npg

ORIGINAL ARTICLE

From Russia with lobe: genetic differentiation in trilobed uncus *Ostrinia* spp. follows food plant, not hairy legs

AN Frolov[1], P Audiot[2], D Bourguet[2], AG Kononchuk[1], JM Malysh[1], S Ponsard[3], R Streiff[2] and YS Tokarev[4]

[1]*Laboratory for Phytosanitary Diagnostics and Forecasts, All-Russian Institute for Plant Protection, St Petersburg-Pushkin, Russia;* [2]*Centre de Biologie pour la Gestion des Populations (CBGP) UMR INRA-IRD-CIRAD-Montpellier SupAgro, Campus International de Baillarguet, Montferrier-sur-Lez Cedex, France;* [3]*CNRS—UMR 5174, Laboratoire Evolution et Diversité Biologique, Université de Toulouse, Toulouse Cedex, France and* [4]*Laboratory for Microbiological Control, All-Russian Institute for Plant Protection, St Petersburg-Pushkin, Russia*

Trilobed uncus taxa of the genus *Ostrinia* (Lepidoptera, Crambidae) illustrate the complex relationship, at early stages of speciation, between reproductive isolation and differentiation in morphology, resource use and genetic variation. On the basis of behaviour and ecology, we recently hypothesized that individuals with small mid-tibiae belong to two distinct species depending on host plant—*O. nubilalis* and *O. scapulalis sensu* Frolov *et al.* (2007) feeding on maize and on a number of dicotyledons, respectively. Individuals with small, medium or massive mid-tibiae would all belong to *O. scapulalis* as long as they feed on these dicotyledons. This contrasts with previous taxonomy, which distinguished three species by male mid-tibia morphology, regardless of host plant. Here, we test our hypothesis by examining the genetic structure of *Ostrinia* populations from regions with mid-tibia polymorphism—Western Russia and Kazakhstan—and comparing it with that of French populations where only small mid-tibiae occur. Results support two predictions: (1) maize- and dicotyledon-collected populations are genetically differentiated from each other like in France, and (2) dicotyledon-collected populations show no genetic evidence of consisting of more than one species. Between-species differentiation was unrelated to geographic distance, despite significant isolation by distance within species. The distinction between two and only two species differing by host plant thus holds at continental scale. Interestingly, one microsatellite locus contributed ~10 times more than the others to differentiation between both taxa. This deserves further investigation, as it might reveal a linkage between this outlier and loci involved in host-plant adaptation and/or reproductive isolation.

Heredity (2012) **108,** 147–156; doi:10.1038/hdy.2011.58; published online 20 July 2011

Introduction

Morphology is a widely used but not always reliable criterion to delineate species. The growing use of molecular tools has revealed that many species actually consist of several cryptic species that cannot be distinguished by morphology (for example, Pfenninger and Schwenk, 2007; Trontelj and Fiser, 2009). The reverse—several 'morphological species' being merged into a single one—is less frequent. One reason may be that the 'burden of proof' is reversed: merging species by demonstrating that traits considered as diagnostic actually have no bearing on reproductive isolation is not easy, as the latter is difficult to assess in the field and easily underestimated in laboratory settings. Another obstacle may be the implicit feeling that, because most diagnostic morphological traits appear complex, they

must be polygenic, so that significant discontinuities in morphological variation result from a long history of separate evolution, and thus reflect meaningful taxonomic boundaries.

However, very clear phenotypic differences are not necessarily determined by a large number of loci. A well-known example is single mutations on development genes causing spectacular morphological differences in *Drosophila*—and as a matter of fact they are not necessarily associated with changes in mating propensity (Lindsey and Zimm, 1992). Similarly, striking, sympatric variations in colour patterns of the Harlequin ladybird *Harmonia axyridis* (Komai, 1956), the peppered moth *Biston betularia* (Kettlewell, 1955, 1956), the oldfield mouse *Peromyscus polionotus* (Hoekstra *et al.*, 2006; Steiner *et al.*, 2007) or the African Mocker swallowtail *Papilio dardanus* (Ford, 1936; Clark *et al.*, 2008) are all controlled by single or very few genes and apparently unrelated to reproductive isolation. However, it also happens that one or a small number of loci control morphological traits directly involved in reproductive isolation (for example, flower colours attracting different pollinators: Schemske and Bradshaw, 1999), and are thus

well suited to define species, following Mayr's (1942) Biological Species Concept.

The genus *Ostrinia* (Lepidoptera, Crambidae) includes another illustration of the sometimes complex relationship between morphology and taxonomy. This genus is subdivided into three groups according to the number (1, 2 or 3) of lobes of the 'uncus', a part of the male genitalia (Mutuura and Monroe, 1970). In the 'trilobed uncus' group, females are indistinguishable, but males can have small and simple, medium-sized and invaginated or massive and hairy mid-tibiae. This conspicuous trait was used by Mutuura and Monroe (1970) as a diagnostic criterion to define three species: *O. nubilalis* Hübner (small mid-tibia), *O. narynensis* Mutuura and Monroe (syn. nov.) (medium mid-tibia) and *O. scapulalis* (Walker) (massive mid-tibia). *O. nubilalis*, the European corn borer, is a maize (*Zea mays* L.) pest reported to also feed on a large number of dicotyledon plants at larval stage (reviewed in Ponsard *et al.*, 2004), whereas *O. narynensis* and *O. scapulalis* are documented on dicotyledon plants alone (Mutuura and Monroe, 1970).

A decade after Mutuura and Monroe's (1970) taxonomic review, Frolov (1981, 1984) showed that, in spite of its apparent complexity, male mid-tibia morphology is inherited in simple Mendelian proportions involving only two loci: one autosomal locus, Mt, for which the dominant allele, Mt, causes males carrying at least one copy to have massive mid-tibiae, and one heterosomal locus, i, of which the recessive allele i causes ii males to have medium mid-tibiae, whereas ii^+ and i^+i^+ males have small mid-tibiae. Small and medium mid-tibiae are only found in Mt^+Mt^+ males, otherwise they have massive mid-tibiae, regardless of genotype at locus i.

In addition, Mutuura and Monroe's (1970) classification of 'trilobed uncus' *Ostrinia* turned out to conflict with Mayr's (1942) Biological Species Concept. Indeed, in the Former Soviet Union (FSU), adults of this group mate freely in laboratory conditions, regardless of male mid-tibia morphology, as long as they all originate from populations collected on dicotyledons (Frolov *et al.*, 2007). Conversely, mating success is significantly reduced between individuals collected on maize versus dicotyledons (Frolov *et al.*, 2007). In France, although all males have small mid-tibiae, trilobed uncus *Ostrinia* populations actually consist of two taxa that are significantly isolated reproductively (Bethenod *et al.*, 2005; Malausa *et al.*, 2005; Pélozuelo *et al.*, 2007) and genetically differentiated (Bourguet *et al.*, 2000; Martel *et al.*, 2003; Bontemps *et al.*, 2004; Malausa *et al.*, 2005, 2007a, 2007b; Leniaud *et al.*, 2006): one feeding mainly on maize and one feeding mainly on two dicotyledons—mugwort (*Artemisia vulgaris* L.) and hop (*Humulus lupulus* L.).

