SCIENTIFIC REPORTS

Assortative mating between two sympatric closely-related specialists: inferred from molecular phylogenetic analysis and behavioral data

Huai-Jun Xue, Wen-Zhu Li & Xing-Ke Yang

Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China.

Host plant shifting of phytophagous insects can lead to the formation of host associated differentiation and ultimately speciation. In some cases, host plant specificity alone acts as a nearly complete pre-mating isolating barrier among insect populations. We here test whether effective pre-mating isolation and host-independent behavioral isolation have evolved under the condition of extreme host specilization using two sympatric flea beetles with incomplete post-mating isolation under laboratory conditions. Phylogenetic analysis and coalescent simulation results showed that there is a limited interspecific gene flow, indicating effctive isolation between these species. Three types of mating tests in the absence of host plant cues showed that strong host-independent behavioral isolation has evolved between them. We conclude that almost perfect assortative mating between these two extreme host specialists results from a combination of reduced encounter rates due to differential host preference and strong sexual isolation.

R eproductive isolation between closely related taxa may occur via multiple isolating barriers that occur either before or after mating[1]. In allopatric and peripatric speciation, both prezygotic and postzygotic reproductive isolation often evolves following geographic isolation due to divergent selection, genetic drift or combined effect[2]. For sympatric populations, one of the key mechanisms ensuring reproductive isolation is assortative mating[3] which may occur due to ecological, temporal or behavioral isolation[4].

Ecological speciation has been posited as an important mechanism creating new species via adaptation to different environments[5–7]. Small host-specific phytophagous insects were suggested to accounting for 25–40% of all animal species[8]. Some of this diversity may be attributed to differential host plant use that can lead to the formation of host associated forms and, ultimately, speciation. In many herbivore species that both feed and mate on the host plant[9,10], reproductively active individuals from populations associated with different host plants species are more likely to encounter potential mates using the same host. In this way, the tendency of adult insects to remain on or return to the larval host plant may result in patterns of host-associated assortative mating that restrict gene flow between populations[10]. In some cases, distinct host specificity alone was suggested to act as a nearly complete pre-mating isolating barrier among insect populations, resulting in reproductive isolation in the absence of behavioral isolation and any post-mating barriers[11–13]. For these herbivorous insects, pre-mating isolation may have evolved as a by-product of adaptation to different ecological environments[14]. However, these environment dependent barriers may not be effective over evolutionary time[15]. In most cases, multiple barriers to reproductive isolation are expected to be needed for speciation[16]. However, the pre-mating isolation barriers between sympatric sibling specialists were only checked in limited systems[11–13].

Pre-mating isolation due to host plant use may be occurring in the flea beetle genus *Altica*, a system that is becoming a model of host plant specialization and ecological speciation[17–20]. Many beetles in *Altica* are specialists, which feed and lay eggs on a limited range of plants. Among them, *Altica fragariae* Nakane (hereafter 'AF') is an oligophagous species with five recorded host plants belonging to the family Rosaceae, and *Duchesnea indica* (Andrews) Focke is the primary host plant in the field. *Altica viridicyanea* (Baly) (hereafter 'AV') is a monophagous species that feeds on *Geranium nepalens* (Sweet) (Geraniaceae)[18–20]. The beetles do not feed and oviposit on each other's host, even under no choice conditions, and the larvae do not develop on alternative host plants, indicating extreme host plant specificity[18,19].

Table 1 | Summary statistics for molecular variation in COI, EF1a and ITS2

		h	Hd	K	π
COI + COII	AF	9	0.32221	0.69710	0.00056
	AV	36	0.95333	7.03263	0.00564
EF1α	AF	1	0	0	0
	AV	8	0.5867	0.371	0.00031
ITS2	AF	1	0	0	0
	AV	1	0	0	0

These two species occur sympatrically both locally and across broad geographic scales. Previous studies have shown that post-mating isolation between them is incomplete. Under laboratory conditions, the hatch rate of hybrid eggs is as high as 65.8% when AV is the mother (although it is lower than the hatch rates in intraspecific crosses: 89.8% in AF, 89.2% in AV), whereas hatch rate is only 1.19% when AF is the mother. The backcross between AF females and F_1 (AV♀ × AF♂) males is inviable, whereas all other three backcrosses (F_1♀ × AF♂, AV♀ × F_1♂, F_1♀ × AV♂) and F_2 generation are viable. Furthermore, these viable offspring can develop on one or both host plants with high survival rates[18–20]. Interestingly, hybrid F_1 and backcrosses feed on both plants under choice conditions[18]. Feeding choice tests of the sympatric field-sampled beetles provided no evidence of heterospecific hybridization (at least no F_1 hybridization) (Xue HJ, unpublished data). Accordingly we speculate effective pre-mating reproductive isolation has evolved between them. These beetles' life histories and the circadian rhythms of sexual behavior are very similar, and their emergence phenology overlaps widely (Xue HJ, personal observation); thus, temporal isolation is likely not an important isolating mechanism. Therefore, habitat isolation, behavioral isolation or their joint action should play a crucial role in preventing interspecific gene flow. In the Beijing area, their host plants often grow side by side, and sympatric beetle populations are common, implying that interspecific hybridization may occur in nature. Considering the host-dependent assortative mating may not be perfect, we predicted behavioral isolation evolved besides habitat barriers. To test whether there is an effective pre-mating isolation between these two host specific lineages, we assessed the pattern of divergence and the level of gene flow between AF and AV using a combination of phylogenetic analyses and coalescent simulations. To test whether host-independent behavioral isolation has evolved between them, three kinds of mating tests that eliminated host plant cues were performed.

Results

Population structure. Summary statistics for molecular variation indicated much higher levels of genetic diversity in AV versus AF for mtDNA and EF1a, and no variation in ITS2 within both species (Table 1). Uncorrected divergence based on mtDNA between species clusters was 3.01% while within species divergence was only 0.27% (AF) and 0.67% (AV), respectively. Similarly, divergence in EF1α was 2.02% between species, only 0.23% within AV, and 0 within AF. Divergence in ITS2 ranged from 0.59% between species to zero divergence within AF and AV. AMOVA analyses showed that most of the genetic variance was attributable to the division between species (83.87% for COI-COII; 97.31% for EF1α and 100% for ITS2), and only a small proportion of the genetic variance was within species (Table 2). The analysis also showed that the genetic variance within populations contributed much more than that of among populations within species (Table 2). Significant Φst values (0.85 for COI-COII, 0.97 for EF1α and 1.00 for ITS2) revealed restricted gene flow between the two species.

