THE ROYAL SOCIETY

Received 16 October 2002
Accepted 14 April 2003
Published online 3 July 2003

Host-based divergence in populations of the pea aphid: insights from nuclear markers and the prevalence of facultative symbionts

J.-C. Simon[1*], S. Carré[2], M. Boutin[1], N. Prunier-Leterme[1], B. Sabater-Muñoz[3], A. Latorre[3] and R. Bournoville[2]

[1]*Unité Mixte de Recherche 'Biologie des Organismes et des Populations appliquée à la Protection des Plantes', Institut National de la Recherche Agronomique B.P. 29, 35653 Le Rheu, Cedex, France*
[2]*Unité de recherche de Zoologie, Institut National de la Recherche Agronomique, 86600 Lusignan, France*
[3]*Instituto Cavanilles de Biodiversidad y Biología Evolutiva, Universitat de València, Edificio Institutos, Campus de Paterna, Apartado de Correos 22085, 46071 Valencia, Spain*

In North America, the pea aphid *Acyrthosiphon pisum* encompasses ecologically and genetically distinct host races that offer an ideal biological system for studies on sympatric speciation. In addition to its obligate symbiont *Buchnera*, pea aphids harbour several facultative and phylogenetically distant symbionts. We explored the relationships between host races of *A. pisum* and their symbiotic microbiota to gain insights into the historical process of ecological specialization and symbiotic acquisition in this aphid. We used allozyme and microsatellite markers to analyse the extent of genetic differentiation between populations of *A. pisum* on pea, alfalfa and clover in France. In parallel, we examined: (i) the distribution of four facultative symbionts; and (ii) the genetic variation in the *Buchnera* genome across host-associated populations of *A. pisum*. Our study clearly demonstrates that populations of *A. pisum* on pea, clover and alfalfa in France are genetically divergent, which indicates that they constitute distinct host races. We also found a very strong association between host races of *A. pisum* and their symbiotic microbiota. We stress the need for phylogeographic studies to shed light on the process of host-race formation and acquisition of facultative symbionts in *A. pisum*. We also question the effects of these symbionts on aphid host fitness, including their role in adaptation to a host plant.

Keywords: *Acyrthosiphon pisum*; host races; ecological specialization; symbiotic bacteria; *Buchnera*; microsatellites

1. INTRODUCTION

Although controversial for a long time, the idea that natural selection can lead to divergence and speciation of sympatric populations has received increasing support during the last decade from both theoretical and empirical studies (Via 2001). Among the biological models in which sympatric speciation can be tested are phytophagous insects. Most insect herbivores are relatively host specific, feeding on only one or a few genera or on a single plant family. This intimacy in insect–plant relationships is a key factor that may promote ecological specialization after host shift in the absence of geographical isolation. For example, the apple maggot, *Rhagoletis pomonella*, is a well-documented model for sympatric speciation in progress (Filchak *et al.* 2000). Host races, feeding on either haw or apple, have been identified and are thought to represent incipient species, in partial reproductive isolation owing to host-associated differentiation (Feder *et al.* 1994). An important element that makes the apple maggot a good candidate for studies of sympatric speciation is the availability of historical records that testify to the formation of the apple race from locally existing haw populations (Bush 1969). Another case study for sympatric speciation is the pea aphid, *Acyrthosiphon pisum*, in which two host races, which feed primarily on alfalfa and red clover, are highly genetically specialized and substantially reproductively isolated in North America (Via 1999). The extensive work of Via *et al.* has allowed the identification of the main ecological and genetic factors that have facilitated divergence and the evolution of reproductive isolation in sympatry (Via 1991; Caillaud & Via 2000; Via *et al.* 2000). Recently, the genetic architecture that underlies specialization and reproductive isolation has been explored through quantitative genetics and quantitative trait loci mapping. This work has shown that these two biological traits are associated through pleiotropy or physical linkage, which creates the potential for rapid speciation, even in the presence of gene flow (Hawthorne & Via 2001). As stressed by Via (2001), an important step in a comprehensive analysis of sympatric divergence in pea aphids is to gather 'historical and phylogenetic data that might reveal whether the initial divergence between populations on the two hosts was sympatric or allopatric'. Pea aphids were introduced to North America from Europe, *ca.* 125 years ago (Thomas 1878). There is, however, no clear evidence as to whether host specialization preceded or followed introduction into North America. In addition, the extent of genetic differentiation between host populations of the pea aphid is not clear because it was assessed with markers differing in their levels of sensitivity. It varies from no

*Author for correspondence (jcsimon@rennes.inra.fr).

mitochondrial divergence between host races in North America (Barrette *et al.* 1994; Boulding 1998) or in the UK (Birkle & Douglas 1999) to highly allozymically differentiated host populations at the local level (Via 1999).

We analysed the level of genetic differentiation between populations of *A. pisum* on alfalfa and red clover from the native range of this species (e.g. France). To check whether the taxonomic distinction of individuals of *A. pisum* feeding on pea as a separate subspecies (i.e. *A. pisum destructor*) is relevant (Hille Ris Lambers 1973), pea aphids sampled from this crop were also included in this study. Genetic divergence was assessed primarily with allozymes, but also with microsatellites on a restricted sample, the latter class of markers being considered as one of the most powerful tools for population-genetics studies (Jarne & Lagoda 1996). To investigate further the genetic relationships between host-associated populations of the pea aphid, we also examined their genetic divergence at the symbiont level. First, genetic differentiation of *Buchnera*, the primary endosymbiont of aphids, was assessed between host races through a single strand conformation polymorphism (SSCP) analysis of a few symbiotic genes. Second, the distribution of secondary or facultative symbionts of *A. pisum* was studied according to host-associated populations. To date, five facultative symbionts that are vertically transmitted have been reported for *A. pisum* (reviewed in Tsuchida *et al.* 2002). Here, four out of the five symbionts were searched for in pea aphids using a PCR detection technique: PASS (pea aphid secondary symbiont; Chen & Purcell 1997) or R-type (Sandström *et al.* 2001); PAUS (pea aphid U-type symbiont; Sandström *et al.* 2001; Tsuchida *et al.* 2002) and PABS (pea aphid *Bemisia*-type symbiont) or T-type (Sandström *et al.* 2001; Darby *et al.* 2001), which are members of the *γ-Proteobacteria*, and PAR (pea aphid *Rickettsia*; Chen *et al.* 1996), a *Rickettsia* member of the *α-Proteobacteria*.

2. MATERIAL AND METHODS

(a) *Aphid collections*

In 1997, individuals of *A. pisum* were collected by sweep netting plants in single adjacent alfalfa (*Medicago sativum*), pea (*Pisum sativum*) and red clover (*Trifolium pratense*) fields (less than 1000 m apart) in April–June in central western France (Poitou). In 1998, pea aphids were again collected in alfalfa, pea and clover fields but from multiple sites in two different regions (Poitou and Languedoc in southern France) in May–June. No clover field was found close to alfalfa fields in Languedoc (table 1). At the time of collection (end of spring), all *A. pisum* sampled were parthenogenetic females and only wingless adults were kept for subsequent analyses.