In search for a unified taxonomic framework consistent with the Biological Species Concept, Frolov *et al.* (2007) suggested a revision of the taxonomic boundaries of the 'trilobed uncus' group arguing that host plant, rather than male mid-tibia morphology, is the best currently available criterion to distinguish species within this group. *O. scapulalis sensu* Frolov *et al.* (2007) would consist of individuals with small, medium and massive mid-tibiae feeding on dicotyledons, whereas individuals feeding on maize—where only small mid-tibiae occur—would belong to *O. nubilalis sensu* Frolov *et al.* (2007) (unless stated otherwise, we will hereafter use the names

O. nubilalis and *O. scapulalis* in this sense). In other words, Frolov *et al.* (2007) hypothesized that male mid-tibia morphology is polymorphic within *O. scapulalis*, with allele frequencies varying across its geographic range. Alleles i^+ and Mt^+ would be fixed in some—but not all—parts of *O. scapulalis'* range (for example, in France), and fixed in *O. nubilalis* across its entire range.

The present work aimed at testing Frolov *et al*'s. (2007) hypothesis by examining the genetic structure of *Ostrinia* sp. populations collected on maize versus dicotyledons in areas of the FSU where male mid-tibiae are polymorphic on dicotyledons, and by comparing it with that of a similar sample collected in France (Malausa *et al.*, 2007b). If Frolov *et al*'s. (2007) hypothesis is correct, at least the two following predictions should hold: (1) as previously documented in France, Russian populations collected on maize and on dicotyledons should be genetically differentiated from each other, and (2) populations collected on dicotyledons in regions of the FSU where male mid-tibiae are polymorphic should show no genetic substructure. In addition, (3) while genetic differentiation between populations feeding on a given host-plant type may increase with geographic distance, the level of differentiation between populations feeding on maize and those feeding on dicotyledons is not, *a priori*, expected to depend on geographic distance (see Feder (1998) or condition 3b in Drès and Mallet (2002)). Finding a different pattern would not necessarily invalidate Frolov *et al*'s. (2007) hypothesis, as more complex scenarios may account for such difference, but finding (3) would strongly support its being valid across North-Western Eurasia.

Materials and methods

Sampling
Samples of *Ostrinia* populations were collected as fifth instars in Western Russia during the summer 2008 (Figure 1 and Table 1). In each locality, we collected up to three sympatric populations on different host plants, within typically <2, and at most 7 km. Populations consisted of 14–42 individuals feeding on maize (*Z. mays* L.), mugwort (*A. vulgaris* L.) and/or hemp (*Cannabis sativa* L.), depending on availability. In all localities, previous collections on dicotyledons in the 1970s and 1980s had yielded populations that were polymorphic for male mid-tibiae (Table 1 and maps Figures 2 and 3 in Frolov *et al.* (2007)). Sampling populations on different hosts at close range was intended to minimize the possibility that an effect of host plant on genetic differentiation might be confounded with that of other factors. We also collected one population on hemp in Kazakhstan (Sorbulak, ~40 km north-west from Almaty, October 2008). In the FSU, we sampled larvae on hemp rather than hop—the second dicotyledon from which *Ostrinia* larvae had been collected in France (Malausa *et al.*, 2007b, see below)—because it was more difficult to locate infested hop, whereas wild hemp was common and often infested by *Ostrinia*. Moreover, hemp was reported to be a common host plant for *O. narynensis* Mutuura and Monroe (syn. nov.) and *O. scapulalis* (Walker, 1859) (Frolov *et al.*, 2007). We found no *Ostrinia* larvae on maize within ~80 km from Almaty (although we did find several maize fields and searched them for
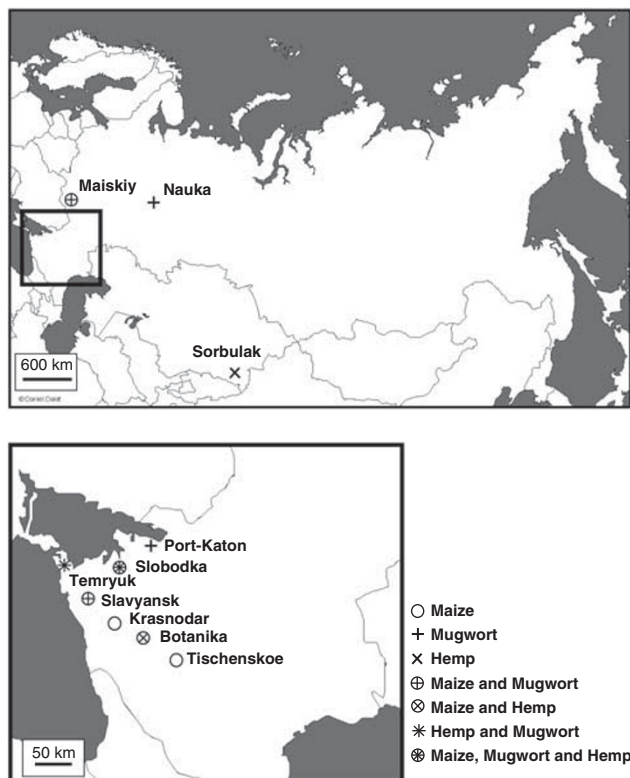
Figure 1 Map of the sampling sites in Russia and Kazakhstan.

Ostrinia larvae), a result consistent with previous observations (Aleksandrova, 1973).

In Sorbulak, we collected > 400 individuals on hemp. We set aside 29 larvae for genetic analysis and reared the others to adult stage, obtaining 32 males with small and 23 males with medium mid-tibiae, that is, an estimate of $\sqrt{(23/55)} = 0.647$ for the frequency of allele *i*. This was very close to our estimate $\sqrt{(10/22)} = 0.674$—based on the proportion of recessive male '*O. narynensis*' with medium mid-tibiae over the total number of males in the collection of the Institute of Zoology (Kazakh Academy of Sciences, Almaty, Kazakhstan), labelled as collected in Kazakhstan in the 1930s—and to Frolov's *et al*. (2007) estimate for the 1970s (0.691, based on $N = 67$ individuals collected on dicotyledons near Almaty). This suggests that, in this location, allele frequencies for allele *i* remained stable over several tens of years.