The parsimony network for mitochondrial genes (COI-COII) showed of 45 mtDNA haplotypes in 144 individuals, and none these were shared by both species (Fig. 1a). The number of haplotypes within AV was much greater than that of AF (36 : 9), although the number of sampled individuals of AV and AF were similar (AV : AF = 76 : 68). In contrast to the mtDNA, we observed only nine haplotypes in EF1α. One of these haplotypes was present in all of the individuals of AF, whereas AV harbored the other eight haplotypes (Fig. 1b). Only two ITS2 haplotypes were recovered: all of the individuals of AV shared one and all of the individuals of AF shared the other (Fig. 1c).

Gene tree estimation. The phylogenetic analysis based on the COI-COII dataset showed that all the individuals of AF formed a well-supported clade that included five individuals of AV (Fig. 2a). The same tree topology was suggested with both phylogenetic reconstruction methods inferred from EF1α and ITS2. AF and AV were reciprocally monophyletic, strongly supported by high bootstrap or posterior probability values in the EF1α dataset (Fig. 2b), and weakly supported in the ITS2 dataset (Fig. 2c).

Coalescent simulations. Four-gametic tests suggested that there is no evidence of recombination for both EF1α and ITS2. The independent runs resulted in very similar parameter estimates. Based on mitochondrial genes, significant unidirectional migration was observed from AV to AF (LLR test ≥ 12.217, $p < 0.001$), but the results did not detect significant migration from AF to AV (LLR test = 0, $p \geq 0.05$) (Fig. 3a,b). In addition, the coalescent analysis based on the combined dataset suggested a similar pattern to that obtained with the mitochondrial dataset (AV to AF: LLR test ≥ 9.778, $p < 0.01$; AF to AV: LLR test = 0.651, $p \geq 0.05$) (Fig. 3c,d). However, individual analysis of the nuclear genes could not reject the zero-migration-rate model (LLR test = 0, $p \geq 0.05$) (Fig. 3e,f).

Strength of behavioral isolation. In the no-choice tests, we primarily observed intraspecific matings. We observed 64 intraspecific

Table 2 | Analysis of molecular variance for COI-COII, EF1a and ITS2 from two *Altica* species, AF and AV

|	Source of variation	Sum of squares	Variance components	Percentage variation	Fixation index (Φst)	P-value
COI-COII	Among species	1250.514	17.32702	83.87396	0.85101	<0.001
	Among populations within species	53.618	0.25344	1.22680		
	Within populations	412.444	3.07794	14.89924		
EF1α	Among species	1003.762	14.06850	97.30613	0.97248	<0.001
	Among populations within species	2.237	−0.00833	−0.05764		
	Within populations	52.909	0.39781	2.75152		
ITS2	Among species	574.222	8.00000	100	1.00000	<0.001
	Among population within species	0	0	0		
	Within populations	0	0	0		
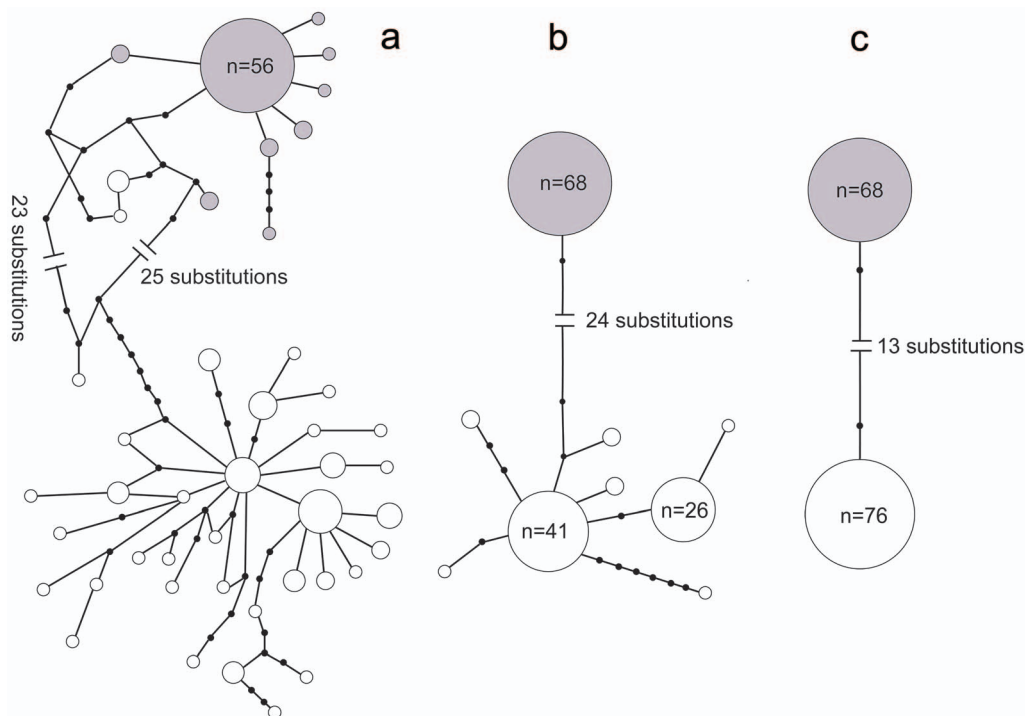
Figure 1 | Maximum parsimony networks for three loci (COI-COII, EF1α and ITS2). *Altica fragariae* (AF) alleles are illustrated in grey, and *A. viridicyanea* (AV) alleles are illustrated in white. The area of each circle is proportional to the number of individuals with that haplotype. Straight lines and small black dots reflect mutations and median vectors, respectively. (a) based on COI-COII; (b) based on EF1α; (c) based on ITS2.

matings (39 AF♀-AF♂ and 25 AV♀-AV♂) as compared to 9 interspecific matings (all of them are AV♀ × AF♂) (Table 3). The overall sexual isolation value (I_{PSI}) was 0.7863 ± 0.0579 ($p < 0.001$). In male-choice tests, we observed only one interspecific mating (AV♀ × AF♂) as compared to 88 intraspecific mating events (42 AF-AF and 46 AV-AV) (Table 4) ($I_{PSI} = 0.9670 \pm 0.0269$, $p < 0.001$). In the multi-choice mating tests, 71 mating events were observed during the trials, 69 of them were intraspecific (42 AF♀-AF♂ and 27 AV♀-AV♂) and only 2 were interspecific (2 AV♀ × AF♂) (Table 5) ($I_{PSI} = 0.9330 \pm 0.0407$, $p < 0.001$).