(b) *Allozyme analysis*

Allozyme analysis was carried out on single individuals and performed on cellulose-acetate gels using standard methods (Hebert & Beaton 1989). Each aphid was assayed for variation in four enzymes: Sdh (sorbitol dehydrogenase), Idh (isocitrate dehydrogenase), Pep-LGG (tripeptidase Leu–Gly–Gly) and Pep-GL (dipeptidase Gly–Leu). These enzymes were chosen for analysis because they were found to be reliably resolved and polymorphic in previous allozyme surveys in French *A. pisum* (Bournoville *et al.* 2000). Since *A. pisum* exhibits pink and green forms, the colour type of each individual was noted prior to electrophoresis. We analysed 21 populations, encompassing more than 1000 individuals (table 1).

(c) *Microsatellite analysis*

DNA was extracted from single individuals using the salting-out method (Sunnucks *et al.* 1996). Microsatellite variation was assessed at seven loci only for pea aphids collected in 1997. The loci Sm10, Sm11, S17b, S23 and S30 were isolated from *Sitobion miscanthi* (Sunnucks *et al.* 1996, 1997; Simon *et al.* 1999; Wilson *et al.* 1999; A. C. Wilson, N. Prunier-Leterme, B. Massonnet, J.-C. Simon and P. Sunnucks, unpublished data). Sa5L was cloned from *S. avenae* (Haack *et al.* 2000; A. C. Wilson, N. Prunier-Leterme, B. Massonnet, J.-C. Simon and P. Sunnucks, unpublished data) while R5.10 was isolated from *Rhopalosiphum padi* (Simon *et al.* 2001). Microsatellite loci were amplified in a final volume of 15 µl using 3 pmol of each primer, 50 µM of each deoxynucleotide triphosphate (dNTP), 1.5 µl of 10× Mg^{++} free reaction buffer (500 mM of KCl, 100 mM of Tris–HCl pH 9.0 and 1.0% Triton X-100), 2 mM of $MgCl_2$, 0.5 U of *Taq* DNA polymerase (Promega) and 1 µl of DNA (5 ng). PCRs were performed on a Hybaid thermocycler with the following programme: one denaturation step at 94 °C for 2 min, followed by 35 cycles of a denaturation step at 94 °C for 20 s, 20 s at a locus-specific annealing temperature (usually 56 °C) and an elongation step at 72 °C for 20 s, then a final cycle at 72 °C for 2 min. Amplification products were resolved on a 6% polyacrylamide–urea electrophoresis gel and visualized after silver nitrate staining as described by Budowle *et al.* (1991). Allele sizes were assigned using the sequence of the pGEM®-3Zf(+) vector (Promega).

(d) *Detection of facultative symbionts*

Detection of the four facultative symbionts (PASS, PAR, PAUS and PABS) was based solely on a PCR assay with specific nucleotide primers to amplify their 16S rDNA. Specific primer pairs are given in Tsuchida *et al.* (2002). Facultative symbionts were amplified separately and the amplification was repeated twice for each sample. Amplifications were performed in a final volume of 25 µl using 4 pmol of each primer, 50 µM of each dNTP, 1.5 µl of 10× Mg^{++} free reaction buffer (500 mM of KCl, 100 mM of Tris–HCl pH 9.0 and 1.0% Triton X-100), 2.5 mM of $MgCl_2$, 0.5 U of *Taq* DNA polymerase (Promega) and 1 µl of aphid genomic DNA (5 ng). PCRs were performed on a Hybaid thermocycler with the following programme: one denaturation step at 95 °C for 5 min, followed by 35 cycles of a denaturation step at 94 °C for 1 min, 1 min at 53 °C and an elongation step at 72 °C for 2 min, then a final cycle at 72 °C for 10 min. Amplification products were resolved in 1.5% agarose gels in 0.5× TBE buffer pH 8.0 (89 mM of Tris, 89 mM of boric acid, 2 mM of ethylene diamine tetraacetic acid (EDTA) pH 8.0) at 100 V, and visualized under ultraviolet light after staining with ethidium bromide.

(e) *Molecular analysis of* Buchnera *diversity*

Molecular diversity among primary endosymbionts of *A. pisum* on pea, clover and alfalfa was assessed by SSCP (Sunnucks *et al.* 2000) analysis of three segments of the *Buchnera* genome. Partial amplification of the genes (i) *aroQ–PheA*, (ii) *GroEL* and (iii) *ffh–rpsP* including the intergenic region (Shigenobu *et al.* 2000) was achieved in 15 different pea aphid genotypes (distinguished after microsatellite analysis) of each host population using the following primers: PheAd4F (5′-ATCCCTCARCCNTTYCARC-3′) and PHEAD-4R (5′-

Table 1. Sampling design and sample sizes in host populations of *Acyrthosiphon pisum* used for genetic characterization and assessment of the prevalence of facultative symbionts.
(*n*, number of populations or individuals tested.)

year	crop	region	population	allozymes	microsatellites	facultative symbionts
1997	alfalfa	Poitou	A1C97	52	16	16
	alfalfa	Poitou	A2C97	46	14	14
	pea	Poitou	P1C97	104	30	30
	clover	Poitou	C1C97	95	30	30
1998	alfalfa	Poitou	A1C98	52	—	30
	alfalfa	Poitou	A2C98	46	—	25
	alfalfa	Poitou	A3C98	46	—	—
	alfalfa	Poitou	A4C98	36	—	—
	alfalfa	Poitou	A5C98	48	—	—
	alfalfa	Languedoc	A1S98	52	—	27
	alfalfa	Languedoc	A2S98	53	—	—
	pea	Poitou	P1C98	53	—	27
	pea	Poitou	P2C98	26	—	29
	pea	Poitou	P3C98	38	—	—
	pea	Poitou	P4C98	43	—	—
	pea	Poitou	P5C98	35	—	—
	pea	Languedoc	P1S98	18	—	40
	pea	Languedoc	P2S98	24	—	—
	pea	Languedoc	P3S98	42	—	—
	clover	Poitou	C1C98	68	—	32
	clover	Poitou	C2C98	51	—	35
overall (*n*)	—	—	21	1028	90	335

TWTTTTCWSWWGGATARCANCC-3′) amplified a 510 base pair (bp) segment of the gene *aroQ–PheA*, GroELR2 (5′-CTGCACCYACTTTRAKWACYGC-3′) and GroELF2 (5′-ATGTTRCAAGATATTTCWATTYTWAC-3′) amplified 795 bp of the *GroEL* gene and ffh-Ap-F1 (5′-CGTTACATTACAG GAAAACC-3′) and RpS16-Ap-GSP1 (5′-CCAGCTTTATAA AATGGACG-3′) amplified 730 bp of the *ffh* and *rpsP* genes. PCR reactions were prepared as for facultative symbionts and subjected to the following cycling conditions: one denaturation step at 94 °C for 2 min, followed by 35 cycles of a denaturation step at 94 °C for 10 s, 30 s at a locus-specific annealing temperature (45 °C for *aroQ-PheA*, 55 °C for *GroEL* and *ffh–rpsP*) and an elongation step at 72 °C for 45 s, then a final cycle at 72 °C for 2 min. PCR products were run using a manual sequencing instrument on a non-denaturing polyacrylamide gel: 10% acrylamide (40 : 2 acrylamide : *bis*-acrylamide) and 0.5× TBE without glycerol. Gel mix (60 ml) was set with 250 µl of ammonium persulphate and 50 µl of tetramethylethylediamine. The gel was run in 0.5× TBE buffer in a cold room (4 °C) at 20 W overnight and silver stained as for microsatellite analysis.