Unfortunately, sampling limitations in Russia prevented us from estimating allele frequencies for *i* and *Mt* in 2008, our year of collection, as we did not type mid-tibia morphology before genotyping. Indeed, when only a few tens of larvae were collected, we did not attempt to raise them to adult stage for fear of losing too many. Thus, allele frequencies at *Mt* and/or *i* in our samples may not be exactly those given in Table 1, for example, due to sampling stochasticity or fluctuations in allele frequencies. However, it is most unlikely that our samples were not polymorphic at all for male mid-tibiae, as they originate from the area between Kharkov (Ukraine), Kazan, Stavropol and Rostov-on-the-Don, where populations were highly polymorphic at both loci in the 1970–1980s (Table 1, and see Figures 2 and 3 in Frolov *et al*. (2007)). Moreover, since then, small,

medium and large mid-tibia males were regularly found in this area during various fieldwork sessions (Frolov and co-workers, pers. obs).

For comparison, we used a similar sample set collected in France in 2006—which consisted of six pairs or triplets of sympatric populations previously analyzed in Malausa *et al*. (2007b)—and one pair of populations collected on maize and hop, respectively, at Pont-du-Château during the winter 2008–2009 (Table 1). The latter population pair fills a gap that existed between a more 'northern' and a more 'southern' group of populations in Malausa *et al*. (2007b). Each of the six pairs or triplets in Malausa *et al*. (2007b) included one population from maize, and one or two populations from hop and/or mugwort.

Finally, three *Ostrinia* populations were collected on maize in the vicinity of Chengdu, Urumqi and Wuhan (China) in Spring 2009. They were used as an external group to root the tree of genetic distances (see below). Given their geographic origin and host plant, they are most likely to be populations of the Asian corn borer, *O. furnacalis* Guénée (Ishikawa *et al*., 1999), a common maize pest in Eastern Asia, which belongs to the trilobed uncus *Ostrinia* spp.

Genotyping
Population samples collected in the FSU and in China were analyzed using a set of eight autosomal microsatellite markers chosen among the 11 loci used in Malausa *et al*. (2007b) and described by Dalecky *et al*. (2006), the seven loci published by Kim *et al*. (2008), as well as one unpublished locus (*Ga54*) developed by Brad S Coates (United States Department of Agriculture) (Supplementary Table S1). This new set of microsatellites was defined in an attempt to exclude loci with the highest proportions of putative null alleles, as the latter can introduce biases in population parameter estimations (Dakin and Avise, 2004; Chapuis and Estoup, 2007). Populations collected in France in 2005–2006, initially analyzed with Malausa *et al*'s. (2007b) loci, were thus genotyped also for the 'missing' loci, so that the data set analyzed here has all French and FSU populations typed at the same eight loci. Consistency of allele scoring was carefully checked between successive analyses by including previously scored samples into new runs.

Genetic analyses
For each population, allele numbers and F_{IS} estimates (Weir, 1996) were calculated using GENEPOP 3.4 (Raymond and Rousset, 1995). Tests for departure from genotypic frequencies expected at Hardy–Weinberg equilibrium (HWE) within any given population were conducted for each locus using Fisher's exact test as implemented in GENEPOP. When a significant deviation from HWE was detected at a particular locus, we used the program micro-checker (van Oosterhout *et al*., 2004) to determine whether the deviation was likely to be due to the presence of null alleles. The conclusion was always that the presence of null alleles could explain the deviation (see Results). Hence, we calculated a maximum likelihood estimate for the frequency of null alleles under this assumption in the different populations and for the different microsatellite loci (Dempster *et al*, 1977) using the FREENA package (Chapuis and Estoup, 2007).
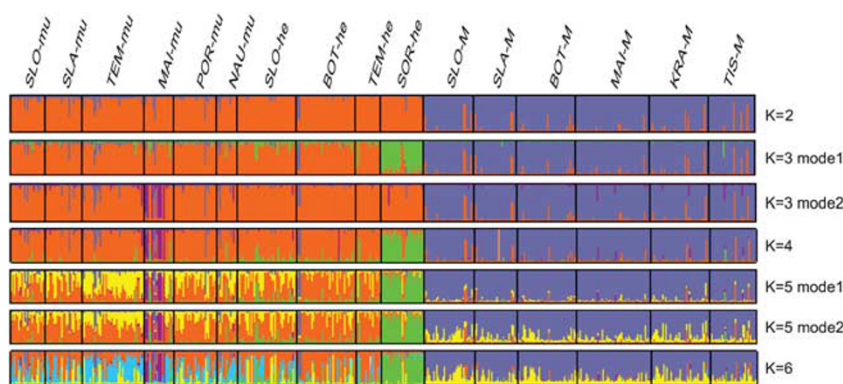
Table 1 *Ostrinia* populations collected in Russia and Kazakhstan

Country/ Republic region	Nearest town	Location	Geographical coordinates	Code	Host plant	N	Frequency of the Mt allele[a]		Frequency of the i allele[a]	
							On maize	On other host plants	On maize	On other host plants
Russia										
Krasnodar	Slavyansk	Slobodka	38°05′E, 45°54′N	SLO	M	34				
					mu	23	0	0–0.005	0	0–0.267
					he	40				
		Slavyansk	38°07′E, 45°15′N	SLA	M	29	0	0–0.005	0	0–0.267
					mu	25				
	Gulkevichi	Botanika	40°42′E, 45°20′N	BOT	M	40	0	0–0.005	0	0.249
					he	40				
	Temryuk	Temryuk	37°23′E, 45°16′N	TEM	mu	42	0	0	0	0–0.267
					he	17				
	Krasnodar	Krasnodar	38°58′E, 45°02′N	KRA	M	40	0	0.005	0	0.267
Belgorod	Belgorod	Maiskiy	36°38′E, 50°33′N	MAI	M	50	0	⩾0.512	0	—
					mu	20				
Rostov	Azov	Port-Katon	38°46′E, 46°52′N	POR	mu	29	0	0.103–0.339	0	0.711–0.822
Stavropol	Izobilnoe	Tischenskoe	41°40′E, 45°27′N	TIS	M	31	0	0	0	0
Tatarstan	Kazan	Nauka	49°07′E, 55°46′N	NAU	mu	14	0	⩾0.632	0	—
Kazakhstan	Almaty	Sorbulak	76°34′E, 43°40′N	SOR	he	29	—	0	—	0.691
France	Boves	Boves	02°38′E, 49°85′N	BOV	mu	38	0	0	0	0
					ho	31	0	0	0	0
					M	36	0	0	0	0
	Châteauneuf sur Isère	Châteauneuf sur Isère	04°94′E, 45°03′N	CHA	ho	33	0	0	0	0
					M	35	0	0	0	0
	Granges sur Lot	Granges sur Lot	00°45′E, 44°38′N	GRA	ho	33	0	0	0	0
					M	33	0	0	0	0
	Grignon	Grignon	01°92′E, 48°85′N	GRI	ho	38	0	0	0	0
					M	38	0	0	0	0
	Lagardelle sur Lèze	Lagardelle sur Lèze	01°44′E, 43°40′N	LAG	ho	33	0	0	0	0
					M	35	0	0	0	0
	Wailly lès Arras	Wailly lès Arras	02°72′E, 50°25′N	WAI	mu	36	0	0	0	0
					M	34	0	0	0	0
	Pont du Château	Pont du Château	03°15′E, 45°48′N	PON	ho	42	0	0	0	0
					M	35	0	0	0	0