Discussion

In contrast to the mito-nuclear concordance observed in a previous study dealing with the same system (but with a third species and small sample sizes, n = 9 for AF and n = 5 for AV)[20] and another complex of specialist *Altica* species[21], the present study showed a small discrepancy between mitochondrial and nuclear gene trees. In the mitochondrial analyses, five AV individuals always fell in the AF clade, suggesting some gene migration. However, in the nuclear gene trees (both ITS2 and EF1α), AV and AF were reciprocally monophyletic (Fig. 2a–c), indicative of complete reproductive isolation. Similarly, IMa analysis based on nuclear data suggested there is no gene migration between the beetle species, while the results inferred from mitochondrial and combined data showed unidirectional gene flow (effective gene flow from AV to AF, $2Nm = 0.8027$ and 0.09338 respectively). Therefore, phylogenetic conclusions that rely on mitochondrial or nuclear marker alone maybe misleading[22–25]. Thus, a combination of markers from both the nuclear and mitochondrial genome provides a more comprehensive view of evolutionary history.

It is difficult to distinguish introgression from incomplete lineage sorting in many studies[26,27]. It was suggested that lineage sorting should proceed faster in mitochondrial genes than in nuclear genes because of the smaller effective population size and uniparental inheritance of mitochondrial genes[28,29]. In the present study, summary statistics for molecular variation indicated levels of genetic diversity in mitochondrial DNA are much higher than in nuclear DNA (Table 1). The absence of shared haplotypes (Fig. 1b,c) and distinct clades (Fig. 2b,c) showed that lineage sorting is complete in the two nuclear genes, while paraphyly of AF was detected in mitochondrial gene (Fig. 1a,2a). Furthermore, considering the incomplete post-mating isolation and sympatric distribution of these two species, we presume that introgression of mtDNA is a more reasonable explanation than incomplete lineage sorting. In fact, insect species in nature are often incompletely isolated for millions of years after their formation[30,31] and gene flow may persist even between non-sibling species[32]. Mitochondrial DNA introgression has been reported in a number of recent studies[33–35].

The mitochondrial genome can be replaced by that of another species as a result of historical hybridization without leaving any trace in the nuclear genome. At least four mechanisms that might result in introgressive hybridization of mtDNA have been suggested: reproductive asymmetry[36], differential selection[37], chance fixation by genetic drift in small populations[38], and paternal leakage in mtDNA[39,40]. In the two *Altica* species, significant asymmetry occurred during the pre-mating stage of isolation. Although the heterospecific mating was not detected for AV males, the males of AF occasionally mated with sibling species females under both no-choice and choice conditions in the laboratory (Table 3–5). Furthermore, previous cross-breeding experiments showed that hybrid F₁ with AV mother and three backcrosses (F₁♀ × AF♂, F₁♀ × AV♂, AV♀ × F₁♂) were viable, while F₁ with AF mother and one backcross (AF♀ × F₁♂) were inviable, suggesting asymmetric post-mating isolation between AF and AV due to cytoplasmic incompatibility[19,20]. These results are consistent with asymmetric mitochondrial introgression estimated by coalescent analysis. Thus, reproductive asymmetry is the most probable scenario in the present system.

The molecular results (high Φst values estimated by AMOVA and low M values estimated by IMa2) suggested very limited gene flow between these two species. Considering the incomplete post-mating
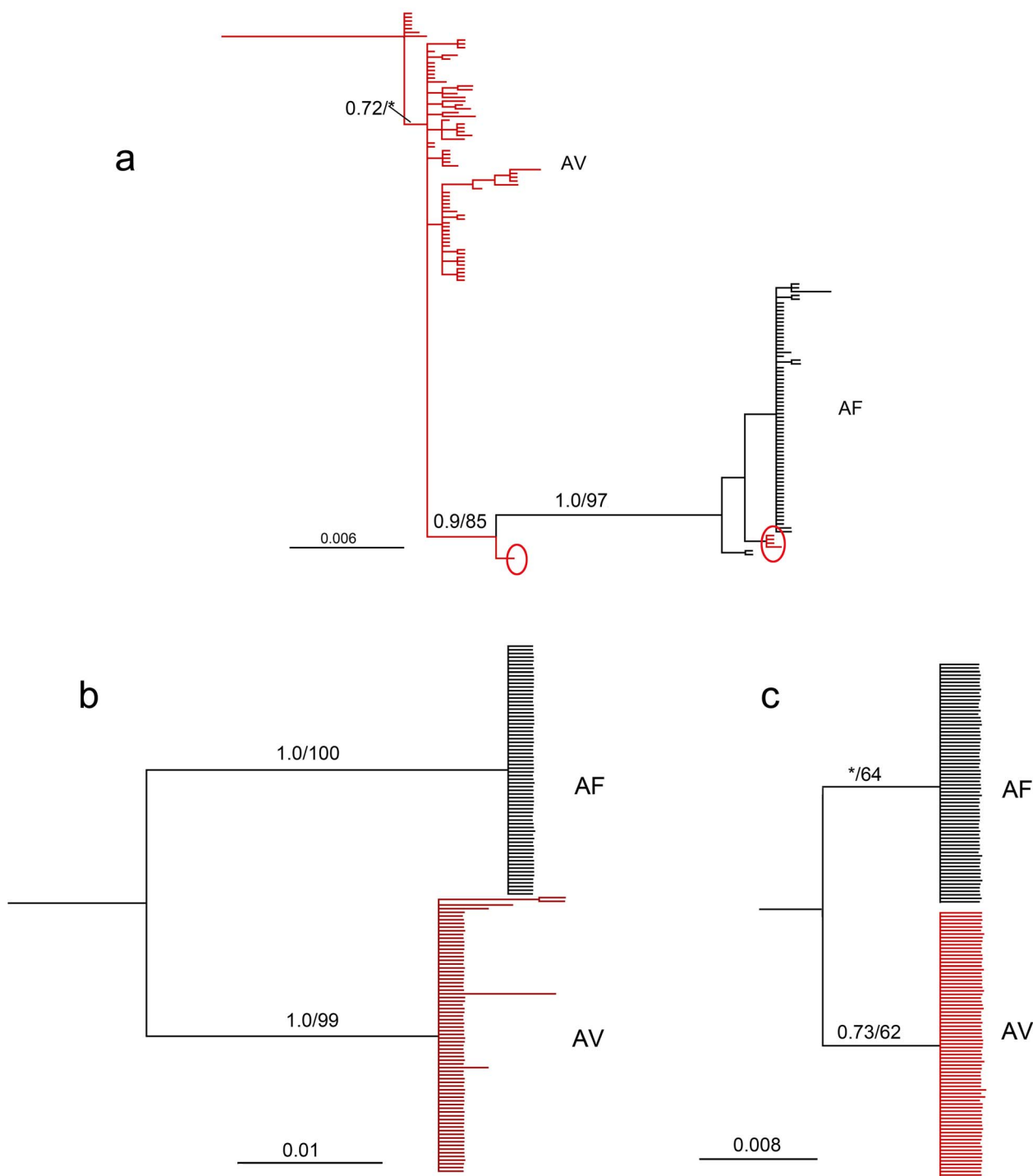
Figure 2 | **Phylogenetic relationships of two *Altica* species.** The numbers above the branches indicate Bayesian posterior probabilities and maximum likelihood bootstrap values. (a) Bayesian consensus tree inferred based on COI-COII dataset. The left oval indicates individual number 0957 and the right oval shows the individuals 1032, 1052, 1055 and 1057. (b) Bayesian consensus trees inferred based on EF1α dataset; (c) Bayesian consensus trees inferred based on ITS2 dataset. The outgroup AK and the terminal labels were removed. The nodes of major branches with ≥50% bootstrap value or posterior probability are labeled. The symbol * is used to represent a node in conflict between the different analyses with less than 50% posterior probability or bootstrap support value.

