Genetic Structure in the Seabuckthorn Carpenter Moth (*Holcocerus hippophaecolus*) in China: The Role of Outbreak Events, Geographical and Host Factors

Jing Tao[1,2], Min Chen[1,2], Shi-Xiang Zong[1,2], You-Qing Luo[1,2]*

1 Beijing Forestry University, Beijing, People's Republic of China, 2 Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing, China

Abstract

Understanding factors responsible for structuring genetic diversity is of fundamental importance in evolutionary biology. The seabuckthorn carpenter moth (*Holcocerus hippophaecolus* Hua) is a native species throughout the north of China and is considered the main threat to seabuckthorn, *Hippophae rhamnoides* L. We assessed the influence of outbreaks, environmental factors and host species in shaping the genetic variation and structure of *H. hippophaecolus* by using Amplified Fragment Length Polymorphism (AFLP) markers. We rejected the hypothesis that outbreak-associated genetic divergence exist, as evidenced by genetic clusters containing a combination of populations from historical outbreak areas, as well as non-outbreak areas. Although a small number of markers (4 of 933 loci) were identified as candidates under selection in response to population densities. *H. hippophaecolus* also did not follow an isolation-by-distance pattern. We rejected the hypothesis that outbreak and drought events were driving the genetic structure of *H. hippophaecolus*. Rather, the genetic structure appears to be influenced by various confounding bio-geographical factors. There were detectable genetic differences between *H. hippophaecolus* occupying different host trees from within the same geographic location. Host-associated genetic divergence should be confirmed by further investigation.

Introduction

Pests with fluctuating population size are of major concern for forest security. Knowledge of a pest's population dynamics and associated influential factors is crucial for forest management. Habitat, weather, natural enemies and heritable traits are considered to play roles in insect population dynamics [1]. Despite many studies, the factors involved in the origin of insect outbreaks remain poorly understood. Multiple explanations have been proposed including: escape from natural enemies [2–5], favorable weather [6], changes in host quality and quantity [7–9], and genetic variation of pests [10–12].

The seabuckthorn carpenter moth, *Holcocerus hippophaecolus* Hua (Lepidoptera: Cossidae) is the main pest of seabuckthorn, *Hippophae rhamnoides* L. (Elaeagnaceae). It usually occurs on seabuckthorn, but can also occur on *Ulmus pumila* L. (Urticales: Ulmaceae) as well as a couple of species of Rosaceae [13]. The larvae seriously obstruct water transportation of seabuckthorn by boring into the trunk and roots. It has one generation every 3–4 years and larval stages occupy most of its life history. It is widely distributed throughout its hosts' range, with most damage being caused to trees more than 5 years old. The adult females have limited dispersal and lay their eggs in masses on nearby plants where the larvae feed gregariously. Berryman [14] has demonstrated that pests with low dispersal properties have short, intense, restricted outbreaks whereas those with high vagility have long, extended outbreaks. Consistent with the former pattern, the

seabuckthorn carpenter moth has limited dispersal ability and exhibits short but intense outbreaks that are geographically restricted [15]. Zhou reported that outbreaks of *H. hippophaecolus* can lead to more than 70% mortality of seabuckthorn in plantations in the Inner Mongolia Autonomous Region [16]. Limited mobility appears to play a role in the spatial restriction of the seabuckthorn carpenter moth. The outbreaks usually continue for one or two years before pest numbers decline [15,17].

Seabuckthorn is native to western and northern China, the northern Himalayas and northwestern Europe, through to central Asia and the Altai Mountains [18]. It is a native in 11 provinces (autonomous region, municipalities) in China, with less than 500 thousand hectares of natural forest in the 1950's [19]. Because of seabuckthorn's nitrogen-fixing symbionts, this plant serves to enrich and protect soils [18]. It has been promoted widely in western and northern China to prevent soil erosion and desertification. There are now 2,000,000 ha of seabuckthorn throughout 22 provinces in China, two-thirds of which are monoculture plantations. *H. hippophaecolus* was firstly reported as a pest of seabuckthorn in 1990 [20]. Today *H. hippophaecolus* is considered to be the main threat to seabuckthorn in China. It infests 133,000 ha of seabuckthorn and killed 67,000 ha during the 1990's. Most of the outbreak events occurred in Seabuckthorn monoculture plantations [16]. Prior to the spread of *H. rhamnoides* plantations in western and northern China, no outbreak events of

H. hippophaecolus had been recorded. *H. rhamnoides* was introduced as a novel plant to Jianping County (Liaoning Province) in the 1950's. Prior to that, *H. hippophaecolus* mainly fed on *U. pumila* [16,19]. Jianping has been cultivating seabuckthorn widely since the 1970's and used to have the largest area of seabuckthorn plantation. However, by 2001, the forests were heavily disturbed by the seabuckthorn carpenter moth.

Molecular markers are widely used for insect population genetic research and are a useful tool to study population structure. Correlation with ecological factors, to identify causes of the observed genetic structuring, is often possible using these techniques [21,22]. The Amplified Fragment Length Polymorphism (AFLP) technique generates a large number of fragments that are distributed throughout the genome, without requiring background knowledge of the genome [23]. In the present study, the AFLP technique was used to examine population genetic patterns of *H. hippophaecolus* to address the questions below.

Are genetic patterns a factor in causing outbreaks?

Natural selection, favoring certain genotypes at low densities and others at high densities, may contribute to the regulation of animal numbers [10]. Many studies have demonstrated that the genetic patterns may play a role in population dynamics. For example, changes of allelic frequencies of different loci have been shown to correlate with population fluctuations in *Microtrous ochgaster* [24]. Simchuk et al [25] suggested that esterase (*Est-4*) and protease (*Pts-4*) loci in *Tortrix viridana* L. were directly related to its population dynamics. Mormon crickets (*Anabrus simplex*) were found to consist of genetically distinct clusters that correspond with gregarious (outbreak) populations and solitary (non-outbreak) populations, respectively [11]. These genetic clades provided evidence that the differences of propensities to outbreak are likely due to genetic polymorphism. The seabuckthorn carpenter moth is a very destructive pest, with outbreaks reported in many areas within its range (outbreak areas), but it exists at low densities in other parts of its range (non-outbreak areas). *H. hippophaecolus* populations may therefore consist of genetically distinct clusters with different propensities for outbreak. Here we describe the genetic variation of *H. hippophaecolus* populations from 10 areas across its range with contrasting historical patterns of outbreak events [15–17,26,27] and assess whether the observed genetic patterns can be explained by outbreak history.