Abbreviations: he, hemp, ho, hop; M, maize, mu, mugwort; *N*, number of individuals.
Information on French populations—except PON—re-analyzed here are from Malausa *et al.* (2007b, Table 1).
[a]For Russian and Kazakh populations, these frequencies are those estimated in the 1970s and 1980s, from appendices 1 and 2 in Frolov *et al.* (2007). In French populations, no medium or large midtibiae have ever been reported (Frolov *et al.*, 2007).
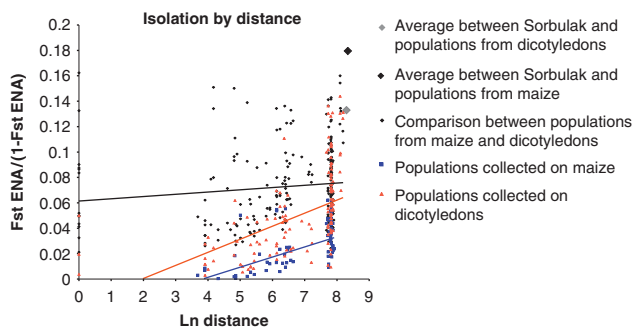

Figure 2 Bayesian clustering analysis of populations collected in Russia and Kazakhstan on maize and dicotyledons, for $K = 2$–6. The two modes for $K = 3$ and $K = 5$ correspond to two equally likely solutions found by the algorithm.

F_{ST} estimates (Weir, 1996) were calculated applying the ENA correction to correct for the bias due to the presence of null alleles. Hereafter, this corrected F_{ST} estimator is referred to as $F_{ST}^{(ENA)}$. As the influence of null alleles on estimates of other parameters is unknown, the ENA correction was only applied to $F_{ST}^{(ENA)}$. All other analyses (including the analysis of molecular variances (AMOVAs), see below) were performed on the original data set.

Prediction no. 1
$F_{ST}^{(ENA)}$ estimates across all loci were calculated between all pairs of populations collected in the FSU and between all pairs of populations collected in France. The genotypic differentiation between pairs of populations for all loci was tested using Fisher's exact test as implemented in GENEPOP. *P*-values were corrected for multiple testing using a Bonferroni correction.

Figure 3 Relation between genetic differentiation and geographic distance within and between *Ostrinia* populations collected on maize or on dicotyledons across France and Western Russia (see text).

Bayesian clustering analyses were conducted among those same populations using the software STRUCTURE (Pritchard *et al.*, 2000). Version 2.3.1 (Hubisz *et al.*, 2009) offers an extension of the previous STRUCTURE model, making it possible to use available information about 'sampling locations' (*sensu lato*, that is, geographic location or host plant). STRUCTURE 2.3.1 can thus detect structure at lower levels of divergence, or with less data, and is not biased towards erroneously detecting structure when it is not present (Hubisz *et al.*, 2009). When analyzing structured data sets for increasing values of K, the most divergent groups are expected to separate into distinct clusters first (Pritchard *et al.*, 2000). We varied the number of putative clusters, K, between 1 and 6, and performed 20 independent runs for each value of K. Each Markov chain was run for 200 000 steps, after a 50 000-step burn-in period, using the admixture model for correlations of allele frequencies across clusters, with the default value for parameter α and the default prior for F_{ST}. Following Evanno *et al.* (2005), we computed the statistic ΔK, which is based on the rate of change in the log posterior probability of the data between successive values of K, to detect the uppermost hierarchical level of structure. Finally, for each value of K we used the 'greedy' algorithm implemented in the CLUMPP software (Jakobsson and Rosenberg, 2007) to detect possible distinct modes among the results of the 20 independent runs. For $K = 3$ and 5, two modes were observed, while for other values of K, all replicates produced a similar structure.

For French populations, both pairwise $F_{ST}^{[ENA]}$ and Bayesian clustering analyses have been published previously (Malausa *et al.*, 2007b), but we updated them to take into account the additional populations and the change in the set of microsatellite loci used.

As 'host plant' appeared to be the major clustering factor according to Bayesian clustering analyses (see Results), we tested for this effect using an AMOVA conducted on the Russian populations data set on one hand, and on the updated data set for French populations on the other hand, both across all loci and across all loci except *OnT2* (see Results). The AMOVAs were performed using the ARLEQUIN 3.11 software (Excoffier *et al.*, 2005).

Prediction no. 2
We computed the pairwise $F_{ST}^{[ENA]}$ estimator between males with small and medium mid-tibiae within the population collected on hemp in Kazakhstan.

We also tested for evidence of a Wahlund effect (consistent deficit in heterozygotes across loci suggesting the presence of more than one taxon) within populations collected on dicotyledons in Russia, and compared the result with that of the same test applied to populations collected on maize in Russia and to populations collected on both host-plant types in France. Although females cannot be recognized by mid-tibia morphology, they can be included when calculating F_{ST} values based on microsatellites, which increases statistical power to detect such possible Wahlund effect.

Prediction no. 3
Isolation-by-distance (IBD) patterns were investigated by assessing the independence of geographic (ln[geographical distance]) and genetic distances ($F_{ST}^{[ENA]}/[1-F_{ST}^{[ENA]}]$) within and between populations collected on each host-plant type—maize and dicotyledons—as described by Rousset (1997). The null hypothesis that genetic and geographic distances were independent of each other was tested against an alternative hypothesis of a positive correlation, as expected in case of significant IBD. To this effect, we used a new GENEPOP option (Rousset, pers. com.) to be included in the next public release of this software. This option calculates bootstrapped estimates for the slope and associated 95% confidence intervals not only, as currently implemented, for population pairs within a group—in our case, a host-plant type—but also between groups—in our case, between two different types.

In addition to the analyses performed to test the three predictions, we constructed a population phenogram based on Saitou and Nei's (1987) neighbour-joining algorithm, using Cavalli-Sforza's genetic distance. This genetic distance has been shown to be less sensitive to the possible presence of null alleles than others (Chapuis and Estoup, 2007). This analysis included all populations collected in the FSU, as well as the updated French data set. The tree was rooted with the three Chinese populations collected on maize (most likely *O. furnacalis*). Calculations and resampling were performed with POPULATIONS 1.2.30 software (http://bioinformatics. org/~tryphon/populations/) and visualized with TREEVIEW 3.2 (Page, 1996).