reproductive isolation[18–20], we suggest near perfect assortative mating has evolved. Differing from many studies in which allochronic factors play an important role in reproductive isolation and speciation[41], allochronic isolation contribute little in the present system. As many other herbivorous specialists, *Altica* flea beetles exhibit high host fidelity[19], and the spatial segregation of the host plant species reduces encounters between the two beetles. Therefore, one explanation for

the reproductive isolation observed between these beetle species is that the preference of beetles for feeding and mating sites leads to a strong assortative meeting, and then assortative mating. However, their host plants often grow side by side, and sympatric beetle populations are common, implying that interspecific meeting may occur in nature. Then considerable gene flow between these two beetles should be detected if host independent behavioral isolation contrib-
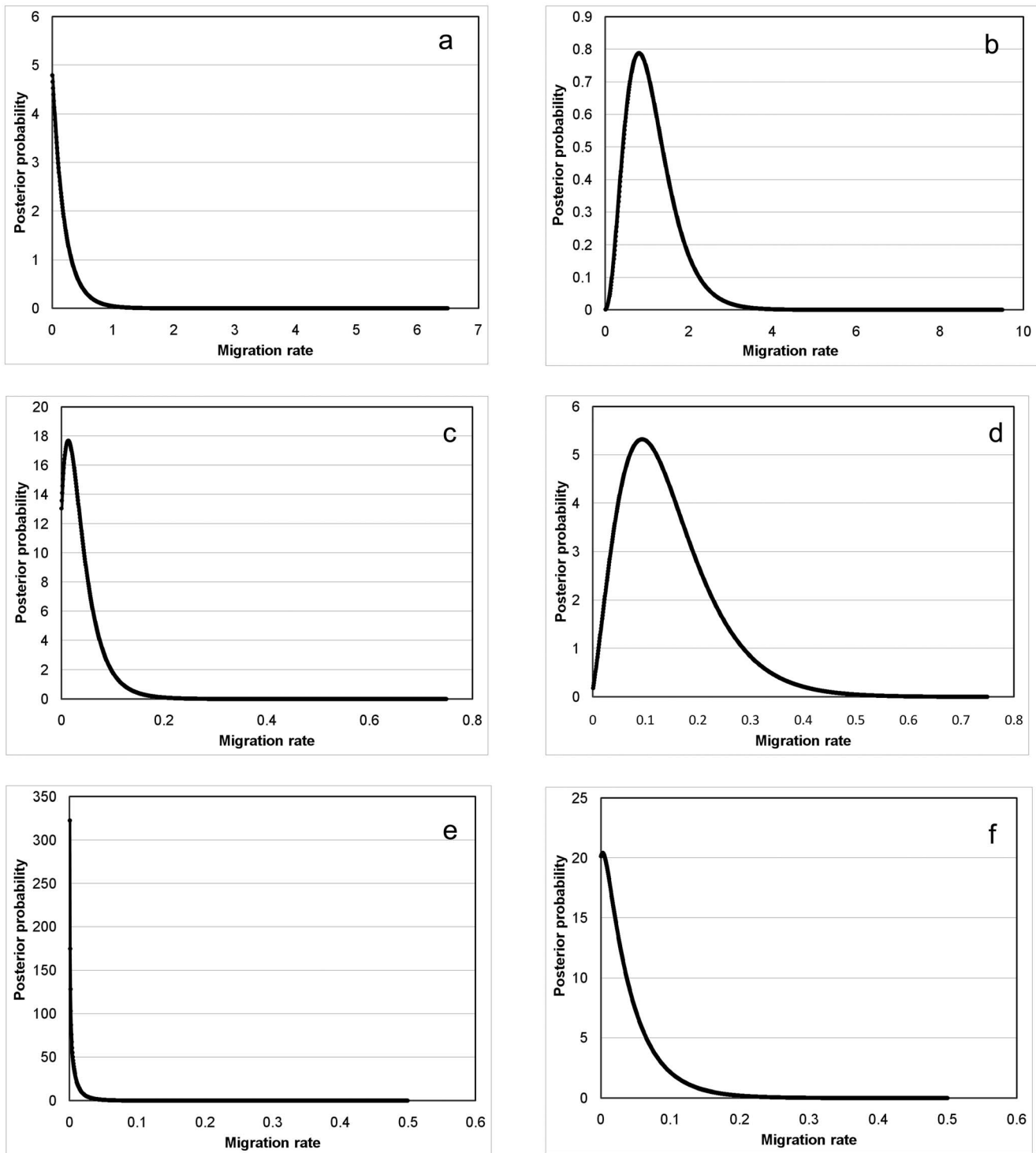
Figure 3 | Migration rates between AF and AV estimates from the IMa2 analysis using three datasets (mitochondrial, combined and nuclear data). (a) from AF to AV, based on mitochondrial data, (b) from AV to AF, based on mitochondrial data, (c) from AF to AV, based on combined data, (d) from AV to AF, based on combined data, (e) from AF to AV, based on nuclear data, (f) from AV to AF, based on nuclear data. Note difference in scale in the Y-axis.