Do geographical distance, outbreaks and drought affect genetic population structure?

Identifying the factors responsible for structuring genetic diversity is important for a better understanding of the insect's evolutionary history. Such fundamental studies help to infer ecological characteristics that are crucial for establishing management strategies. Geographical distance, outbreaks, and drought were considered in this study. First, in the case of species with restricted dispersal abilities, we expect to observe a positive correlation between the genetic and geographical distance. Moreover, in the historical outbreak regions, *H. hippophaecolus* undergoes outbreak events that are followed by population decline when food plants become unavailable. Such population fluctuations may have large effects on genetic diversity within populations. Furthermore, non-genetic studies indicate recent outbreak events were related to a lack of rainfall [16]. Heavy periodic rainfalls could restrict *H. hippophaecolus* movement and genetic exchange within populations. Therefore, we might also expect lower genetic differentiation within populations during drought.

Does host-associated diversity exist in a common locality?

Herbivores and their host plants maintain an intimate relationship in feeding, oviposition, mate finding and predator avoidance. Distribution, availability, longevity and chemistry of host plants are major factors that affect the genetic differentiation of herbivorous populations [28]. Seabuckthorn was a novel host plant for insects following its introduction into Jianping. Herbivore populations can suffer from disruptive selection following shifts to novel host plants [29]. How did they adapt to the changes? Is there detectable genetic diversity among *H. hippophaecolus* populations feeding on different host trees in a common location? To answer these questions, we sampled larvae from sympatric populations of four host trees, sebuckthorn, *U. pumila*, *Prunus armeniaca* L, (Rosales: Rosaceae), *Pyrus pyrifolia* (Burman) Nakai (Rosales: Rosaceae), in Jianping.

Materials and Methods

Sample collection and DNA extraction

Individuals (n = 217) were collected from 10 locations across the carpenter moth range during the summer of 2008 (Table 1) by directly sampling under the bark of infested trees and byusing light and pheromone traps. Sampling locations represented two contrasting patterns of historical outbreak events, based on a literature survey and unpublished data (J. Zong, personal communication) (Figure 1). Populations from some areas have experienced outbreaks while in others population densities have been consistently low. In Jianping, a further 24 insects were collected from different hosts (*U. pumila* (JPY, n = 7), *Prunus armeniaca* (JPX, n = 8), *Pyrus pyrifolia* (JPL, n = 9)). Individuals were transported alive to the laboratory, and then kept at −80°C. Prior to DNA extraction, insects were washed in 80% ethanol. Total genomic DNA was isolated using the SDS-method of Zhang and Hewitt [30]. After extraction, DNA was dissolved in TE buffer and stored at −20°C until further use.

AFLP protocol

Amplified fragment length polymorphism (AFLP) analysis was used to assess genetic diversity among sampled populations of *H. hippophaecolus*. The AFLP procedure followed Vos et al. [23] with minor modifications. Genomic DNA was digested with *Eco*RI and *Mse*I restriction enzymes (New England Biolabs) and double stranded adapters were ligated to the sticky ends of the fragments. After 4 h incubation at 37°C, each sample was diluted 1:9 with H_2O and a two-step amplification strategy was used. Pre-selective amplification was performed for 3 min at 94°C, then 30 cycles of 30 s at 94°C, 30 s at 56°C and 1 min 72°C. A 20-μl Pre-selective amplification PCR mixture consisted of 30 mM $MgCl_2$, 4.5 mM dNTP, 0.6 U Taq DNA polymerase, 30 ng *Eco*RI-C and *Mse*I-A primer. In the selective amplification, we used the following nine primer combinations selected from 100 tested combinations [31]: *Eco*RI-AAC/*Mse*I-CAA, *Eco*RI-AAC/*Mse*I-CAC, *Eco*RI-AAC/*Mse*I-CCT, *Eco*RI-AAC/*Mse*I-CTT, *Eco*RI-AAG/*Mse*I-CCA, *Eco*RI-AAG/*Mse*I-CTG, *Eco*RI-CA/*Mse*I-CAA, *Eco*RI-CA/*Mse*I-CAC, *Eco*RI-CA/*Mse*I-CCT. The *Eco*RI primers were labeled with IRD-700. Selective amplification was performed with the following touchdown thermal profile: 3 min at 94°C; 12 touchdown cycles at 94°C for 30 s, 65°C for 30 s (decreasing the temperature by 0.7°C per cycle), and 72°C for 60 s; 30 cycles at 94°C for 30 s, 56°C for 30 s, 72°C for 1 min; 5 min at 72°C. The 10 μl PCR mixture contained 15 $MgCl_2$, 1.5 ng *Mse*I and *Eco*RI primer, 2 mM dNTP), 2 μl diluted (1:9) Pre-amplified

Table 1. Geographical location, average annual rainfall, and hosts for populations of *H. hippophaecolus*.

No.	Location	Coordinates	Average annual rainfall/mm	Population indentifier	Host plant	Sample Size
1	Liaoning, Jianping	119.71E/41.84N	478.37	JPS	*H.rhamnoides*	26
				JPY	*U. pumil*	7
				JPX	*P. armeniaca*	8
				JPL	*P. pyrifolia*	9
2	Inner mongolia, Linxi	118.23E/43.61N	378.98	LX	*H.rhamnoides*	16
3	Hebei, Fengning	116.61E/41.21N	535.16	FN	*H.rhamnoides*	25
4	Shanxi, Youyu	112.39E/39.96N	377.24	YY	*H.rhamnoides*	26
5	Inner mongolia, Dongsheng	111.25E/39.87N	410.52	DS	*H.rhamnoides*	15
6	Shanxi, Wuzhai	111.87E/38.88N	431.54	WZ	*H.rhamnoides*	18
7	Shanxi, Yulin	109.76E/38.28N	397.43	YL	*H.rhamnoides*	19
8	Shanxi, Wuqi	108.26E/36.90N	534.83	WQ	*H.rhamnoides*	29
9	Nixia, Yanchi	107.48E/37.89N	290.24	YC	*H.rhamnoides*	22
10	Nixia, Pengyang	106.50E/35.82N	498.67	PY	*H.rhamnoides*	21

doi:10.1371/journal.pone.0030544.t001

DNA. All PCRs were conducted on a GeneAmp PCR System 9700 (USA Applied Biosystems).

