MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 2676–2692 doi: 10.1111/j.1365-294X.2011.05129.x

Genetic differentiation across North America in the generalist moth *Heliothis virescens* and the specialist *H. subflexa*

A. T. GROOT,*[1] A. CLASSEN,* O. INGLIS,† C. A. BLANCO,‡ J. LÓPEZ JR,§ A. TÉRAN VARGAS,¶ C. SCHAL,† D. G. HECKEL* and G. SCHÖFL**

*Department of Entomology, Max Planck Institute for Chemical Ecology, Hans-Knöll Strasse 8, 07745 Jena, Germany, †Department of Entomology and W. M. Keck Center for Behavioral Biology, North Carolina State University, Raleigh, NC 27695-7613, USA, ‡USDA-APHIS Biotechnology Regulatory Services, 4700 River Road, Riverdale, MD 20737, USA, §Areawide Pest Management Research Unit, USDA-ARS SPARC, 2771 F and B Road, College Station, TX 77845, USA, ¶Campo Experimental Las Huastecas, Carretera Tampico-Mante, km. 55, Villa. Cuauhtemoc, Tamaulipas, CP 89610, Mexico, **Leibniz Institute for Natural Product Research and Infection Biology, Beutenbergstrasse 11a, 07745 Jena, Germany

Abstract

The two moth species *Heliothis virescens* (Hv) and *H. subflexa* (Hs) are closely related, but have vastly different feeding habits. Hv is a generalist and an important pest in many crops in the USA, while Hs is a specialist feeding only on plants in the genus *Physalis*. In this study, we conducted a comparative population genetic analysis to assess whether and how generalist and specialist life styles are reflected in differences in population structures. In Hv 98% of the total variation occurred within populations. The overall differentiation (F_{ST}) between regions was 0.006 and even lower between years (0.0039) and hosts (0.0028). Analyses of population structure suggest that all individuals form one genetically homogeneous population, except for at most 12 individuals (6%) that diverged from this cluster. Population homogeneity likely results from the high mobility of Hv and its generalist feeding behaviour. Hs exhibited substantially more population structure. Even though 96% of the total variation was attributable to within-population variability, F_{ST}-values between Hs populations were 10 times higher than between Hv populations. Hs populations showed significant isolation by distance. Analyses of Hs population structure suggest at least two subpopulations and thus some degree of metapopulation structure. We speculate that the patchy distribution of *Physalis* – the exclusive food source of Hs – contributes to differences in population structure between these closely related species. The finding that the specialist shows more population differentiation than the generalist corroborates the notion that host specialization is not an evolutionary dead end but a dynamic trait.

Keywords: host plant distribution, metapopulation, population structure, structurama

Received 26 January 2011; revision received 20 March 2011; accepted 4 April 2011

Introduction

The process of speciation is a central theme in evolutionary biology. One of the most straightforward scenarios is that reduced or limited gene flow between populations can promote divergence between populations, either through natural selection, i.e. adaptation to local environments, or through drift (e.g. Dobzhansky & Pavlovsky 1957; Lande 1976; Slatkin 1985; Lynch & Hill 1986; Orr 1998; Doebeli & Dieckmann 2003; Nosil & Crespi 2004; Pathan *et al.* 2007; Raesaenen & Hendry 2008). Therefore, unravelling the structure of natural populations is an important step in understanding the presence and level of population divergence.

Correspondence: AT Groot, Fax: +49 3641 571502;
E-mail: agroot@ice.mpg.de, a.t.groot@uva.nl
[1]Current address: University of Amsterdam, IBED, Science Park 904, 1098 XH Amsterdam, the Netherlands

Among other factors, speciation may be accompanied by or driven by a change in niche breadth and the diversity of resources utilized. The overwhelming presence of specialist herbivores compared to generalists (Schoonhoven 2005) has given rise to the classical concept that host plant specialization is the derived evolutionary form of generalism and an evolutionary dead end (Mayr 1963). More recent work has provided convincing evidence against this generalization (Nosil & Mooers 2005). Several studies on insect herbivores have demonstrated that generalized lineages are often derived from specialists, so that host specialization appears to be a dynamic trait (Janz *et al.* 2001; Nosil 2002). Whether population divergence is more prevalent in generalist or specialist species is thus an interesting question requiring comparative studies. Moreover, generalist and specialist herbivores likely have different population structures. For example, specialists may show a more patchy distribution following the distribution of their host plants, while generalists may be subject to divergent host-related selection pressures.

The noctuid moths *Heliothis virescens* (F.) (Hv) and *H. subflexa* Guenée (Hs) are closely related species, but not considered to be sister species, that separated about 2.5 Mya (Mitter 1993; Fang *et al.* 1997; Cho *et al.* 2008). Both species occur sympatrically throughout North and South America and have overlapping activity times at night (Heath *et al.* 1991), but differ vastly in their host plant use. Hv, the tobacco budworm, is a generalist, feeding on more than 37 plant species from 14 different families (Waldvogel & Gould 1990; Sheck & Gould 1993; Blanco *et al.* 2007), and attaining pest status on important agricultural crops such as tobacco, cotton, soybean, tomato, sunflower and chickpea. These mainly cultivated annual plants represent an environment that is constantly changing over time and space, which is dynamically exploited due to the high mobility of Hv on both local and regional scales (Fitt 1989).

In contrast to Hv, Hs is a specialist, restricted to host plants in the genus *Physalis* (Solanaceae) (McElvare 1941; Sheck & Gould 1995). The commercial cultivation of *Physalis philadelphica* Lam. throughout Mexico favours the distribution of Hs (Groot *et al.* 2007), where it has become so abundant that it is an important pest in this area. In the USA, *Physalis* grows primarily as a weedy plant that is mostly patchily distributed, e.g. in roadside ditches or between crop fields. Due to the patchy distribution of the host plants, larval populations of Hs probably occur in small pockets throughout the USA.

Since Hv is one of four major heliothine pest species in North America, its ecology has been the subject of a vast number of studies (reviewed in Farrow & McDonald 1987; Fitt 1989). Mark–recapture studies and radar observations have shown high mobility of Hv (e.g. Hendricks *et al.* 1973; Wolf *et al.* 1986; Farrow & Daly 1987; Westbrook *et al.* 1994; Schneider 1999). For example, mark-recapture studies with pheromone traps confirmed that released sterile Hv males disperse in the range of 10 km (Schneider 1989, 1999), and even up to 113 km within 4–5 days, although with rapidly decreasing abundance (Hendricks *et al.* 1973; Raulston *et al.* 1982). Hv has also been found to be facultatively migratory (e.g. Schneider 1989) and passive movements of migrations from Mexico northward have been documented (Raulston *et al.* 1982, 1986).

Studies have also been conducted on population structure and gene flow in Hv (Sluss & Graham 1979; Korman *et al.* 1993; Mallet *et al.* 1993; Roehrdanz *et al.* 1994; Han & Caprio 2004). All but one of these studies (Sluss & Graham 1979) used pheromone traps to sample populations, so that only adult males were included in the analyses. Since males are attracted to the long-range pheromone of females, they are more likely to move over large distances, which may reduce a signature of population differentiation. For example, Han & Caprio (2004) showed in *Helicoverpa zea* (Boddie) that the genetic differentiation was significantly higher between populations when eggs were collected from the field, than when adult males were collected using pheromone traps. The one study that did show considerable allozyme polymorphism genetic structure in Hv used larvae collected from North Carolina to California instead of adult males (Sluss & Graham 1979). However, misinterpretation of genetic polymorphisms in the study of Sluss and Graham might have led to an overestimation of genetic differentiation (Mallet *et al.* 1993).

In contrast to the large number of population studies in Hv, so far no studies have been conducted on migration rates or gene flow between Hs populations. Due to the ephemeral nature of *Physalis* patches in the USA, it is likely that Hs exists as a metapopulation with limited gene exchange between (some) populations. Also, natural barriers, such as the two mountain ranges in Mexico, may reduce gene flow between western Mexico populations and populations from Texas or North Carolina. In this study we aim to assess the population structure of Hs and compare it to the population structure of Hv.

Even though there is no specific information on population structure in Hs, generally in moths short-range movements (up to approx. 10 km) can be distinguished from long-distance (passive) movements (Hendricks *et al.* 1973; Farrow & Daly 1987; Schneider 1989): In short-range movements moths orient at heights up to 10 m to new feeding, oviposition or mating sites. The flight is still controlled in speed and orientation and not necessarily downwind (Fitt 1989). Short-range

movements also include the attraction of males via sex pheromones that is usually oriented upwind (Kennedy 1983; Baker *et al.* 1988). In contrast, long-distance migratory movement in moths is largely independent from external stimuli, once it is triggered. Crossing the flight boundary layer, heliothine moths can be transported downwind by synoptic-scale wind systems and can thus cover hundreds of kilometres (Drake & Farrow 1985; Drake *et al.* 1981; Raulston *et al.* 1986; Westbrook *et al.* 1994).

Especially in the Southern part of the USA, heliothine moths like Hv have also developed another strategy to cope with seasonal environmental changes. By entering a facultative diapause as pupae, they can overcome periods of drought, cold or lack of food and oviposition sites (Schneider 2003). Migration and long distance movements often result in genetic homogeneity, as shown for many other mobile insects (Pashley 1989; Brower & Boyce 1991; Chapco *et al.* 1992; Zehnder *et al.* 1992; Bogdanowiez *et al.* 1993), but see also: Zhang *et al.* (2009). In contrast, overwintering diapause probably increases differentiation (Korman *et al.* 1993), as it constitutes a period where allele frequencies could change due to local selection or drift, but no dispersal takes place.