(f) *Analysis of genetic data*

Two kinds of genetic data were analysed: allozyme and microsatellite data. Because of their great sensitivity, microsatellites allow the detection of clonal copies in parthenogenetic organisms, whereas allozymes usually do not (e.g. Delmotte *et al.* 2002). Thus, since clonal 'amplification' of genotypes could affect data interpretation, analyses were carried out on all individuals for the allozyme dataset and on one copy per genotype for the microsatellite dataset (see Sunnucks *et al.* 1997 for more details). GENETIX v. 3.3 (http://www.univ-montp2.fr/~genetix/genetix/genetix.html) was used to calculate allelic frequencies, the mean number of alleles per locus, unbiased estimates of gene diversity and the expected heterozygosity under Hardy–Weinberg assumptions (Nei 1978). Genetic differentiation between pairs of subpopulations was analysed using exact tests provided by GENEPOP v. 3.1 (Raymond & Rousset 1995). *F*-statistic values (Weir & Cockerham 1984) were estimated and tested for significance with the FSTAT program (Goudet 1995).

Cavalli-Sforza & Edwards (1967) chord distance was calculated for the allozyme data of *A. pisum* on alfalfa, clover and pea, while the allele shared distance or D_{AS} (Chakraborty & Jin 1993) was computed for the microsatellite data of individual genotypes found on the three host plants. Pairwise genetic distance matrices were then subjected to cluster analysis using either neighbour-joining (NJ) or unweighted pair group method with arithmetic mean (UPGMA) algorithms. Bootstrap values were computed by resampling loci and are given as percentages over 1000 replications. Calculation of the genetic distances and tree reconstruction were carried out using POPULATIONS v. 1.2 (http://www.pge.cnrs-gif.fr/bioinfo/populations).

3. RESULTS

(a) *Genetic differentiation revealed by allozymes*

Populations of *A. pisum* showed low allozyme diversity with an average of 2.8 alleles per locus and a mean observed heterozygosity of 0.377 over all loci. There were 48 four-locus genotypes among the 1028 individuals examined for allozyme variation. Adding colour variation led to the characterization of a further 18 genotypes. Colour differences were found between host populations with almost no pink individuals in pea fields (table 2). Pea aphids from pea, clover and alfalfa fields also exhibited striking gene-frequency differences (table 2). Populations on clover showed near-fixation at the *Idh* locus and a private allele at the *Pep-GL* locus. Similarly, populations on pea showed near-fixation at the loci *Pep-LGG* and

Table 2. Colour distribution and allozyme variation between populations of *Acyrthosiphon pisum* in alfalfa, pea and clover fields. (H_{obs}, observed heterozygosity; H_{exp}, expected heterozygosity; s.e., standard error. Pairwise F_{ST} values of each host-population comparison are indicated on the bottom row (A–P, alfalfa–pea; P–C, pea–clover; C–A, clover–alfalfa).)

		population		
		alfalfa ($n = 431$)	pea ($n = 383$)	clover ($n = 214$)
colour				
green		0.789	0.995	0.658
pink		0.211	0.005	0.342
locus/allele				
Idh	A1	0.000	0.000	0.003
	A2	0.875	0.699	0.997
	A3	0.125	0.301	0.000
Sdh	A1	0.976	0.669	0.983
	A2	0.025	0.331	0.017
Pep-LGG	A1	0.580	0.062	0.416
	A2	0.419	0.931	0.585
	A3	0.000	0.007	0.000
Pep-GL	A1	0.000	0.000	0.306
	A2	0.221	0.076	0.112
	A3	0.779	0.924	0.582
H_{obs} (s.e.)		0.354 (0.325)	0.359 (0.308)	0.419 (0.460)
H_{exp} (s.e.)		0.275 (0.187)	0.283 (0.172)	0.270 (0.290)
within-host population F_{ST}		0.071	0.005	0.054
pairwise F_{ST} values		F_{ST} (A–P) = 0.270	F_{ST} (P–C) = 0.272	F_{ST} (C–A) = 0.098

Pep-GL. All except one test of pairwise single-locus genic differentiation between host populations were significant ($p < 0.05$). Only alfalfa and clover populations at locus *Sdh* were not significantly different ($p = 0.649$). Single-locus F_{ST} values were significant for all loci ($p < 0.001$), and the multilocus F_{ST} estimate was also highly significant ($F_{ST} = 0.199$, $p < 0.001$), indicating profound genetic differentiation between host populations. In all cases, genetic differentiation within host populations (F_{ST} values for populations of the same host) was much lower than genetic differentiation between host populations (table 2). The UPGMA clustering analysis on the chord distances of Cavalli-Sforza & Edwards (1967) revealed two major groups: the first corresponding to all populations found on pea and the second comprising the rest of the samples. However, an additional subdivision between clover and alfalfa populations was supported by a high bootstrap value (93%), indicating that these two host populations are also genetically differentiated from each other (figure 1).

(b) *Genetic differentiation revealed by microsatellites*

Microsatellites exhibited a higher allelic diversity than allozymes, with an average of 6.4 alleles per locus. The mean observed heterozygosity ranged from 0.238 to 0.738 per locus and averaged 0.593 over all loci. Thus, heterozygosity at microsatellite loci was 65% higher than at allozyme loci. There were 50 seven-locus genotypes among the 90 individuals of *A. pisum* examined for microsatellite variation (18 on alfalfa, 15 on pea and 17 on clover), indicating that several individuals shared the same genotype, i.e. were clonal copies. Genotypes in several copies were never found on two different host plants but were frequent within host populations.

As in allozymes, allelic-frequency distributions at microsatellite loci showed strong differences between host populations (table 3). Each of the three host populations possessed five private alleles out of a total of 45. In the alfalfa population, the frequency of one such private allele ($S30^{169}$) reached more than 55%. Similarly, one private allele ($S23^{129}$) of the pea population was present at a frequency of 25%. This strong genetic differentiation between host populations was confirmed by highly significant single-locus F_{ST} values (ranging from 0.070 for Sm11 to 0.499 for Sm10) and multilocus F_{ST} estimate ($F_{ST} = 0.231$, $p < 0.001$). An NJ tree using D_{AS} distances between individuals (genotypes) placed all but one (A8) of the sampled genotypes from the same host in the same group, reinforcing the idea of restricted gene flow between the three host populations (figure 2).