Amplification products were separated on 6% polyacrylamide gels for 2.5 h on a LI-COR 4300 DNA Analyzer (LI-COR Biosciences, USA), using LI-COR 50–700 bp (Labeled with IRD-700) as a size standard. Fragments from 100–700 bp in size were scored as present (1) or absent (0) using SAGA MX (LI-COR Biosciences, USA), and exported for data analysis.

A blank control was carried out along with each set of DNA extractions and PCR amplifications to monitor any possible cross contamination. Poor-quality DNA samples that did not amplify were excluded from further analysis.

Data analysis

Genetic variation and structure of *H. hippophaecolus* populations. The diversity of geographic populations was assessed by estimating the percentage of polymorphic loci (%P) and Nei's heterozygosity. Percentage of polymorphic loci estimates were based on 99% criteria and heterozygosity estimates were made using the software TFPGA [32].

The genetic structure was examined by an analysis of molecular variance (AMOVA) performed by the software ARLEQUIN 3.1 [33]. This method was used to partition the genotypic variance among and within populations. Two separate analyses were performed to test the hypotheses of genetic structure attributable to variation: among individuals across the different localities feeding on *H. rhamnoides* and among individuals across different host plants in Jianping. An additional analysis of individuals feeding on *H. rhamnoides* compared to the group combining three other host plants in Jianping was also performed. Genetic differentiation coefficients between populations (both geographic and host-associated) were calculated as F_{ST}, with 95% confidence

intervals (CI) obtained by bootstrapping 1000 replicates over loci. The TFPGA software was also applied to calculate Nei's genetic distance (D) [34]. Neighbor-joining (NJ) trees were constructed based on D using MEGA4 [35].

Identification of candidate outbreak loci under selection. Outlier loci were identified using the Dfdist approach [36,37] in Mcheza program [38] (available at http://popgen.eu/soft/mcheza/). Allele frequencies are estimated in Dfdist based on Zhivotovsky's [39] Bayesian approach.

Because of our particular interest in outbreak-associated divergence, the Dfdist was run for two groups of populations (outbreaking population vs non-outbreaking population). A total of 50000 realizations were performed and maximum allowable allele frequency was 0.99. We chose the 0.995 confidence interval and set the significance level at 99%. The Benjamini and Hochberg false discovery rate (FDR) correction method was used to correct for the occurrence of false positives in loci identified as under selection [40]. Loci with significant *P*-values at FDR threshold of 50% were identified using the Benjamini and Hochberg method.

Testing outbreaks and environmental factors driving genetic structure. The following analysis tested outbreaks and environmental factors that potentially influenced genetic population structure. The effect of geographical distance was assessed using linear map distances between *H. hippophaecolus* populations. Secondly, outbreak patterns were scored with 1 indicating populations from areas where outbreaks had occurred and 0 representing populations in non-outbreaking areas. Finally, an index for the "degree of drought," represented by the average annual rainfall collected over 50 years was obtained (1955–2007, China meteorological data sharing service system http://cdc.cma.gov.cn/). Mantel tests were conducted with the software TFPGA to test the correlation between Euclidean distances for all the factors and genetic distances.
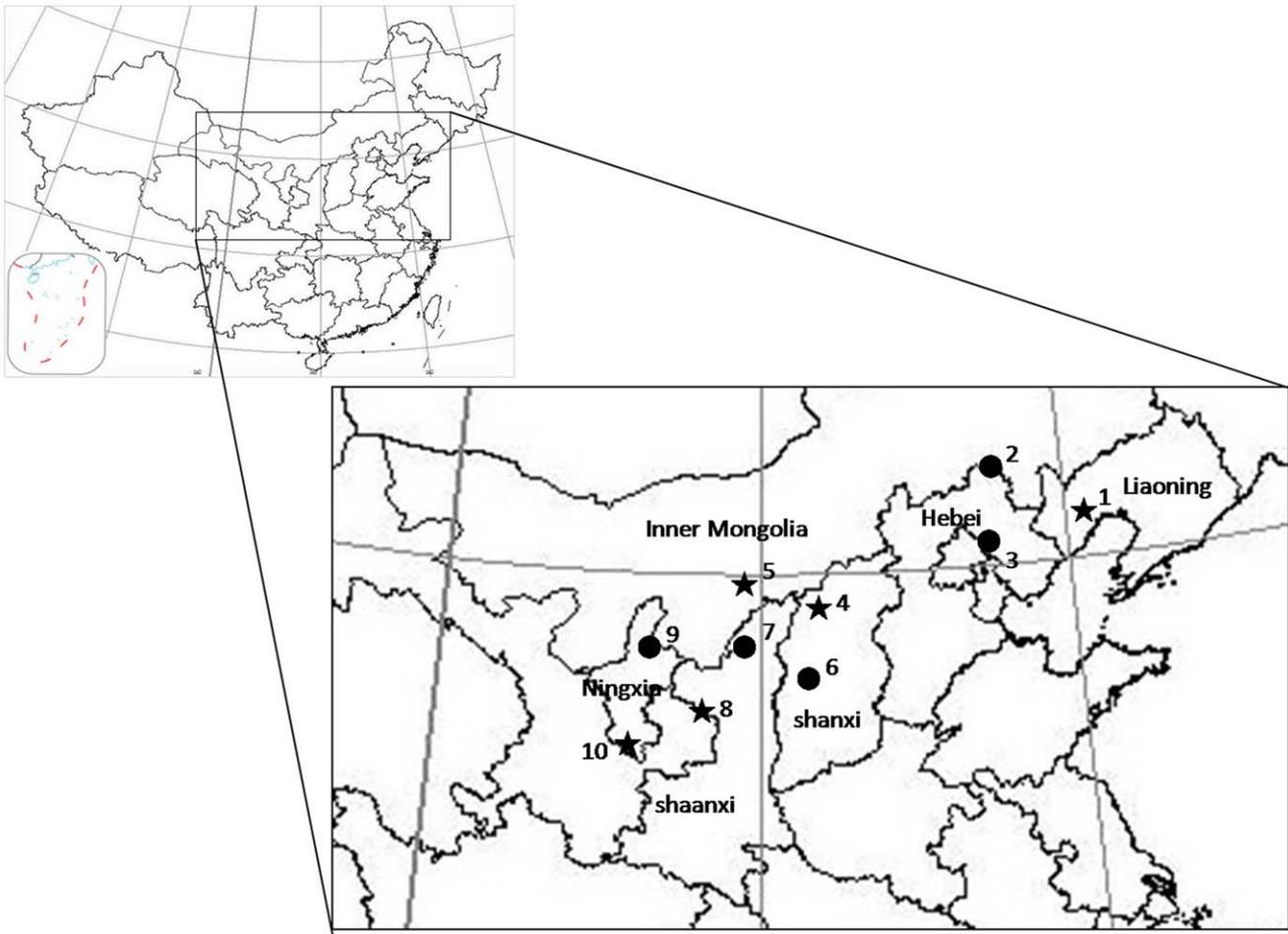
Figure 1. Locations of 10 *H. hippophaecolus* sites with known outbreak (star) or non-outbreak (circle) patterns. Outbreak patterns based on Luo et al. [16], population 1; Wang [26], poulation 4; Zhou [15], population 5; You [27], population 7; Zong et al. [17] population 10; Zong, personal communication, population 2,3,6,7,9. Identities and locations of populations see Table 1.
doi:10.1371/journal.pone.0030544.g001

