Genetic structure and gene flow in French populations of two Ostrinia taxa: host races or sibling species?

Thibaut Malausa, Ambroise Dalecky, Sergine Ponsard, Philippe Audiot, Réjane Streiff, Yannick Chaval, Denis D. Bourguet

HAL Id: hal-02662428
https://hal.inrae.fr/hal-02662428

Submitted on 12 Aug 2020

Genetic structure and gene flow in French populations of two *Ostrinia* taxa: host races or sibling species?

T. MALAUSA,*† A. DALECKY,* S. PONSARD,† P. AUDIOT,* R. STREIFF,* Y. CHAVAL* and D. BOURGUET*

*Centre de Biologie et de Gestion des Populations (CBGP), UMR INRA-IRD-CIRAD-SupAgro, Institut National de la Recherche Agronomique, Campus International de Baillarguet, 34 988 Montferrier/Lez, France, †Laboratoire Dynamique de la Biodiversité, Université P. Sabatier — Toulouse III, UMR CNRS 5172, 31 062 Toulouse cedex 09, France

Abstract

Most models of ecological speciation concern phytophagous insects in which speciation is thought to be driven by host shifts and subsequent adaptations of populations. Despite the ever-increasing number of studies, the current evolutionary status of most models remains incompletely resolved, as estimates of gene flow between taxa remain extremely rare. We studied the population genetics of two taxa of the *Ostrinia* genus — one feeding mainly on maize and the other on mugwort and hop — occurring in sympatry throughout France. The actual level of divergence of these taxa was unknown because the genetic structure of populations had been investigated over a limited geographical area and the magnitude of gene flow between populations had not been estimated. We used 11 microsatellite markers to investigate the genetic structure of populations throughout France and the extent of gene flow between the two *Ostrinia* taxa at several sites at which they are sympatric. We observed clear genetic differentiation between most populations collected on the typical respective hosts of each taxon. However, populations displaying intermediate allelic frequencies were found on hop plants in southern France. Individual assignments revealed that this result could be accounted for by the presence of both taxa on the same host. Gene flow, estimated by determining the proportion of hybrids detected, was low: probably < 1% per generation, regardless of site. This indicates that the two *Ostrinia* taxa have reached a high level of genetic divergence and should be considered sibling species rather than host races.

Keywords: Bayesian approaches, ecological speciation, European corn borer, gene flow, host race, hybridization, microsatellites, *Ostrinia nubilalis*, *Ostrinia scapulalis*, sympatric speciation

Introduction

An increasing number of animal taxa have been described as candidate cases of sympatric and/or ecological speciation (Schluter 2001; Via 2001; Berlocher & Feder 2002; Drès & Mallet 2002; Rundle & Nosil 2005) in which genetic divergence is thought to result mainly from natural selection via adaptation. Some of the best-documented models of sympatric speciation concern taxonomic groups such as African cichlid fishes (Kornfield & Smith 2000;

Barluenga *et al.* 2006) and *Gasterosteus* sticklebacks (McKinnon & Rundle 2002; McKinnon *et al.* 2004; Gow *et al.* 2006), but most such models concern phytophagous insects (Berlocher & Feder 2002; Drès & Mallet 2002). Phytophagous insect taxa currently divided into two genetically differentiated groups of populations using different host plants may be undergoing speciation. They may represent different stages of the speciation process, according to the current extent of gene flow between them. They may constitute a study case of ecological speciation, assuming that the process was initiated by a host shift and ultimately results in separation into two different species as a result of divergent natural selection and assortative mating (Berlocher & Feder 2002; Drès & Mallet 2002).

Correspondence: Thibaut Malausa, Fax: 0033 1492 386401, E-mail: thibaut.malausa@sophia.inra.fr

Studies of phytophagous insect host races over the last two decades have highlighted both the mechanisms involved in reproductive isolation (Wood & Keese 1990; Feder *et al.* 1994; Via 1999; Emelianov *et al.* 2001, 2003) and the genetic architecture of key adaptive traits (Sezer & Butlin 1998; Craig *et al.* 2001; Via & Hawthorne 2002; Dambroski *et al.* 2005). However, despite the additional knowledge obtained in these studies, the current evolutionary status of most, if not all models remains unclear, as direct estimates of gene flow between host races — that is estimates not inferred from the level of genetic differentiation — remain extremely rare. Moreover, most of the few direct estimates available were obtained from laboratory experiments or from combinations of results from laboratory and field experiments (see review in Drès & Mallet 2002). Thus, although several phytophagous insect species have been described as consisting of several host-specific taxa, the current level of gene flow between these taxa — and therefore their taxonomic status, as defined by Drès & Mallet (2002) — generally remains an open question. Even in 'text book' examples of current speciation as a result of a host shift, such as *Acyrthosiphon pisum* (Via 1999) and *Rhagoletis pomonella* (Bush 1994), the actual occurrence and frequency of hybridization in natural populations is not known precisely. One of the rare exceptions in this respect is *Zeiraphera diniana*: Emelianov *et al.* (2004) used amplified fragment length polymorphic (AFLP) markers to provide an estimate of current gene flow between the larch and pine host races of this lepidopteran species.

In Western Europe, allozyme frequencies and female sex pheromone data suggest that the European corn borer (ECB) — *Ostrinia nubilalis* Hübner (Lepidoptera: Crambidae) — is divided into at least two taxa, one infesting mainly maize (*Zea mays* L.) and one feeding on hop (*Humulus lupulus* L.) and mugwort (*Artemisia vulgaris* L.) (Bourguet *et al.* 2000a; Martel *et al.* 2003; Thomas *et al.* 2003; Bontemps *et al.* 2004). In France, these two taxa also differ in terms of the composition of the long-range pheromone emitted by females and recognized by males (Thomas *et al.* 2003; Bontemps *et al.* 2004; Pélozuelo *et al.* 2004): populations feeding on maize use a 97:3 blend of Z : E isomers of 11-tetradecenyl acetate (11–14: OAc) — referred to as the 'Z' pheromone blend — whereas populations collected on hop and mugwort use a 1:99–4:96 blend of Z : E isomers of 11–14: OAc — referred to as the 'E' pheromone blend. According to Frolov *et al.* (2007), the two *Ostrinia* taxa probably correspond to two sibling species, the taxon feeding on maize being *O. nubilalis* and the taxon feeding on mugwort and hop being *Ostrinia scapulalis*.

Until recently, descriptions of the genetic structure of populations in Western Europe were based exclusively on samples collected in northern France. Leniaud *et al.* (2006) and Malausa *et al.* (2007) investigated the genetic structure of French populations, by studying allozymes and nuclear gene sequences of individuals collected from maize, mugwort, hop and five other plants. They showed that the maize Z-race also infested various other cultivated and noncultivated plants. They also found that populations collected from hop in southern France could not be unambiguously assigned to the expected *O. scapulalis* genetic cluster, raising doubts about the consistency of the genetic structure of populations, with division into two well-separated taxa over France and Europe. In addition, although the frequency of assortative mating in the field has been estimated at more than 95% in northern France (Malausa *et al.* 2005), the level of gene flow between the two *Ostrinia* taxa has never before been estimated, making it difficult to determine their taxonomic status as defined by Drès & Mallet (2002).

