substitutions/site
cox2
0.005 substitutions/site
act
0.0005 substitutions/site
arr
0.005 substitutions/site
awd2
0.001 substitutions/site
EF1α
0.001 substitutions/site
fer
0.001 substitutions/site
mlc
0.001 substitutions/site
lop1
0.001 substitutions/site
sdhB
0.001 substitutions/site
anon
0.005 substitutions/site

Figure 4

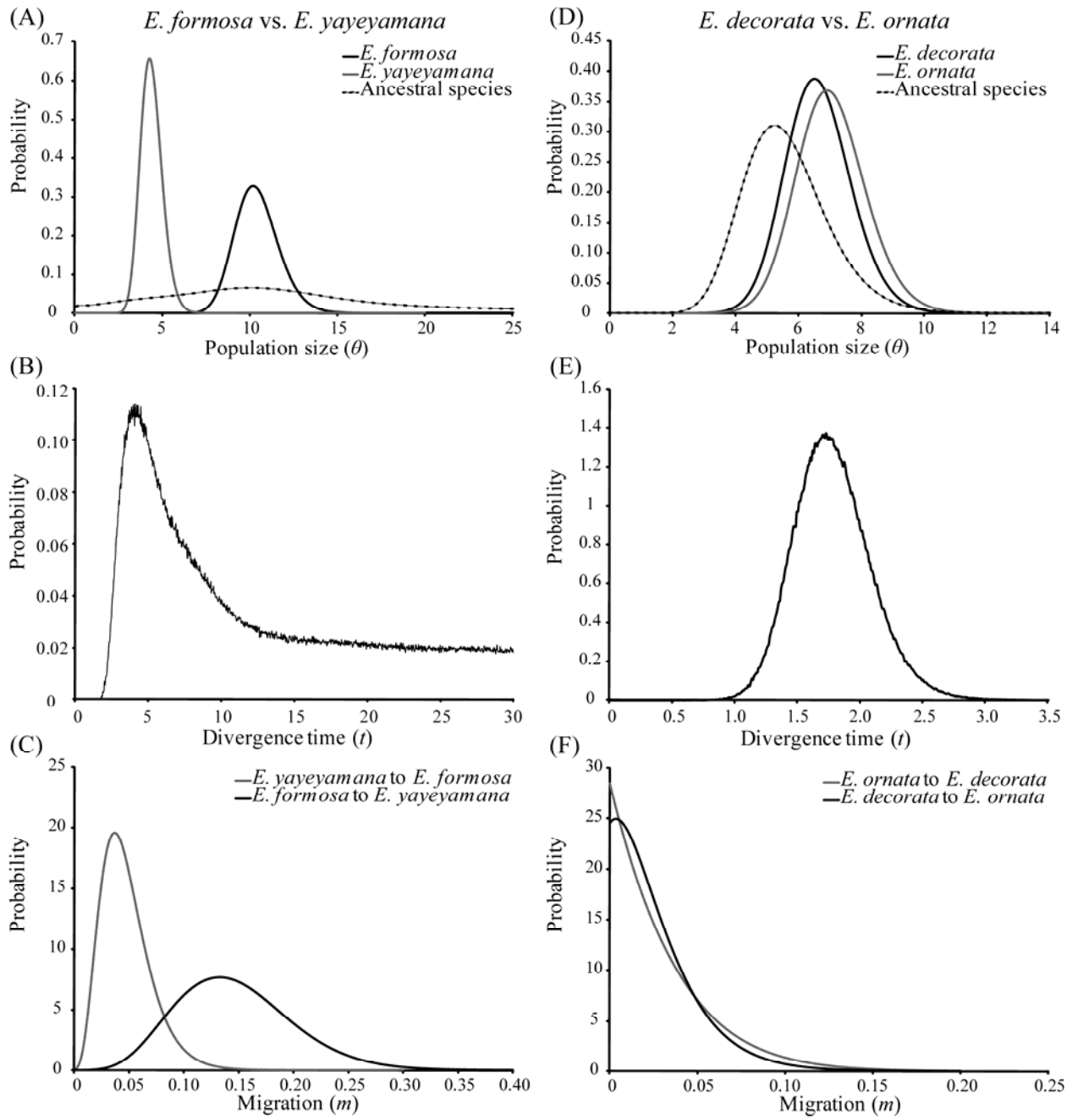

(A)
E. formosa vs. E. yayeyamana
—E. formosa
—E. yayeyamana
---Ancestral species
Probability
Population size (θ)
(B)
Probability
Divergence time (t)
(C)
—E. yayeyamana to E. formosa
—E. formosa to E. yayeyamana
Probability
Migration (m)
(D)
E. decorata vs. E. ornata
—E. decorata
—E. ornata
---Ancestral species
Probability
Population size (θ)
(E)
Probability
Divergence time (t)
(F)
—E. ornata to E. decorata
—E. decorata to E. ornata
Probability
Migration (m)