The general linear models (GLM) method was also used to test the effect of outbreak and drought on the genetic differentiation between populations. In this analysis the factor "drought" was defined as locations with less than 400 mm average annual rainfall. Values of 1 were used for drought locations (YY, YL, YC, LX) and 0 for other locations (PY, JP, WQ, DS, WZ, FN). The outbreak factor was standardized, as previously, for an outbreak area of 1 and a non-outbreak area of 0. We performed a GLM analysis of the heterozygosity with outbreak and drought as fixed factors. A *P*-value of <0.05 was used to indicate statistical significance. GLM was implemented using SPSS 16.0.

Results

Genetic variation and structure of *H. hippophaecolus* populations

The nine primer combinations produced a total of 933 bands. The global Θ among the 10 sites was 0.2106 (95% CI 0.1981–0.2230). Nei's heterozygosity for each geographical population was moderate and ranged from 0.1505~0.2042 (Table 2).

AMOVA conducted on AFLP markers confirmed the presence of moderate genetic differentiation showing that 22.54% of total variability was due to the variation among geographic populations

Table 2. Percentage of polymorphic loci (%P) and Nei's heteozygosity of *H. hippophaecolus* populations.

	Population indentifier									
	JP	LX	FN	YY	DS	WZ	YL	WQ	YC	PY
%Polymorphic loci (p)	72.56	51.23	77.60	58.41	74.49	68.27	70.85	66.34	56.38	61.74
Heterozygosity (H)	0.1854	0.1529	0.1702	0.1495	0.2042	0.1505	0.1872	0.1749	0.1604	0.1679

doi:10.1371/journal.pone.0030544.t002

Table 3. Analysis of molecular variance (AMOVA) of *H. hippophaecolus* populations.

Source of variation	d.f.	SS	Percentage of variation (%)	P
Geographical grouping				
Among localities	9	5831.247	22.54	<0.0001
Individuals within localities	206	18395.396	77.46	<0.0001
Host-plant grouping				
Among host plants in Jianping	3	1537.343	31.73	<0.0001
Individuals within host plants in Jianping	30	3895.897	68.27	<0.0001
Among two host groups in Jianping	1	1249.317	34.82	<0.0001
Individuals within groups in Jianping	48	4183.923	65.18	<0.0001

doi:10.1371/journal.pone.0030544.t003

($F_{ST} = 0.2254$, $P < 0.0001$) (Table 3). The pair-wise comparisons between populations were characterized by values of F_{ST} ranging from 0.0424–0.3663 (Table 3). Most of the populations showed highly significant differences ($P < 0.0001$) with the exception of the YY and LX populations ($P = 0.0182$). This result indicates that most of the 10 sampled populations represent differentiated populations.

The Neighbor Joining phenogram (Figure 2) indicates that the clusters comprised populations with a mixture of outbreak patterns. For instance, populations from Dongsheng and Youyu were in two distinct NJ genetic clusters, although they have the same intensity of outbreak events.

Examination of the AFLP data using Dfdist in Mcheza sought to determine whether there was evidence of any highly differentiated loci. F_{ST} is plotted against heterozygosity in Figure 3. The outbreak and non-outbreak population comparison performed with Dfist resulted in four markers out of 993 (loci 93, 188, 223, 390) showing more differentiation than expected at the 99.5% confidence level. All these loci were detected as potential positive outliers at the 50% FDR threshold (Figure 3).

Testing outbreaks and drought as factors driving *H. hippophaecolus* genetic structure

The Mantel test based on the 10 localities gave an *r* value of 0.0554 ($P = 0.3460$, for 10000 randomizations), indicating no correlation between geographic and genetic differences. The Nei's genetic distances between populations were not significantly correlated to outbreak differences in the Mantel test ($r = 0.2516$, $P = 0.0740$). The interaction between Euclidean distances for average annual rainfall and genetic distances was also not significant (Mantel test $r = 0.1271$, $P = 0.2070$). GLM analysis showed that the factors of outbreak and drought, and their interaction, did not have a significant effect on heterozygosity ($F_{1,10} = 0.053$, $P = 0.826$, $F_{1,10} = 1.329$, $P = 0.293$ and $F_{1,10} = 2.904$, $P = 0.139$ respectively).

Host-associated diversity

The host plant was found to have a larger effect on the genetic structure among populations than geographic location. The global Θ value among different hosts was 0.2785 (95% CI 0.2548–0.3024), higher than the value among 10 sites (0.2106). AMOVA
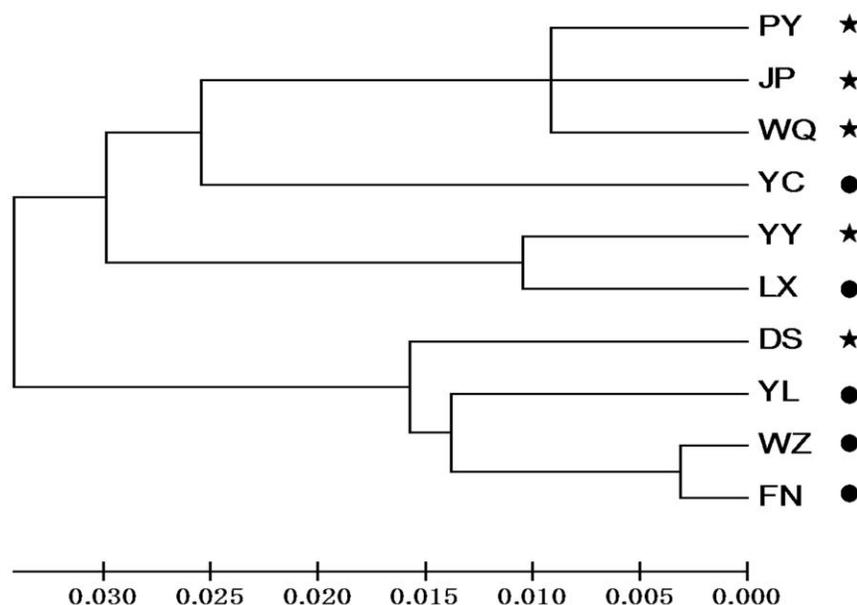

