Biological Journal of the Linnean Society, 2007, **91**, 687–702. With 5 figures

Microsatellite DNA and behavioural studies provide evidence of host-mediated speciation in *Myzus persicae* (Hemiptera: Aphididae)

JOHN T. MARGARITOPOULOS[1], GAVIN MALARKY[2], JOHN A. TSITSIPIS[1] and ROGER L. BLACKMAN[2]*

[1]*Laboratory of Entomology and Agricultural Zoology, Department of Agriculture Crop Production and Rural Environment, Fytokou Str., 384 46, Nea Ionia, Magnesia, Greece*
[2]*Department of Entomology, The Natural History Museum, London SW7 5BD, UK*

Received 3 April 2006; accepted for publication 2 October 2006

In many parts of the world, the tobacco specialist *Myzus persicae nicotianae* is isolated from the generalist *Myzus persicae s.s*. because either or both taxa reproduce parthenogenetically. Here we investigated how the genomic integrity of the tobacco specialist is maintained in Greece, where both taxa have a bisexual generation on peach. Microsatellite DNA analysis revealed greatest genetic divergence between populations in tobacco-growing regions and those in a region where tobacco is not cultivated. This was irrespective of reproductive mode, which has an important effect on population structure. Bayesian clustering and admixture analyses split the aphid genotypes into three groups, corresponding with *persicae*, bisexual *nicotianae* and unisexual *nicotianae*, respectively. Genetic distance parameters showed strong regional differentiation but marked year-on-year stability, indicating low interpopulation migration. Assortative mating between taxa is promoted by differences in the daily rhythm of female signalling behaviour, with peak activity coinciding with periods of consubspecific male searching activity. Males showed greater attraction to the sex pheromone of their own subspecies. Thus, despite relatively low overall genetic differentiation, processes are in place facilitating further genomic divergence and eventual speciation.

ADDITIONAL KEYWORDS: cyclical parthenogenesis – mating behaviour – *nicotianae* –tobacco aphid.

The classical Darwinian view that differential selection, and adaptation to different environmental factors (both biotic and abiotic), leads to divergence of populations and speciation has recently regained broader acceptance among zoologists (Wu, 2001; Via, 2002). Theoretical and empirical studies have shown how environmental differences may interact with genetic traits to initiate the processes of divergence, leading to incipient speciation (reviewed by Schluter, 2001; Via, 2001). Particular attention has been paid to phytophagous insects showing divergent evolution through adaptation to different hosts, and numerous examples of host-adapted races or biotypes have been demonstrated (reviewed by Drès & Mallet, 2001; Berlocher & Feder, 2002). In some cases specific 'genetic architectures' have been found that may facilitate the rapid evolution of specialization and reproductive isolation, even in the face of gene flow: e.g. the pea aphid *Acyrthosiphon pisum* (Harris) (Via & Hawthorne, 2002) and the apple maggot *Rhagoletis pomonella* (Walsh) (Berlocher & Feder, 2002; Feder *et al*., 2003). Such studies are usually claimed to provide evidence of sympatric speciation (Drès & Mallet, 2001; Via, 2001), but contrary views have been expressed (Futuyama & Mayer, 1980; Claridge, 1988). There are well-documented examples of host shifters that conform either to sympatric (see above) or to allopatric, e.g. butterflies *Euphydryas editha*, speciation hypotheses (see Berlocher & Feder, 2002 for a review).

Aphids are ideal models for studying the evolution of insect–plant relationships. They show a high degree of host specificity, with 99% of all species restricted to one or a few closely related plant species (Eastop,

*Corresponding author. E-mail: r.blackman@nhm.ac.uk

1973), and many examples have been published demonstrating host-adapted aphid biotypes, races or subspecies (Blackman, 1990; Via, 1991; Sunnucks *et al*., 1997; Vanlerberghe-Masutti & Chavigny, 1998; Haack *et al*., 2000). Aphids typically have cyclical parthenogenesis, in which an annual bisexual generation through the winter alternates with a succession of parthenogenetic generations during spring and summer. The bisexual generation may be lost either totally (genotypes with obligate parthenogenesis) or partially (genotypes that reproduce mainly parthenogenetically, but retain the ability to produce a few mating females and/or males), and the complete range of reproductive strategies can occur sympatrically within species (reviewed by Hales *et al*., 1997; Simon, Rispe & Sunnucks, 2002). The complexities of the aphid life cycle need to be fully considered in studies of their population genetics, as the method of reproduction determines how genetic variation is produced and organized. Work with microsatellite DNA markers on several aphid species has shown that the mode of reproduction significantly affects various aspects of the genetic structure of populations, such as levels of genic and genotypic variability, heterozygosity, and genetic differentiation between populations (Simon *et al*., 1999; Delmotte *et al*., 2002; Guillemaud, Mieuzet & Simon, 2003; Vorburger, Lancaster & Sunnucks, 2003). The evolutionary consequences of these differences are still unclear, and contrary views have been expressed about whether adaptation to new hosts is more likely to be favoured by cyclical or by obligate parthenogenesis (Lynch, 1984; Hales *et al*., 1997).