Host plant associated differentiation can be an additional driver of diversification in phytophagous insects (e.g. Ehrlich & Raven 1965; Jermy 1984; Pashley 1986; Thompson 1988; Jaenike 1990; Pashley *et al.* 1992; Stireman *et al.* 2005; Feder & Forbes 2008). Diminished gene flow between host races may result in (sympatric) speciation (e.g. Feder *et al.* 1990; Menken *et al.* 1992; Dres & Mallet 2002; Via & West 2008). To our knowledge, host plant associated genetic differences have not been assessed in Hv so far (e.g. Sluss & Graham 1979). However, host plant preference for oviposition was detected for a Hv population from the Virgin Islands that was to some extent distinguishable from Arizona and Mississippi strains (Waldvogel & Gould 1990). Additionally, Blanco *et al.* (2008) recently discovered that two Hv lab strains respond differently to lyophilized plant compounds. One of the strains grows well on a chickpea diet, which completely inhibits the development of another strain that grows well on a cotton diet. Hence, populations of the generalist Hv could be subdivided according to different host plant associations. Since Hs is a specialist, only feeding on *Physalis* plants, such a subdivision is not expected.

From knowledge on the ecology of both Hv and Hs, we hypothesize that (i) Hv is one panmictic population with perhaps some genetic differentiation based on host plant availability and (ii) Hs consists of two or more subpopulations based on the patchy distribution of their host plants.

Materials and methods

Populations sampled for population genetics

Population samples. We analysed the DNA of 184 Hv individuals from eight collections from different years, regions and host plants (see Fig. 1 and Table 1). Each Hv population sample consisted of 12 males and 11 females. All adults from North Carolina were derived from field-collected eggs, while Hv from Mississippi (MS), Texas (TX) and East Mexico (MXE) were collected as larvae. For Hs, a total of 179 individuals from 9 collections from different years and regions were analysed (see Table 2). All Hs individuals had been collected as larvae from *Physalis* plants. Individuals collected in the

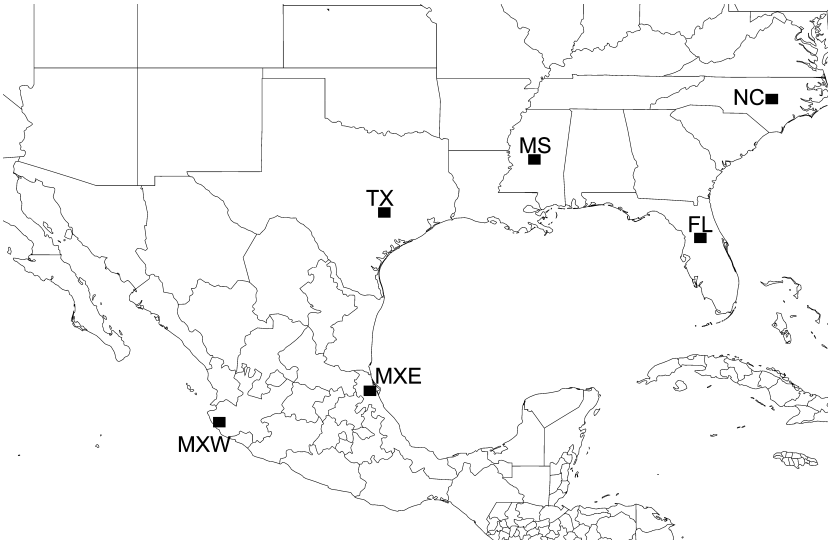

Fig. 1 Map showing the locations of the collection sites. For explanations of abbreviations and coordinates of collection sites, see Tables 1 and 2.

Table 1 Sites, years and host plants from which Hv eggs or larvae were collected. Each population sample consisted of 23 individuals (11 females and 12 males). North Carolina (NC), Mississippi (MS), Texas (TX), Eastern Mexico (MXE). tobacco (T), garbanzo (G) and cotton and garbanzo (CG)

	Regions			
	NC	MS	TX	MXE
2005	23 (T)		23 (G)	23 (CG)
2007	23 (T)		23 (G)	
2008	23 (T)	23 (T) + 23 (G)		

same year, on the same host plant and in the same region are consistently referred to as populations. Moths were reared to adulthood in the laboratory at North Carolina State University (NCSU) (Raleigh, NC, USA) and frozen at −80 °C until DNA extraction.

Generation of AFLP markers

DNA was extracted from half of an adult thorax (~20 mg tissue), using the QIAGEN DNeasy 96 Blood and Tissue Kit (QIAGEN, Valencia, CA, USA). Genomic DNA from each individual was digested with the two restriction enzymes *Eco*RI (6U) and *Mse*I. To the restricted DNA, *Eco*RI and *Mse*I adapters were added, which generated template DNA for the amplification of DNA fragments by polymerase chain reaction. The core sequences of both adapters are: *Eco*RI adapter 5′-CTCGTAGACTGCGTACC,5′-AATTGGTACGCAGTC-TAC; *Mse*I adapter 5′-GACGATGAGT-CCTGAG,5′-TA-CTCAGGACTCAT. After pre-amplifying the ligation product selective amplifications were conducted using 10–12 different primer combinations (see Table 3). The selectively amplified samples were loaded on 96-well polyacrylamide gels. Two 96-well plates per species were used. The plate layouts of the two 96-well microplates were designed such that samples of one population were equally divided onto the two plates, as well as within the plates. Five to 10 individuals were the same on both plates to ensure scoring consistency.

In Hv a total of 250 markers were scored using 10 primer combinations, in Hs a total of 243 markers were scored using 12 primer combinations (Table 3). Similar primer combinations were run for both species. However, scoring of gels occurred randomly, which resulted in scoring three overlapping primer combinations in both species. The markers were scored using the AFLP-Quantar™ *Pro* 1.0 software (KeyGene Products, Wageningen, The Netherlands). An internal control was the presence/absence pattern of the individuals present on both plates. Identical patterns of these individuals of same-sized bands on both gels indicated the same markers. Only those markers that were scored consistently on both gels were used for subsequent analysis. Scoring results were exported to a Microsoft® Office Excel file and transformed as required for statistical analysis.

Statistical analysis

Genetic diversity and population differentiation. Basic population genetic parameters were computed for both data sets in *AFLPsurv* 1.0 (Vekemans *et al*. 2002). Frequencies of the null allele were estimated using a Bayesian method with non-uniform prior allele frequency distributions (Zhivotovsky 1999) and assuming Hardy–Weinberg genotypic proportions. Expected heterozygosities were then calculated using the approach of Lynch & Milligan (1994) for dominant markers. Within populations or groups of populations, we calculated the percentage of polymorphic loci at the 5% level (i.e. considering only those loci with a major allele frequency of < 0.95 to be polymorphic).

To estimate genetic differentiation between populations or groups of populations, overall and pairwise F_{ST}-values and pairwise genetic distances (Nei's *D* after Lynch & Milligan 1994) were calculated in *AFLPsurv*.

Table 2 Sites and years from which larvae of Hs were collected. Between brackets the number of females (f) and males (m) included in each sample

	Regions				
	NC	FL	TX	MXE	MXW
2004	16 (16 f, 0 m)	12 (12 f, 0 m)	19 (19 f, 0 m)		
2005	27 (15 f, 12 m)		19 (12 f, 7 m)	17 (11 f, 6 m)	24 (9 f, 15 m)
2006	22 (12 f, 10 m)				
2007	23 (5 f, 18 m)				

NC: Clayton, North Carolina: 35° 39′ 58″ N, 78° 30′ 36″ W
FL: Gainesville, Florida: 29° 39′ 06″ N, 82° 19′ 29″ W
MS: Stoneville, Mississippi: 33° 25′04″ N, 90° 54′37″ W
TX: College Station, Texas: 30° 38′ 22″ N, 96° 21′ 39″ W
MXE: (eastern Mexico) Tampico, Tamaulipas, Mexico: 22° 13′ 02″ N, 97° 50′ 40″ W
MXW: (western Mexico) Chamela, Jalisco, Mexico: 19° 31′ 49″ N, 105° 03′ 47″ W.

Table 3 AFLP primer combinations used in both species. In bold are the primer combinations that were scored in both species

Hv primer combinations used			Hs primer combinations used		
Mse	EcoR	No. markers	Mse	EcoR	No. markers
ACA	AAG (700)	16	AAG	AAG (700)	4
ACG	ACG (800)	6	AAG	ACG (800)	5
AGG	ACG (800)	34	AAG	ACC (700)	31
AGG	ACC (700)	28	AAG	ACT (800)	21
AGG	ACT (800)	46	ACG	AAG (700)	26
			ACA	ACG (800)	30
CGA	**ACG (800)**	**7**	**CGA**	**ACG (800)**	**25**
CTG	ACC (700)	36	CTG	ACG (800)	20
CTG	ACT (800)	16	CAT	AAG (700)	25
ACG	**ACC (700)**	**37**	**ACG**	**ACC (700)**	**13**
CAT	**ACG (800)**	**24**	**CAT**	**ACG (800)**	**24**
			CTG	AAG (700)	19

Associations between geographic distance and genetic divergence matrices were tested using the Mantel test implemented in the R package Ecodist (Goslee & Urban 2007). One-tailed P values for the Mantel coefficient r (H_0: $r \leq 0$) were obtained by 10 000 permutations. Geographic distances between populations were obtained from Google maps, and $F_{ST}/(1–F_{ST})$ was used as the measure of genetic divergence.