The genetic relatedness between host populations of *A. pisum* assessed with microsatellites is consistent with that revealed by allozymes, and both types of data suggest that: (i) pea lineages are monophyletic; (ii) clover lineages are monophyletic; while (iii) alfalfa lineages appear to be paraphyletic (figures 1 and 2).

(c) *Prevalence of facultative symbionts across populations of* A. pisum

Most pea aphids (86%) harboured facultative symbionts but the prevalences of the various symbiotic types were very different between host populations (figure 3). In 1997, all pea genotypes of *A. pisum* harboured PASS, while only 7% and 3% of aphids, respectively, on alfalfa and clover were infected with this bacterium (figure 3*a*).

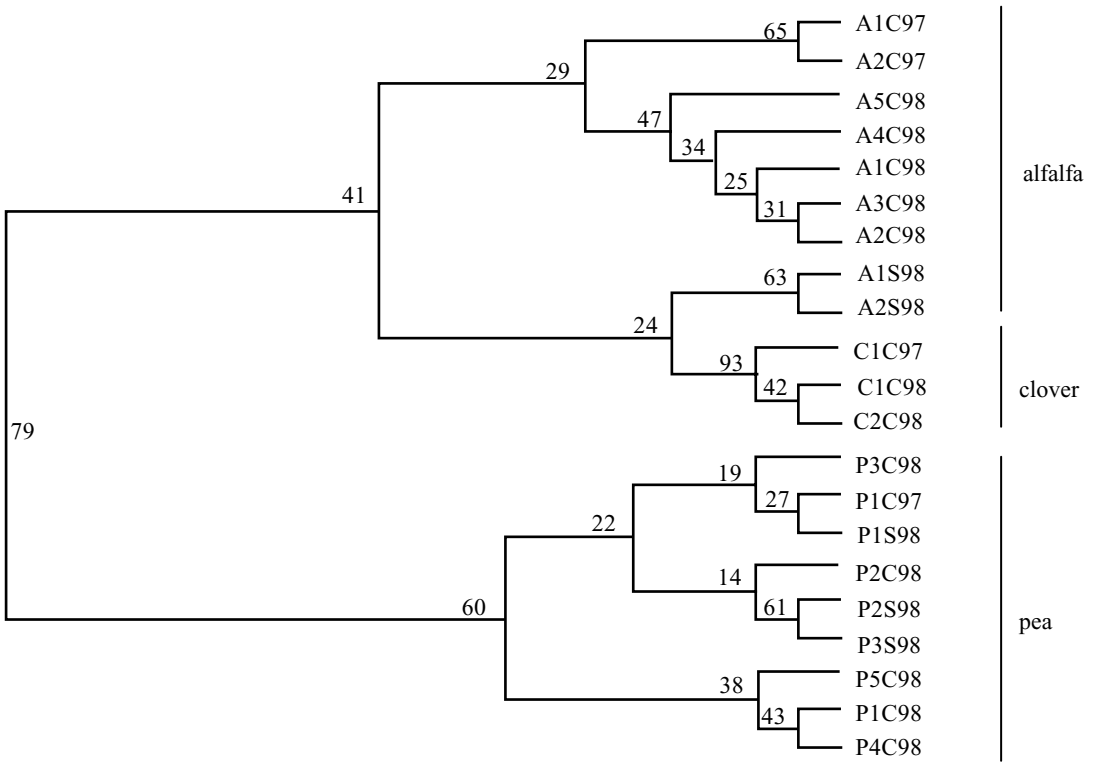

Figure 1. Unrooted UPGMA tree based on the chord distance of Cavalli-Sforza & Edwards (1967) calculated with four allozyme loci for populations of *Acyrthosiphon pisum* collected on pea, clover and alfalfa in 1997 and 1998 in two regions of France. The numbers at the nodes are bootstrap values given as percentages. Each population is coded by its plant source (A, alfalfa; C, clover; P, pea), field number, geographical origin (C, central western France; S, southern France) and year of collection (97, 1997; 98, 1998).

In addition, pea genotypes were infected with PAR at a rate of 30% and none of them harboured PABS or PAUS. By contrast, 70% of alfalfa genotypes were infected with PABS, while 71% of clover genotypes were infected with PAUS. This pattern of a predominance of PASS and PAR in pea genotypes, PABS in alfalfa genotypes and PAUS in clover genotypes was confirmed in 1998 for the two regions (Poitou and Languedoc). An average of 90% of pea genotypes were infected with PASS and 37% with PAR (figure 3*b*), with no significant differences between regions. By contrast, only 8% and 10% of alfalfa genotypes harboured PASS and PAR, respectively, and none of the clover aphids was infected with either bacterium (figure 3*b*). PABS was found only in alfalfa genotypes, and PAUS was the only facultative symbiont of clover genotypes.

Co-infected aphids represented 17.2% of the sample. PASS + PAR infection was the most common (54% of co-infected aphids) and exclusive to pea genotypes. PABS + PAUS (25%), PAR + PABS (9%) and PASS + PABS (6%) infections were restricted to alfalfa genotypes, and PASS + PAUS (6%) infection was restricted to clover genotypes. Aphids with more than two types of facultative symbiont were not found.

(**d**) *SSCP analysis of* Buchnera *genome*

SSCP analysis did not detect any variation between amplified portions of the genes *aroQ-PheA* and *GroEL* of *Buchnera* harboured by the various host populations of *A. pisum* (data not shown). By contrast, all clover genotypes of *A. pisum* examined with SSCP showed a clearly distinguishable conformation pattern of the *ffh–rpsP* genes of their associated *Buchnera*, probably caused by a deletion in the intergenic region (figure 4). Pea and alfalfa genotypes of *A. pisum* were not discriminated by SCCP of their *Buchnera* genes.

4. DISCUSSION

(**a**) *Genetic differentiation between host-associated populations of the pea aphid*

The genetic data presented here and the biological data on performances and feeding behaviour presented in a parallel study (Bournoville *et al.* 2003) together demonstrate that French populations of *A. pisum* consist of genetically divergent host-associated races. As Via (1999) found for pea aphids on clover and alfalfa in North America, our results indicate that genetic differentiation between sympatric races of pea aphids on pea, alfalfa and clover is greater than between subpopulations in geographically distant fields of the same crop. The present study suggests that host specialization in pea aphids may have preceded their introduction into North America. Alternatively, it may be that only one host-adapted race was introduced into North America, where the diversification between alfalfa and clover races took place secondarily. A comparison of the genetic relatednesses between and within host races of *A. pisum* in France and North America is needed to test whether host specialization evolved once or many times.