In this study, we investigated the genetic structure of 26 *Ostrinia* populations collected across France, by genotyping individuals at 11 microsatellite loci. We investigated the levels of gene flow between the two *Ostrinia* taxa by collecting larvae from several sites in maize fields and from hop and/or mugwort stands separated by less than 2 km. In each population, we were able to assign individual genotypes to either one of the two *Ostrinia* taxa or to a 'hybrid' category, using the Bayesian approach developed by Anderson & Thompson (2002). This provided estimates of hybridization rates — and thus an upper limit for gene flow — between the two taxa in each geographical area. Moreover, this work made it possible, for the first time, to assign individuals to one of the host races without using the host plant of origin as the only criterion.

Materials and methods

Sampling

Diapausing larvae were collected from plant stems at several sites across France (Fig. 1) during the winters of 2004–2005 and 2005–2006. In 2004–2005, samples were collected on two host plants — maize and mugwort — around Paris (Ile de France, France, 48°46′N, 02°04′E), a region in which the rate of assortative mating within the two taxa has been estimated (Malausa *et al.* 2005). In this region, we collected 197 individuals from maize and 129 from mugwort and analysed them (Table 1). In 2005–2006, sampling was expanded to both northern and southern France and was extended to a third host plant: hop. Larvae were collected from six sites at which host plants of both *Ostrinia* taxa (maize and hop or mugwort) could be found in close sympatry (< 2 km apart) allowing geographical mixing of the two taxa during the period of reproduction. In southern France, where mugwort is not infested by *Ostrinia* (Martel *et al.* 2003), populations were collected from maize and hop only. At each site, we collected and analysed 31–38 individuals from each host plant type (Table 1).
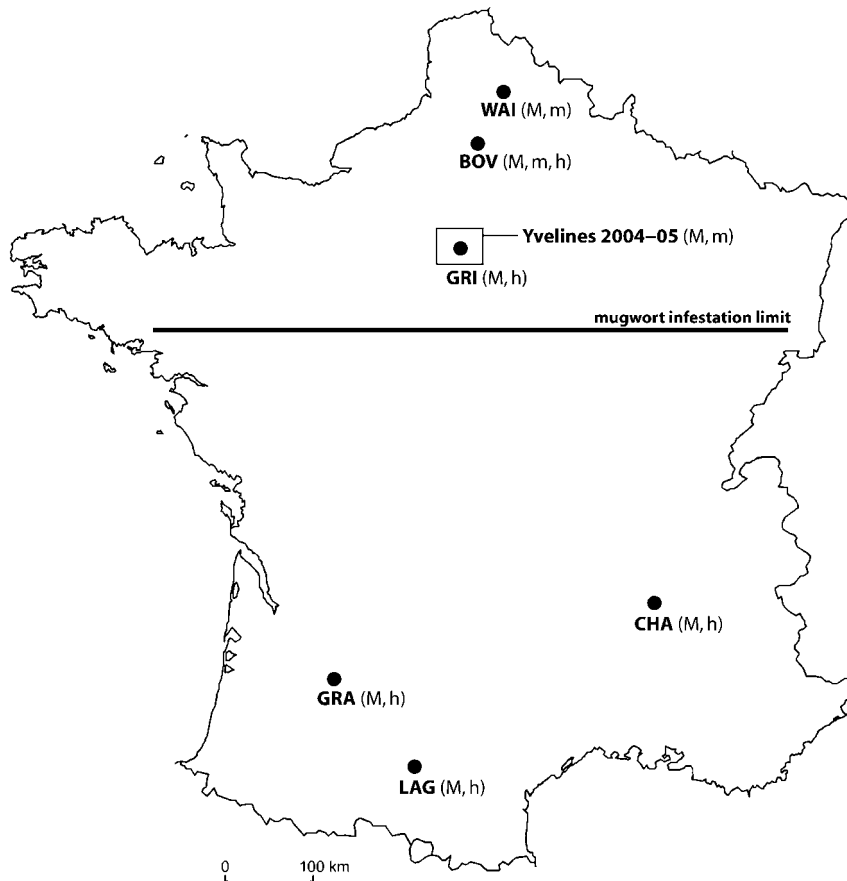
Fig. 1 Location of the sampling sites across France. The 'Yvelines' sample consists of several small populations collected at short distances from each other (see Table 1) and of isolated individuals. 'M' designates *Ostrinia* populations collected from maize, 'm' from mugwort and 'h' from hop. The approximate latitudinal limit of mugwort infestation by *Ostrinia* is indicated by a bold line (mugwort is not infested in southern France).

Genotyping

Genomic DNA was isolated from the head of the larvae, using (i) the PureGene extraction kit (Gentra Systems) for samples collected during the 2004–2005 winter, or (ii) a standard phenol–chloroform procedure (Sambrook *et al.* 1989) for samples collected during the 2005–2006 winter. We followed the protocol described by Dalecky *et al.* (2006) to genotype each individual at 11 microsatellite loci (D9, D25, D27, D29, D63, D65, D77, D145, D243, D282 and T81). For each sample, three multiplex fluorescent polymerase chain reactions (PCR) were performed, using the QIAGEN Multiplex Kit (QIAGEN). PCR products were diluted (1:5–1:20 depending on the concentration of the PCR products) and typing was performed with an ABI PRISM 310 sequencing machine (Applied Biosystems). Alleles were scored using GENESCAN 3.1 and GENOTYPER 2.5 software (Applied Biosystems). The repeatability of our genotyping procedure was checked by amplifying and reading microsatellite profiles for ~100 of the 779 individuals twice, independently, using the same DNA extracts.

Data analysis

Within-population genetic variability. Linkage disequilibrium between pairs of loci and deviation from Hardy–Weinberg equilibrium (HWE) for each locus were tested, using Fisher's exact tests implemented in GENEPOP 3.4 software (Raymond & Rousset 1995). We corrected for multiple testing, using the false discovery rate (FDR) approach (Benjamini & Hochberg 1995). When deviation from HWE was detected, we used the MICRO-CHECKER program (van Oosterhout *et al.* 2004) to determine whether the departure from panmixia was due to the presence of null alleles or to genotyping errors. MICRO-CHECKER displays allele length and frequencies and indicates whether the departure observed is more likely to be due to scoring errors or to null alleles. The presence of null alleles (alleles not amplified by PCR because of nucleotide variations in the regions to which primers were designed to bind) was considered the most probable explanation in all cases (see the 'Results' section), so a maximum-likelihood estimate was calculated for the frequency of null alleles (Dempster *et al.* 1977), using the FREENA package (Chapuis & Estoup 2007). Finally,