In addition to the well-documented work on host races of *A. pisum* (Via, 2001), closely related aphid taxa with different host adaptations in the genus *Cryptomyzus* (Guldemond & Dixon, 1994) and in the *Aphis fabae* group (Raymond *et al*., 2001) have also been put forward as good candidates for sympatric speciation. Sympatric speciation in host-alternating aphids may involve a reinforcement process resulting from selection against hybrids between the two host-specialized subspecific forms (Guldemond & Mackenzie, 1994; Mackenzie & Guldemond, 1994). These hybrids are less able to use either of the two parental secondary hosts because of host-associated fitness trade-offs. Hybrid dysfunction occurs widely among aphids (e.g. Müller, 1985; Guldemond, 1990; Via, Bouck & Skillman, 2000), and performance trade-offs have been demonstrated (e.g. Via, 1991a; Mackenzie, 1996; Douglas, 1997). Strong selection for behaviour that favours assortative mating has been shown to occur where closely related taxa reproduce sexually on the same host (Guldemond & Dixon, 1994; Thieme & Dixon, 1996), or even on different hosts (Raymond *et al*., 2001).

We have been investigating the genetics and ecology of *Myzus persicae* populations from various regions of mainland Greece. This species is extremely polyphagous (Blackman & Eastop, 2000), but there is evidence of genetic variation with respect to host-plant adaptation. It has been long suspected that populations colonizing tobacco *Nicotiana tabacum* L. are different from those on other crops (De Jong, 1929; Brain, 1942). This was confirmed by multivariate morphometric studies (Blackman, 1987), which have also subsequently shown that the tobacco-feeding form retains its host-related properties through time, even in regions (Greece and Japan) where there is a bisexual generation on the shared primary host, the peach *Prunus persica* L. (Margaritopoulos *et al*., 2000, 2003, 2007). Analysis of clonal lineages reared under controlled conditions on the same host plant proved that these differences have a genetic basis. The adaptation to tobacco involves negative trade-offs that reduce the performance of tobacco aphids on other crops (Nikolakakis, Margaritopoulos & Tsitsipis, 2003), and there is selection against transfers between hosts that has presumably led to the improved host recognition by winged female migrants (Margaritopoulos *et al*., 2005). These studies also showed correlations between acceptance and performance traits, but the genetic basis has yet to be confirmed by breeding experiments. Subspecies status (as for *Myzus persicae nicotianae* Blackman) has been proposed for the tobacco-feeding form of *M. persicae* (Margaritopoulos *et al*., 2003; Eastop & Blackman, 2005). Nevertheless, DNA studies on *M. persicae*, including some aimed specifically at comparing tobacco- and non-tobacco-feeding populations, have failed to find consistent diagnostic genetic markers (Margaritopoulos, Mamuris & Tsitsipis, 1998; Clements *et al*., 2000a, b). DNA sequence data indicate that adaptation to tobacco must have occurred very recently, and that there is some introgression of genes from non-tobacco-feeding populations (Field *et al*., 1994; Clements *et al*., 2000b). However, there are no comparative studies of *persicae* and *nicotianae* using microsatellite DNA markers, which are highly polymorphic codominant markers suitable for population genetic analysis.

In a recent study (Blackman *et al*., 2007), we used microsatellite markers to examine the distribution of common genotypes (i.e. those found at more than one time and place, which are therefore less likely to be products of recent sexual reproduction) in relation to host plant and regional life cycle differences in Greece and in southern Italy. Common genotypes comprised 49% of all the lineages analysed, and as expected their distribution matched that of anholocyclic clones, i.e. those capable of continuous year-round parthenogenesis. Most of the more common genotypes in tobacco-growing regions were sampled from both tobacco and

other crops, although some of the genotypes found on weeds and in a non-tobacco-growing region were never collected from tobacco.