Although our genetic data also confirm earlier biological observations of host-associated populations of *A. pisum* on

Table 3. Allele frequencies and mean observed (H_{obs}) and expected (H_{exp}) heterozygosities at seven microsatellite loci in populations of *Acyrthosiphon pisum* on alfalfa, pea and clover (one-copy-per-genotype dataset).
(Microsatellite alleles are given as numbers of base pairs; s.e. standard error. Pairwise F_{ST} values of each host-population comparison are indicated on the bottom row (A–P, alfalfa–pea; P–C, pea–clover; C–A, clover–alfalfa).)

	population		
locus	alfalfa ($n = 18$)	pea ($n = 15$)	clover ($n = 17$)
S30			
159	0.079	0.125	0.526
161	0.237	0.125	0.211
163	0.000	0.125	0.000
165	0.000	0.000	0.026
167	0.000	0.594	0.026
169	0.553	0.000	0.000
171	0.132	0.031	0.132
182	0.000	0.000	0.026
187	0.000	0.000	0.053
Sm10			
151	0.526	0.000	0.806
153	0.368	1.000	0.194
156	0.079	0.000	0.000
189	0.026	0.000	0.000
S23			
125	0.000	0.063	0.000
129	0.000	0.250	0.000
131	0.526	0.031	0.762
134	0.316	0.063	0.071
136	0.053	0.406	0.048
138	0.000	0.125	0.024
140	0.105	0.000	0.095
142	0.000	0.063	0.000
R5.10			
238	0.028	0.219	0.184
254	0.028	0.000	0.158
263	0.083	0.000	0.053
268	0.528	0.000	0.395
270	0.139	0.031	0.211
273	0.194	0.750	0.000
S5.L			
203	0.579	0.688	0.333
204	0.105	0.063	0.024
205	0.316	0.219	0.310
207	0.000	0.031	0.333
S17b			
213	0.342	0.167	0.476
215	0.158	0.567	0.024
218	0.500	0.233	0.500
219	0.000	0.033	0.000
Sm11			
137	0.000	0.233	0.175
139	0.111	0.300	0.000
141	0.000	0.000	0.025
144	0.222	0.367	0.450
146	0.083	0.000	0.050
150	0.083	0.100	0.175
152	0.222	0.000	0.075
176	0.000	0.000	0.050
197	0.194	0.000	0.000
199	0.083	0.000	0.000
H_{obs} (s.e.)	0.718 (0.119)	0.857 (0.330)	0.614 (0.187)
H_{exp} (s.e.)	0.636 (0.090)	0.502 (0.250)	0.577 (0.160)
pairwise F_{ST} values	F_{ST} (A–P) = 0.244	F_{ST} (P–C) = 0.326	F_{ST} (C–A) = 0.104

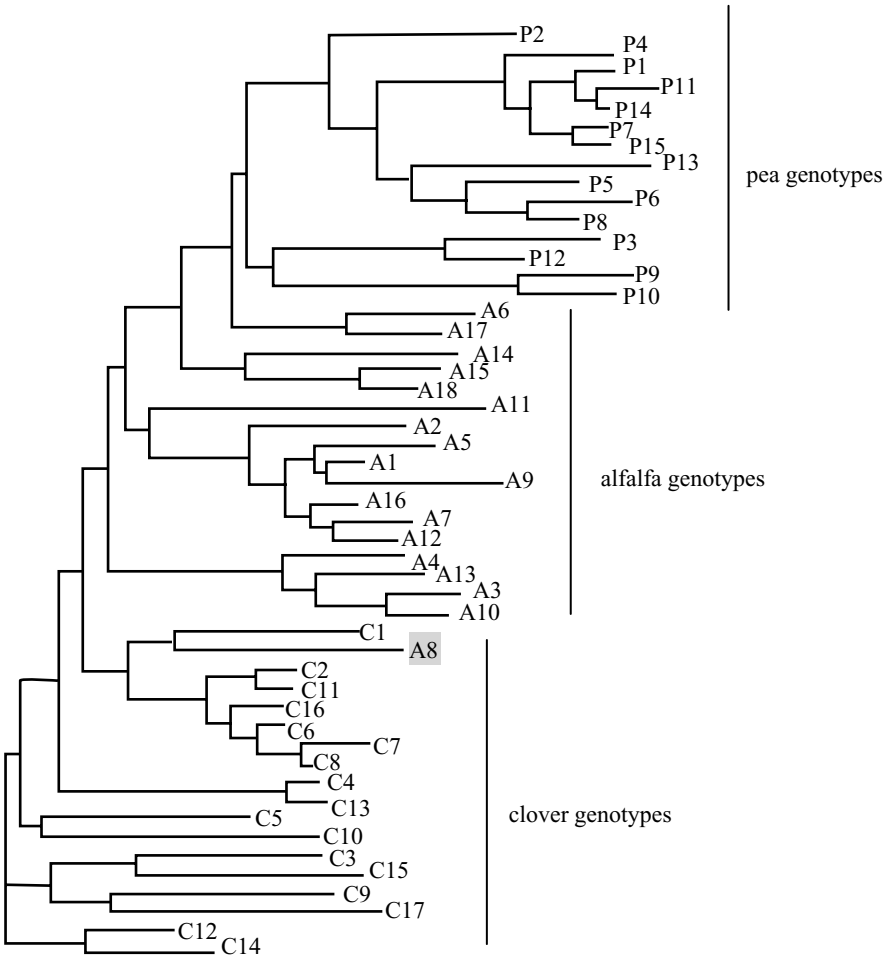

Figure 2. Unrooted NJ tree based on allele shared distance (D_{AS}) calculated with seven microsatellite loci for multilocus genotypes of *Acyrthosiphon pisum* collected on pea, clover and alfalfa in 1997. Genotypes are designated with a letter corresponding to their host-plant source (A, alfalfa; C, clover; P, pea). Genotype A8 (highlighted on the tree), which was collected on alfalfa but clusters into the clover group, presumably corresponds to a migrant that just landed on a bad host (see § 4).

several herbaceous Fabaceae in Europe (Müller 1980), they do not give conclusive support to the taxonomic distinction of pea aphids on pea as *A. p. destructor* (Hille Ris Lambers 1973). The extent of gene flow (as measured by F_{ST}) between pea aphids on pea and those on clover and alfalfa is more than twice as great as that between pea aphids on clover and alfalfa (tables 2 and 3). However, it appears from the SSCP analysis of the primary endosymbiont that pea aphids on pea and alfalfa have closer *Buchnera* genomes than those on clover. Whatever their species status, these divergent host-associated races represent three distinct 'genotypic clusters' (Mallet 1995), which are close to the species boundary (Via 1999).