Table 1 Characteristics of the *Ostrinia* samples: year of sampling, location, location code on Fig. 1, host plant, number (*N*) of diapausing larvae analysed and mean expected heterozygosity (H_E) over all loci, with the standard deviation indicated in brackets. 'M' designates *Ostrinia* populations collected from maize, 'm' from mugwort and 'h' from hop

Year	Location	Location code	Latitude (°E)	Longitude (°N)	Host-plant	N	H_E
2004–2005	Bourdonné	Yvelines	1.65	48.76	m	5	0.627 (0.326)
					M	23	0.648 (0.197)
	Gambais	Yvelines	1.67	48.78	m	31	0.682 (0.239)
					M	30	0.685 (0.202)
	Gazeran	Yvelines	1.78	48.63	m	11	0.648 (0.231)
	Jouars-Pontchartrain	Yvelines	1.89	48.79	m	22	0.616 (0.260)
					M	31	0.625 (0.216)
	Le Mesnil St-Denis	Yvelines	1.96	48.73	M	21	0.624 (0.243)
	Le Perray en Yvelines	Yvelines	1.83	48.71	M	29	0.653 (0.237)
	Les Loges en Josas	Yvelines	2.13	48.77	m	29	0.605 (0.235)
	Thiverval	Yvelines	1.96	48.84	m	26	0.599 (0.226)
					M	30	0.661 (0.217)
	Villiers le Bâcle	Yvelines	2.14	48.73	M	33	0.644 (0.214)
	isolated individuals	Yvelines	—	—	m	10	—
2005–2006	Boves	BOV	2.38	49.85	m	38	0.668 (0.193)
					h	31	0.659 (0.223)
					M	36	0.615 (0.222)
	Châteauneuf sur Isère	CHA	4.94	45.03	h	33	0.679 (0.259)
					M	35	0.644 (0.239)
	Granges sur Lot	GRA	0.45	44.38	h	33	0.660 (0.177)
					M	33	0.661 (0.227)
	Grignon	GRI	1.92	48.85	h	38	0.623 (0.245)
					M	38	0.667 (0.214)
	Lagardelle sur Lèze	LAG	1.44	43.40	h	33	0.654 (0.185)
					M	35	0.651 (0.229)
	Wailly les Arras	WAI	2.72	50.25	m	36	0.627 (0.237)
					M	34	0.620 (0.218)

we used the mean expected heterozygosity (H_E, Nei 1987) as a measure of genetic variability within populations, because the presence of null alleles has only a limited effect on this statistic (Chapuis & Estoup 2007).

Genetic structure and phylogenetic relationships. We calculated pairwise estimators of F_{ST} (Weir 1996) and applied the ENA correction provided by the FREENA package to correct for bias due to the presence of null alleles at several loci (see the 'Results' section). This corrected F_{ST} estimator is referred to hereafter as $F_{ST}^{\{ENA\}}$. We evaluated the genotypic differentiation between pairs of populations for all loci, using Fisher's exact tests implemented in GENEPOP. We corrected for multiple testing, using the FDR approach. We tested for isolation by distance with Mantel tests implemented in GENEPOP, using a semimatrix containing pairwise $F_{ST}^{\{ENA\}}$ and a semimatrix containing geographical distances converted using $F_{ST}^{\{ENA\}}/(1 - F_{ST}^{\{ENA\}})$ and log(geographical distance), as recommended by Rousset (1997). This was carried out for (i) all populations, (ii) populations from maize, and (iii) populations from hop and mugwort. We further

examined the genetic structure of the populations, using a Bayesian clustering analysis implemented in BAPS 4.13 (Corander *et al*. 2004) to determine clusters of populations by minimizing Hardy–Weinberg and linkage disequilibrium within clusters. This conservative method for the identification of different genetic structures was very robust to the presence of null alleles (M.P. Chapuis, M. Lecoq, A. Loisean, S. Piry, Y. Michalakis & A. Estoup, unpublished data). We therefore used BAPS to define clusters of populations in two situations: (i) situations in which the number of clusters was predefined: two clusters, corresponding to the two *Ostrinia* taxa; (ii) situations in which the number of clusters was not predetermined, but with an upper limit set to 50 (i.e. more than the number of sampled populations), to make it possible to identify more detailed patterns across geographical areas and host plants.

Phylogenetic relationships between all the populations sampled in 2004–2005 and 2005–2006 were examined by calculating a neighbour-joining (NJ) dendrogram based on chord distances D_C (Cavalli-Sforza & Edwards 1967), using the POPULATION program (Olivier Langella, http://

bioinformatics.org/~tryphon/populations). Null alleles have little effect on this distance estimate (Chapuis & Estoup 2007), particularly if data corrected with the FREENA package are used. Dendrogram node stability was assessed by running 10 000 bootstrap replications (over loci). The figure was then drawn with TREEVIEW (Page 1996).

Migration and gene flow. The global data set was divided into two data subsets corresponding to the two sampling years. The Yvelines data subset contained all individuals collected in 2004–2005 on maize and mugwort around Paris, whereas the other data set contained data from all sites at which two or three closely sympatric (< 2 km apart) host populations were collected in 2005–2006 (see Fig. 1): Lagardelle, Granges, Châteauneuf and Grignon all included populations from both maize and hop; Wailly included populations from maize and mugwort; Boves included populations from maize, mugwort and hop. We then used a Bayesian approach to test for the presence of (i) individuals displaying the typical multilocus genotypes of one species but found on host plants of the other species, providing evidence of *Ostrinia* 'migration' from host to host; and (ii) individuals displaying hybrid multilocus genotypes, providing evidence of effective gene flow between the two *Ostrinia* taxa. The Bayesian genetic clustering approach implemented in NEWHYBRIDS 1.1 (Anderson & Thompson 2002) sorts individuals into six genetic classes (two F_0 classes, F_1 and F_2 hybrids, two classes of backcrosses) and calculates, by Markov chain Monte Carlo simulations, the corresponding posterior probabilities of each individual belonging to each class. However, Anderson & Thompson (2002) pointed out that this program requires diagnostic alleles for distinguishing between the specific genotype frequency classes of hybrid individuals (F_1 vs. F_2 and backcrosses). As none of the loci studied here was completely diagnostic for separation of the two *Ostrinia* taxa (see 'Results' section), and as only a limited number of loci were available, it proved impossible to distinguish between the different hybrid classes. Thus, we summed the posterior probability values for all hybrid classes for an individual and used this sum as a single estimate, as recommended by Vähä & Primmer (2006), who tested the efficiency of NEWHYBRIDS for detecting hybrids in simulated data sets. Uniform priors (for allele frequencies and mixing proportion) were used rather than Jeffreys priors. Indeed, using the latter overestimates the posterior probabilities associated to hybrid classes when low-frequency alleles are abundant (Anderson 2003), which proved to be the case in our data set (Table 2, also see Supplementary material). The probabilities used to assign individuals to one of the two taxa or to a hybrid class were obtained from the mean of five runs of 1 million iterations, following 100 000 burn-in iterations.