Analysis of molecular variance (AMOVA). Genetic structure was further tested on different hierarchical levels (e.g. region of origin, year of collection, host-plant) using the AMOVA framework (Excoffier *et al.* 1992). In the case of dominant markers, AMOVA can only partition the genotypic variance, and not estimate the hierarchical variance of allele frequencies. For Hv, we examined three different population structure models with AMOVA. First, we grouped populations according to their region of origin (irrespective of year or host plant); second according to year, and finally according to host plant. For Hs, we examined two different population structure models, one where the populations were grouped according to region and one according to year. We used the program Arlequin 3.1 (Excoffier *et al.* 2005) to perform locus-by-locus AMOVA. One thousand permutations were performed to test the significance of variance components.

Bayesian analyses of population structure. Non-hierarchical genotypic clustering of individuals independent of the sampling regime was performed using two Bayesian model-based approaches. These models account for the presence of HW and linkage disequilibrium among alleles at each of the marker loci by introducing population groupings that minimize deviations from equilibrium within clusters. We used the methods implemented in the programs *Structure v2.3* (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007) and *Structurama* (Huelsenbeck & Andolfatto 2007) to determine the most likely number of genetically homogeneous clusters (K) needed to explain the observed data and the assignment of individuals into these clusters.

Using *Structure*, the most likely number of clusters for both data sets was estimated by determining the change in the marginal likelihood of the data $\Pr(X \mid K)$ when K was fixed to different values ($K = 1, 2, 3, \ldots, 8$). We used an ancestry model that allowed for admixture and correlated allele frequencies between populations. Under this model, individuals are fractionally assigned to clusters using a membership coefficient (interpretable as the fraction of the genome with membership in that cluster). We ran eight replicate Markov chains for each K for 4×10^5 iterations after a burn-in of 1×10^5 iterations. We also implemented the ΔK method of Evanno *et al.* (2005) to detect the level of structuring beyond which a further subdivision does not substantially improve the fit of the admixture model. ΔK is the second-order rate of change of the marginal likelihood function and takes into account both the gain in posterior probabilities over a range of K-values and the variance between independent runs at given values of K.

Memberships coefficients were averaged over all runs after 'label switching' heterogeneity had been accounted for using the software *CLUMPP* (Jakobsson & Rosenberg 2007) and visualized using the software *Distruct* (Rosenberg 2004).

Using *Structurama* (Huelsenbeck & Andolfatto 2007), the number of clusters and the assignment of individuals to those clusters were estimated simultaneously by applying a Dirichlet process prior, which treats both the assignment of individuals to populations and the number of populations as random variables. *Structurama* allows the user to run analyses where the concentration parameter α of the Dirichlet process prior (which shapes the prior probability of the number of clusters) is either set to a fixed value or itself treated as a random variable drawn from a Gamma hyperprior (see the program's documentation, http://fisher.berkeley.edu/structurama/manual.html). We used both approaches, first setting α such that the number of populations had a prior mean of $E(K) = 2$, $E(K) = 5$ and $E(K) = 10$, while under the second approach, for each run α was drawn from a differently parametrized gamma distribution with shape parameter k set to 1.8 or 2.4 and scale parameter θ set to 0.4 or 0.6. For each data set we thus performed seven analyses consisting of a single Markov-chain run for 5.0×10^6 cycles. Samples were drawn

from the chain every 125th cycle. The first 20 000 of the resulting 40 000 samples were removed as burn-in prior to analysis. The posterior probabilities of the number of populations given the data Pr($K \mid X$) were averaged across runs. For population assignment of individuals the mean partition, i.e. the partition of the sample that minimizes the squared distance to all partitions of the sample visited during an MCMC-run, was calculated (Huelsenbeck & Andolfatto 2007). The final placement of an individual in a cluster was based on the majority assignment across all replicate runs and visualized using *Distruct* (Rosenberg 2004).

Results

Genetic diversity

In Hv, between 40% and 55% of the 252 loci were polymorphic across populations and a moderate number (between 0 and 17) of alleles were found to be specific to given populations (i.e. the number of private alleles, see Table 4). Across different populations the estimated levels of expected heterozygosity were stable at 0.13–0.14 (Table 4). In contrast, in Hs between 80% and 95% of the 243 loci were polymorphic and no private alleles were detected in any of the populations (Table 4). The estimated levels of expected heterozygosity were twice as high compared to Hv, ranging between 0.29 and 0.33 (Table 4).

Genetic differentiation

In Hv, the genetic differentiation between any two groups was extremely low. F_{ST} values were lowest between the two MS groups collected from different host plants (F_{ST} = 0.0005) (Table 5) and highest between the two TX groups collected in different years (F_{ST} = 0.0140) (Table 5). Overall F_{ST} was significantly different from zero and marginally higher among regions (0.006) than among years (0.0039) and hosts (0.0028). Nei's genetic distance showed a trend similar to the F_{ST} values: It was lowest between the groups sampled in Mississippi from garbanzo and tobacco, and highest between populations from Texas sampled in 2005 and 2007 (Table 5). No significant isolation-by-distance signal was observed when plotting the genetic distance ($F_{ST}/(1–F_{ST})$) between all Hv populations against the geographic distances between each of the eight populations (Mantel r = 0.075, P = 0.34; Figure 2a).

In Hs, the genetic differentiation between any two populations was low as well, but F_{ST} values were consistently 10x higher than in Hv (Table 6). Overall F_{ST} was significantly different from zero when we compared all nine populations, the five geographic regions, and the four years. The lowest F_{ST} value was between TX and MXE, and between MXE and MXW (F_{ST} = 0), while the highest F_{ST} was between NC07 and FL04. A significant isolation-by-distance signal was observed in Hs when plotting the genetic distance ($F_{ST}/(1–F_{ST})$) for all populations against the distances between each of the nine populations (Mantel r = 0.48, P < 0.05; Figure 2b).

Partitioning of genetic variance

For Hv, no significant variation among groups of populations, irrespective of whether they were grouped according to region, year or host plants, was detected (Table 7). The variation among populations within each of these groupings explained approximately 2% of the total variation, respectively, and was highly significant (each P < 0.0001). The vast majority of detected genotypic variation (~98%) was assigned to the variation among individuals within populations (each P < 0.0001). Thus, although there was significant variation among populations, this variance component can neither be explained by geographic distance, temporal distance nor host plant origins (Table 7).

In Hs, the variation among geographic regions explained about 3% of the total variation, which was significant (P < 0.01), while the variation among populations within regions explained 1% of the variation (P < 0.01) (Table 8). Also in this case the vast majority of detected variation (96%) was due to the variation among individuals within populations. When comparing variation among years, only 0.5% of the total variance was explained by year, which was not significant. Thus, in Hs a small but significant portion of the total variance can be explained by geographic origin.

Bayesian analysis of population structure

Algorithms that infer structural patterns in sampled populations without prior assignment of individuals into groups gave inconclusive results. For Hv, using *Structure v2.3* we performed eight independent runs for each of the models with K (number of populations) = 1 to K = 8. An informal pointer to the 'true' number of K proposed in the *Structure* manual (the 'true' K is often found where the posterior probabilities for larger K begin to plateau) suggests K = 2 or 3 for Hv (Fig. 3a). Evanno's method (Evanno *et al.* 2005) found that K = 2 was the best value (Fig. 3a). The assignment of individuals into two clusters shows one cluster containing the vast majority of individuals while membership in the second cluster was restricted to a few samples spread evenly across populations (Fig. 4). In contrast to *Struc-*

Table 4 Genetic diversity within subpopulations of *Heliotis virescens* (Hv) and *H. subflexa* (Hs). *n* is the number of individuals, *P* is the percentage of polymorphic loci, H_e is the expected heterozygosity (± standard deviation), and #PA is the number of private alleles in the different subpopulations. Populations are denoted by site of origin, year of collection, and host plant (e.g. NC2005T denotes individiuals collected in NC in 2005 on tobacco)

	Hv populations							
	NC2005T	NC2007T	NC2008T	MS2008T	MS2008G	TX2005G	TX2007G	MXE2005CG
n	22	23	23	23	19	23	23	23
P (%)	55.2	48.0	42.1	40.5	50.0	43.3	44.8	42.5
H_e (± SD)	0.135 (0.008)	0.142 (0.008)	0.140 (0.008)	0.131 (0.008)	0.128 (0.007)	0.135 (0.008)	0.136 (0.009)	0.132 (0.008)
H_{eB} (± SD)	0.158 (0.009)	0.163 (0.010)	0.162 (0.010)	0.160 (0.010)	0.159 (0.009)	0.159 (0.010)	0.158 (0.010)	0.157 (0.009)
#PA	0	0	0	0	0	0	0	4

	Regions				Years			Hosts	
	NC	MS	TX	MXE	2005	2007	2008	Tobacco	Garbanzo
n	68	42	46	23	68	46	65	91	65
P (%)	47.6	46.8	46.0	42.5	44.0	44.4	42.9	46.0	43
H_e (± SD)	0.136 (0.008)	0.128 (0.008)	0.134 (0.009)	0.132 (0.009)	0.132 (0.009)	0.138 (0.008)	0.131 (0.009)	0.134 (0.008)	0.130 (0.008)
H_{eB} (± SD)	0.164 (0.013)	0.158 (0.012)	0.161 (0.012)	0.161 (0.010)	0.162 (0.013)	0.165 (0.012)	0.162 (0.13)	0.167 (0.015)	0.165 (0.014)
#PA	7	0	0	4	6	0	2	17	3