While aphids in general do not exhibit much allozyme polymorphism, the variation present may still indicate significant differences in host utilization (Hales *et al.* 1997). Earlier allozyme studies on pea aphids failed to detect substantial between-population variation (Simon *et al.* 1982; Dalmasso & Bournoville 1983). However, broader screenings of enzymatic systems and the use of sensitive allozyme techniques have revealed several polymorphic loci in *A. pisum* (Via 1999; Bournoville *et al.* 2000). Combining the results of Via (1999) with the present results revealed alleles specific to host-associated populations, which can be used for race diagnosis (e.g. the slow *Pep-GL* allele, which is restricted to clover genotypes in both North American and French populations). Microsatellites are even more powerful diagnostic tools as they correctly assigned most individual genotypes to their host race (figure 2). Only genotype A8 was wrongly assigned; it could be the offspring of a migrant that recently landed on its bad host, so the effect of selection has not had time to be completed. Alternatively, A8 could be a generalized genotype able to feed on both clover and alfalfa. Although laboratory F_1 hybrids between clover and alfalfa races have a lower average fitness than their specialized parents, there is still significant genetic variation in fitness traits among those hybrids for a generalized pea aphid to evolve (Via *et al.* 2000). However, the first explanation is preferred because (i) aphid sampling was done quite early in the cropping season and (ii) generalized genotypes of the pea aphid have not yet been detected in the field.

Interestingly, the extent of genetic differentiation between clover and alfalfa races in North America ($F_{ST} = 0.212$; Via 1999) is very similar to that found in France between the three races for both allozyme ($F_{ST} = 0.199$) and microsatellite ($F_{ST} = 0.231$) markers (present study). This level of genetic divergence, which falls within the range of values usually measured for geographical or chromosomal races, still offers the possibility

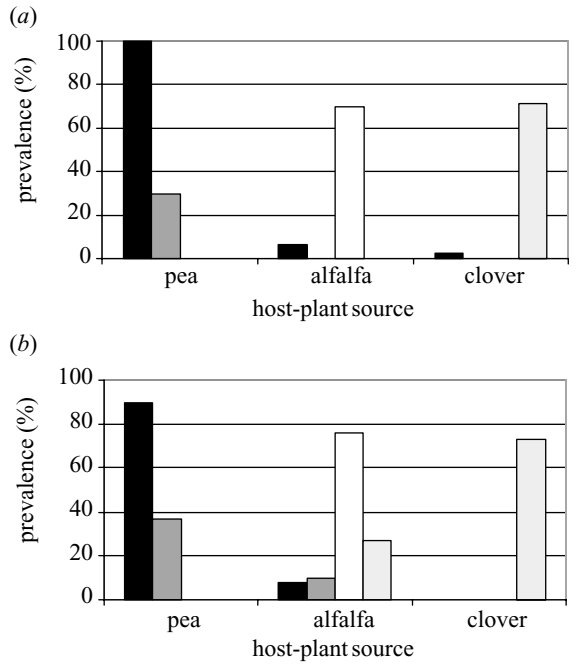

Figure 3. Prevalences of four facultative symbionts of *Acyrthosiphon pisum* in populations sampled in pea, clover and alfalfa fields in (*a*) 1997 and (*b*) 1998 in two regions of France. Black bars, PASS; dark-grey bars, PAR; open bars, PABS; light-grey bars, PAUS.

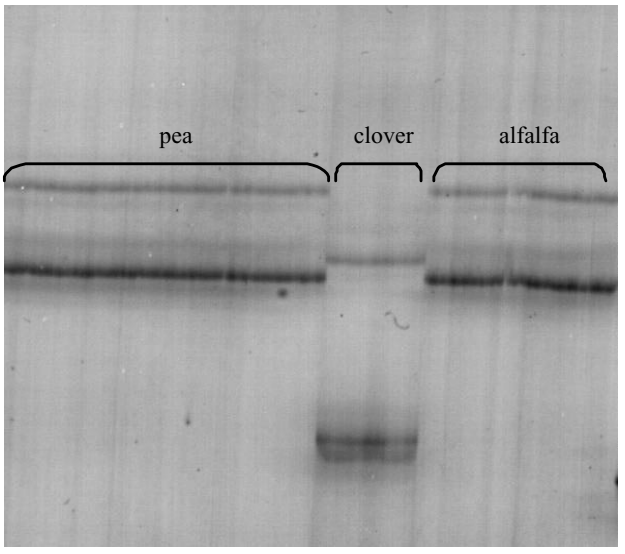

Figure 4. Molecular variation between the *Buchnera* genomes of host-associated populations of *Acyrthosiphon pisum* (SSCP variation at the *ffh–rpsP* genes).

for gene exchange (Via 1999). By contrast, it is intriguing that, despite highly significant genetic differences in allozyme and microsatellite markers, previous mitochondrial DNA surveys failed to detect any variation associated with host adaptation in either North American (Barrette *et al.* 1994; Boulding 1998) or British (Birkle & Douglas 1999) populations of *A. pisum*. This may indicate that divergence between host races is too recent for the detection of a phylogenetic signal on the mitochondrial genome.

(b) *Diversity and distribution of symbionts across host races of the pea aphid*

The main results of our analysis of the microbial community residing in French *A. pisum* are threefold: (i) besides the primary endosymbiont, *Buchnera*, most individual aphids are infected with one or two facultative symbionts; (ii) there is a very strong association between host races of *A. pisum* and their facultative symbionts; and (iii) the *Buchnera* genome may differ between host-associated populations. In summary, this survey of symbiotic diversity in *A. pisum* further confirms the high genetic divergence between clover, alfalfa and pea races revealed by nuclear markers.

French populations of pea aphids are characterized by a great diversity in their facultative symbionts. The four main facultative symbionts of *A. pisum* have been detected in our survey, all at high rates. Only a small fraction of individuals (14%) had *Buchnera* only. This fraction of aphids could harbour other symbionts, such as *Spiroplasma*, which infected a small percentage of Japanese *A. pisum* (Fukatsu *et al.* 2001; Tsuchida *et al.* 2002), or other bacteria that have not yet been described. Alternatively, facultative symbionts could have been lost from these aphids under certain environmental conditions or after failure of vertical transmission (Darby *et al.* 2001; Tsuchida *et al.* 2002). Variation between clones of *A. pisum* in their facultative-symbionts repertoire has already been reported for PASS, PAR and PAUS in North America (Chen *et al.* 1996; Chen & Purcell 1997; Sandström *et al.* 2001), for PABS in the UK (Darby *et al.* 2001, 2003) and for PASS, PAR, PAUS and *Spiroplasma* in Japan (Tsuchida *et al.* 2002). None of these previous studies made a clear link between host-associated populations of pea aphids and the distribution of their facultative symbionts, possibly because in most cases it was not explicitly searched for.

Double-infected individual aphids represented a substantial fraction (17.2%) of French populations of *A. pisum*. Previous studies found much lower co-infection rates, for example 0.5% in Japanese pea aphids (Tsuchida *et al.* 2002). It was suggested that this rarity of multiple-symbiont infections resulted from conflicts for resources and space in the same host (Sandström *et al.* 2001). This seems not to be the case for French populations of pea aphids, in which at least some combinations of facultative symbionts are relatively common (e.g. PASS–PAR infection, which is found in 54% of co-infected pea aphids). These differences in double-infection rates between French and Japanese populations of *A. pisum* might be explained by environmental or historical factors. For example, it could be that extreme temperatures, which are more frequent in Japan than in France, may affect the stability of symbiont infection. Alternatively, a higher rate of multiple-symbiont infections may suggest a longer history of symbiont–pea aphid interactions and the potential for numerous cases of horizontal transfer.