Results

Allele numbers, F_{IS} values, tests for departure from genotypic frequencies expected at HWE and estimated frequency of null alleles for each microsatellite locus are given in Supplementary Table S1. Detailed allele frequencies in each population are available from the corresponding author upon request.

Prediction no. 1
The first and very obvious division into two clusters among the *Ostrinia* larvae collected in Russia was between populations collected on maize on one hand and on hemp and mugwort on the other hand (Figure 2; Evanno *et al.* (2005), ΔK statistic being maximum for $K = 2$). When increasing the number of clusters to up to $K = 6$, populations collected on maize were consistently recognized as belonging to a single cluster, regardless of geographic origin, and this cluster did not include any populations or individuals collected on

Table 2 Pairwise $F_{ST}^{(ENA)}$ (below the diagonal) and *P*-values—with Bonferroni correction—of Fisher's exact tests for genotypic differentiation between populations (above diagonal)

	SLO-he	BOT-he	TEM-he	SOR-he	SLO-mu	SLA-mu	TEM-mu	MAI-mu	POR-mu	NAU-mu	SLO-M	SLA-M	BOT-M	KRA-M	MAI-M	TIS-M
SLO-he	—	*	*	**	NS	NS	NS	**	NS	***	**	**	**	**	**	**
BOT-he	0.020	—	**	**	NS	**	**	**	***	**	**	**	**	**	**	**
TEM-he	0.016	0.043	—	**	NS	**	**	**	NS	**	**	**	**	**	**	**
SOR-he	0.070	0.080	0.105	—	**	**	**	**	**	**	**	**	**	**	**	**
SLO-mu	0.004	0.014	0.027	0.098	—	NS	n.s	*	NS	*	**	**	**	**	**	**
SLA-mu	0.003	0.009	0.032	0.081	−0.003	—	NS	**	***	***	**	**	**	**	**	**
TEM-mu	0.011	0.028	0.048	0.077	0.012	0.008	—	**	*	*	**	**	**	**	**	**
MAI-mu	0.047	0.062	0.100	0.092	0.043	0.048	0.032	—	**	*	**	**	**	**	**	**
POR-mu	0.006	0.008	0.026	0.049	0.010	0.008	0.010	0.046	—	NS	**	**	**	**	**	**
NAU-mu	0.013	0.042	0.045	0.049	0.038	0.028	0.018	0.044	0.022	—	**	**	**	**	**	**
SLO-M	0.032	0.036	0.058	0.129	0.021	0.026	0.027	0.070	0.029	0.071	—	***	NS	NS	***	NS
SLA-M	0.044	0.056	0.073	0.126	0.033	0.040	0.041	0.069	0.041	0.076	0.008	—	NS	NS	NS	NS
BOT-M	0.045	0.048	0.081	0.142	0.033	0.035	0.042	0.081	0.047	0.078	0.004	0.005	—	NS	***	NS
KRA-M	0.037	0.047	0.067	0.129	0.030	0.033	0.042	0.085	0.037	0.083	0.000	−0.001	0.003	—	NS	NS
MAI-M	0.050	0.059	0.074	0.138	0.038	0.042	0.045	0.077	0.049	0.088	0.012	0.009	0.023	0.013	—	NS
TIS-M	0.041	0.041	0.059	0.122	0.026	0.029	0.036	0.073	0.031	0.076	0.001	0.003	0.007	−0.001	−0.001	—

Abbreviations: he, hemp; M, maize; mu, mugwort; NS, not significant.
*$0.01 < P < 0.05$.
**$0.001 < P < 0.01$.
***$P < 0.001$.

dicotyledons (Figure 2). For $K = 3$, two equally likely clustering patterns (two modes) were identified: the global dichotomy between maize and dicotyledons remained, and the third group could consist either of nearly all individuals sampled in Sorbulak on hemp or of seven individuals sampled in Maiskiy on mugwort. For $K > 3$, these two groups remained and no further group emerged. The overall dichotomy between samples collected on maize on one side and on dicotyledons on the other side was similar to the pattern already documented in France (Malausa *et al*., 2007b), despite using a partly different set of microsatellite loci (Supplementary Figure S1). It also held when the French and Russian data sets were analyzed together (Supplementary Figure S2).

$F_{ST}^{(ENA)}$ values were consistent with this clustering analysis. Indeed, even in sympatry (< 10 km), Russian populations collected from maize were substantially differentiated from populations collected on hemp and mugwort in Russia ($F_{ST}^{(ENA)}$ values ranged from 0.021 to 0.088, averaging 0.051; Table 2). This level of differentiation was similar to—although slightly lower than—the differentiation documented in France between population pairs collected on maize and on dicotyledons ($F_{ST}^{(ENA)}$ values ranged from 0.034 to 0.140, averaging 0.078; Supplementary Table S2). Differentiation between pairs of populations collected on two different dicotyledons was low both in Russia between hemp and mugwort (0.003–0.100, averaging 0.029; Table 2) and in France between hop and mugwort (0.014–0.064, averaging 0.029; Supplementary Table S2).

The AMOVA confirmed that the influence of host plant on genetic differentiation was significant both in Russian ($P < 10^{-4}$) and in French ($P < 10^{-4}$) populations (Table 3); this result did not change when the population from Kazakhstan was included in the analysis, details not shown). Locus-by-locus AMOVAs revealed a much stronger host-plant effect at one particular locus—*OnT2*—compared with other loci (details not shown). Indeed, $F_{ST}^{(ENA)}$ values between populations collected on

maize and dicotyledons were one order of magnitude higher at *OnT2* (mean = 0.228 and 0.317 in the FSU and in France, respectively; Table 4) than at any other locus (mean = 0.018–0.038 and 0.012–0.057 in the FSU and in France, respectively; Table 4).

When the AMOVA was conducted across all loci except *OnT2*, the host-plant effect remained significant both in Russia (mean overall $F_{ST} = 0.023$, $P = 0.022$) and in France (mean overall $F_{ST} = 0.032$, $P < 10^{-4}$), even though, logically, this remaining host-plant effect was one order of magnitude weaker (Table 3). STRUCTURE Bayesian clustering conducted excluding *OnT2* still detected the structure by host plant in France, but no longer did so in Russia (results not shown).

Finally, in the FSU, the number of *OnT2* alleles in populations collected from dicotyledons was twice as high as that found in maize-collected populations (Supplementary Table S1). This trend was also found in France, although to a lesser extent (Supplementary Table S1). Note that the decrease in the number of *OnT2* alleles in populations collected in maize versus dicotyledons is spectacular when compared with the other seven loci (Supplementary Table S1).