	Hs populations								
	FL04	MXE05	NC04	NC05	NC06	NC07	TX04	TX05	MXW05
n	11	16	15	26	21	22	18	16	19
P (%)	87.7	95.5	79.4	83.5	87.2	86.4	93.4	93.0	88.1
H_e (± SD)	0.305 (0.011)	0.329 (0.009)	0.304 (0.011)	0.294 (0.010)	0.305 (0.010)	0.288 (0.010)	0.319 (0.009)	0.311 (0.010)	0.313 (0.010)
H_{eB} (± SD)	0.307 (0.009)	0.321 (0.009)	0.304 (0.008)	0.299 (0.008)	0.305 (0.008)	0.300 (0.008)	0.321 (0.010)	0.315 (0.010)	0.313 (0.009)
#PA	0	0	0	0	0	0	0	0	0

	Regions					Years			
	FL	NC	TX	MXE	MXW	2004	2005	2006	2007
n	11	86	35	16	19	46	79	21	22
P (%)	87.7	86.4	94.7	95.5	88.1	91.8	90.5	87.2	86.4
H_e (± SD)	0.305 (0.011)	0.288 (0.010)	0.309 (0.009)	0.329 (0.009)	0.313 (0.010)	0.303 (0.010)	0.301 (0.009)	0.305 (0.010)	0.288 (0.010)
H_{eB} (± SD)	0.311 (0.010)	0.300 (0.009)	0.325 (0.013)	0.325 (0.010)	0.317 (0.010)	0.317 (0.011)	0.317 (0.012)	0.305 (0.008)	0.300 (0.008)
#PA	0	0	0	0	0	0	0	0	0

Table 5 Genetic differentiation (F_{ST}) (lower diagonal), and Nei's genetic distance (upper diagonal) between groups of *H. virescens*. Populations derive from North Carolina (NC), Mississippi (MS), Texas (TX), and Mexico (MXE), collected on either tobacco (T) garbanzo (G) or cotton and garbanzo (CG) in years 2005, 2007 and/or 2008

Populations								
	NC2005 T	NC2007 T	NC2008 T	MS2008 T	MS2008 G	TX2005 G	TX2007 G	MXE2005 CG
NC2005 T		0.0003	0.0003	0.0016	0.0012	0.0010	0.0013	0.0016
NC2007 T	0.0018		0.0006	0.0012	0.0009	0.0018	0.0015	0.0004
NC2008 T	0.0017	0.0035		0.0003	0.0004	0.0015	0.0008	0.0004
MS2008 T	0.0105	0.0074	0.0020		0.0001	0.0015	0.0017	0.0006
MS2008 G	0.0077	0.0061	0.0026	0.0005		0.0018	0.0013	0.0005
TX2005 G	0.0064	0.0113	0.0096	0.0096	0.0118		0.0022	0.0013
TX2007 G	0.0081	0.0094	0.0052	0.0111	0.0083	0.0140		0.0013
MXE2005 CG	0.0101	0.0029	0.0027	0.0037	0.0035	0.0085	0.0085	
Overall F_{ST}	**0.0067 ($P < 0.0001$)**							
Regions								
	NC	MS	TX	MXE				
NC		0.0010	0.0008	0.0010				
MS	0.0067		0.0011	0.0006				
TX	0.0054	0.0073		0.0009				
MX	0.0064	0.0041	0.0060					
Overall F_{ST}	**0.0060 ($P < 0.0001$)**							
Years								
	2005	2007	2008					
2005		0.0005	0.0007					
2007	0.0030		0.0007					
2008	0.0043	0.0044						
Overall F_{ST}	**0.0039 ($P < 0.0020$)**							
Hosts								
	Tobacco	Garbanzo						
Tobacco		0.0004						
Garbanzo	0.0028							
Overall F_{ST}	**0.0028 ($P < 0.007$)**							

ture the algorithm implemented in *Structurama* (Huelsenbeck & Andolfatto 2007) explicitly models the most likely number of underlying groups. *Structurama* showed a posterior probability maximum for three distinct population groups (Fig. 3a). The partitioning of Hv individuals into groups, however, revealed only one major cluster containing all but four individuals distributed into two separate groups. The four individuals that were excluded from the main cluster by *Structurama* grouped in the main cluster inferred by *Structure*.

For Hs, across eight replicate *Structure* runs the posterior probability of K plateaued after $K = 3$ to 4, and Evanno's ΔK showed a peak at $K = 3$ (Fig. 3b). The assignment of individuals into three or four clusters suggests a major division between all North Carolina populations (NC2004, NC2005, NC2006, NC2007) and the remainder (Florida, Texas and Mexico) (Fig. 5). Mainly within the latter group several individuals showed variable degrees of co-ancestry in two further source populations without a clear geographic connection (although the 'red' group seems to be over-represented in the FL2004 and TX2004 populations, Fig. 5). *Structurama* supported $K = 5$ in 5 runs and $K = 6$ in two runs (Fig. 3b). However, most groups inferred by *Stucturama* consist of a few isolated individuals; the meaningfully interpretable clusters are similar to those inferred by *Structure* (Fig. 5).

Discussion

The structure of populations from different regions depends on the extent of gene flow between geographically separated populations (Slatkin 1987). Five classical models associated with contrasting gene flow events have been recognized (Tero *et al.* 2003), ranging from panmictic population structure to stepping-stone (Kimura & Weiss 1964), sink-source (Pulliam 1986; Gaggiotti & Smouse 1996) and metapopulations (e.g. Levins 1970; Tero *et al.* 2003) to completely separated populations.

Overall, the eight populations of Hv were found to be largely genetically uniform and thus resemble one panmictic population, similar to what has been found

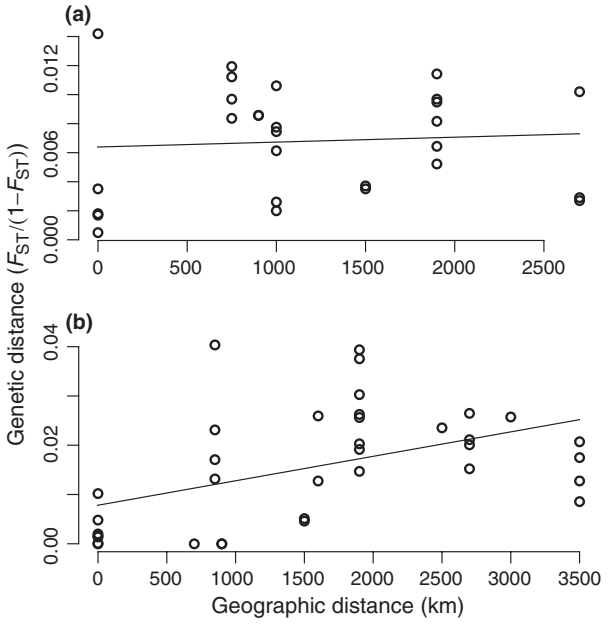

Fig. 2 Isolation by distance plotted for the Hv populations (a) and the Hs populations (b).

previously in this species (Sluss & Graham 1979; Korman *et al.* 1993; Mallet *et al.* 1993; Roehrdanz *et al.* 1994; Han & Caprio 2004). There was no genetic isolation due to geographical or temporal distance, or differentiation based on host plant association. A small but significant percentage of the variance was explained by differentiation among populations within groups. Since this part of the variance was independent of whether we grouped populations according to regions, years or host plants, we lack a straightforward biological explanation for this variance component. Stochastic effects due to the small sample sizes within populations may be a possible interpretation.

The absence of population structure in Hv might be the result of a bottleneck event in the past with subsequent expansion of one population in North America within a short evolutionary time (Roehrdanz *et al.* 1994). Alternatively, the populations may either expand annually from a source population and/or there is free gene flow between overwintering populations and migrants (Roehrdanz *et al.* 1994). Both scenarios are supported by the high mobility of Hv and its polyphagy, as these traits allow expansion. Extensive agricultural activity in the USA might have boosted the spread of Hv, as e.g. with every crop harvest the organism has to switch to adjacent or more distant patches.

In contrast to Hv, in Hs we did find population structure. Even though the overall and pairwise F_{ST} values were low, they were 10 times higher than in Hv. In Hs we found significant isolation by geographic distance but not temporal distance. Thus, the overall F_{ST} between regions was significant, while the F_{ST} between years was not. The Hs populations can be meaningfully clustered in at least two groups: a North Carolina cluster and a group comprising the Florida, Texas and Mexico populations. The Florida and Texas populations from 2004 seem to be linked by a number of individuals that share co-ancestry in a distinct source population (red segments in Fig. 5) and the same is true for the Texas and western Mexico 2005 populations (yellow segments in Fig. 5). This indicates some degree of metapopulation structure in Hs. The fact that we found only two population groups is likely due to the relatively low number of populations sampled.

Comparing the level of population differentiation between Hv and Hs thus indicates that the specialist Hs shows more population divergence than the generalist Hv, suggesting that the specialist is more likely to diversify. This corroborates the notion that host specialization appears to be a dynamic trait (Janz *et al.* 2001; Nosil 2002). The generalist does not seem to be subject to divergent host-related selection pressures, although surprisingly we found 17 alleles that were exclusively present in 11 Hv individuals that were collected from tobacco. These private alleles may be directly or indirectly involved in the ability to feed on commercial tobacco plants containing high levels of nicotine. If so, such an adaptation might be in an early evolutionary stage, because despite these private alleles the overall differentiation between the two putative strains is very low. Selection pressure on different hosts might drive divergence either at loci that encode phenotypically important information or at loci that are linked to such loci. Future studies will focus on possible host plant differentiation in Hv.