The strong association between symbionts and races of *A. pisum* reported here raises several questions regarding (i) its origin(s) and maintenance; and (ii) its consequences for host fitness. It has been previously shown from molecular phylogenetic analyses that facultative symbionts of pea aphids had independent origins (Chen *et al.* 1996; Chen & Purcell 1997; Fukatsu *et al.* 2001; Sandström *et*

al. 2001). Thus, two scenarios can be proposed to explain the specific symbiotic composition of pea aphid races. First, it could be that different facultative symbionts infected some lineages of *A. pisum* before ecological specialization and genetic differentiation of the host races. Subsequently, these symbionts could have played a part in plant adaptation. However, this scenario contradicts recent findings that the genetic architecture of pea aphids *per se* determines plant specialization, with no evidence for the involvement of maternal factors, including obligate and facultative symbionts (Caillaud & Via 2000; Hawthorne & Via 2001). The second, more probable, scenario postulates that facultative-symbiont infections occurred after ecological specialization in *A. pisum*, their current distribution reflecting the neutral genetic divergence between host-associated populations. Acquisitions of facultative symbionts might have occurred independently in the distinct host races through either aborted parasitoid attacks or the oral route (Darby *et al.* 2001; Sandström *et al.* 2001). For example, the host specialization found in a wasp parasitoid that parasitizes the pea aphid (Henter 1995) might have been responsible for the pattern of association of symbionts with host races. By contrast, it is plausible that facultative symbionts were acquired through ingestion from plant surfaces or from the interior of plants, which would readily explain the specific symbiotic microbiota of *A. pisum* races. Whatever their origin, the relationships between host races and their facultative symbionts are not strict, which suggests some lability in symbiotic interactions or, more probably, occasional horizontal-transmission events. The latter may be possible either through exchange between host races that are still genetically connected or through the parasitic or oral routes mentioned above.

While it is now well established that *Buchnera*, the obligate symbiont of aphids, supplies its host with essential amino acids that are absent from the insect diet (Baumann *et al.* 1995; Douglas 1998; Shigenobu *et al.* 2000), the effects of facultative symbionts on host fitness are still poorly understood. PASS can have both positive and negative effects depending on pea aphid genotype and temperature (Chen *et al.* 2000), while PAR and PABS seem to confer resistance to parasitic wasps (Oliver *et al.* 2003). Thus, we expect that infections with these bacteria will have consequences for the physiology and ecology of their aphid hosts (Sandström *et al.* 2001; Oliver *et al.* 2003). It is possible that adaptive variation even exists between the *Buchnera* genomes of pea aphids and could facilitate host utilization. For example, clones of the pea aphid can differ in their copy number of *trpEG* genes located on a plasmid that overproduces tryptophan (Baumann *et al.* 1995). Whether this copy number is related to host races requires further investigation.

In conclusion, our study clearly demonstrates that French *A. pisum* on pea, clover and alfalfa are highly genetically divergent and that each race possesses a specific composition of facultative symbionts. Whether this association between host races of *A. pisum* and their symbionts is universal remains to be explored through phylogeographic studies, in particular in North America where alfalfa and clover races coexist. This would help us to understand the origin, and the routes and age of acquisition of facultative symbionts of pea aphids. In addition, a very important task would be to elucidate the effect of facultative symbionts on their aphid hosts, including their role in host adaptation. Finally, the bacterial community residing in pea aphids shows a high level of complexity, which requires a thorough investigation of symbiotic interactions and their dynamics.

The authors thank Tsutomu Tsuchida from Takema Fukatsu's group in Japan for sharing information on the PCR detection of facultative symbionts in the pea aphid, Francisco J. Silva for providing some of the *Buchnera* primers, and Alex Wilson and Paul Sunnucks for sharing unpublished microsatellites of *Sitobion miscanthi*. Claude Rispe, Paul Sunnucks and three anonymous reviewers made helpful comments on this paper. The authors thank the Département 'Santé des Plantes et Environnement' of the Institut National de la Recherche Agronomique for partial financial support of this work.

REFERENCES

Barrette, R. J., Crease, T. J., Hebert, P. D. N. & Via, S. 1994 Mitochondrial DNA diversity in the pea aphid, *Acyrthosiphon pisum*. *Genome* **37**, 858–865.

Baumann, P., Baumann, L., Lai, C. Y., Rouhbakhsh, D., Moran, N. A. & Clark, M. A. 1995 Genetics, physiology and evolutionary relationships of the genus *Buchnera*: intracellular symbionts of aphids. *A. Rev. Microbiol.* **49**, 55–94.

Birkle, L. M. & Douglas, A. E. 1999 Low genetic diversity among pea aphid (*Acyrthosiphon pisum*) biotypes of different plant affiliation. *Heredity* **82**, 605–612.

Boulding, E. G. 1998 Molecular evidence against phylogenetically distinct host races of the pea aphid (*Acyrthosiphon pisum*). *Genome* **41**, 769–775.

Bournoville, R., Simon, J.-C., Badenhausser, I., Girousse, C. & André, S. 2000 Clones of pea aphid, *Acyrthosiphon pisum* (Hemiptera: Aphididae) distinguished using genetic markers, differ in their damaging effect on a resistant alfalfa cultivar. *Bull. Entomol. Res.* **90**, 33–39.

Bournoville, R., Carré, S., Badenhausser, I., Simon, J.-C., Hennis, C. & Greze, C. 2003 Host-races of the pea aphid, *Acyrthosiphon pisum*: biological criteria and feeding behaviour of clones originating from legumes. In *Proc. VIth Int. Symp. on Aphids* (ed. J.-C. Simon, C.-A. Dedryver, C. Rispe & M. Hullé). Paris: INRA Editions. (In the press.)

Budowle, B., Chakraborty, R., Giusti, A. M., Eisenberg, A. J. & Allen, R. C. 1991 Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am. J. Hum. Genet.* **48**, 137–144.

Bush, G. L. 1969 Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera: Tephritidae). *Evolution* **23**, 237–251.

Caillaud, C. M. & Via, S. 2000 Specialized feeding behavior influences both ecological specialization and assortative mating in sympatric host races of pea aphids. *Am. Nat.* **156**, 606–621.

Cavalli-Sforza, L. L. & Edwards, A. W. 1967 Phylogenetic analysis: models and estimation procedures. *Am. J. Hum. Genet.* **19**, 233–257.

Chakraborty, R. & Jin, L. 1993 A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. In *DNA fingerprinting: state of the science* (ed. S. D. J. Pena, R. Chakraborty, J. T. Epplen & A. J. Jeffreys), pp. 153–175. Basel: Birkhauser.

Chen, D. Q. & Purcell, A. H. 1997 Occurrence and transmission of facultative endosymbionts in aphids. *Curr. Microbiol.* **34**, 220–225.

Chen, D. Q., Campbell, B. C. & Purcell, A. H. 1996 A new

rickettsia from a herbivorous insect, the pea aphid *Acyrthosiphon pisum* (Harris). *Curr. Microbiol.* **33**, 123–128.