ute negligibly. Both no choice and choice mating tests without host cues showed only limited heterospecific mating. The close, artificial conditions of these experiments were likely an overestimate of the probability of heterospecific mating, further indicating almost perfect behavioral isolation has evolved between these species.

We thus suggested almost perfect assortative mating between these two flea beetles resulted from a combination of reduced encounter rates due to differential host preference and strong sexual

isolation. One explanation for the behavioral reproductive isolation is a by-product of adaptation to the different hosts, even in the absence of direct selection[42]. And another, but more reasonable explanation is, behavioral isolation resulted from host-independent sexual selection acting or indirect effects of host-associated selection on sexually selected traits. For these two species, divergent contact sex pheromones seem to play a pivotal role in mating behavior (Xue HJ, unpublished data). Previous studies have suggested that some

Table 3 | Results of no-choice mating test. One male and one female were placed in a 9-cm Petri dish lined with moistened filter papers and given two hours to mate. Four comparisons and 60 replicates for each combination were conducted

Combines	Number of trials	Successful mating number during the test
AF♂ + AF♀	60	39
AV♂ + AV♀	60	25
AF♂ + AV♀	60	9
AV♂ + AF♀	60	0

Table 5 | Results of multi-choice mating test. Five pairs of AF and AV (5AF♂ + 5AF♀ + 5AV♂ + 5AV♀) were placed in a glass jar simultaneously and allowed to mate freely for three hours, 18 replicates were conducted; thus, a total of 180 mating events were possible

Male	Female	
	AF♀	AV♀
AF♂	42	2
AV♂	0	27

phytophagous insects may acquire bioactive chemicals or required chemical precursors of sex pheromones from host plant material[43]. Therefore, sex pheromone composition may change following a host shift, subsequently leading to host plant mediated sexual isolation[44]. Further chemical ecological studies are needed to estimate whether and how host plants affect host-associated divergence and speciation in these closely related *Altica* flea beetles.

Methods

Potential interspecific gene flow estimation. *Field collections and identification.* Adults of *A. fragariae* (n = 68) and *A. viridicyanea* (n = 76) were sampled from five sympatric locations in the Beijing area (Table 6). These two species are too similar to be distinguished by morphological characters, especially for the females. Furthermore, just using host plant during collection as a way to identify the beetle is not safe because there may be some sporadic wandering individuals in the non-host plant, and some beetles may jump to the ground under mixed plants when they were disturbed, therefore we identified individuals using a 5-day feeding-choice tests[18].

DNA extraction, amplification, and sequencing. Total genomic DNA was extracted from whole beetles using nondestructive DNA extraction method with the TIANamp Genomic DNA Kit (Tiangen, Shanghai, China). After extraction, beetle specimens were retained as vouchers and were deposited in the Institute of Zoology, Chinese Academy of Sciences, Beijing, China.

Three gene regions were used in the present study: a mitochondrial region spanning parts of cytochrome oxidase I and II (COI-COII, 1368 bp), an internal transcribed spacer of the nuclear ribosomal RNA cluster (ITS2, 354 bp), and a fragment of the nuclear protein-coding gene elongation factor 1-alpha (EF1α, ~1221 bp with a ~525 bp intron). PCR protocols and primers were the same as Xue *et al.* (2011)[20]. The PCR products were sent to Beijing Sunbiotech Co. Ltd. for sequencing with both forward and reverse primers. The sequence data were aligned and edited using CodonCode Aligner 3.7.1 (CodonCode, Dedham, MA, USA). Variable sites were identified using the automated mutation detection function in CodonCode Aligner, followed by manual inspection of electropherograms and additional re-sequencing of ambiguous bases. Successful PCR amplification and sequencing of target genes were confirmed by aligning the resulting sequences to other closely related species in the Chrysomelidae, verifying the correct reading frame without stop codons. All EF1α and ITS2 sequences were apparently homozygous, as judged on the basis of a lack of double peaks in chromatograms from both directions.

Population structure. After assembling aligned DNA sequence data matrices, we used DNaSP 5.1[45] to calculate the following summary statistics for each gene in the total sample of individuals of each species separately: number of haplotypes (*h*); haplotypic diversity (*H*d, Nei 1987); average number of nucleotide differences (*K*, Tajima 1983); and nucleotide diversity (π, Nei 1987). To estimate genetic divergence within and between two species, we analyzed the genetic distances with haplotypes data using MEGA 4.0[46]. To characterize genetic differentiation between species, among populations within species and within populations, and examine how these patterns varied among loci, we also assessed molecular variance (locus-by-locus AMOVA, with 1000 permutations) and calculated Φst values between species using Arlequin 3.5[47]. The distribution of haplotypes between these two beetles was also summarized using statistical parsimony networks inferred in TCS 1.2.1[48]. All haplotypes sequences are deposited in GenBank under the accession numbers KJ803199-KJ803254.