Figure 2. Neighbor-joining phenogram of Nei's genetic distances between *H. hippophaecolus* populations at different collection sites. Site names and outbreak patterns correspond to those in figure 1. Populations are depicted as outbreaking (star) or non-outbreaking (circle).
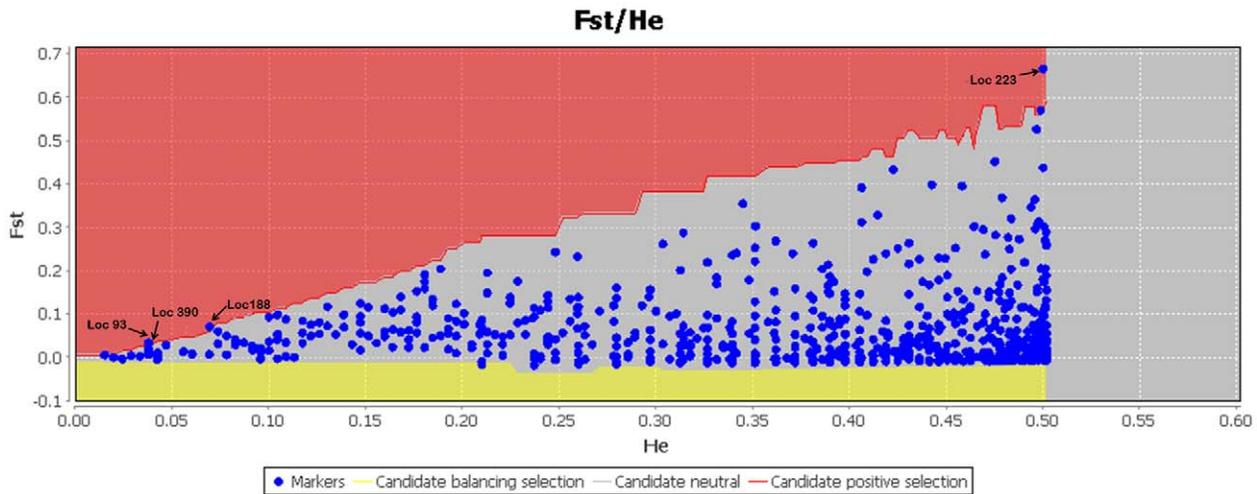doi:10.1371/journal.pone.0030544.g002

Figure 3. Result of Dfdist analyses in Mcheza. F_{ST} against heterozygosity plot is provided. Loci were identified as significant, corresponding to a 99% critical frequency and 99.5% confidence interval. Red and yellow lines represent 0.995th and 0.005th quantiles of the conditional distribution obtained from Dfdist simulations respectively. Each blue dot indicates AFLP marker; dots in gray areas represent neutral loci falling below the 0.995 quantile's broad line, whereas outlier loci in red areas are pointed out by arrows and accompanied by the locus number.
doi:10.1371/journal.pone.0030544.g003

with ARLEQUIN found greater variation among populations in host-plant groupings (31.73%) than populations in geographical groupings (22.54%) (Table 3). Pairwise F_{ST} statistics between JPS and each other location population ranged from 0.0856 to 0.2978 (Table 4), while the genetic divergences were all highly significant 0.3510~0.3773 in the host-associated analysis (Table 5).

In Jianping, individuals feeding on *H. rhamnoides* had a great separation from individuals feeding on other host plants. When combined individuals feeding on *U. pumila*, *P. armeniaca* and *P. pyrifolia* as a group, compared to individuals feeding on *H. rhamnoide*, the variation among two groups rose up to 34.82% by AMOVA with ARLEQUIN. Pairwise comparisons of F_{ST} values between all host plant combinations further supported the pattern of genetic structure. F_{ST} values were much greater in comparisons between the *H. rhamnoides* feeders (0.3510–0.3773) and each other host-plant feeders (0.0527–0.1180) (Table 5). *H. rhamnoides* feeders showed strongly significant differences ($P<0.0001$) with the moth on other host plants (Table 5).

Discussion

Genetic patterns associated with outbreak events of *H. hippophaecolus*

Genetic clustering did not support distinct outbreak-associated genetic clades in *H. hippophaecolus*. NJ genetic population clusters contained a combination of populations from historical outbreak areas as well as non-outbreak areas (Figure 2). The outbreak effect may have been difficult to detect among different geographical populations due to various confounding biogeographical factors that also shape genetic structure in *H. hippophaecolus*. In addition, one cannot exclude the possibility that the outbreak and non-outbreak patterns are associated with a single genotype, but depend on the expression of different phenotypes in different environments.

Indeed, our results support the notion that outbreak events were likely to be endemic population changes from latent to epidemic rather than being due to insects with an outbreak-associated

Table 4. Nei's genetic distance and F_{ST} value between all geographic combinations.

	PY	JP	YY	WQ	DS	YL	WZ	YC	LX	FN
PY	—	0.0902	0.2409	0.0683	0.1589	0.2179	0.2929	0.2779	0.2601	0.2298
JP	0.0191	—	0.2682	0.0856	0.1969	0.2204	0.2978	0.2754	0.2737	0.2467
YY	0.0465	0.0587	—	0.2654	0.1115	0.2938	0.3663	0.3477	0.1047	0.2715
WQ	0.0170	0.0196	0.0583	—	0.1813	0.2048	0.2914	0.2404	0.2856.	0.2454
DS	0.0401	0.0442	0.0423	0.0468	—	0.1550	0.1828	0.2511	0.1089	0.0986
YL	0.0530	0.0559	0.0888	0.0527	0.0362	—	0.1129	0.1642	0.2520	0.1298
WZ	0.0816	0.0780	0.1247	0.0869	0.0375	0.0254	—	0.2998	0.3131	0.0424
YC	0.0538	0.0541	0.0774	0.0447	0.0617	0.0437	0.0871	—	0.3387	0.2819*
LX	0.0456	0.0578	0.0209	0.0606	0.0343	0.0678	0.0944	0.0731	—	0.2126
FN	0.0669	0.0652	0.0969	0.0740	0.0207	0.0298	0.0062	0.0827	0.0705	—

Nei's genetic distances are below the diagonal. F_{ST} value and their significance level are above the diagonal. Significance level of associated F_{ST} value are shown as:
*0.01<P<0.05, unmarked mean P<0.0001.
doi:10.1371/journal.pone.0030544.t004

Table 5. Pairwise comparisons of genetic divergence estimates (F_{ST}) between all host plants combinations.