As null alleles were detected in our data set (see the 'Results' section), and the robustness of the models implemented in NEWHYBRIDS was unknown, we examined the stability of our results by analysing two data sets for each study site: the original data set and a data set corrected for null alleles using the method implemented in the FREENA package.

Results

Within-population genetic variability

Exact tests for genotypic linkage disequilibrium resulted in only one significant value ($P < 10^{-3}$, D9 and T81 loci) for

Table 2 Number of alleles, size range, number of populations of the 26 in which deviations from HWE were detected and mean estimated proportion of null alleles in *Ostrinia* populations (standard deviation over populations is indicated within brackets) collected from maize and hop or mugwort, for each of the 11 microsatellite loci for which they were genotyped

Locus	No. of alleles	Size range of alleles	No. of deviations from HWE	Mean estimated proportion of null alleles	
				Maize	Hop or mugwort
D9	41	252–373	23	0.24 (0.07)	0.22 (0.11)
D25	38	63–161	25	0.25 (0.08)	0.27 (0.08)
D27	12	167–187	20	0.32 (0.07)	0.20 (0.14)
D29	14	216–258	20	0.58 (0.06)	0.55 (0.18)
D63	12	96–118	1	0.02 (0.03)	0.05 (0.06)
D65	14	252–277	7	0.08 (0.06)	0.06 (0.07)
D77	16	111–123	22	0.40 (0.09)	0.36 (0.10)
D145	6	160–186	0	0.04 (0.05)	0.08 (0.09)
D243	8	316–335	1	0.09 (0.10)	0.05 (0.08)
D282	8	230–253	25	0.40 (0.10)*	0.62 (0.09)*
T81	25	104–182	0	0.04 (0.05)	0.03 (0.03)

*The mean estimated proportion of null alleles was significantly ($P < 0.05$) lower in populations collected on maize than in populations collected on hop and mugwort.

only one population (Châteauneuf, collected on hop). Thus, the various microsatellite loci investigated in this study appeared to segregate independently. At six loci, a significant deviation from HWE was detected in more than 19 populations. For the other five loci, less than eight populations showed such deviation (Table 2). In each population and for each locus where such deviation was detected, the presence of null alleles was the most probable explanation for heterozygote deficiency. The mean frequency of null alleles varied between 2% and 62% (Table 2), depending on the locus and the group of populations considered (populations from maize or populations from hop and mugwort). The microsatellite markers were set up using DNA from populations collected on maize, but null alleles were more frequent in the populations collected on mugwort or hop for only one locus (D282, Table 2). Within-population variability, estimated using mean expected heterozygosity, was consistent over populations ($H_E \sim 0.6$), regardless of their host plants of origin (Table 1).

Genetic structure and phylogenetic relationships

Significant genotypic differentiation between populations ($P < 0.05$) was detected in 266 of 325 pairwise comparisons, with $F_{ST}^{\{ENA\}}$ values varying from –0.001 to 0.084 (Table 3). All 169 comparisons involving populations from maize on the one hand and populations from mugwort or hop on the other hand displayed significant differentiation ($0.015 \leq F_{ST}^{\{ENA\}} \leq 0.084$, see Table 3). Most populations collected from mugwort or hop also displayed differentiation from each other: $F_{ST}^{\{ENA\}}$ values ranged between 0.007 and 0.067 and more than 90% of the 78 pairwise comparisons showed significant genotypic differentiation. In contrast, comparisons between populations collected on maize revealed much lower levels of differentiation ($-0.001 \leq F_{ST}^{\{ENA\}} \leq 0.019$). Accordingly, only 33% of the 78 pairwise comparisons between these populations revealed significant genotypic differentiation. A significant genetic pattern of isolation by distance was detected by Mantel tests, in populations from hop and mugwort ($P < 10^{-3}$), but not in populations from maize ($P = 0.370$) or when all populations were considered ($P = 0.563$).

The BAPS clustering analysis confirmed the trends observed with pairwise comparisons. Indeed, when no a priori number of clusters was imposed, the highest posterior probability was obtained for a structure with six separate clusters, one pooling all populations collected on maize, and five consisting of various populations collected from hop and mugwort. Two of these five clusters included geographically separated populations from northern France, whereas each of the other three clusters corresponded to one of the three populations collected on hop in southern France. When constraining the number of clusters to $K = 2$, all but one of the populations from hop and mugwort (collected from hop at Granges) were assigned to the same cluster, whereas all populations collected on maize belonged to the second cluster.

The intermediate allele frequencies of populations collected from hop in southern France were highlighted by the NJ unrooted dendrogram based on chord distances D_C (Fig. 2). This dendrogram was weakly supported, but displayed two clusters — one containing all populations from maize and the other consisting of populations collected in northern France, from mugwort and hop. The populations collected on hop in southern France could not be unambiguously pooled within either of these clusters. Noteworthy, when host plants of both *Ostrinia* taxa (maize and hop or mugwort) were sampled in close sympatry, these populations never clustered together, but rather each of them clustered with populations of their respective host-plant. This is highly suggestive of low gene flow between taxa.

Migration and gene flow

The results obtained with the Bayesian genetic assignment method of NEWHYBRIDS were used to estimate the proportion of F_0 migrants at each site — the proportion of individuals assigned to a given F_0 genetic class but collected on the typical host plant of the other F_0 class. These estimates varied from 0 to 33% and the highest proportions of migrants were observed at the three sites in southern France ($\geq 10\%$), with all estimated proportions $\leq 7\%$ in northern France (Table 4). In Granges (Southern France), the proportion of migrants detected exceeded 30%; indeed, most individuals were assigned to the genetic cluster composed primarily of individuals collected from maize. Very similar results were obtained for different runs.

The proportion of hybrid genotypes estimated by NEWHYBRIDS ranged from 0 to 1%. Correcting the data set for the presence of null alleles prevented NEWHYBRIDS from producing estimates for the 2005–2006 data set, as all individuals were assigned to a single genetic cluster. However, results with and without FREENA correction gave similar trends for the Yvelines data set (Table 4). FREENA assumes that any deviation from HWE is caused by null alleles and that there is only one null allele for each locus in all populations. It then corrects the frequencies of visible allelic states based on these assumptions. This correction decreased, and may have underestimated the differences in allele frequency between the two species at each site, thereby reducing the performance of the Bayesian approach. This loss of power has been shown to result in overestimation of the proportion of hybrid genotypes (Vähä & Primmer 2006). It therefore remains unclear whether FREENA correction improved the reliability of results by correcting the bias induced by null alleles or

Table 3 Pairwise F_{ST}^{ENA} (below the diagonal) and P value of Fisher's exact tests for genotypic differentiation between populations (above the diagonal)