Gene tree estimation. Gene trees were estimated with the following three combinations: COI-COII (along with the intervening leucine tRNA, COI + Leu +

COII), EF1α (exon + intron) and ITS2. The best-fit model of nucleotide substitution for each partition or each combination was selected using the Akaike Information Criterion (AIC) in jMODELTEST 0.1.1[49]. *Altica koreana* was used as the outgroup, and the associated sequences are under the following accession numbers in GenBank: DQ865067 (COI-COII), JN903082 (EF1α) and JN903103 (ITS2). Maximum likelihood (ML) analyses were performed with Garli 2.0[50], 1000 bootstrap replications were used to obtain a set of trees, and the bootstrap consensus tree was constructed with PAUP* 4.0[51]. Bayesian analyses were carried out with the program MRBAYES version 3.1[52,53]. The settings were two simultaneous runs (each with two Markov chains) of the Markov chain Monte Carlo (MCMC) for 6–15 × 10^6 generations (making sure the average standard deviation of split frequencies was less than or approaching 0.01), sampling every 100 generations. The first 25% of the generations were discarded as the burn-in. Log likelihood plots of trees from the Markov chain samples were examined in TRACER version 1.5[54] to determine convergence to a stable log likelihood value. Posterior probability (PP) values represent the proportion of MCMC samples that contain a particular node. The default branch length priors in MrBayes may lead to unreasonable estimates of substitution rates and the bias of PPs; thus, empirical priors for this parameter were defined based on the mean branch lengths recovered from maximum likelihood topologies[55] inferred using Garli 2.0.

Coalescent simulations. We estimated gene flow between AF and AV using IMa2, which allows for analysis of divergence and gene flow between two or more populations[56]. Because the IM model assumes that all sequences are free from intra-locus recombination, before using IMa2, we tested nuclear genes for recombination[57] using DNaSP 5.1. Preliminary simulations were run with a wide range of priors to explore the sensitivity of parameter estimates to different upper bounds. We set the maximum values for q = 1000 in all the final runs. The other two parameters (t, the splitting time parameter; and m, the migration parameter) in the analysis were set to the following values: t = 100 and m = 2 for mitochondrial DNA, t = 600 and m = 2 for nuclear DNA, and t = 600 and m = 3 for the total dataset. The HKY model was chosen for each locus. We performed at least five independent runs of twenty million steps (recording every 100 steps), with a burn-in period of two million steps per run, using different random seeds with 10 or 20 Metropolis coupled chains and geometric heating schemes with low heating parameters (a = 0.96, b = 0.9). The analysis was considered to have converged upon a stationary distribution if the independent runs showed effective sample size (ESS) values above 100[58]. Log-likelihood-ratio (LLR) statistics were used to test whether contemporary gene migration rates were greater than zero[59].

Strength of sexual isolation. As in many other beetles, both sexes of *Altica* species mate multiple times, the courtship behavior is simple, and mating typically lasts several hours. Mating is initiated by the male by first antennating a female, then mounting the female's back, and finally, attempting to insert his intromittent organ, the aedeagus. As a consequence, male mate choice is an important stage of pre-mating reproductive isolation in *Altica* species. For sympatric populations, choice mating experiments seem to be more relevant, however, when beetles occur at low densities, a male encountering females of closely-related species alone is also a possible scenario, thus the no-choice mating experiments also should be considered. To gain a comprehensive estimation of the strength of behavioral isolation, we used three types of mating assays that excluded host plants: no-choice, male-choice, and multi-choice experiments. We wanted to use both choice and no-choice assays because it has been shown in *Drosophila* that the strength of behavioral isolation may depend strongly on the experimental design, especially the possibility of choice: stronger sexual isolation was suggested when flies were allowed to choose between conspecific and heterospecific mates than those in no-choice trials[60].

Table 4 | Results of male-choice mating test. One male and two females were placed in a 9-cm Petri dish arena lined with moistened filter papers for three hours. Two comparisons and 100 replicates for each combination were carried out

Combines	Number of trials	Successful mating during tests	Intraspecific mating	Interspecific mating
AF♂ + AF♀AV♀	100	43	42	1
AV♂ + AF♀AV♀	100	46	46	0

Table 6 | Location, sample size and sampling date of the *Altica* beetles used in population genetic study

Location	Geographic coordinates	Samples	Sampling date
Taotiaogou, Yanqing	40.635′N, 116.537′E	AF (n = 15): 0924–0938 AV (n = 19): 0939–0957	2009-IX-25
Beikouzi, Yanqing	40.539′N, 116.417′E	AF (n = 12): 0961–0970, 1014–1015, AV (n = 14): 0958–0960, 1003–1013	2009-IX-26, 2010-VI-10
Sijiashui, Mentougou	40.091′N, 115.948′E	AF (n = 16): 1016–1019, 1124–1129, 1131–1136 AV (n = 14): 1020–1028, 1137–1141	2010-IX-2, 2011-VIII-11
Matao, Mentougou	40.142′N, 115.787′E	AF (n = 16): 1034–1048, 1059 AV (n = 15): 1029–1033, 1049–1058	2010-IX-9, 2010-IX-15
Qianjiadian, Yanqing	40.697′N, 116.383′E	AF (n = 9): 1101–1109 AV (n = 14): 1110–1123	2011-VIII-1

All mating tests were carried out between 13:00 and 18:00 hours in an air-condition room at 25–27°C under natural light conditions. In all experiments, a successful mating was scored whenever pairing occurred for longer than 5 minutes.

No-choice mating tests. More than 30 over-wintered adults of *A. fragariae* and *A. viridicyanea* were collected from field populations on their host plants at Matao, Mentougou (40.14°N, 115.79°E) and Sijiashui, Mentougou (40.09°N, 115.95°E) of the Beijing area respectively, and subsequently maintained in the laboratory. The insects were fed their respective host plants (*D. indica* fed to AF and *G. nepalens* fed to AV) in growth chambers at 16 : 8 LD and 25°C[18–20]. To eliminate the possible effects due to different physiology, virgin captive-bred offspring were used for no-choice tests. Shortly after eclosion, adults were sexed and kept separately for 10 days after eclosion to ensure sexual maturity. One male and one female were placed in a 9-cm Petri dish lined with moistened filter papers and given two hours to mate. Four pairwise comparisons (AF♀ × AF♂; AV♀ × AV♂; AF♀ × AV♂; AV♀ × AF♂) and 60 replicates for each combination were conducted. The number of successfully copulating pairs was recorded. Each individual was tested only once.