In order to understand the evolution of host specialization, and the extent of genetic differentiation and gene flow between tobacco- and non-tobacco-feeding populations, it is necessary to examine genotypes that participate in sexual reproduction. In the present paper we therefore analysed the genetics of populations in mainland Greece containing various proportions of common and 'unique' genotypes. 'Unique' genotypes are those collected only at one site and on one occasion, and are therefore more likely to be the products of recent genetic recombination. We also investigated how the two taxa can retain their host-associated differences, even when reproducing sexually on the same host, by studying their mating behaviour.

MATERIAL AND METHODS

Aphid samples used in microsatellite DNA analysis

The populations investigated were a subset of those used in a previous study (Blackman *et al*., 2007), and comprised 219 clonal lineages originating from peach, tobacco and other crops in tobacco-growing regions of northern, north-eastern and central Greece, and from peach and pepper in a non-tobacco-growing region in eastern central Greece (Fig. 1; Table 1). The life cycle categories of most of these lineages (Table 1) had been tested by Margaritopoulos *et al*. (2002). Because the number of individuals was low in some cases, we pooled samples collected from the same region and host. Samples from crops other than tobacco were also combined (termed here 'other crops'), and for some analyses we also pooled data between years. We included both unique and multicopy genotypes in the analysis, but with only one copy of each multicopy genotype per population, because clonal 'amplification' is not equal over all genotypes, leading to deviations from the Hardy–Weinberg linkage genetic equilibria within populations, and distorted estimates of allele frequencies (Sunnucks *et al*., 1997). This procedure greatly reduced the number of individuals per sample, and it was not possible to incorporate into the analysis samples from tobacco in southern Greece and Italy, because they consisted almost entirely of a few multicopy genotypes.

Allele frequencies and Hardy–Weinberg equilibrium

We examined the data for all seven loci for anomalies and the presence of null alleles using Microchecker version 2.2.3 (Oosterhout *et al*., 2004; http://www.microchecker.hull.ac.uk). A null allele was found to be present at one locus, myz3, in all but three populations, so we decided to exclude it from the analysis. There were also uncertainties about the interpretation of allele sizes at another locus, myz9, so this too was excluded. The final analysis was based on five loci (myz2, myz25, M35, M37, and M40), which are all on different chromosomes (Sloane *et al*., 2001).

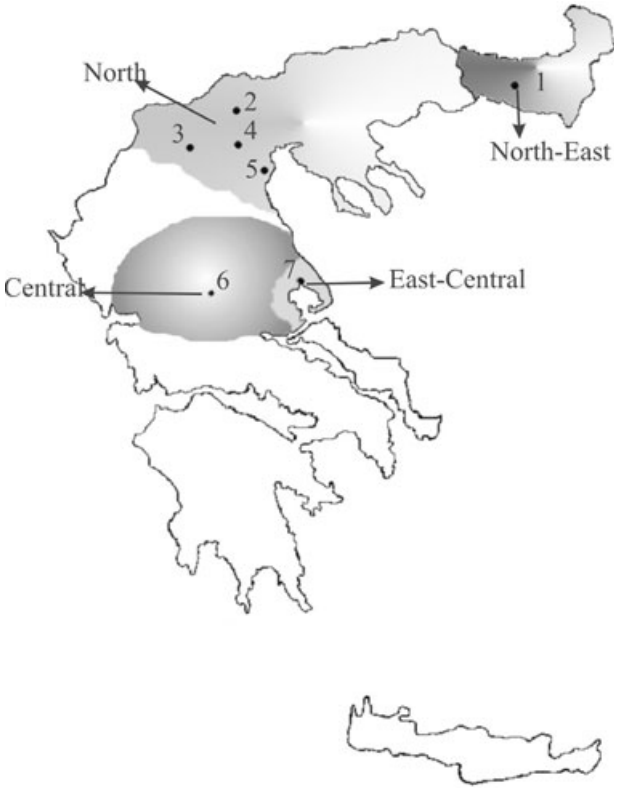

Figure 1. Sampling sites in different regions of mainland Greece. Tobacco-growing regions: Xanthi (1), Aridea (2), Ptolemaida (3), Meliki (4), Katerini (5), and Karditsa (6). Non-tobacco-growing region: Lehonia (7).