	BOV-h	CHA-h	GRA-h	GRI-h	LAG-h	BOU-m	BOV-m	GAM-m	GAZ-m	JOU-m	LOG-m	THI-m	WAI-m	BOU-M	BOV-M	CHA-M	GAM-M	GRA-M	GRI-M	JOU-M	LEM-M	LEP-M	LAG-M	THI-M	VIL-M	WAI-M
BOV-h	—	***	***	***	***	NS	NS	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
CHA-h	0.022	—	***	***	***	*	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
GRA-h	0.041	0.029	—	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
GRI-h	0.026	0.043	0.049	—	***	NS	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
LAG-h	0.044	0.029	0.039	0.065	—	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
BOU-m	0.010	0.037	0.059	0.018	0.050	—	NS	NS	NS	**	*	*	NS	***	***	***	**	***	***	***	***	***	***	***	***	***
BOV-m	0.007	0.025	0.035	0.027	0.036	0.021	—	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
GAM-m	0.033	0.024	0.042	0.023	0.042	0.009	0.029	—	*	*	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
GAZ-m	0.037	0.041	0.055	0.027	0.063	0.017	0.035	0.017	—	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
JOU-m	0.041	0.042	0.045	0.035	0.053	0.047	0.032	0.008	0.030	—	*	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
LOG-m	0.049	0.060	0.055	0.033	0.060	0.042	0.037	0.012	0.029	0.015	—	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
THI-m	0.051	0.056	0.067	0.032	0.067	0.023	0.048	0.013	0.035	0.021	0.025	—	***	***	***	***	***	***	***	***	***	***	***	***	***	***
WAI-m	0.019	0.036	0.035	0.031	0.045	0.016	0.024	0.032	0.052	0.033	0.043	0.038	—	***	***	***	***	***	***	***	***	***	***	***	***	***
BOU-M	0.071	0.042	0.024	0.068	0.065	0.083	0.062	0.055	0.074	0.070	0.080	0.070	0.062	—	NS	*	NS	NS	NS	NS	*	NS	NS	NS	NS	*
BOV-M	0.082	0.051	0.034	0.075	0.071	0.077	0.065	0.054	0.084	0.057	0.073	0.059	0.058	0.015	—	***	**	**	NS	NS	***	NS	**	NS	**	NS
CHA-M	0.059	0.030	0.016	0.055	0.038	0.069	0.045	0.036	0.064	0.043	0.057	0.051	0.042	0.001	0.012	—	**	NS	**	NS	NS	NS	NS	NS	NS	**
GAM-M	0.042	0.020	0.020	0.054	0.046	0.053	0.039	0.038	0.045	0.044	0.060	0.048	0.039	0.009	0.011	0.009	—	NS	NS	NS	NS	NS	NS	*	*	*
GRA-M	0.054	0.032	0.015	0.049	0.043	0.056	0.042	0.032	0.058	0.035	0.044	0.048	0.039	0.005	0.006	−0.001	0.006	—	NS	NS	NS	NS	NS	NS	NS	***
GRI-M	0.058	0.030	0.028	0.053	0.061	0.047	0.050	0.032	0.060	0.053	0.059	0.046	0.048	0.008	0.004	0.011	0.007	0.005	—	NS	*	NS	NS	NS	NS	NS
JOU-M	0.065	0.034	0.033	0.066	0.057	0.076	0.057	0.042	0.071	0.050	0.073	0.057	0.052	0.006	0.010	0.005	0.001	0.005	0.003	—	NS	NS	NS	NS	NS	NS
LAG-M	0.057	0.031	0.028	0.060	0.041	0.068	0.045	0.041	0.070	0.048	0.068	0.052	0.046	0.011	0.012	0.001	0.006	0.005	0.009	0.003	—	NS	NS	*	***	***
LEM-M	0.044	0.025	0.016	0.054	0.051	0.074	0.037	0.040	0.067	0.046	0.061	0.065	0.048	0.002	0.013	0.004	0.002	0.004	0.009	0.007	0.014	—	NS	**	NS	***
LEP-M	0.054	0.028	0.026	0.063	0.045	0.069	0.045	0.040	0.063	0.048	0.067	0.054	0.052	0.006	0.019	−0.001	0.000	0.003	0.012	0.004	0.002	0.002	—	NS	NS	*
THI-M	0.067	0.033	0.029	0.065	0.054	0.060	0.052	0.035	0.071	0.047	0.060	0.053	0.053	0.014	0.004	0.006	0.007	0.002	0.004	0.006	0.008	0.013	0.007	—	*	*
VIL-M	0.053	0.032	0.030	0.043	0.053	0.056	0.040	0.030	0.066	0.040	0.049	0.044	0.035	0.014	0.015	0.006	0.013	0.006	0.006	0.009	0.009	0.009	0.011	0.010	—	**
WAI-M	0.069	0.045	0.037	0.079	0.064	0.074	0.057	0.054	0.080	0.062	0.082	0.063	0.062	0.014	0.002	0.013	0.006	0.012	0.006	0.012	0.009	0.012	0.010	0.009	0.015	—

P values of exact tests for genotypic differentiation after correction using the FDR approach: NS, not significant; *$0.01 < P < 0.05$; **$0.001 < P < 0.01$; ***$P < 0.001$.

Table 4 Estimates of the proportions of hybrid genotypes and migrants obtained with NEWHYBRIDS. *N*, number of individuals; 'n.d.' corresponds for each population within each site to the number of individuals that could not be assigned to one of the two clusters. The 95% confidence intervals for the proportions of migrants and of hybrid genotypes were calculated using the modified WALD method (Agresti & Coull 1998)

Site	Host plant (*N*)		NEWHYBRIDS results		Population 1 (maize)				Population 2 (mugwort/hop)				Prop.of interm.genot.		Prop. of migrants	
	Pop. no. 1	Pop. no. 2	Ind. in cluster no. 1	Ind. in cluster no. 2	Cluster no. 1	Cluster no. 2	Interm. genot.	n.d.	Cluster no. 1	Cluster no. 2	Interm. genot.	n.d.	Mean	95% CI	Mean	95% CI
Data uncorrected for null alleles																
Northern France																
Yvelines	M (197)	m (129)	184	130	179	10	0	8	5	120	1	3	0.00	0.000–0.019	0.050	0.028–0.075
Wailly	M (34)	m (36)	37	25	33	0	0	1	4	25	0	7	0.00	0.000–0.064	0.057	0.019–0.143
Boves	M (36)	h, m (69)	41	64	35	1	0	0	6	63	0	0	0.00	0.000–0.044	0.067	0.031–0.134
Grignon	M (38)	h (38)	39	33	35	1	0	2	4	32	0	2	0.00	0.000–0.059	0.066	0.026–0.149
Southern France																
Chateauneuf	M (35)	h (33)	42	39	33	1	0	1	9	22	0	2	0.00	0.000–0.066	0.147	0.081–0.253
Granges	M (33)	h (33)	55	7	33	0	0	0	22	7	0	4	0.00	0.000–0.068	0.333	0.232–0.454
Lagardelle	M (35)	h (33)	43	19	30	3	0	2	13	16	1	3	0.01	0.000–0.088	0.235	0.150–0.350
Data corrected for null alleles																
Northern France																
Yvelines	M (197)	m (129)	188	118	183	3	1	10	5	115	0	9	0.00	0.000–0.019	0.025	0.012–0.049