Male-choice mating tests. Wild collected adults were used for a male-choice test. In order to increase mating motivation and ensure sexual maturity, the sexes and species were kept in separate glass jars for 10 days before mating tests. Because the species are difficult to distinguish based on appearance, we marked them with different colour enamel paint on their elytra to enable subsequent identification[61]. A pilot experiment showed that the enamel paint does not alter mating propensity in this system (Xue HJ, unpublished data). Males were provided with a choice between a female from each species. One male and two females (AF♂ + AF♀ + AV♀; AV♂ + AF♀ + AV♀) were placed in a 9-cm Petri dish arena lined with moistened filter papers. The number of successfully copulating pairs was recorded over a period of three hours. Both combinations were carried out for 100 replicates, respectively.

Multi-choice mating tests. Wild collected adults were also used for multi-choice tests. Five males and five females of each species (20 individuals total) marked with different colored enamel paint on the elytra were placed in a glass jar (11.5 cm tall by 12 cm in diameter) lined with moistened filter papers. The beetles were allowed to mate freely for three hours. Copulating pairs were removed from the Petri dish and the paired combination was recorded. Eighteen replicates were run with different individuals, thus, a total of 180 mating events were possible.

Statistics. To assess the degree of sexual isolation between AF and AV, we used the software JMATING 1.0.8[62] to estimate *IPSI*, an estimator proposed by Rolán-Alvarez & Caballero[63]. The index of sexual isolation, *IPSI*, ranges from −1 to 1, where −1 is complete disassortative mating, 0 is random mating, and 1 is complete assortative mating. Standard deviations (*SD*) and tests of significance for total *IPSI* were obtained by bootstrapping with 10000 bootstrap iterations[62].

1. Bailey, R. I., Thomas, C. D. & Butlin, R. K. Premating barriers to gene exchange and their implications for the structure of a mosaic hybrid zone between *Chorthippus brunneus* and *C. jacobsi* (Orthoptera: Acrididae). *J. Evol. Biol.* **17**, 108–119 (2004).
2. Coyne, J. A. & Orr, H. A. Patterns of speciation in *Drosophila. Evolution* **43**, 362–381 (1989).
3. Malausa, T. *et al.* Assortative mating in sympatric host races of the European corn borer. *Science* **308**, 258–260 (2005).
4. Funk, D. J., Filchak, K. E. & Feder, J. L. Herbivorous insects: model systems for the comparative study of speciation ecology. *Genetica* **116**, 251–267 (2002).
5. Nosil, P., Crespi, B. J. & Sandoval, C. P. Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* **417**, 441–443 (2002).
6. Rundle, H. D. & Nosil, P. Ecological speciation. *Ecol. Lett.* **8**, 336–352 (2005).
7. Schluter, D. Evidence for ecological speciation and its alternative. *Science* **323**, 737–741 (2009).
8. Berlocher, S. H. & Feder, J. L. Sympatric speciation in phytophagous insects: moving beyond controversy? *Annu. Rev. Entomol.* **47**, 773–815 (2002).
9. Matsubayashi, K. W., Ohshima, I. & Nosil, P. Ecological speciation in phytophagous insects. *Entomol. Exp. Appl.* **134**, 1–27 (2010).
10. Feder, J. L. *et al.* Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *PNAS* **91**, 7990–7994 (1994).
11. Katakura, H., Shioi, M. & Kira, Y. Reproductive isolation by host specificity in a pair of phytophagous ladybird beetles. *Evolution* **43**, 1045–1053 (1989).
12. Via, S. Reproductive isolation between sympatric races of pea aphids. I. Gene flow and habitat choice. *Evolution* **53**, 1446–1457 (1999).
13. Matsubayashi, K. W., Kahono, S. & Katakura, H. Divergent host plant specialization as the critical driving force in speciation between populations of a phytophagous ladybird beetle. *J. Evol. Biol.* **24**, 1421–1432 (2011).
14. McKinnon, J. S. *et al.* Evidence for ecology's role in speciation. *Nature* **429**, 294–298 (2004).
15. Futuyma, D. J. [Macroevolutionary consequences of speciation: inference from phytophagous insects] *Speciation and Its Consequences* [Otte, D. & Endler, J. A. (eds.) [557–579] (Sinauer Associates, Sunderland, 1989).
16. Matsubayashi, K. W. & Katakura, H. Contribution of multiple isolating barriers to reproductive isolation between a pair of phytophagous ladybird beetles. *Evolution* **63**, 2563–2580 (2009).
17. Xue, H. J. & Yang, X. K. Host plant use in sympatric closely related flea beetles. *Environ. Entomol.* **36**, 468–474 (2007).
18. Xue, H. J., Li, W. Z. & Yang, X. K. Genetic analysis of feeding preference in two related species of Altica (Coleoptera: Chrysomelidae: Alticinae). *Ecol. Entomol.* **34**, 74–80 (2009).
19. Xue, H. J., Magalhães, S., Li, W. Z. & Yang, X. K. Reproductive barriers between two sympatric beetle species specialized on different host plants. *J. Evol. Biol.* **22**, 2258–2266 (2009).
20. Xue, H. J., Li, W. Z., Nie, R. E. & Yang, X. K. Recent speciation in three closely related sympatric specialists: inferences using multi-locus sequence, post-mating isolation and endosymbiont data. *PLoS ONE* **6**, e27834 (2011).
21. Jenkins, T. M. *et al.* Insights into flea beetle (Coleoptera: Chrysomelidae: Galerucinae) host specificity from concordant mitochondrial and nuclear DNA phylogenies. *Ann. Entomol. Soc. Am.* **102**, 386–395 (2009).
22. Shaw, K. L. Conflict between mitochondrial and nuclear DNA phylogenies of a recent species radiation: what mitochondrial DNA reveals and conceals about modes of speciation in Hawaiian crickets. *PNAS* **99**, 16122–16127 (2002).
23. Segraves, K. A. & Pellmyr, O. Testing the out-of-Florida hypothesis on the origin of cheating in the yucca–yucca moth mutualism. *Evolution* **58**, 2266–2279 (2004).
24. Linnen, C. R. & Farrell, B. D. Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. *Evolution* **61**, 1417–1438 (2007).
25. Sequeira, F. *et al.* Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. *BMC Evol. Biol.* **11**, 264 (2011).
26. Buckley, T. R., Cordeiro, M., Marshall, D. C. & Simon, C. Differentiating between hypotheses of lineage sorting and introgression in New Zealand Alpine cicadas (*Maoricicada* Dugdale). *Syst. Biol.* **55**, 411–425 (2006).
27. Koblmüller, S., Egger, B., Sturmbauer, C. & Sefc, K. M. Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Mol. Phylogenet. Evol.* **55**, 318–334 (2010).
28. Palumbi, S. R., Cipriano, F. & Hare, M. P. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution* **55**, 859–868 (2001).
29. Peters, J. L., Zhuravlev, Y., Fefelov, I., Logie, A. & Omland, K. E. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (Anas spp.). *Evolution* **61**, 1992–2006 (2007).
30. Mallet, J. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **20**, 229–237 (2005).
31. Kronforst, M. R. Gene flow persists millions of years after speciation in *Heliconius* butterflies. *BMC Evol. Biol.* **8**, 98 (2008).
32. Bull, V. *et al.* Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Evol. Biol.* **4**, 11 (2006).
33. Sota, T. & Vogler, A. P. Incongruence of mitochondrial and nuclear gene trees in the Carabid beetles *Ohomopterus. Syst. Biol.* **50**, 39–59 (2001).