Figure 5

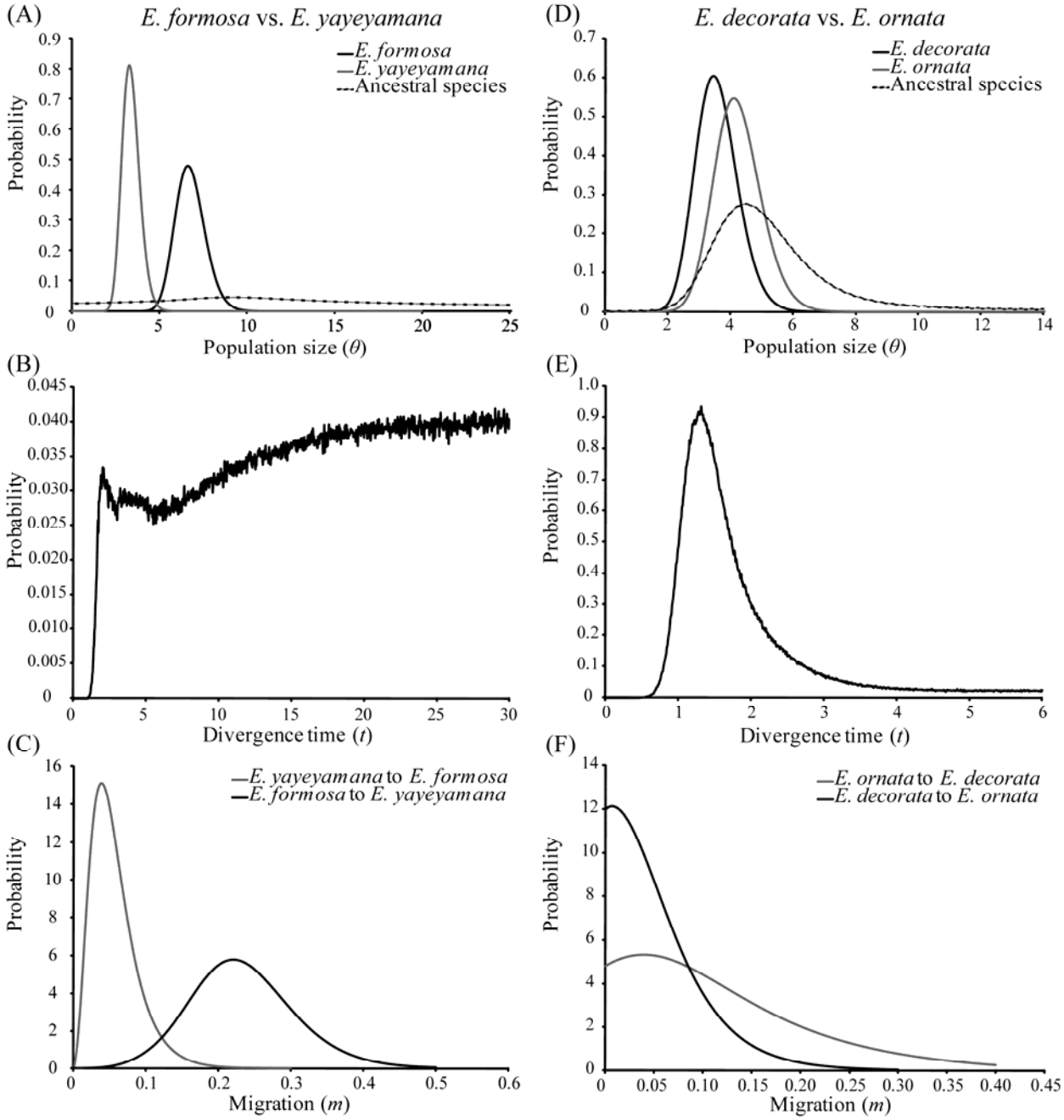

(A)
E. formosa vs. E. yayeyamana
—E. formosa
—E. yayeyamana
---Ancestral species
Probability
Population size (θ)
(B)
Probability
Divergence time (t)
(C)
—E. yayeyamana to E. formosa
—E. formosa to E. yayeyamana
Probability
Migration (m)
(D)
E. decorata vs. E. ornata
—E. decorata
—E. ornata
---Ancestral species
Probability
Population size (θ)
(E)
Probability
Divergence time (t)
(F)
—E. ornata to E. decorata
—E. decorata to E. ornata
Probability
Migration (m)