Prediction no. 2
There was no evidence of genetic differentiation between males with small and medium mid-tibiae within the population collected in Kazakhstan ($F_{ST}^{(ENA)} = -0.0002$, *P*-value for genotypic differentiation = 0.468). Across all populations collected from dicotyledons in the FSU, we detected no consistent deficit in heterozygotes across microsatellite loci: for some loci (for example, *OnT2* and *OnT3*) F_{IS} values were high, whereas for others (for example, *D145* and *D243*) they were consistent with HWE expectations in all populations (details not shown). In addition, for most loci and across all loci, F_{IS} values were of similar magnitude (Supplementary Table S1) within groups where all individuals belong to only one species according to Mutuura and Monroe's (1970)

Table 3 Results of the AMOVA for French and Russian populations, with and without locus *OnT2*

Source of variation	d.f.	Sum of squares	Variance components	% of variation	P-value
All loci					
Russia					
Among host-plant groups	1	60.55	0.12	4.62	<0.0001
Among populations within host-plant groups	13	67.83	0.05	1.83	<0.0001
Within populations	933	2194.28	2.35	93.54	<0.0001
Total	947	2322.66	2.51		
France					
Among host-plant groups	1	89.19	0.16	6.60	<0.0001
Among populations within host-plant groups	13	93.60	0.07	3.04	<0.0001
Within populations	1035	2223.90	2.15	90.36	<0.0001
Total	1049	2406.69	2.38		
All loci except OnT2					
Russia					
Among host-plant groups	1	7.27	0.01	0.33	<0.0001
Among populations within host-plant groups	13	50.13	0.03	1.45	<0.0001
Within populations	933	1875.55	2.01	98.22	<0.0001
Total	947	1932.95	2.05		
France					
Among host-plant groups	1	20.42	0.03	1.51	<0.0001
Among populations within host-plant groups	13	59.56	0.04	1.88	<0.0001
Within populations	1035	2010.79	1.94	96.62	<0.0001
Total	1049	2090.77	2.01		

Abbreviation: AMOVA, analysis of molecular variances.

Table 4 $F_{ST}^{\{ENA\}}$ values between *Ostrinia* populations collected on maize versus dicotyledons in the FSU (60 comparisons) and in France (56 comparisons)

Locus	Russia–Kazakhstan			France		
	Min	Max	Mean	Min	Max	Mean
Ga54	−0.021	0.339	0.033	−0.012	0.226	0.047
OnT2	0.098	0.408	0.228	0.143	0.568	0.317
OnT3	−0.017	0.135	0.020	−0.012	0.120	0.018
D145	−0.017	0.291	0.038	−0.013	0.120	0.027
D243	−0.013	0.110	0.020	−0.013	0.115	0.034
D63	−0.012	0.082	0.020	−0.004	0.097	0.026
D65	−0.010	0.090	0.018	−0.009	0.081	0.012
T81	−0.012	0.127	0.024	−0.002	0.151	0.057

Abbreviation: FSU, Former Soviet Union.

classification (populations collected from maize both in France and in Russia, and on dicotyledons in France) and where they belong to up to three different species according to this classification (populations collected from dicotyledons in the FSU).

Accordingly, in all populations and for all loci displaying a significant departure from HWE, the program MICRO-CHECKER identified the presence of null alleles as a likely cause of the observed heterozygote deficit (details not shown). We estimated the frequency of null alleles at each locus within each population using MICRO-CHECKER: on average across populations, they ranged from 0 to 23% depending on locus (Supplementary Table S1).

Prediction no. 3
Across France and Russia, we found a positive and highly significant correlation between genetic differentiation and geographic distance (IBD) when considering population pairs collected either both on maize

(slope = 0.0079, 95% confidence interval (0.0018, 0.0306), that is, not including zero) or both on dicotyledons (slope = 0.0103, 95% confidence interval (0.0038, 0.0299)) (Figure 3). In contrast, the genetic differentiation between population pairs collected on two different host-plant types was independent of geographic distance (slope = 0.0018, 95% confidence interval (−0.0001, 0.0062), including zero) and remained higher than within host-plant type at similar distance over the entire range considered in this study—that is, ca. 3000 km (Figure 3). These calculations were conservatively made without figures for the population collected on hemp in Kazakhstan (for which average genetic differentiation and geographic distance with maize- and with dicotyledon-collected populations is simply plotted in Figure 3), but results are essentially the same when including it. We also checked whether the IBD pattern held even when excluding locus *OnT2* from the analysis: it did (data not shown).

The population phenogram (Supplementary Figure S3) illustrates the clear dichotomy between populations sampled on maize on one hand and on dicotyledons on the other hand, as well as the geographic structure—probably resulting at least in part from IBD—between the Russian and French populations among those collected from a particular host-plant type.

Discussion

The host-plant-based distinction between *O. scapulalis* and *O. nubilalis* probably holds across at least North-Western Eurasia
We found host plant to have a major effect on genetic structure among Russian populations. Indeed, these populations were most likely subdivided into two genetic clusters corresponding very clearly to one feeding on maize and one feeding on dicotyledons (Figure 2).

In addition, for any given geographic distance across a range of ca. 3000 km (Figure 3), genetic differentiation between population pairs both sampled on the same host-plant type was lower than between population pairs sampled each on a different host-plant type (see Drès and Mallet (2002) and Feder (1998, p 139–140)).

Hence, the distinction between *O. nubilalis* feeding on maize on one hand and *O. scapulalis* feeding on several dicotyledons—mugwort, hop and hemp—on the other hand probably holds not only locally in France, but also over a larger geographic scale including a significant part—maybe even most—of the range where both species coexist (Frolov *et al.*, 2007). The average differentiation between *O. nubilalis* and *O. scapulalis* is relatively low, but significant and remarkably stable over this range. This supports our earlier suggestion (Frolov *et al.*, 2007; Malausa *et al.*, 2007b) to consider these two taxa as sibling species *sensu* Drès and Mallet (2002).

O. scapulalis, a single species with polymorphic male mid-tibiae

All populations collected in the FSU on dicotyledons were genetically very similar to each other. There was no evidence of differentiation between populations collected on hemp versus mugwort and we found no evidence of a consistent Wahlund effect across loci—that is, no evidence of a significant deficit in heterozygotes resulting from the presence of two or more species. Regardless of geographic distance, divergence among populations collected on dicotyledons remained lower than between them and populations collected on maize (Figure 3).

Despite this very low level of genetic differentiation among populations collected on dicotyledons, two equally likely candidate subclusters were identified within this group (Figure 2). One includes most individuals collected at Sorbulak. Yet, this population contained small- and medium-tibia males in approximately equal proportions. These tibia-based groups showed no genetic differentiation from each other, and no significant departure from HWE was detected in the Sorbulak population as a whole. Therefore, the slightly higher differentiation of the Sorbulak population is most likely due to intra-specific IBD (especially as it is separated from the other populations by the Ural mountains) rather than differences in male mid-tibia morphology. In sum, there is no evidence in our data for differentiation, within dicotyledon-feeding individuals, between small- and medium-tibia individuals (*O. nubilalis* and *O. narynensis sensu* Mutuura and Monroe 1970, respectively). This fully supports our earlier suggestion—based on ecological and behavioural arguments—to include them both into a single species feeding on dicotyledons, regardless of tibia morphology (*O. scapulalis*).