Temporal differentiation

Even though in Hv pairwise comparisons between regions did not reveal an increase of F_{ST} with increasing distance, such a tendency was detectable when comparing single populations that were collected in 2005. For example, the population differentiation between NC 2005 and MX 2005 was slightly higher (0.01) than between NC 2005 and TX 2005 (0.006) (see Table 5), resembling a stepping-stone population model. That these tendencies did not appear when the populations were grouped according to region might be a consequence of the unbalanced sampling among regions, years and hosts. A positive correlation between genetic and geographic distance might disappear because e.g. the larvae from MS included only populations that were sampled in 2008, whereas the larvae from NC also included populations collected in 2005 and 2007.

Table 6 Genetic differentiation (F_{ST}) (lower diagonal), and Nei's genetic distance (upper diagonal) between groups of *H. subflexa*. Populations derive from Florida (FL), North Carolina (NC), Texas (TX), Eastern Mexico (Tampico, Tamaulipas) (MXE), and Western Mexico (Chamela, Jalisco) (MXW), collected in years 2004, 2005, 2006 and/or 2007

Populations									
	FL04	NC04	NC05	NC06	NC07	TX04	TX05	MXE05	MXW05
FL 2004		0.0058	0.0099	0.0075	0.0171	0.0058	0.0116	0.0109	0.0116
NC 2004	0.013		0	0.0009	0.0042	0.0066	0.0085	0.0097	0.0057
NC 2005	0.0226	0.0001		0	0.0007	0.0115	0.0088	0.0066	0.0037
NC 2006	0.0168	0.002	0		0.002	0.0117	0.0135	0.0092	0.0079
NC 2007	0.0388	0.0101	0.0016	0.0048		0.0163	0.0169	0.0115	0.0088
TX 2004	0.0126	0.0145	0.0256	0.025	0.0362		0.0007	0	0.0024
TX 2005	0.0253	0.0188	0.0199	0.0294	0.0379	0.0014		0	0.0021
MXE 2005	0.023	0.0207	0.015	0.0197	0.0258	0	0		0
MXW2005	0.0251	0.0126	0.0085	0.0172	0.0203	0.0051	0.0046	0	
Overall F_{ST}	**0.0152 (*P* < 0.0001)**								
Regions									
	FL	NC	TX	MXE	MXW				
FL		0.0111	0.0083	0.0109	0.0116				
NC	0.0257		0.0122	0.0102	0.0071				
TX	0.0183	0.0281		0	0.0018				
MXE	0.023	0.0233	0		0				
MXW	0.0251	0.0167	0.0039						
Overall F_{ST}	**0.0157 (*P* < 0.0001)**								
Years									
	2004	2005	2006	2007					
2004		0.0034	0.0046	0.0095					
2005	0.0079		0.0048	0.0061					
2006	0.0104	0.011		0.002					
2007	0.0222	0.0145	0.0048						
Overall F_{ST}	**0.0118 (*P* < 0.0001)**								

Table 7 AMOVA analyses (Arlequin3.1) partitioning genetic variance in *H. virescens* among regions, years, and hosts. The category 'regions' contains the groups North Carolina, Mississippi, Texas, Mexico; the category 'years' consists of the groups 2005, 2007, 2008 and 'hosts' includes the groups tobacco and garbanzo. Further information is given in the text. d.f. = degree of freedom; Number of permutations: 1000

Source of variation	d.f.	Sum of squares	Variance	Percentage of variance	*P*
(a) Regions					
Among regions	3	111.098	−0.031	−0.13	0.5699
Among populations, within single regions	3	107.202	0.532	2.21	< 0.0001
Within populations	172	4060.007	23.604	97.92	< 0.0001
Total	178	4278.307	24.105	100	
(b) Years					
Among years	2	77.532	0.042	0.18	0.2112
Among populations, within single years	4	140.768	0.473	1.96	< 0.0001
Within populations	172	4060.007	23.605	97.86	< 0.0001
Total	178	4278.307	24.120	100	
(c) Hosts					
Among hosts	1	38.655	0.039	0.16	0.2375
Among populations, within single hosts	5	177.805	0.520	2.14	< 0.0001
Within populations	153	3623.271	23.681	97.69	< 0.0001
Total	159	3839.731	24.240	100	

Table 8 AMOVA analyses (Arlequin3.1) partitioning genetic variance in *H. subflexa* among regions and years. The category 'regions' contains the groups Florida, North Carolina, Texas, Eastern Mexico and Western Mexico; the category 'years' consists of the groups 2004, 2005, 2006 and 2007. Further information is given in the text. d.f. = degree of freedom; Number of permutations: 1000

Source of variation	d.f.	Sum of squares	Variance	Percentage of variance	*P*
(a) Regions					
Among regions	4	306.865	1.0834	2.86	0.00978
Among populations, within single regions	4	180.445	0.412	1.09	0.00196
Within populations	169	6155.876	36.425	96.06	< 0.0001
Total	177	6643.185	37.921	100	
(b) Years					
Among years	3	213.915	0.198	0.50	0.22178
Among populations, within single years	5	299.615	1.259	3.20	< 0.0001
Within populations	160	6054.246	37.839	96.29	< 0.0001
Total	168	6567.775	39.297	100	

Genetic distance between the adjacent populations of NC and MS might thus be higher than expected, because temporal distance increases the differentiation. This idea is supported by the fact that genetic distance between North Carolina and Mississippi also increased over time in the single population comparison, whereas an increase in F_{ST} was not detectable among the groups in different years. A significant between-year F_{ST} indicates that Hv populations in consecutive years are formed from largely unrelated individuals, most likely due to exchange of migrating individuals.

In Hs, the fact that the populations from different regions clustered irrespective of their collection years suggests that overwintering occurs locally without migration to one area where panmixis may occur. Over multiple years, Gould and coworkers have tried to find overwintering pupae in North Carolina, by digging in the soil as well as by setting up large field tents over infested areas in the Fall to assess moth emergence in the Spring, without any success (F. Gould, personal communication; N. Benda, A.T. Groot, J. Petzold and F. Gould personal observation). However, pheromone traps catch Hs males in the Spring, before *Physalis* plants begin fruiting, suggesting that males come from local emergence sites (A.T. Groot and F. Gould personal observation). This observation, together with the evidence of geographic population structure irrespective of collection years, suggests that Hs overwinters locally and does not migrate to overwintering sites.

Use of AFLP markers

May the use of AFLP markers have biased our results? On average, 46% of the AFLP loci in the *H. virescens* populations were polymorphic, i.e. 15% less than Mallet and co-workers (1993) found when investigating 23 enzyme loci in Hv. Also, there was a slightly lower average expected heterozygosity (0.135 or 0.159, depending on the method), compared to Mallet *et al.* (1993; (0.172)), which is surprising as AFLP analyses tend to estimate higher levels of genetic variation than allozyme analyses (McMichael & Prowell 1999).

In Hs, we found on average 90% polymorphic AFLP loci and 2-fold higher heterozygosity than in Hv (see Table 4). This is unexpected in light of the higher degree of gene flow among Hv populations than among Hs populations and the generally much larger population sizes of Hv in the USA (e.g. Fitt 1989; Williams 2000) as compared to Hs. However, this may point to one or more bottlenecks occurring in the demographic past of Hv populations as suggested by fluctuations in Hv population sizes in the US over the past 40 years. When DDT was introduced and widely used in cotton fields to control the boll weevil, Hv populations in these cotton fields expanded greatly due to the elimination of parasitoids that had previously held these populations in check. The rapid evolution of DDT resistance in Hv necessitated a switch to organophosphorus compounds, to which Hv then developed resistance (Heckel *et al.* 1998). Then pyrethroids were used to control Hv populations, which was successful to a certain extent and led to a severe reduction of Hv populations until pyrethroid resistance developed (Park & Taylor 1997; Park *et al.* 1997; Zhao *et al.* 2000; Cho *et al.* 2008). Currently, with the wide use of Bt-cotton, Hv populations appear to be declining again (Caprio *et al.* 2004; Micinski *et al.* 2008).

Hs populations are much smaller, at least in the USA, but likely also more stable. This may explain the relatively high genetic diversity in Hs, which is similar to that found in the highly mobile corn pest *Ostrinia nubilalis* (Lepidoptera: Pyralidae) collected in the Midwestern United States (mean percentage of polymorphic loci = 83%; mean H_e = 0.31) (Krumm *et al.* 2008). As

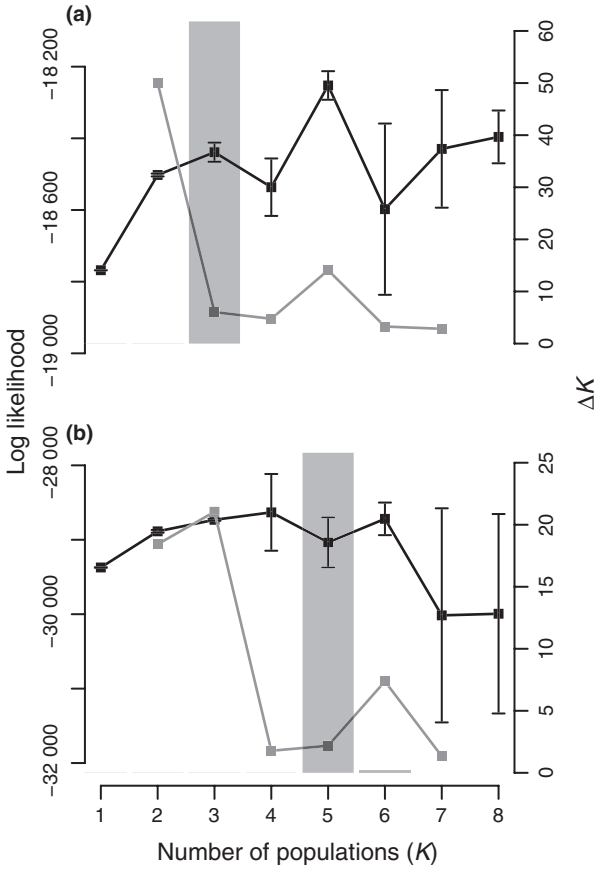

Fig. 3 Estimates of the most likely number of (a) *Heliothis virescens* (Hv) populations and (b) *H. subflexa* (Hs) populations. Black squares show the marginal log likelihoods of the data Pr(X | *K*) when the number of populations (*K*) is fixed to different values averaged over eight *structure* runs. The grey squares denote Δ*K*, an *ad hoc* indicator of the uppermost hierarchical level of structure detected, based on the rate of change in Pr(X | *K*) between successive *K* values. The grey bar denotes the posterior probability distributions Pr(*K* | X) for the number of populations averaged over seven *stucturama* runs where *K* is treated as a random variable.