Figure 6

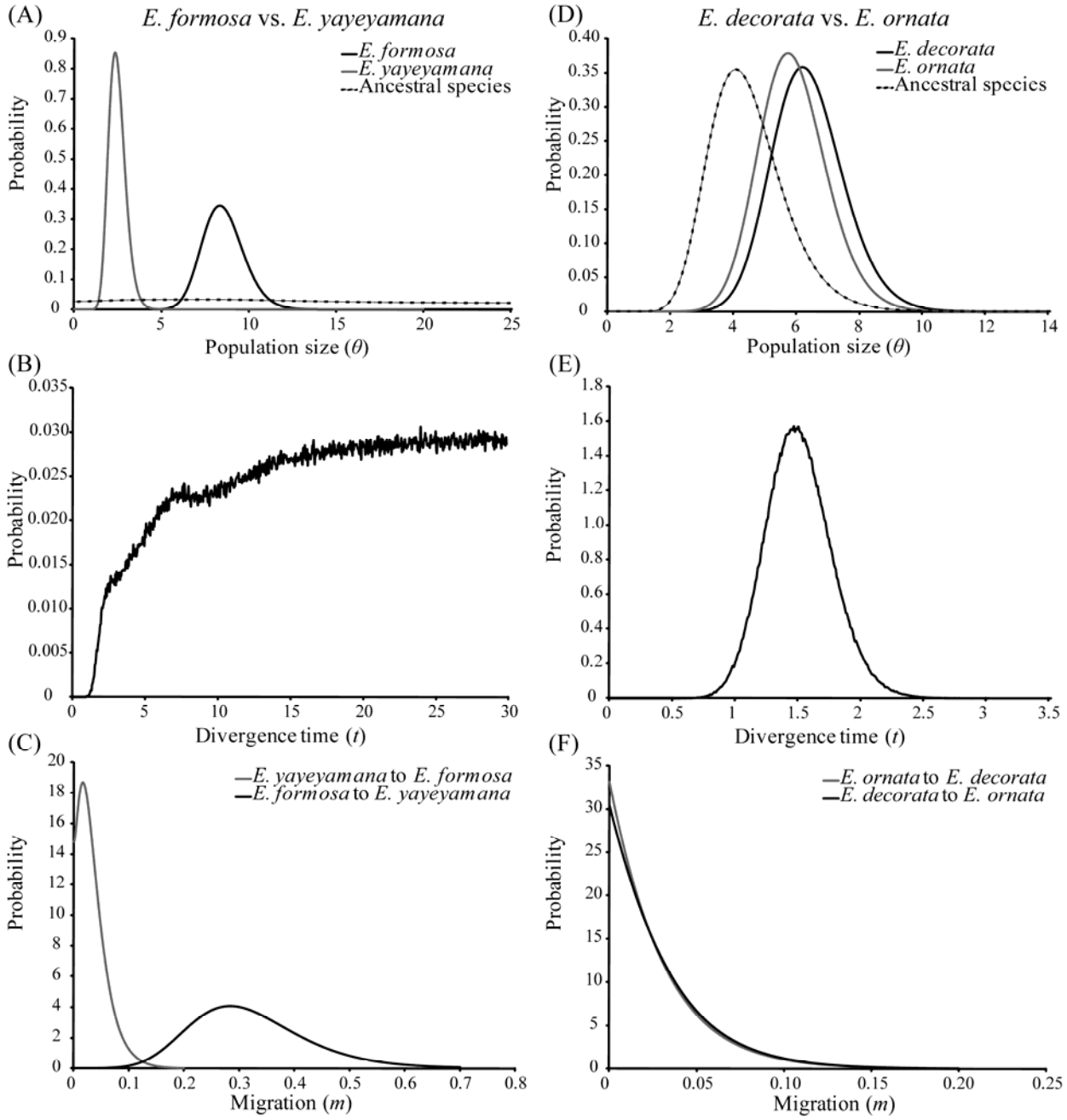
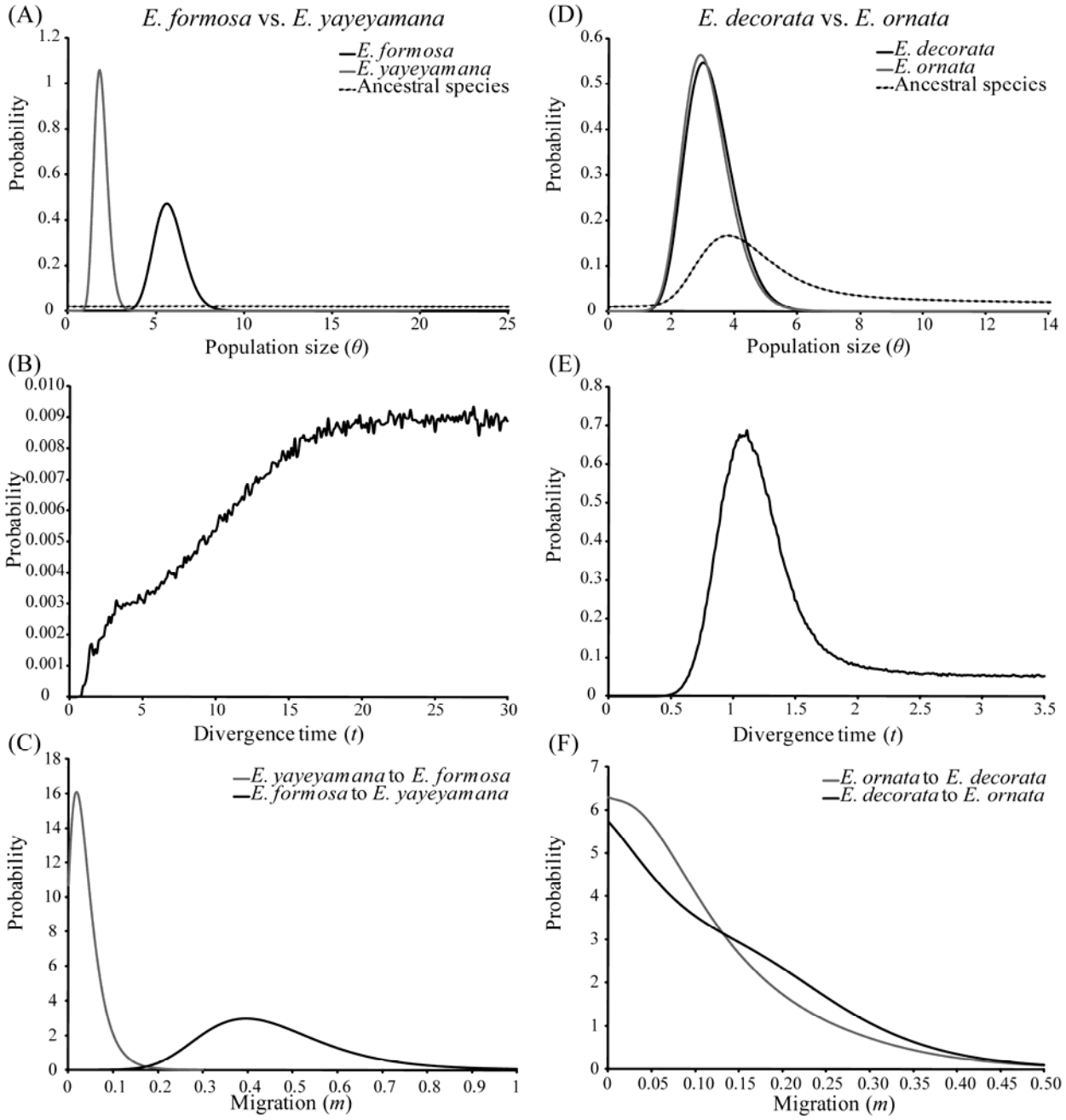

(A)
E. formosa vs. E. yayeyamana
E. formosa
E. yayeyamana
Ancestral species
Probability
Population size (θ)
(B)
Probability
Divergence time (t)
(C)
E. yayeyamana to E. formosa
E. formosa to E. yayeyamana
Probability
Migration (m)
(D)
E. decorata vs. E. ornata
E. decorata
E. ornata
Ancestral species
Probability
Population size (θ)
(E)
Probability
Divergence time (t)
(F)
E. ornata to E. decorata
E. decorata to E. ornata
Probability
Migration (m)

Figure 7


(A)
E. formosa vs. E. yayeyamana
E. formosa
E. yayeyamana
Ancestral species
Probability
Population size (θ)
(B)
Probability
Divergence time (t)
(C)
E. yayeyamana to E. formosa
E. formosa to E. yayeyamana
Probability
Migration (m)
(D)
E. decorata vs. E. ornata
E. decorata
E. ornata
Ancestral species
Probability
Population size (θ)
(E)
Probability
Divergence time (t)
(F)
E. ornata to E. decorata
E. decorata to E. ornata
Probability
Migration (m)