34. Bossu, C. M. & Near, T. J. Gene trees reveal repeated instances of mitochondrial DNA introgression in Orangethroat Darters (Percidae: *Etheostoma*). *Syst. Biol.* **58**, 114–129 (2009).
35. Marshall, D. C., Hill, K. B. R., Cooley, J. R. & Simon, C. Hybridization, mitochondrial DNA phylogeography, and prediction of the early stages of reproductive isolation: lessons from New Zealand cicadas (genus *Kikihia*). *Syst. Biol.* **60**, 482–502 (2011).
36. Rohwer, S., Bermingham, E. & Wood, C. Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. *Evolution* **55**, 405–422 (2001).
37. Gerber, A. S., Loggins, R., Kumar, S. & Dowling, T. E. Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? *Annu. Rev.Genet.* **35**, 539–566 (2007).
38. Kawakami, T. *et al.* Differential gene flow of mitochondrial and nuclear DNA markers among chromosomal races of Australian morabine grasshoppers (Vandiemenella, viatica species group). *Mol. Ecol.* **16**, 5044–5056 (2007).
39. Andolfatto, P., Scriber, J. M. & Charlesworth, B. No association between mitochondrial DNA haplotypes and a female-limited mimicry phenotype *in Papilio glaucus*. *Evolution* **57**, 305–316 (2003).
40. Fontaine, K. M., Cooley, J. R. & Simon, C. Evidence for paternal leakage in hybrid periodical cicadas (Hemiptera: *Magicicada spp.*). *PLoS ONE* **2**, e892 (2007).
41. Ording, G. J., Mercader, R. J., Aardema, M. L. & Scriber, J. M. Allochronic isolation and incipient hybrid speciation in tiger swallowtail butterflies. *Oecologia* **162**, 523–531 (2010).
42. Dodd, D. M. B. Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. *Evolution* **54**, 1308–1311 (1989).
43. Landolt, P. J. & Phillips, T. W. Host plant influences on sex pheromone behavior of phytophagous insects. *Annu. Rev. Entomol.* **42**, 371–391 (1997).
44. Geiselhardt, S., Otte, T. & Hilker, M. Looking for a similar partner: host plants shape mating preferences of herbivorous insects by altering their contact pheromones. *Ecol. Lett.* **15**, 971–977 (2012).
45. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
46. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599 (2007).
47. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
48. Clement, M., Posada, D. & Crandall, K. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1660 (2000).
49. Posada, D. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256 (2008).
50. Zwickl, D. J. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion* (PhD dissertation, The University of Texas at Austin, 2006).
51. Swofford, D. L. *PAUP*. Phylogenetic analysis using parsimony (*and other methods), Version 4.0.* (Sunderland: Sinauer Associates, Massachusett, 2003).
52. Huelsenbeck, J. P. & Ronquist, F. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
53. Ronquist, F. R. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
54. Rambaut, A. & Drummond, A. J. Tracer v1.5. (2009) Available at http://tree.bio.ed.ac.uk/software/tracer/ [Date of access:30/09/2010].
55. Drummond, C. S., Xue, H. J., Yoder, J. B. & Pellmyr, O. Host–associated divergence and incipient speciation in the yucca moth *Prodoxus coloradensis* (Lepidoptera: Prodoxidae) on three species of host plants. *Heredity* **105**, 183–196 (2010).
56. Hey, J. Isolation with migration models for more than two populations. *Mol. Biol. Evol.* **27**, 905–920 (2010).
57. Hudson, R. R. & Kaplan, N. L. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**, 147–164 (1985).
58. Kuhner, M. K. & Smith, L. P. Comparing likelihood and bayesian coalescent estimation of population parameters. *Genetics* **175**, 155–165 (2007).
59. Hey, J. & Nielsen, R. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *PNAS* **104**, 2785–2790 (2007).
60. Coyne, J. A., Elwyn, S. & Rolán-Alvarez, E. Impact of experimental design on *Drosophila* sexual isolation studies: direct effects and comparison to field hybridization data. *Evolution* **59**, 2588–2601 (2005).
61. Wood, T. K., Tilmon, K. J., Shantz, A. B. & Harris, C. K. The role of host-plant fidelity in initiating insect race formation. *Evol. Ecol. Res.* **1**, 317–332 (1999).
62. Carvajal-Rodríguez, A. & Rolán-Alvarez, E. JMATING: a software for the analysis of sexual selection and sexual isolation effects from mating frequency data. *BMC Evol. Biol.* **6**, 40 (2006).
63. Rolán-Alvarez, E. & Caballero, A. Estimating sexual selection and sexual isolation effects from mating frequencies. *Evolution* **54**, 30–36 (2000).

Acknowledgments

Author contributions

H.-J.X. and X.-K.Y. designed the research; H.-J.X. performed the molecular work; H.-J.X. and W.Z.L. performed the behavioral experiments; all authors discussed the results; H.-J.X. wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xue, H.-J., Li, W.-Z. & Yang, X.-K. Assortative mating between two sympatric closely-related specialists: inferred from molecular phylogenetic analysis and behavioral data. *Sci. Rep.* **4**, 5436; DOI:10.1038/srep05436 (2014).