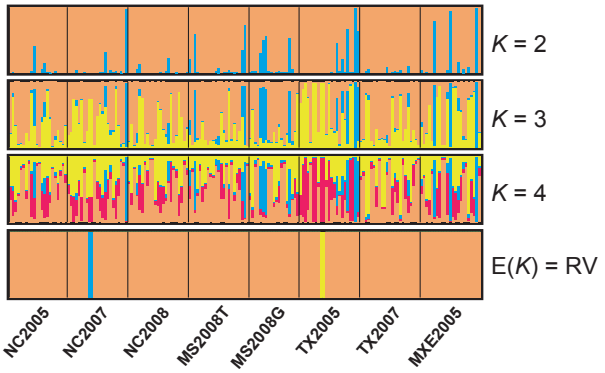

Fig. 4 Estimated population structure of *Heliothis virescens* (Hv). Analyses were performed in *structure* under an admixture model with the numbers of populations fixed (*K* = 2, 3 and 4), and in *structurama* under a no-admixture model treating the number of populations as a random variable (*E*(*K*) = RV). Individuals are represented by vertical lines partitioned in *K* coloured segments that represent the individuals' membership coefficient averaged over eight structure runs at each *K* and seven *structurama* runs under different prior expectations of *K*.

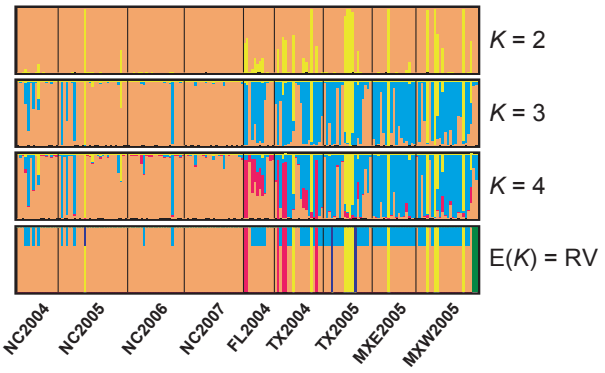

Fig. 5 Estimated population structure of *Heliothis subflexa* (Hs). Analyses were performed in *structure* under an admixture model with the numbers of populations fixed (*K* = 2, 3 and 4), and in *structurama* under a no-admixture model treating the number of populations as a random variable (*E*(*K*) = RV). Individuals are represented by vertical lines partitioned in *K* coloured segments that represent the individuals' membership coefficient averaged over eight structure runs at each *K* and seven *structurama* runs under different prior expectations of *K*. The assignment of individuals by *structurama* generated two distinct outcomes: in two runs *K* = 6 populations but two major clusters were supported, in five runs *K* = 5 populations but only one major cluster was supported.

genetic polymorphism can be maintained in the face of differential selection in heterogeneous environments (Bossard & Scriber 1995; Yeaman & Jarvis 2006), these results may hint to the possibility that (i) the environment of Hs is heterogeneous and/or (ii) Hs undergoes differential selection pressures in different environments and/or through time. A component of the heterogeneous environment may be the variable presence and abundance of Hv at different times and in different regions, which may cause differential selection pressures due to communication interference in the sex pheromone channel (Groot *et al.* 2009, 2010).

Compared to other Lepidoptera whose population structure was also analysed with AFLPs, the genetic diversity that we found in both Hv and Hs is high. For example, in Japanese populations of *Pieris rapae* and *P. melete* (Lepidoptera: Pieridae) the percentage of polymorphic loci ranged from 9.5% to 22% while H_e ranged from 0.03 to 0.06 (Takami *et al.* 2004). As mentioned above, low (expected) heterozygosity means low genetic variability, which is expected when population sizes are

small or have been (repeatedly) reduced by bottlenecks. *Pieris rapae* was introduced in Japan about 2 kya and may thus have gone through a bottleneck, while *P. melete* does not migrate over long distances and population sizes are very small (Takami *et al.* 2004).

Population structure detection

The two methods of analysis we used to infer levels of substructure led to conflicting results, to some extent. The ΔK criterion (Evanno *et al.* 2005), which uses the posterior likelihoods of K provided by *structure* to infer the most likely number of subpopulations, supported one or two subpopulations fewer than the clustering algorithm implemented in *Structurama*. This may be reasonable since the ΔK is an indicator of the uppermost hierarchical level of structure supported by the data (Evanno *et al.* 2005). However, the additional populations introduced by *Structurama* included only very few individuals both in Hv and Hs. Moreover, these individuals were generally not recognized as distinct populations by *Structure* even if the analysis was run fixing K to a higher number. In both cases the arguably most likely numbers of biologically meaningful subgroupings were best inferred by Evanno's ΔK. However, the drawback of ΔK is that it cannot support a single panmictic cluster (which we would suggest as the most likely scenario for Hv), because it is based on the rate of change in the log probability of the data *between* successive K values. It thus seems that the final interpretation of whether an inferred population structure is real should take into account whether the actual assignments can be meaningfully interpreted and not necessarily take the inferred numbers of populations at face value.

Evolutionary potential of generalists vs. specialists

Hv and Hs belong to an insect group with a typical distribution of specialists and generalists: the noctuid subfamily Heliothinae, which contains about 365 species worldwide and originated about 20 Mya (Cho *et al.* 2008). Larvae of most species feed mainly on flowers and fruits. The majority are host plant specialists, but some genera are very diverse, containing polyphagous, oligophagous, or monophagous species. Comparison of food plant lists with phylogenetic analysis indicates that the highest incidence of polyphagy occurs in the 'Heliothis group' containing the genera *Australothis*, *Heliocheilus*, *Helicoverpa*, and *Heliothis*; the latter two account for several significant agricultural pests (Cho 1997). A recent comprehensive molecular phylogeny has shown that in groups diverging before the 'Heliothis group', host-plant specialization is the rule and generalists derive from them very rarely (Cho *et al.* 2008). However, the 'Heliothis group' not only has a much higher incidence of polyphagy, shifts to or from monophagy are more frequent and at least six must be invoked to explain food habits of the 23 species with known hosts. In an analysis of the '*Heliothis virescens* subgroup' containing Hv and Hs, these two were shown to be closely related, but not sister species as each is more closely allied with different species restricted to South America (Poole *et al.* 1993). Yet host plant records for the rest of this subgroup are insufficient for any confident statement about whether their common ancestor was restricted to Solanaceae like Hs, or polyphagous like Hv, despite the opinion expressed favouring the latter idea (Poole *et al.* 1993).

Cho *et al.* (2008) have elaborated on a hypothesis of Hardwick (1965) to explain why the frequency of niche breadth shifts should be so much greater in the 'Heliothis group' compared to the other Heliothinae. They suggested that life-history differences imposed different types of constraints on the changeability of host plant range. Adults of groups such as the large mostly monophagous genus *Schinia* lay a few large eggs, often with specialized ovipositors for placement deep within host flowers, and remain on or near the host plant, protected by cryptic colouration. These constraints may provide fewer 'accidental opportunities' to broaden host range, than traits of adults of the 'Heliothis group' which typically are larger, lay many more eggs with unspecialized ovipositors on different exposed parts of the host plant, are not cryptic in a host-specific manner, and are highly vagile or migratory. This suite of traits is conducive of, but not obligatory for, host-range expansion; and niche contraction could evolve if favoured by other selective agents. In the case of Hs, the release from parasitoid and predator pressure due to the enemy-free space provided by the calyx that covers the developing fruit of the host *Physalis* has been cited as one such counter-selective pressure.

The idea that specialization is an evolutionary 'dead end' is partly based on the supposition that a successful response to certain selection pressures promoting specialization limits future evolvability. These could include competition for resources, predator avoidance, lower costs of information processing, and lower costs of searching for suitable habitat and mates. Response to strong directional selection results in a reduction of heritable genetic variation which can further restrict evolvability; for example in chrysomelid beetles of the genus *Ophraella*, absence of genetic variability in traits enabling the use of novel host plants can hinder the evolution of a broader host range (Futuyma *et al.* 1995). In the Hv–Hs comparison, however, the greater specialization of Hs is not accompanied by a decrease in

genetic variation, at least relative to the generalist Hv. If Hs evolved from a generalist ancestor like Hv, a reduction in dispersal promoted by the patchy nature of its more restricted host plant distribution could still help to maintain high levels of intra-population variation, at least temporarily. Additional comparisons of closely related generalist/specialist pairs across lineages with measurable rates of niche expansion or contraction will be required to determine whether, among specialists, there is a positive correlation between genetic variability and the tendency to evolve polyphagy.