No estimates could be produced with the 2005–2006 data corrected by FREENA.
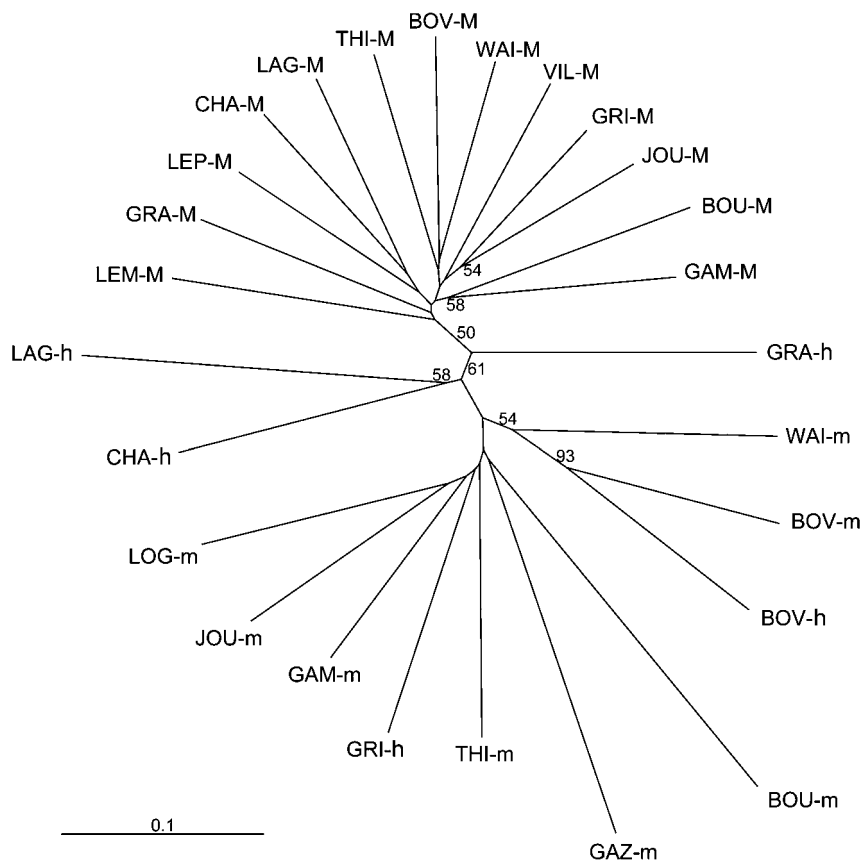
Fig. 2 Unrooted dendrogram based on Cavalli-Sforza chord distances (D_C). Populations are named using the first three letters of the sampling site and the plant on which they were collected, coded as in Fig. 1.

introduced a major bias by considering there to be only one allelic state for null alleles. If we used only the five loci that did not depart from the HW expectations (including one almost monomorphic locus), the signal was too weak and all the individuals were again assigned to same genetic cluster.

Discussion

Of all the markers used to date to investigate the genetic structure of populations of French *Ostrinia* taxa, including allozymes and mitochondrial and nuclear DNA sequences, microsatellites proved to be the most powerful, despite the presence of null alleles in our data set. They clarified certain trends weakly detected with other markers in previous studies that required confirmation. The genetic differentiation between populations collected on maize and mugwort (and hop in northern France) previously detected for allozyme (Bourguet *et al.* 2000b; Martel *et al.* 2003; Thomas *et al.* 2003; Bontemps *et al.* 2004; Malausa *et al.* 2005; Leniaud *et al.* 2006) and nuclear loci (Malausa *et al.* 2007) was also clearly demonstrated in this study (Fig. 2; Table 3). The lack of any pattern of genetic isolation by distance among populations collected on

maize, observed for allozyme loci throughout France (Bourguet *et al.* 2000a; Martel *et al.* 2003; Leniaud *et al.* 2006), and the isolation by distance detected by Martel *et al.* (2003) among populations feeding on mugwort were also confirmed in this study. Finally, we confirmed the trend detected by Malausa *et al.* (2007) for four nuclear sequences, suggesting that populations collected on hop in southern France display intermediate allele frequencies (Fig. 2).

Microsatellites also made it possible, for the first time, to achieve sufficiently high power for the identification of genetic clusters of individuals regardless of the plants on which they fed as larvae. Indeed, although female moths emerging from maize preferred to lay eggs on maize, they were able to lay a small proportion of their eggs on hop and mugwort and vice versa (Bethenod *et al.* 2005; Malausa *et al.* in press), suggesting that individuals of both *Ostrinia* taxa are found, at least occasionally, on both types of plant. The results of Bontemps *et al.* (2004) suggest that this situation probably also occurs in the field. Therefore, groups of individuals collected from any given plant cannot be considered to be a sample of a 'pure' population of a given *Ostrinia* taxon: they correspond instead to a mixture of different taxa (Thomas *et al.* 2003; Pélozuelo

et al. 2004; Malausa *et al.* 2005). The microsatellite-based individual assignment presented here provides the first estimate of the proportion of migrants between the two host plants.

Moreover, any sample may include hybrid genotypes resulting from crosses between the two taxa. The Bayesian individual clustering approach of NEWHYBRIDS made it possible, for the first time, to estimate the proportion of hybrid genotypes. These estimates can be viewed as the upper limit of the extent of gene flow between taxa because (i) the estimated numbers of hybrid genotypes probably include F_1, F_2 and backcrosses, that is hybrids produced over at least two generations; (ii) hybrids may be more numerous among larvae than among adults, particularly if hybrid fitness costs occur at later stages of development; and (iii) the Bayesian approach we used tends to overestimate the proportion of hybrid genotypes (Vähä & Primmer 2006). The effect of the presence of null alleles on these hybrid frequency estimates remains unknown. However, as differences in the frequency of null alleles were observed at only one locus, null alleles probably did not artificially increase the estimated differentiation between populations, probably resulting instead in an overall loss of data and thus of discriminating power. This loss of power probably also leads to an overestimation of hybrid frequency.