The second potential subcluster among the populations collected on dicotyledons in Russia includes seven individuals within the populations collected at Maiskyi (those individuals appear in pink colour in Figure 2 ($K = 3$, mode 2)). Even if those individuals all had massive mid-tibia, their genetic differentiation from small- and medium-tibia individuals feeding on dicotyledons would be of equal magnitude as that found at intra-specific level between the Sorbulak population and the rest of the dicotyledon-collected populations.

Only a genetic analysis of individuals raised to adult stage (and thus formally documented to have massive male mid-tibiae) could definitively settle this point, but in our view, either our overall sample did not include any massive tibia individuals—which is most unlikely given previous records in our sampling location (see Materials and methods)—or it did, and they showed no evidence of genetic differentiation from other individuals feeding on dicotyledons. This, again, supports our earlier suggestion to also merge massive mid-tibia individuals into *O. scapulalis*.

Hence, as hypothesized by Frolov *et al.* (2007), the differences in male mid-tibiae that led Mutuura and Monroe (1970) to split trilobed uncus *Ostrinia* feeding on dicotyledons into three distinct species most likely correspond to intra-specific polymorphism. Amusingly, this is not the first time morphology has been misleading for taxonomy within the trilobed uncus *Ostrinia* spp.: back in the eighteenth century, even males and females of what is now called *O. nubilalis* were initially described as two different species, *Pyrausta nubilalis* and *P. silacealis*, respectively (reviewed in Beck, 1987).

OnT2, a marker possibly linked to a locus under selection

Both in France and in the FSU, the contribution of locus *OnT2* to the global genetic differentiation between populations collected from maize and dicotyledons was an order of magnitude larger than that of any of the other loci examined here or in previous studies (Malausa *et al.*, 2007b). At this locus, the number of different alleles in populations collected from dicotyledons was nearly twice that in populations collected from maize, sample sizes being comparable. This pattern is typical for selection, suggesting that *OnT2* could be located in the vicinity of one or several loci involved in adaptation to host plant and/or in reproductive isolation, although data for a larger number of microsatellite loci would be needed to ensure that F_{ST} values at *OnT2* are significantly above the range usually found at other loci.

Various studies aimed at explaining the genetic differentiation between *O. nubilalis* and *O. scapulalis* or among different races or subspecies within *O. nubilalis* have produced two sets of candidate genes/factors that might be—or be closely linked to—genes/factors responsible for reproductive isolation. The first set is sex-linked, as often in lepidopteran incipient species (Prowell, 1998). It includes genes affecting post-diapause development time (*Pdd*: Glover *et al.*, 1992; Dopman *et al.*, 2005, 2010), male behavioural response to different female sex pheromones (*Resp*: Glover *et al.*, 1990; Dopman *et al.*, 2004, 2005, 2010), differential success rate in larval development on either host plant (Calcagno *et al.*, 2007) and/or other genes located in the vicinity of *Tpi* (Bourguet *et al.*, 2000; Martel *et al.*, 2003; Bontemps *et al.*, 2004; Dopman *et al.*, 2004, 2005). The second set of candidates is autosomal and includes genes responsible for differences in female sex pheromones (*Pher*: Roelofs *et al.*, 1987; Dopman *et al.*, 2004, 2010; Lassance *et al.*, 2010), an unidentified close-range mechanism ensuring assortative mating in laboratory settings (*AssMat*: Pélozuelo *et al.*, 2007) and/or other genes located in the vicinity of *Mpi* (Bourguet *et al.*, 2000; Martel *et al.*, 2003; Bontemps *et al.*, 2004).

It is worth noting that the two loci causing male mid-tibiae to differ from the 'small and simple' morphology

fixed in *O. nubilalis* show one heterosomal (*i*) and one autosomal (*Mt*) location. In both cases, crossing experiments have shown that they are not, as such, responsible for assortative mating (Frolov et al., 2007). Nevertheless, they should be at linkage disequilibrium with loci involved in host-plant adaptation and/or reproductive isolation, as they never introgress into populations feeding on maize despite incomplete reproductive isolation (discussed in Frolov et al., 2007). This points to at least two different (and maybe more) genomic regions being involved in reproductive isolation between maize- and dicotyledon-feeding *Ostrinia* sp. Although *i* belongs to the sex-linked region, the present data set shows that *OnT2* is not sex-linked. Indeed, in Lepidoptera, the heterogametic sex is the female, and females heterozygous at locus *OnT2* were common in our data set. Hence, it would be interesting to examine whether *OnT2* is at linkage disequilibrium with *Mt*, *Mpi*, *Pher* or *AssMat*—that is, whether it belongs to the (or one of the) autosomal region(s) suspected to contribute to reproductive isolation between trilobed uncus *Ostrinia* spp.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

References

Aleksandrova AE (1973). The stem borer *Ostrinia nubilalis* Hb. in the irrigated zone of the Almaty district. *Proc Kazakh Inst Agric Sci* **11**: 114–128.

Beck SD (1987). Developmental and seasonal biology of *Ostrinia nubilalis*. *Agric Zool Rev* **2**: 59–96.

Bethenod M-T, Thomas Y, Rousset F, Frérot B, Pélozuelo L, Genestier G 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 *Bt* resistance in natural populations of the European corn borer, *Ostrinia nubilalis*: host plants, host race and pherotype of adult males at aggregation sites. *Proc R Soc Lond Ser B* **271**: 2179–2185.

Bourguet D, Bethenod M-T, Trouvé C, Viard F (2000). Host-plant diversity of the European corn borer *Ostrinia nubilalis*: what value for sustainable transgenic insecticidal *Bt* maize? *Proc R Soc Lond Ser B* **267**: 1177–1184.

Calcagno V, Thomas Y, Bourguet D (2007). Sympatric host races of the European corn borer: adaptation to host plants and hybrid performance. *J Evol Biol* **20**: 1720–1729.

Chapuis M-P, Estoup A (2007). Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* **24**: 621–631.

Clark R, Brown SM, Collins SC, Jiggins CD, Heckel DG, Alfried P (2008). Colour pattern specification in the Mocker swallowtail *Papilio dardanus*: the transcription factor invected is a candidate for the mimicry locus H. *Proc R Soc London Ser B* **275**: 1181–1188.

Dakin EE, Avise JC (2004). Microsatellite null alleles in parentage analysis. *Heredity* **9**: 504–509.

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* Hübner (Lepidoptera: Crambidae). *Mol Ecol Notes* **6**: 945–947.

Dempster AP, Laird NM, Rubin DB (1977). Maximum Likelihood Estimation from incomplete data via the EM algorithm (with discussion). *J R Stat Soc Series B* **39**: 1–38.