In conclusion, comparing the population structures of Hv and Hs, we found that Hv is one panmictic population while Hs seems to be a metapopulation with geographic isolation by distance between populations. These population structures are consistent with our results on the intraspecific variation of the sex pheromone of Hv and Hs. In Hv most of the intraspecific variation in sex pheromone disappeared when rearing these populations for 4–5 generations in the laboratory (A.T. Groot and O. Inglis unpublished result), indicating that this variation is largely determined by environmental and not genetic factors. In contrast, in Hs the geographic differences between the sex pheromone remained and we are currently mapping the genetic variation (Groot *et al*. in preparation). Our finding that Hv is one panmictic population while Hs is a metapopulation consisting at least two subpopulations predicts that also for other traits intraspecific variation in Hv is likely due to environmental variance while in Hs such variation is likely due to genetic variation.

Acknowledgments

We thank Rob Meagher for his help with the collection of Hs in Florida in 2004, and two anonymous reviewers for suggestions that improved the manuscript. This research was supported by the National Research Initiative of the USDA Cooperative State Research, Education and Extension Service (CSREES), grant # 2005–00896, the W. M. Keck Center for Behavioral biology and the Max Planck Gesellschaft.

References

Baker TC, Hansson BS, Löfstedt C, Löfqvist J (1988) Adaptation of antennal neurons in moths is associated with cessation of pheromone-mediated upwind flight. *Proceedings of the National Academy of Sciences, USA*, **85**, 9826–9830.

Blanco CA, Terán-Vargas AP, López JD, Kauffman JV, Wei XK (2007) Densities of *Heliothis virescens* and *Helicoverpa zea* (Lepidoptera: Noctuidae) in three plant hosts. *Florida Entomologist*, **90**, 742–750.

Blanco CA, Terán-Vargas AP, Abel CA *et al.* (2008) Plant host effect on the development of *Heliothis virescens* F. (Lepidoptera: Noctuidae). *Environmental Entomology*, **37**, 1538–1547.

Bogdanowiez SM, Wallner WE, Bell J, Odell TM, Harrison RG (1993) Asian gypsy moths (Lepidoptera: Lymantriidae) in North America: evidence from molecular data. *Annals of the Entomological Society of America*, **86**, 710–715.

Bossard JL, Scriber JM (1995) Maintainance of ecologically significant genetic variation in the tiger swallowtail butterfly through differential selection and gene flow. *Evolution*, **49**, 1163–1171.

Brower AVZ, Boyce TM (1991) Mitochondrial DNA variation in monarch butterflies. *Evolution*, **45**, 1281–1286.

Caprio MA, Faver MK, Hankins G (2004) Evaluating the impacts of refuge width on source-sink dynamics between transgenic and nontransgenic cotton. *Journal of Insect Science*, **4**, Available online: insectscience.org/4:3.

Chapco W, Kelln RA, McFadyen DA (1992) Intraspecific mitochondrial DNA variation in the migratory grasshopper *Melanoplus sanguinipes*. *Heredity*, **69**, 547–557.

Cho S (1997) *Molecular Phylogenetics of the Heliothinae (Lepidoptera: Noctuidae) Based on the Nuclear Genes for Elongation Factor-1a and Dopa Decarboxylase*. PhD thesis, University of Maryland, College Park.

Cho S, Mitchell A, Mitter C *et al.* (2008) Molecular phylogenetics of heliothine moths (Lepidoptera: Noctuidae: Heliothinae), with comments on the evolution of host range and pest status. *Systematic Entomology*, **33**, 581–594.

Dobzhansky T, Pavlovsky O (1957) An experimental study of interaction between genetic drift and natural selection. *Evolution*, **11**, 311–319.

Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. *Nature*, **421**, 259–264.

Drake VA, Farrow RA (1985) A radar and aerial trapping study of an early spring migration of moths (Lepidoptera) in inland South Wales *Australian*. *Journal of Ecology*, **10**, 223–235.

Drake VA, Helm KF, Readshaw JL, Reid DG (1981) Insect migration across bass strait during Spring – a radar study. *Bulletin of Entomological Research*, **71**, 449–466.

Dres 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.

Ehrlich PR, Raven PH (1965) Butterflies and plants: a study in coevolution. *Evolution*, **18**, 586–608.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure: extensions to linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Fang QQ, Cho S, Regier JC *et al.* (1997) A new nuclear gene for insect phylogenetics: dopa decarboxylase is informative

of relationships within heliothinae (Lepidoptera: Noctuidae). *Systematic Biology*, **46**, 269–283.

Farrow RA, Daly JC (1987) Long-range movements as an adaptive strategy in the genus *Heliothis* (Lepidoptera, Noctuidae) – a review of its occurrence and detection in 4 pest species. *Australian Journal of Zoology*, **35**, 1–24.

Farrow RA, McDonald G (1987) Migration strategies and outbreaks of noctuid pests in Australia. *Insect Science and Its Application*, **8**, 531–542.

Feder JL, Forbes AA (2008) Host fruit odor discrimination and sympatric host-race formation. In: *Specialization, Speciation and Radiation: The Evolutionary Biology of Herbivorous Insects* (ed Tilmons KJ). pp. 101–106, University Press, Berkely, CA.

Feder JL, Chilcote CA, Bush GL (1990) Regional, local and microgeographic allele frequency variation between apple and hawthorn populations of *Rhagoletis pomonella* in Western Michigan. *Evolution*, **44**, 595–608.

Fitt GP (1989) The ecology of *Heliothis* species in relation to agroecosystems. *Annual Review of Entomology*, **34**, 17–52.

Futuyma DJ, Keese MC, Funk DJ (1995) Genetic constraints on macroevolution – the evolution of host affiliation in the leaf beetle genus *Ophraella*. *Evolution*, **49**, 797–809.

Gaggiotti OE, Smouse PE (1996) Stochastic migration and maintainance of genetic variation in sink populations. *American Naturalist*, **147**, 919–945.

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**, 1–19.

Groot AT, Santangelo RG, Ricci E, Brownie C, Gould F (2007) Differential attraction of *Heliothis subflexa* males to synthetic pheromone lures in eastern US and western Mexico. *Journal of Chemical Ecology*, **33**, 353–368.

Groot AT, Estock ML, Horovitz JL *et al.* (2009) QTL analysis of sex pheromone blend differences between two closely related moths: insights into divergence in biosynthetic pathways. *Insect Biochemistry and Molecular Biology*, **39**, 568–577.

Groot AT, Blanco CA, Classen A *et al.* (2010) Variation in sexual communication of the tobacco budworm, *Heliothis virescens*. *Southwestern Entomologist*, **35**, 367–372.

Han Q, Caprio MA (2004) Evidence from genetic markers suggests seasonal variation in dispersal in *Heliothis virescens* (Lepidoptera: Noctuidae). *Environmental Entomology*, **33**, 1223–1231.

Hardwick DF (1965) The corn earworm complex. *Memoirs of the Entomological Society of Canada, No. 40*. The Entomological Society of Canada, Ottawa, Ontario.

Heath RR, McLaughlin JR, Proshold F, Teal PEA (1991) Periodicity of female sex pheromone titer and release in *Heliothis subflexa* and *H. virescens* (Lepidoptera, Noctuidae). *Annals of the Entomological Society of America*, **84**, 182–189.

Heckel DG, Bryson PK, Brown TM (1998) Linkage analysis of insecticide-resistant acetylcholinesterase in *Heliothis virescens*. *Journal of Heredity*, **89**, 71–78.

Hendricks DE, Graham HM, Raulston JR (1973) Dispersal of sterile tobacco budworms from release points in Northeastern Mexico and Southern Texas. *Environmental Entomology*, **2**, 1085–1088.

Huelsenbeck JP, Andolfatto P (2007) Inference of population structure under a Dirichlet process model. *Genetics*, **175**, 1787–1802.

Jaenike J (1990) Host specialization in phytophagous insects. *Annual Review of Ecology and Systematics*, **21**, 243–274.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Janz N, Nyblom K, Nylin S (2001) Evolutionary dynamics of host–plant specialization: a case study of the tribe Nymphalini. *Evolution*, **55**, 783–796.

Jermy T (1984) Evolution of insect host plant relationships. *American Naturalist*, **124**, 609–630.

Kennedy JS (1983) Zigzagging and casting as a programmed response to wind-borne odor – a review. *Physiological Entomology*, **8**, 109–120.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**, 561–576.

Korman AK, Mallet J, Goodenough JL *et al.* (1993) Population structure in *Heliothis virescens* (Lepidoptera: Noctuidae): an estimate of gene flow. *Annals of the Entomological Society of America*, **86**, 182–188.

Krumm JT, Hunt TE, Skoda SR *et al.* (2008) Genetic variability of the European corn borer, *Ostrinia nubilalis*, suggests gene flow between populations in the Midwestern United States. *Journal of Insect Science*, **8**, 1–12.

Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution*, **30**, 314–334.

Levins R (1970) Extinction. In: *Lecture on Mathematics in the Life Science* (ed Gerstenhaber M). pp. 77–107, AMS, Providence, RI.

Lynch M, Hill WG (1986) Phenotypic evolution by neutral mutation. *Evolution*, **40**, 915–935.

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.

Mallet J, Korman A, Heckel DG, King P (1993) Biochemical genetics of *Heliothis* and *Helicoverpa* (Lepidoptera, Noctuidae) and evidence for a founder event in *Helicoverpa zea*. *Annals of the Entomological Society of America*, **86**, 189–197.