	JPY	JPL	JPX	JPS
JPY	0.00000			
JPL	0.07590*	0.00000		
JPX	0.05271*	0.11808**	0.00000	
JPS	0.36080***	0.37730***	0.35109***	0.00000

Significance level of associated F_{ST} value are shown as:
*0 01<P<0.05;
**0.001<P<0.01;
***P<0.0001.
doi:10.1371/journal.pone.0030544.t005

genotype spreading to outbreak areas. This conclusion is also consistent with the poor dispersal ability of *H. hippophaecolus*, which has been observed in non-genetic studies. Zong et al. [17] indicated female moths usually choose a nearby tree for mating after emergence, and that the body weight of fertilized female moths is too heavy for long-distance migration. Males are not attracted to sex pheromone traps located too far away from the infested forests (<100 m) [13]. Furthermore, adults live for only several days, which limits the degree of dispersal. Young seedlings (1–2 years) and seeds of seabuckthorn are often used for its introduction. However, *H. hippophaecolus* only harms seabuckthorn plants that are more than 5 years old. Therefore, *H. hippophaecolus* would not be dispersed long-distances by artificial movement of host plants.

We rejected the hypothesis of genetic difference associated with outbreaks in the seabuckthorn carpenter moth. However, habitat, weather, natural enemies are also considered as main factors affecting insect population dynamics. An outbreak occurs when the physiological state of the plant permits a herbivore phenotype with a high reproductive capacity to become dominant. Agricultural and forest monocultures consisting of extensive plantings of hosts with narrow genetic variability are havens for pest outbreaks [1]. Seabuckthorn monoculture plantations are optimal sites for survival of *H. hippophaecolus*, especially for those that were introduced as an exotic species growing under unfavorable conditions [41]. Weather and climatic conditions significantly affect population fluctuation. Unusual weather is known to have strong effects on the dynamics of insect populations [42,43]. Several authors associated the initiation of outbreaks in Jianping and Dongsheng with consecutive dry years before outbreak [15,19,]. Seabuckthorn plantations have plenty of nutrients, relatively few natural enemies and are highly vulnerable to drought or human disturbance, which may explain why outbreak events happen there.

The outlier analysis revealed that although differentiation for the majority of markers did not significantly deviate from neutral expectations, a small number of markers (n = 4) were identified as outlier loci. The false discovery rate test supported the conclusion that all identified outlier loci are under selection. These results do not allow us to reject the hypothesis that specific genome regions or genes are associated with outbreak events. Having indentified outlier loci in *H. hippophaecolus*, it will be necessary to try to find candidate genes that could correspond to these AFLP markers. Then we can characterize these genes with functional genomics analyses.

Factors influencing the population structure of *H. hippophaecolus*

We found evidence of limited gene flow among samples collected from 10 locations. The poor correlation between genetic and geographical distances is an unexpected result given the previous assumption that populations are isolated-by-distance.

These results suggest that multiple factors other than simple geographic distance are influencing the genetic composition of populations. The Mantel analysis failed to support the idea that outbreak and drought are pertinent factors underlying genetic structure in *H. hippophaecolus*. Compared to other observations in insects, Nei's heterozygosity values for each population are moderate [44,45,46]. We found similar genetic diversity within outbreak and non-outbreak populations of the seabuckthorn carpenter moth. Similarly, GLM analysis showed that outbreaks had no significant effects on heterozygosity. Genetic diversity within outbreak populations was not strongly affected by increases in population size during outbreak periods. This result could be consistent with the theoretical prediction that long-term fluctuating populations correspond to the harmonic mean size over time, and should thus be closer in size to that during the remission period than during an outbreak [47,48]. We rejected the hypothesis that drought populations have lower genetic variation. GLM analysis showed that drought had no significant effects on heterozygosity.

Comparative studies of population structure in phytophagous insects show that genetic structure is mostly determined by the ability of the species to disperse [49]. The probability of successful dispersal is largely determined by habitat availability [50,51]. Louy et al. [52] have shown in experiments with three skipper species that dispersal ability and habitat availability determine the genetic structure of species. Whether a habitat is available for phytophagous insects strongly depends on the existence of host trees. Several forest insects have pronounced geographical structure that follows the distribution of their host tree species [46,53].

Louy et al. [52] suggest that limited habitat availability in combination with low dispersal capacity result in independent genetic structure with relative high genetic differentiations and low gene flow among populations. *H. hippophaecolus* is a specialist, which only feeds on a few plants (*U. pumila*, *P.armeniaca*, *P.pyrifolia*). Moreover, most of the principal host in China, *H. rhamnoides*, is in single species plantations which are suffering from human interference. Could this independent population structure of *H. hippophaecolus* be explained by low dispersal capacity and habitat fragmentation? Combining habitat landscape and population genetic analysis might answer this question in the future.

The role of the plant

Host races are genetically differentiated sympatric populations of parasites that use different hosts and between which there is limited gene flow [54]. Our analyses uncovered very high F_{ST} values (0.3510–0.3773) between JPS and other non-seabuckthorn populations. It is indicated that *H. rhamnoides* constitutes a barrier to gene flow between *H. hippophaecolus* populations from other host plants in Jianping. *H. hippophaecolus* feeding on *H. rhamnoides* in Jianping are more genetically differentiated than those from other hosts in sympatric rather than other geographically distant populations of seabuckthorn. Host races might therefore exist in seabuckthorn and other host plant used by *H. hippophaecolus*. Factors favoring host race formation include correlations between host choice and mate choice. Although host fidelity and assortative mating has not been fully explored in *H. hippophaecolus*, tests using both artificial and natural methods suggest female host preferences may exist. Adult emergence from the seabuckthorn roots confirmed oviposition preference on *H. rhamnoides*, rather than on *U. pumila* and *P. armeniaca* [55].