Overall, our results suggest that there may be less than 1% gene flow between taxa for each generation, at all sites, regardless of the local proportion of migrants from host to host and the relative proportions of individuals assigned to each taxon: indeed, the Bayesian assignment method used detected only two hybrids among the 778 individuals tested. The estimated proportion of hybrids in larval populations did not differ significantly from the proportion of hybrid matings estimated by Malausa *et al.* (2005) (two-tailed Fisher's exact test, $P = 0.334$) and from the proportion of adult females found by Pélozuelo *et al.* (2004) to emit pheromone blends (two-tailed Fisher's exact test, $P = 0.171$). Indeed, around Paris, Malausa *et al.* (2005) detected three 'hybrid' matings among 361 assessed, and Pélozuelo *et al.* (2004) found that three of 345 females emitted sex pheromone blends intermediate between the pure 'Z' and 'E' pheromone blends, these intermediate blends being emitted by hybrid females (Roelofs *et al.* 1987). All three estimates confirm that the two *Ostrinia* taxa have already reached a high level of reproductive isolation, even in close sympatry. Moreover, the similarity between the estimate of assortative mating reported by Malausa *et al.* (2005) and our estimate of the proportion of hybrids among fifth instar larvae suggests that postzygotic barriers play little role in the reproductive isolation between the two *Ostrinia* taxa. The two *Ostrinia* taxa may thus be considered as two sibling species, as defined by Drès & Mallet (2002): the taxon typically feeding on maize may be considered to be *Ostrinia nubilalis* and that typically feeding on mugwort and hop may be considered to be *Ostrinia scapulalis*, as proposed by Frolov *et al.* (2007). Indeed, current estimates for gene flow between these two taxa are at the lower end of those generally reported for host-race models. Indeed, in *Lochmaea capreae*, Kreslavsky *et al.* (1981) estimated gene flow at between 1% and 3%, based on assortative mating and host choice estimates. In *Zeiraphera diniana*, Emelianov *et al.* (2004) identified three hybrids among the 92 individuals collected in areas in which the two host races were sympatric. Finally, based on a combination of release–recapture experiments, Feder *et al.* (1994) estimated the gene flow between *Rhagoletis pomonella* host races to be ~6% per generation.

Individual assignment based on microsatellite data provided an additional explanation for the previously reported divergence of populations collected on hop in southern France from other populations collected on hop elsewhere in France (Leniaud *et al.* 2006; Malausa *et al.* 2007). Indeed, populations collected on hop displayed the typical allele frequencies of *O. scapulalis* when collected in northern France but genetic profiles intermediate between those of *O. scapulalis* and *O. nubilalis* when collected in southern France (see Fig. 2). This finding may be accounted for by a higher proportion of *O. nubilalis* individuals feeding on hop in southern France than in northern France (Table 4). One of the populations we collected from hop (in Granges, Southern France) even seemed to consist of a majority of *O. nubilalis* individuals. This high proportion of *O. nubilalis* on hop is not really surprising, because *O. nubilalis* females lay significant proportions of eggs on hop in semicontrolled experiments (Malausa *et al.* in press) and because hop patches are surrounded by large amounts of cultivated maize in the areas sampled in southern France. However, this high proportion of migrants on both maize and hop contrasts with the results of Pélozuelo *et al.* (2004), who found no individual, throughout France, on maize using the 'E' pheromone blend (typical of *O. scapulalis*) and no individual on mugwort or hop using the 'Z' pheromone blend (typical of *O. nubilalis*) in a study of 176 individuals on maize and 169 on hop or mugwort. However, differences in sampling may account for the difference between the results of these studies. Indeed, we sampled in sites where maize was very abundant and thus increased the probability of finding *O. nubilalis* individuals on any host plant occasionally chosen by this taxon. The higher level of F_0 migration onto hop reported in our study is therefore not entirely surprising. The simultaneous determination of microsatellite multilocus data and pheromone blend use would have been useful to check whether individuals genetically assigned to a given taxon also use the typical pheromone blend of this taxon. However, the sample size needed for such combined study

would have been much larger. Indeed, phenotype determination is possible only at the adult stage, larval survival is generally low in the laboratory (~20%), and another ~20% individual phenotypes typically remain undetermined. Moreover, our genetic analyses needed to be conducted on sympatric populations collected from different host plants (including hop where population densities are low), so that a high mortality in only one of two (or two of three) populations collected at a given site would have prevented us from using the remaining population as well. Therefore, rearing the larvae we collected until adult stage instead of freezing them immediately for DNA extraction would have implied that we collect 20 times more larvae per site and that we sample more sites than we did to ensure a similar sample size for microsatellite analyses.

An additional explanation for the intermediate genetic profile of southern populations collected on hop may also be a possible introgression of *O. nubilalis* genes into *O. scapulalis*, which might be facilitated by *O. nubilalis* being present in much larger numbers than *O. scapulalis* in southern France. Such introgression has been reported in other model taxa used for studies of sympatric speciation, for example the frequency of hybrids between the benthic and limnetic taxa of *Gasterosteus* sticklebacks has been found to vary locally from ~5 to ~20%, the highest estimates being obtained for lakes in which the proportions of 'pure-type' individuals were the most unbalanced (Gow *et al*. 2006). At a site with ~20% hybrids, Taylor *et al*. (2006) showed that both taxa were likely to collapse into a single hybrid 'swarm' as a result of introgression. However, in *Ostrinia*, the proportions of hybrids were not higher in southern than in northern France. Thus, individuals collected from hop in southern France nonetheless constituted two distinct genetic clusters, matching the genetic profile of the two species at the scale of the whole territory (Table 4). As our sampling scheme did not include populations from central France, future studies should address the issue of whether there is a latitudinal cline of *O. nubilalis* allele frequencies in *O. scapulalis* populations.

Based on Mayr's Biological Species Concept, Frolov *et al*. (2007) hypothesized that the two *Ostrinia* taxa found in France may constitute two different species — *O. nubilalis* and *O. scapulalis*. Our study supports this hypothesis, as it suggests that there is a very low level of gene flow (< 1%) between the two taxa. This low level of gene flow would confer these taxa the status of sibling species, as defined by Drès & Mallet (2002). Our results further show that individuals of the two *Ostrinia* taxa occasionally share a same host plant. As a consequence, using the host plant on which the insect was collected as the sole criterion to define taxa would have lead to erroneous conclusions. This highlights the need for individual assignment based on genetic multilocus criteria especially if closely related taxa, for example host-races, are studied.

References

Agresti A, Coull BA (1998) Approximate is better than 'exact' for interval estimation of binomial proportions. *American Statistician*, **52**, 119–126.

Anderson EC (2003) *User's Guide to the Program* NEWHYBRIDS, Version 1.1. beta. April 7th, 2003. URL: http://ib.berkley.edu/labs/slatkin/eriq/software/software.htm

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217–1229.

Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, **439**, 719–723.

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.

Berlocher S, Feder J (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Review of Entomology*, **47**, 773–815.

Bethenod M-T, Thomas Y, Rousset F *et al*. (2005) Genetic isolation between two sympatric host plant races of the European corn borer, *Ostrinia nubilalis* Hübner. II. Assortative mating and host plant preferences for oviposition. *Heredity*, **94**, 264–270.

Bontemps A, Bourguet D, Pélozuelo L, Bethenod M-T, Ponsard S (2004) Managing the evolution of *Bacillus thuringiensis* resistance in natural populations of the European corn borer, *Ostrinia nubilalis*: host plant, host race and phenotype of adult males at aggregation sites. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 2179–2185.