Mayr E (1963) *Animal Species and Evolution*. Belknap Press, Cambridge, MA.

McElvare RR (1941) Validity of the species *Heliothis subflexa* (Gn.) (Lepidoptera). *Bulletin of the Brooklyn Entomological Society*, **36**, 29–30.

McMichael M, Prowell DP (1999) Differences in amplified fragment–length polymorphisms in fall armyworm (Lepidoptera: Noctuidae) host strains. *Annals of the Entomological Society of America*, **92**, 175–181.

Menken SBJ, Herrebout WM, Wiebes JT (1992) Small ermine moths (Yponomeuta): their host relations and evolution. *Annual Review of Entomology*, **37**, 41–66.

Micinski S, Blouin DC, Waltman WF, Cookson C (2008) Abundance of *Helicoverpa zea* and *Heliothis virescens* in pheromone traps during the past twenty years in northwestern Louisiana. *Southwestern Entomologist*, **33**, 139–149.

Mitter C (1993) Biosystematics of the Heliothinae (Lepidoptera: Noctuidae). *Annual Review of Entomology*, **38**, 207–225.

Nosil P (2002) Transition rates between specialization and generalization in phytophagous insects. *Evolution*, **56**, 1701–1706.

Nosil P, Crespi BJ (2004) Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and

sexual isolation in *Timema cristinae* walking-sticks. *Evolution*, **58**, 102–112.

Nosil P, Mooers AO (2005) Testing hypotheses about ecological specialization using phylogenetic trees. *Evolution*, **59**, 2256–2263.

Orr HA (1998) Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics*, **149**, 2099–2104.

Park Y, Taylor MFJ (1997) A novel mutation L1029H in sodium channel gene *hscp* associated with pyrethroid resistance for *Heliothis virescens* (Lepidoptera: Noctuidae). *Insect Biochemistry and Molecular Biology*, **27**, 9–13.

Park Y, Taylor MFJ, Feyereisen R (1997) A valine421 to methionine mutation in IS6 of the *hscp* voltage-gated sodium channel associated with pyrethroid resistance in *Heliothis virescens* F. *Biochemical and Biophysical Research Communications*, **239**, 688–691.

Pashley DP (1986) Host-associated genetic differentiation in fall armyworm (Lepidoptera, Noctuidae) – a sibling species complex. *Annals of the Entomological Society of America*, **79**, 898–904.

Pashley DP (1989) Host-associated differentiation in army worms (Lepidoptera: Noctuidae): an allozymic and mitochondrial DNA perspective. In: *Electrophoretic Studies on Agricultural Pests, Systematics Associatio* (eds Loxdale HD, Den Hollander J). pp. 103–114, Clarendon, Oxford.

Pashley DP, Hammond AM, Hardy TN (1992) Reproductive isolating mechanisms in fall armyworm host strains (Lepidoptera, Noctuidae). *Annals of the Entomological Society of America*, **85**, 400–405.

Pathan AAK, Devi KU, Vogel H, Reineke A (2007) Analysis of differential gene expression in the generalist entomopathogenic fungus *Beauveria bassiana* (Bals.) Vuillemin grown on different insect cuticular extracts and synthetic medium through cDNA–AFLPs. *Fungal Genetics and Biology*, **44**, 1231–1241.

Poole RW, Mitter C, Huettel MD (1993) A revision and cladistic analysis of the *Heliothis virescens* species-group (Lepidoptera: Noctuidae) with a preliminary morphometric analysis of *H. virescens*. Mississippi Agricultural and Forestry Experiment Station Technical Bulletin No. 185, Mississippi Entomological Museum Publication Series No. 4.

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

Pulliam HR (1986) Sources, sinks, and population regulation. *American Naturalist*, **132**, 652–661.

Raesaenen K, Hendry AP (2008) Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, **11**, 624–636.

Raulston JR, Wolf WW, Lingren PD, Sparks AN (1982) Migration as a factor in *Heliothis* management. In: *Proceedings, International Workshop on Heliothis Management*, Patancheru, India, pp. 61–73.

Raulston JR, Pair SD, Pedraza Martinez FA *et al.* (1986) Ecological studies indicating the migration of *Heliothis zea*, *Spodoptera frugiperda*, and *Heliothis virescens* from Northeastern Mexico and Texas. In: *Insect flight: dispersal and migration* (ed Danthanarayana W.). pp. 204–220, Springer-Verlag, Berlin/Heidelberg.

Roehrdanz RL, Lopez JD, Loera J, Hendricks DE (1994) Limited mitochondrial DNA polymorphism in North-American populations of *Heliothis virescens* (Lepidoptera, Noctuidae). *Annals of the Entomological Society of America*, **87**, 856–866.

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Schneider JC (1989) Role of movements in evaluation of area-wide insect pest management tactics. *Environmental Entomology*, **18**, 438–446.

Schneider JC (1999) Dispersal of a highly vagile insect in a heterogeneous environment. *Ecology*, **80**, 2740–2749.

Schneider JC (2003) Overwintering of *Heliothis virescens* (F.) and *Helicoverpa zea* (Boddie) (Lepidoptera: Noctuidae) in cotton fields of northeast Mississippi. *Journal of Economic Entomology*, **96**, 1433–1447.

Schoonhoven LM (2005) Insect–plant relationships: the whole is more than the sum of its parts. *Entomologia Experimentalis Et Applicata*, **115**, 5–6.

Sheck AL, Gould F (1993) The genetic basis of host range in *Heliothis virescens* – larval survival and growth. *Entomologia Experimentalis Et Applicata*, **69**, 157–172.

Sheck AL, Gould F (1995) Genetic analysis of differences in oviposition preferences of *Heliothis virescens* and *Heliothis subflexa* (Lepidoptera, Noctuidae). *Environmental Entomology*, **24**, 341–347.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393–430.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Sluss TP, Graham HM (1979) Allozyme variation in natural populations of *Heliothis virescens*. *Annals of the Entomological Society of America*, **72**, 317–322.

Stireman JO, Nason JD, Heard SB (2005) Host-associated genetic differentiation in phytophagous insects: general phenomenon or isolated exceptions? Evidence from a goldenrod insect community. *Evolution*, **59**, 2573–2587.

Takami Y, Koshio C, Ishii M *et al.* (2004) Genetic diversity and structure of urban populations of *Pieris* butterflies assessed using amplified fragment length polymorphism. *Molecular Ecology*, **13**, 245–258.

Tero N, Aspi J, Siikamaki P, Jakalaniemi A, Tuomi J (2003) Genetic structure and gene flow in a metapopulation of an endangered plant species, *Silene tatarica*. *Molecular Ecology*, **12**, 2073–2085.

Thompson JN (1988) Evolutionary ecology of the relationship between oviposition preference and performance of offspring in phytophagous insects. *Entomologia Experimentalis Et Applicata*, **47**, 3–14.

Vekemans X, Beauwens T, Lemaire M, Roldán-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139–151.

Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology*, **17**, 4334–4345.

Waldvogel M, Gould F (1990) Variation in oviposition preference of *Heliothis virescens*. *Evolution*, **44**, 1326–1337.

Westbrook JK, Wolf WW, Lingren PD, Raulston JR (1994) Tracking tetroons to evaluate tobacco budworm and bollworm migration. *Proceedings Beltwide Cotton Conference*, **2**, 791–793.

Williams MR (2000) Cotton insect pheromone traplines in Mississippi – 1999. *Proceedings Beltwide Cotton Conferences*, **2**, 1252–1254.

Wolf WW, Sparks AN, Pair SD, Westbrook JK, Truesdale FM (1986) Radar observations and collections of insects in the Gulf of Mexico. In: *Insect Flight: Dispersal and Migration* (ed Danthanarayana W). pp. 221–234, Springer Verlag, Berlin.

Yeaman S, Jarvis A (2006) Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proceedings Royal Society B*, **273**, 1578–1593.

Zehnder GW, Sandall L, Tisler AM, Powers TO (1992) Mitochondrial DNA diversity among 17 geographic populations of *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *Annals of the Entomological Society of America*, **85**, 234–240.

Zhang D-X, Yan L-N, Ji Y-J, Hewitt GM, Huang Z-S (2009) Unexpected relationships of substructured populations of Chinese *Locusta migratoria*. *BMC Evolutionary Biology*, **9**, 144–155.

Zhao Y, Park Y, Adams ME (2000) Functional and evolutionary consequences of pyrethroid resistance mutations in S6 transmembrane segments of a voltage-gated sodium channel. *Biochemical and Biophysical Research Communications*, **278**, 516–521.

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

A.T.G. is an evolutionary ecologist interested in chemical communication, A.C. is a student in ecology, O.I. is interested in soil ecology, C.A.B. is a research entomologist interested in host specialization and detoxification mechanisms in moths, J.L.Jr. is a research entomologist interested in insect pest ecology with emphasis on toxicity of insecticides, A.T.V. is a research entomologist interested in insect pest management, C.S. is a chemical ecologist and urban entomologist, D.G.H is interested in the genetic basis of complex life-history traits in insects and insect-plant coevolution, G.S. is interested in population structure of natural populations and their effects on evolutionary processes.

Data accessibility

For AFLP scores and collection site data for all individuals (Supporting Information).

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 AFLP scores of the 252 markers of all 179 Hv individuals used in the analyzes.

Table S2 AFLP scores of the 243 markers of all 179 Hs individuals used in the analyzes.

Table S3 Legends of all abbreviations used in Tables S1 and S2.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.