Dopman EB, Bogdanowicz SM, Harrison RG (2004). Genetic mapping of sexual isolation between E and Z pheromone strains of the European corn borer (*Ostrinia nubilalis*). *Genetics* **167**: 301–309.

Dopman EB, Perez L, Bogdanowicz SM, Harrison RG (2005). Consequences of reproductive barriers for genealogical discordance in the European corn borer. *Proc Natl Acad Sci USA* **102**: 14706–14711.

Dopman EB, Robbins PS, Seaman A (2010). Components of reproductive isolation between North American pheromone strains of the European corn borer. *Evolution* **64**: 881–902.

Drès M, Mallet J (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Philos Trans R Soc Lond Ser B* **357**: 471–492.

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

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

Feder LJ (1998). The Apple Maggot Fly, *Rhagoletis pomonella*. In: Howard DJ, Berlocher SH (eds). *Endless Forms: Species and Speciation*. Oxford University Press: Oxford, United Kingdom. pp 309–319.

Ford EB (1936). The genetics of *Papilio dardanus* Brown (Lep). *Trans R Entomol Soc Lond* **85**: 435–466.

Frolov AN (1981). Genetic analysis of 'large' tibia—the taxonomic character of brush-leg borer *Ostrinia scapulalis* Wlk. (Lepidoptera, Pyraustidae). *Genetika* **17**: 2160–2166. (in Russian).

Frolov AN (1984). Biotaxonomic analysis of harmful species of the genus *Ostrinia* Hbn. *Ethologiya Nasekomykh, Trudy Vsesoyuznogo Entomologitsheskogo Obshchestva* **VI**: 4–100. (in Russian).

Frolov AN, Bourguet D, Ponsard S (2007). Reconsidering the taxonomy of several *Ostrinia* species in the light of reproductive isolation: a tale for E. Mayr *Biol J Linn Soc* **91**: 49–72.

Glover TJ, Campbell M, Robbins PS, Roelofs WL (1990). Sex-linked control of sex pheromone behavioral responses in European corn borer moths (*Ostrinia nubilalis*) confirmed with TPI marker gene. *Arch Insect Biochem Physiol* **15**: 67–77.

Glover TJ, Robbins PS, Eckenrode CJ, Roelofs WL (1992). Genetic control of voltinism characteristics in European corn borer races assessed with a marker gene. *Arch Insect Biochem Physiol* **20**: 107–117.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. *Mol Ecol Res* **9**: 1322–1332.

156

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP (2006). A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**: 101–104.

Ishikawa Y, Takanashi T, Kim C-G, Hoshizaki S, Tatsuki S, Huang Y-P (1999). *Ostrinia* spp. in Japan: their host plants and sex pheromones. *Entomol Exp Appl* **91**: 237–244.

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.

Kettlewell H (1955). Selection experiments on industrial melanism in the Lepidoptera. *Heredity* **9**: 323–342.

Kettlewell H (1956). Further selection experiments on industrial melanism in the Lepidoptera. *Heredity* **10**: 287–301.

Kim KS, Coates BS, Hellmich RL, Sumerford DV, Sappington TW (2008). Isolation and characterization of microsatellite loci from the European corn borer, *Ostrinia nubilalis* (Hübner) (Insecta: Lepidoptera: Crambidae). *Mol Ecol Res* **8**: 409–411.

Komai T (1956). Genetics of ladybeetles. *Adv Genet* **8**: 155–15. 188.

Lassance J-M, Groot AT, Liénard MA, Antony B, Borgwardt C, Anderson F *et al.* (2010). Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature* **466**: 486–491.

Leniaud L, Audiot P, Bourguet D, Frérot B, Genestier G, Lee SF *et al.* (2006). Genetic structure of European and Mediterranean maize borer populations on wild and cultivated host plants. *Entomol Exp Appl* **20**: 51–62.

Lindsey DL, Zimm G (1992). *The Genome of Drosophila melanogaster*. Academic Press Inc: San Diego, CA.

Malausa T, Bethenod M-T, Bontemps A, Bourguet D, Cornuet J-M, Ponsard S (2005). Assortative mating in sympatric host races of the European corn borer. *Science* **308**: 258–260.

Malausa T, Leniaud L, Martin J-F, Audiot P, Bourguet D, Ponsard S *et al.* (2007a). Molecular differentiation at nuclear loci in French host races of the European corn borer (*Ostrinia nubilalis*). *Genetics* **176**: 2343–2355.

Malausa T, Dalecky A, Ponsard S, Audiot P, Streiff R, Chaval Y *et al.* (2007b). Genetic structure and gene flow in French populations of two *Ostrinia* taxa: host races or sibling species? *Mol Ecol* **16**: 4210–4222.

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.

Mayr E (1942). *Systematics and the Origin of Species*. Columbia University Press: New York, NY.

Mutuura A, Monroe E (1970). Taxonomy and distribution of the European corn borer and allied species: genus *Ostrinia* (Lepidoptera: Pyralidae). *Mem Entomol Soc Can* **71**: 1–112.

Page RDM (1996). TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**: 357–358.

Pélozuelo L, Meusnier S, Audiot P, Bourguet D, Ponsard S (2007). Sex pheromones are for meeting not for mating. *PLoS One* **2**: e555.

Pfenninger M, Schwenk K (2007). Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol Biol* **7**: 121.

Ponsard S, Bethenod M-T, Bontemps A, Pélozuelo L, Souqual M-C, Bourguet D (2004). Carbon stable isotopes: a tool for studying the mating, oviposition, and spatial distribution of races of the European corn borer (*Ostrinia nubilalis* Hübner, Lepidotera: Crambidae) among host plants in the field. *Can J Zool* **82**: 1177–1185.

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

Prowell DP (1998). Sex linkage and speciation in Lepidoptera. In: Howard DJ, Berlocher SH (eds). *Endless Forms: Species and Speciation*. Oxford University Press : Oxford, UK. pp 309–319.

Raymond M, Rousset F (1995). GENEPOP (Version 1.2): population genetics software for exact tests and eucumenism. *J Hered* **86**: 248–249.

Roelofs WL, Glover TJ, Tang X-H, Sreng I, Robbins P, Eckenrode C *et al.* (1987). Sex pheromone production and perception in European corn borer moths is determined by both autosomal and sex-linked genes. *Proc Natl Acad Sci 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.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Schemske DW, Bradshaw Jr HD (1999). Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). *Proc Natl Acad Sci USA* **96**: 11910–11915.

Steiner CC, Weber JN, Hoekstra HE (2007). Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol* **5**: e219.

Trontelj P, Fiser C (2009). Cryptic species diversity should not be trivialised. *System Biodiv* **7**: 1–3.

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

Weir BS (1996). *Genetic Data Analysis II*. Sinauer Associates: Sunderland, MA.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)