Bourguet D, Bethenod M-T, Pasteur N, Viard F (2000a) Gene flow in the European corn borer *Ostrinia nubilalis*: implications for the sustainability of transgenic insecticidal maize. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **267**, 117–122.

Bourguet D, Bethenod M-T, Trouvé C, Viard F (2000b) Host-plant diversity of the European corn borer *Ostrinia nubilalis*: what value for sustainable transgenic insecticidal *Bt* maize? *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **267**, 1177–1184.

Bush G (1994) Sympatric speciation in animals: new wine in old bottles. *Trends in Ecology & Evolution*, **9**, 285–288.

Cavalli-Sforza L, Edwards A (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics*, **19**, 233–257.

Chapuis M, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Corander J, Waldmann P, Marttinen P, Sillanpää MJ (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics*, **20**, 2363–2369.

Craig TP, Horner JD, Itami JK (2001) Genetics, experience, and host-plant preference in *Eurosta solidaginis*. Implications for host shifts and speciation. *Evolution*, **55**, 773–782.

Dalecky A, Bogdanowicz SM, Dopman EB, Bourguet D, Harrison RG (2006) Two multiplex sets of eight and five microsatellite markers for the European corn borer, *Ostrinia nubilalis* Hubner (Lepidoptera: Crambidae). *Molecular Ecology Notes*, **6**, 945–947.

Dambroski HR, Linn C, Berlocher SH *et al*. (2005) The genetic basis for fruit odor discrimination in *Rhagoletis* flies and its significance for sympatric host shifts. *Evolution*, **59**, 1953–1964.

Dempster A, Laird N, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B, Methodological*, **39**, 1–38.

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.

Emelianov I, Drès M, Baltensweiler W, Mallet J (2001) Host-induced assortative mating in host races of the larch budmoth. *Evolution*, **55**, 2002–10.

Emelianov I, Simpson F, Narang P, Mallet J (2003) Host choice promotes reproductive isolation between host races of the larch budmoth *Zeiraphera diniana*. *Journal of Evolutionary Biology*, **16**, 208–218.

Emelianov I, Marec F, Mallet J (2004) Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 97–105.

Feder J, Opp S, Wlazlo B *et al*. (1994) Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Sciences, USA*, **91**, 7990–7994.

Frolov A, Bourguet D, Ponsard S (2007) Reconsidering the taxonomy of several *Ostrinia* species in the light of reproductive isolation: a tale for Ernst Mayr. *Biological Journal of the Linnean Society*, **91**, 49–72.

Gow JL, Peichel CL, Taylor EB (2006) Contrasting hybridization rates between sympatric three-spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. *Molecular Ecology*, **15**, 739–752.

Kornfield I, Smith PF (2000) African cichlid fishes: model systems for evolutionary biology. *Annual Review of Ecology and Systematics*, **31**, 163–196.

Kreslavsky A, Mikheev A, Solomatin V, Gritzenko V (1981) Genetic exchange and isolating mechanisms in sympatric races of *Lochmaea capreae* (Coleoptera, Scolytidae). *Zoologicheskii Zhurnal*, **72**, 62–68.

Leniaud L, Audiot P, Bourguet D *et al*. (2006) Genetic structure of European and Mediterranean maize borer populations on several wild and cultivated host plants. *Entomologia Experimentalis et Applicata*, **120**, 51–62.

Malausa T, Bethenod M-T, Bontemps A *et al*. (2005) Assortative mating in sympatric host races of the European corn borer. *Science*, **308**, 258–260.

Malausa T, Pélissié B, Piveteau V *et al*. (in press) Differences in oviposition behaviour of two sympatric sibling species of the *Ostrinia* genus. Bulletin of Entomological Research.

Malausa T, Leniaud L, Martin JF *et al*. (2007) Molecular Differentiation at several Nuclear Loci in French Host Races of the European Corn Borer (*Ostrinia nubilalis*). *Genetics*, **176**, 2343–2355.

Martel C, Réjasse A, Rousset F, Bethenod M-T, Bourguet D (2003) Host-plant-associated genetic differentiation in northern French populations of the European corn borer. *Heredity*, **90**, 141–149.

McKinnon JS, Rundle HD (2002) Speciation in nature: the threespine stickleback model systems. *Trends in Ecology & Evolution*, **17**, 480–488.

McKinnon JS, Mori S, Blackman BK *et al*. (2004) Evidence for ecology's role in speciation. *Nature*, **429**, 294–298.

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

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Page R (1996) TREEVIEW: an application to display phylogenetic tress on personal computers. *Computer Applications in the Biosciences*, **12**, 357–358.

Pélozuelo L, Malosse C, Genestier G, Guenego H, Frérot B (2004) Host-plant specialization in pheromone strains of the European corn borer *Ostrinia nubilalis* in France. *Journal of Chemical Ecology*, **30**, 335–352.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Roelofs W, Glover T, Tang X-H *et al*. (1987) Sex pheromone production and perception in European corn borer moths is determined by both autosomal and sex-linked genes. *Proceedings of the National Academy of Sciences, USA*, **84**, 7585–7589.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rundle H, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336–352.

Sambrook J, Fritsch E, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Schluter D (2001) Ecology and the origin of species. *Trends in Ecology & Evolution*, **16**, 372–380.

Sezer M, Butlin RK (1998) The genetic basis of oviposition preference differences between sympatric host races of the brown planthopper (*Nilaparvata lugens*). *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **265**, 2399–2405.

Taylor EB, Boughman JW, Groenenboom M *et al*. (2006) Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (Gasterosteus aculeatus) species pair. *Molecular Ecology*, **15**, 343–355.

Thomas Y, Bethenod M-T, Pélozuelo L, Frérot B, Bourguet D (2003) Genetics isolation between two sympatric host-plant races of the European corn borer, *Ostrinia nubilalis* Hübner. I. Sex pheromone, moth emergence timing, and parasitism. *Evolution*, **57**, 261–273.

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63–72.

Via S (1999) Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, **53**, 1446–1457.

Via S (2001) Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution*, **16**, 381–390.

Via S, Hawthorne DJ (2002) The genetic architecture of ecological specialization: correlated gene effects on host use and habitat choice in pea aphids. *American Naturalist*, **159**, S76–S88.

Weir BS (1996) *Genetic Data Analysis II*. Sinauer & Associates, Sunderland, Massachusetts.

Wood TK, Keese MC (1990) Host-plant induced assortative mating in *Enchenopa* treehoppers. *Evolution*, **44**, 619–628.

Supplementary material

The following supplementary material is available for this article:

Table S1. Allele frequencies for each locus and each population.

Table S2. Distribution of assignment probabilities of individuals to each genetic class, for all sites sampled. Probabilities were computed using the Newhybrids program (see methods). The results for sites other than Yvelines with corrected data are not displayed here because the program failed to separate individuals into two genetic clusters.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/
10.1111/j.1365-294X.2007.03457.x
(This link will take you to the article abstract).