Seabuckthorn was an endemic perennial, sporadically growing in Inner Mongolia, Shanxi and areas of Liaoning province before

it was widely promoted. The timing of host shifting of *H. hippophaecolus* in Jianping is likely due to the introduction of *H. rhamnoides*. However, how did host shifting occur in *H. hippophaecolus* in Jianping? When did host-associated genetic divergence initially occur in *H. hippophaecolus*? Data from many host utilization systems gave rise to a possible scenario that host shifts occur as a result of host plant's increased abundance and availability as a potential resource following human-mediated plant community changes [56,57]. If this is the case, our data suggests a local host shift and genetic differentiation of *H. hippophaecolus* following the introduction of seabuckthorn in Jianping. Though a rapid range expansion of *H. hippophaecolus* following human-mediated changes is possible, it does seem unlikely given the wide extent of genetic divergence observed during such a brief time. This scenario was also rejected by Sword et al in the *Hesperotettix viridis* host utilization system [58]. Another possibility is a genetic divergence of moth between *H. rhamnoides* and other hosts, prior to the host shift. Feder et al. [59] found genetic divergence between apple and hawthorne host races of *Rhagoletis pomonella* L. pre-dating the introduction of the apple to North America. Given the long life history of *H. hippophaecolus* and brief planting history of *H. rhamnoides* in Jianping, we suppose the latter scenario is the case. Seabuckthorn is native to parts of western and northern China although records for the historical host plant use by *H. hippophaecolus* are lacking. Our results indicate that an *H. hippophaecolus* lineage might have adapted to utilize *H. rhamnoides* in China prior its spread. The possibilities of an ancestral host shift and stable host-associated genetic divergence in seabuckthorn carpenter moth are suggested.

We found no fixed diagnostic differences in AFLP data between the different host-associated forms. Host-associated genetic divergence should also be further demonstrated by sampling additional populations feeding on different host plants in more locations. In future studies, more different genetic markers are recommended in this system. They should include co-dominant markers such as microsatellites (not currently available for this species) and incorporation of variable regions of the mitochondrial genome. Microsatellites are highly polymorphic, locus-specific and can show co-dominant inheritance. They may recover higher levels of variability than other markers, particularly if following a population bottleneck associated with host shift. Mitochondrial sequences can be analyzed to determine patterns of evolutionary relationships between different haplotypes. This may provide information on the historical evolution of host-associated forms in the seabuckthorn moth.

Acknowledgments

Author Contributions

Conceived and designed the experiments: JT MC S-XZ Y-QL. Performed the experiments: JT MC. Analyzed the data: JT MC S-XZ Y-QL. Contributed reagents/materials/analysis tools: JT MC S-XZ Y-QL. Wrote the paper: JT. Collected samples: JT S-XZ.

References

1. Wallner WE (1987) Factors affecting insect population dynamics: differences between outbreak and non-outbreak species. Annual Review of Entomology 32: 317–340.
2. Price PW (1987) The role of natural enemies in insect populations, In: Barbosa P, Schultz J, eds. Insect outbreaks. New York: Academic Press. pp 208–312.
3. Myers JH (1988) Can a general hypothesis explain population cycles of forest Lepidoptera. San Diego: Academic Press Inc. pp 179–242.
4. Walsh PJ (1990) Site factors, predators and pine beauty moth mortality. In: Watt AD, Leather SR, Hunter MD, Kidd NAC, eds. Population dynamics of forest insects. Andover: Intercept. pp 242–252.
5. John LM, Susan H, Mary G (2001) Origin of an insect outbreak: escape in space or time from natural enemies? Oecologia 126: 595–602.
6. Martinat PJ (1987) The role of climatic variation and weather in forest insect outbreaks, In: Barbosa P, Schultz J, eds. Insect outbreaks. New York: Acedemic Press. pp 241–268.
7. White TCR (1984) The abundance of invertebrate herbivores in relation to the availability of nitrogen in stressed food plants. Oecologia 63: 90–105.
8. Mattson WJ, Haack RA (1987) The role of drought in outbreaks of plant-eating insects. Bio Science 37: 110–118.
9. Rossiter M (1992) The impact of resource variation on population quality in herbivorous insects: a critical aspect of population dynamics. In: Hunter MD, Ohguhi T, Price PW, eds. Effects of resource distribution on animal-plant interactions. New York: Academic Press. pp 13–42.
10. Chitty D (1971) The natrual selection of self-regulatory behavior in animal populations. In McLaren IA, ed. Natrual Regulation of Animal Populations. New York: Atherton Press.
11. Bailey NW, Gwynne DT, Ritchie MG (2005) Are solitary and gregarious Mormon crickets (*Anabrus simplex*, Orthoptera, Tettigoniidae) genetically distinct? Heredity 95: 166–173.
12. Chapuis MP, Estoup A, Auge SA, Foucart A, Lecoq M, et al. (2008) Genetic variation for parental effects on the propensity to gregarise in *Locusta migratoria*. BMC Evolutionary Biology 8: 37.
13. Zong SX, Luo YQ, Lu CK, Xu ZC, Zhang LS (2006) Preliminary Study on Biological Characteristic of *Holcocerus hippophaecolus*. Scientia Silvae Sinicae 41: 79–84.
14. Berryman AA (1978) Towards a theory of insect epidemiology. Research in Population Ecology 19: 181–196.
15. Luo YQ, Lu CK, Xu ZC (2003) Control strategies on a new serious forest pest insect seabuckthorn carpenterworm, *Holcocerus hippophaecolus*. Forest Pest Disease 5: 25–28.
16. Zhou ZY (2002) The causes of death and control strategies to deal with *Holcocerus hippophaecolus* in the east of Erdos City, Inner Mongolia Autonomous Region. Hippophae 15: 7–11.
17. Zong SX, Luo YQ, Xu ZC, Yao GL, Wang YY (2006) Harm characteristics and population dynamics of two important borers in *Hippophae rhamnoidea*. Forest Pest and Disease 25: 7–10.
18. Stewart PE, Pearson MC (1967) Nodulation and nitrogen-fixation by *Hippophae rhamnoides* L. in the field. Plant and Soil 26: 348–360.
19. Luo YQ, Lu CK, Xu ZC (2003) Control strategies on a new serious forest pest insect seabuckthorn carpenterworm, *Holcocerus hippophaecolus*. Forest Pest Disease 5: 25–28.
20. Hua BZ, Zhou Y, Fang DQ (1990) Chinese Cossidae. Beijing: Tianze Press. pp 56–57.
21. Behura SK (2006) Molecular marker systems in insects: current trends and future avenues. Molecular Ecology 15: 3087–3113.
22. Foll M, Gaggiotti O (2006) Identifying the Environmental Factors That Determine the Genetic Structure of Populations. Genetics 174: 857–891.
23. Vos P, Hogers R, Bleeker M, Reijans M, Lee Tvd, et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407–4414.
24. Gaines MS, McClenaghan LR, Rose RK (1978) Temporal patterns of allozymic variation in fluctuating populations of *Microtus ochrogaster*. Evolution 32: 723–739.
25. Simchuk APV, Ivashov A, Companiytsev VA (1999) Genetic patterns as possible factors causing population cycles in oak leafroller moth, *Tortrix Viridana* L. Forest Ecology and Management 113: 35–49.
26. Wang DS (2007) Research on the regularity for the change of the population of *Holcocerus hippophaecolus* in Youyu County, Shuozhou City. Sci-Tech Information Development & Economy 17: 170.
27. You Y (2000) The present situation of forest diseases and insect pests on Yulin and the control measures. Shaanxi Forest Science and Technology 4: 1–3.
28. Mopper S (1996) Adaptative genetic structure in phytophagous insect populations. Tree 11: 235–238.
29. Funk DJ, Filchak KE, Feder JL (2002) Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116: 251–267.
30. Zhang DX, Hewitt GM (1998) Isolation of animal cellular total DNA. In: Karp A, Ssac PG, Ingram DS, eds. Molecular Tools for Screening Biodiversity: Plant and Animals. London: Chapman & Hall. pp 5–9.
31. Chen M, Tao J, Ma CD, Yin WL, Zhu YY, et al. (2008) Selection of primers and establishment of AFLP analysis system on *Holcocerus hippopaecolu*. Journal of Beijing Forestry University 30: 116–120.
32. Miller MP (1997) Tools for population genetic analysis (TFPGA) 1.3:A Windows program for the analysis of allozyme and molecular population genetic data. Department of Biological Sciences, Northern Arizona University: AZ, USA.
33. Schneider S, Roessli D, Excoffier L (2000) *Arlequin, Version 2.0: a Software for Population Genetic Data Analysis*. Genetics and Biometry Laboratory. University of Geneva: Switzerland.