Chen, D. Q., Montflor, C. B. & Purcell, A. H. 2000 Fitness effects of two facultative endosymbiotic bacteria on the pea aphid, *Acyrthosiphon pisum* and the blue alfalfa aphid, *A. kondoi*. *Entomol. Exp. Appl.* **95**, 315–323.

Dalmasso, A. & Bournoville, R. 1983 Isoenzymes d'une population d'*Acyrthosiphon pisum*. *Agronomie* **3**, 891–896.

Darby, A. C., Birkle, L. M., Turner, S. L. & Douglas, A. E. 2001 An aphid-borne bacterium allied to the secondary symbionts of whitefly. *FEMS Microbiol. Ecol.* **36**, 43–50.

Darby, A. C., Tosh, C. R., Walters, K. F. A. & Douglas, A. E. 2003 The significance of a facultative bacterium to natural populations of the pea aphid *Acyrthosiphon pisum*. *Ecol. Entomol.* **28**, 145–150.

Delmotte, F., Leterme, N., Gauthier, J. P., Rispe, C. & Simon, J.-C. 2002 Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Mol. Ecol.* **11**, 711–723.

Douglas, A. E. 1998 Nutritional interactions in insect–microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *A. Rev. Entomol.* **43**, 17–37.

Feder, J. L., Opp, S. B., Wlazlo, B., Reynolds, K., Go, W. & Spisak, S. 1994 Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proc. Natl Acad. Sci. USA* **91**, 7990–7994.

Filchak, K. E., Roethele, J. B. & Feder, J. L. 2000 Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature* **407**, 739–742.

Fukatsu, T., Tsuchida, T., Nikoh, N. & Koga, R. 2001 *Spiroplasma* symbiont of the pea aphid, *Acyrthosiphon pisum* (Insecta: Homoptera). *Appl. Environ. Microbiol.* **67**, 1284–1291.

Goudet, J. 1995 FSTAT (version 1.2). A computer program to calculate *F*-statistics. *J. Hered.* **86**, 485–486.

Haack, L., Simon, J.-C., Gauthier, J. P., Plantegenest, M. & Dedryver, C. A. 2000 Evidence for predominant clones in a cyclically parthenogenetic organism provided by combined demographic and genetic analyses. *Mol. Ecol.* **9**, 2055–2066.

Hales, D. F., Tomiuk, J., Wöhrmann, K. & Sunnucks, P. 1997 Evolutionary and genetic aspects of aphid biology: a review. *Eur. J. Entomol.* **94**, 1–55.

Hawthorne, D. J. & Via, S. 2001 Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* **412**, 904–907.

Hebert, P. D. N. & Beaton, M. J. 1989 *Methodologies for allozyme analysis using cellulose acetate electrophoresis*. Beaumont, TX: Helena Laboratories.

Henter, H. J. 1995 The potential for coevolution in a host–parasitoid system. II. Genetic variation within a population of wasps in the ability to parasitize an aphid host. *Evolution* **49**, 439–445.

Hille Ris Lambers, D. 1973 Wirtspflanzen und uberwinterung eines erbsen lebenden biotyps von *Acyrthosiphon pisum* (Harris) (Homoptera: Aphididae). *Deutsche entomol. Zeitschr.* **20**, 321–328.

Jarne, P. & Lagoda, P. J. 1996 Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **11**, 424–429.

Mallet, J. 1995 A species definition for the modern synthesis. *Trends Ecol. Evol.* **10**, 294–299.

Müller, F. P. 1980 Wirtspflanzen, generationenfolge und reproduktive isolation intraspezifischer formen von *Acyrthosiphon pisum*. *Entomol. Exp. Appl.* **28**, 145–157.

Nei, M. 1978 The estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590.

Oliver, K. M., Russell, J. A., Moran, N. A. & Hunter, M. S. 2003 Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc. Natl Acad. Sci. USA* **100**, 1803–1807.

Raymond, M. & Rousset, F. 1995 GENEPOP, a population genetics software for exact tests and oecumenicism. *J. Hered.* **86**, 248–249.

Sandström, J., Russel, J. A., White, J. P. & Moran, N. A. 2001 Independent origins and horizontal transfer of bacterial symbionts of aphids. *Mol. Ecol.* **10**, 217–228.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. & Ishikawa, H. 2000 Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* **407**, 81–86.

Simon, J.-C., Baumann, S., Sunnucks, P., Hebert, P. D. N., Pierre, J. S., Le Gallic, J. F. & Dedryver, C. A. 1999 Reproductive mode and population genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Mol. Ecol.* **8**, 531–545.

Simon, J.-C., Leterme, N., Delmotte, F., Martin, O. & Estoup, A. 2001 Isolation and characterization of microsatellite loci in the aphid species, *Rhopalosiphum padi*. *Mol. Ecol. Notes* **1**, 4–5.

Simon, J. P., Parent, M. A. & Auclair, J. L. 1982 Isozyme analysis of biotypes and field populations of the pea aphid, *Acyrthosiphon pisum*. *Entomol. Exp. Appl.* **32**, 186–192.

Sunnucks, P., England, P. R., Taylor, A. C. & Hales, D. F. 1996 Microsatellite and chromosome evolution of parthenogenetic *Sitobion* aphids in Australia. *Genetics* **144**, 747–756.

Sunnucks, P., De Barro, P. J., Lushai, G., Maclean, N. & Hales, D. F. 1997 Genetic structure of an aphid studied using microsatellites: cyclic parthenogenesis, differentiated lineages, and host specialization. *Mol. Ecol.* **6**, 1059–1073.

Sunnucks, P., Wilson, A. C., Beheregaray, L. B., Zenger, K., French, J. & Taylor, A. C. 2000 SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Mol. Ecol.* **9**, 1699–1710.

Thomas, C. 1878 A list of the species of the tribe Aphidini, family Aphidae, found in the United States, which have been heretofore named with descriptions of some new species. *State Lab. Nat. Hist. Bull.* **1**, 1–16.

Tsuchida, T., Koga, R., Shibao, H., Matsumoto, T. & Fukatsu, T. 2002 Diversity and geographical distribution of secondary endosymbiotic bacteria in natural populations of the pea aphid, *Acyrthosiphon pisum*. *Mol. Ecol.* **11**, 2123–2135.

Via, S. 1991 The genetic structure of host plant adaptation in a spatial patchwork: demographic variability among reciprocally transplanted pea aphid clones. *Evolution* **45**, 827–852.

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 Ecol. Evol.* **16**, 381–390.

Via, S., Bouck, A. C. & Skillman, S. 2000 Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution* **54**, 1626–1637.

Weir, B. S. & Cockerham, C. C. 1984 Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

Wilson, A. C. C., Sunnucks, P. & Hales, D. F. 1999 Microevolution, low clonal diversity and genetic affinities of parthenogenetic *Sitobion* aphids in New Zealand. *Mol. Ecol.* **8**, 1655–1666.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.