34. Nei M (1978) Estimation of heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

35. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596–1599.

36. Beaumont MA, Nichols RA (1996) Evaluating Loci for Use in the Genetic Analysis of Population Structure. Proceedings: Biological Sciences 263: 1619–1626.

37. Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology 13: 969–980.

38. Antao T, Beaumout MA (2011) Mcheza: A workbench to detect selection using dominant markers. Bioinformatics 27: 1117–1118.

39. Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. Molecular Ecology 8: 907–913.

40. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 57: 289–300.

41. Zhou ZY, Yin WL, Liang HJ, Yu JM, Zhang Q (2007) Mechanism and the stand conditions of *Hippophae rhamnoides* resistance to *Holcocerus hippophaecolus*. Journal of Beijing Forestry University 29: 50–56.

42. Crister S (1991) Unusual weather and insect population dynamics: *Lygaeus equestris* during an extinction and recovery period. Oikos 60: 343–350.

43. Kennedy GG, Storer NP (2000) Life systems of polyphagous arthropod pests in temporally unstable cropping systems. Annual Review of Entomology 45: 467–493.

44. Salvato P, Battisti A, Concato S, Masutti L, Patarnello T, et al. (2002) Genetic differentiation in the winter pine processionary moth (*Thaumetopoea pityocampa - wilkinsoni* complex), inferred by AFLP and mitochondrial DNA markers. Molecular Ecology 11: 2435–2444.

45. Alessandro G, Sanna B, Leena L, Anne L, Johanna M (2005) The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations. In Molecular ecology. pp 4207–4219.

46. Mock KE, Bentz BJ, O'neill EM, Chong JP, Orwin J, et al. (2007) Landscape-scale genetic variation in a forest outbreak species, the mountain pine beetle (*Dendroctonus ponderosae*). Molecular Ecology 16: 553–568.

47. Motro U, Thomson G (1982) On heterozygosity and the effective size of populations subject to size changes. Evolution 36: 1059–1066.

48. Chapuis MP, Loiseau A, Michalakis Y, Lecoq M, Franc A, et al. (2009) Outbreaks, gene flow and effective population size in the migratory locust, *Locusta migratoria*: a regional-scale comparative survey. Molecular Ecology 18: 792–800.

49. Peterson MA, Denno RF (1998) Life-history strategies and the genetic structure of phytophagous insect populations, In: Mopper S, Strauss SY, eds. Genetic Structure and Local Adaptation in Natural Insect Population New York, Chapman & Hall. pp 236–322.

50. Arnaud JF (2003) Metapopulation genetic structure and migration pathways in the land snail *Helix aspersa*: influence of landscape heterogeneity. Landscape Ecology 18: 333–346.

51. Sander AC, Purtauf T, Wolters V, Dauber J (2006) Landscape genetics of the widespread ground-beetle *Carabus auratus* in an agricultural region. Basic and Applied Ecology 7: 555–564.

52. Louy D, Habel J, Schmitt T, Assmann T, Meyer M, et al. (2007) Strongly diverging population genetic patterns of three skipper species: the role of habitat fragmentation and dispersal ability. Conservation Genetics 8: 671–681.

53. Cognato AI, Harlin AD, Fisher ML (2003) Genetic structure among pinyon pine beetle populations (Scolytinae: *Ips confusus*). Environmental Entomology 32: 1262–1270.

54. Drès M, Mallet J (2002) Host races in plant-feeding insects and their importance in sympatric speciation. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 357: 471–492.

55. Wang ZZ, Wen JB, Yao GL, Zong SX, Luo YQ (2010) Oviposition preference of *Holcocerus hippophaecolus* to different tree species. Journal of Beijing Forestry University 32: 130–135.

56. Janike J (1990) Host specialization in phytophagous insects. Annual Review of Ecology and Systematics 21: 243–273.

57. Bernays EA, Chapman RF (1994) Host-plant selection by phytophagous Insects. New York: Chapman & Hall.

58. Sword GA, Joern A, Senior LB (2005) Host plant-associated genetic differentiation in the snakeweed grasshopper, *Hesperotettix viridis* (Orthoptera: Acrididae). Molecular Ecology 14: 2197–2205.

59. Feder JL, Berlocher SH, Roethele JB, Dambroski H, Smith JJ, et al. (2003) Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. Proceedings of the National Academy of Sciences of the United States of America 100: 10314–10319.