Allele frequencies, mean number of alleles per locus, and heterozygosity were calculated using GENEPOP version 3.4 (Raymond & Rousset, 1995a; see also http://genepop.curtin.edu.au). Differences in the average observed heterozygosity over all loci between the populations were examined using the STRUC program, which computes an unbiased estimate of the exact P value of a probability test of homogeneity on $R \times C$ contingency tables using a Markov chain method (Raymond & Rousset, 1995b). To test for deviation from random mating, the 'exact Hardy–Weinberg test' of Guo & Thompson (1992) was applied separately to each locus. A Markov chain method was used for the unbiased estimation of the exact P value of this test, and a global test

Table 1. Observed (H_O) and expected (H_{EXP}) heterozygosity over all loci, mean number of alleles per locus (N_a), and percentage of anholocyclic lineages (An) and cyclical parthenogens (CP) in the nine populations of *Myzus persicae* examined (N_i, number of genotypes used in population genetic analysis as one copy per population)

								Life cycle category*		
Region/Locality	Host plant	Abbr.	Collection date	N_i	N_a	H_O	H_{EXP}	An (%)	CP (%)	NT
E.C. Greece	Other crops	ECC	–	11	6.0	0.673	0.727	80.0	20.0	1
Lehonia	Pepper	–	vi. 1999	–	–	–	–	–	–	–
E.C. Greece	Peach	ECP	–	26	5.0	0.698	0.698	0.0	100	3
Lehonia	–	–	iv–v. 1998, vs. 2000	–	–	–	–	–	–	–
N. Greece	Other crops	NC	–	30	6.6	0.605	0.632	19.2	80.8	4
Katerini	Pepper	–	viii. 1998	2	–	–	–	–	–	–
Meliki	Cabbage	–	ix. 1998	7	–	–	–	–	–	–
Meliki	Pepper	–	vi. 1999	6	–	–	–	–	–	–
Meliki	Potato	–	vi. 1999	10	–	–	–	–	–	–
Ptolemaida	Pepper	–	viii. 1998	5	–	–	–	–	–	–
N. Greece	Peach	NP	–	46	5.8	0.622	0.637	5.3	94.7	4
Aridea	–	–	v. 1998	5	–	–	–	–	–	–
Katerini	–	–	iv–v. 1998	11	–	–	–	–	–	–
Meliki	–	–	iv.–v. 1998–99	30	–	–	–	–	–	–
N. Greece	Tobacco†	NKT	–	21	5.2	0.495	0.549	31.6	68.4	2
Katerini	–	–	v–vi. 1998	–	–	–	–	–	–	–
N. Greece	Tobacco	NMT	–	45	5.6	0.524	0.538	9.3	90.7	2
Meliki	–	–	vi., ix. 1998, vi. 1999	–	–	–	–	–	–	–
N.E. Greece	Tobacco	NET	–	11	4.2	0.618	0.537	100	0	0
Xanthi	–	–	vi. 1999	–	–	–	–	–	–	–
C. Greece	Tobacco	CT	–	13	5.4	0.462	0.624	84.6	15.4	0
Karditsa	–	–	vi. 1998	–	–	–	–	–	–	–
C. Greece	Other crops	CC	–	13	5.8	0.462	0.611	90.0	10.0	3
Karditsa	Pepper	–	vii. 1999	–	–	–	–	–	–	–
N.E. Greece‡	Other crops	NEC	–	3	–	–	–	66.7	33.3	0
Xanthi	Pepper	–	vii. 1999	–	–	–	–	–	–	–

*Percentages of genotypes with different life cycle categories were calculated according to the number of individuals examined (NT = no. individuals not tested). †Three lineages were collected from tobacco seedbeds. ‡These lineages were included in the analysis when a comparison between geographical regions was performed.

across loci was also performed (Raymond & Rousset, 1995b).

Between populations genetic analysis

Population structure was assessed by calculating multilocus F_{ST} values (Weir & Cockerham, 1984) for pairwise comparisons of samples using Arlequin version 2.0 (Schneider, Roessli & Excoffier, 2000). The null distribution of pairwise F_{ST} values under the hypothesis of no difference between the populations is obtained by permuting diploid multilocus genotypes between populations. The P value of the test is the proportion of 100 000 permutations leading to an F_{ST} value larger than or equal to the observed one. The structure of the data was also investigated by analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) using Arlequin version 2.0. A permutation non-parametric approach was used for the significance of fixation indices described in Excoffier *et al.* (1992). Allelic differentiation between populations was examined using GENEPOP version 3.4. An unbiased estimate of the P value of the Fisher exact test was made using a Markov chain method described in Raymond & Rousset (1995b).

To investigate the genetic relationship between populations, an UPGMA tree based on the allele shared distance (DAS; Chakraborty & Jin, 1993) was constructed using the software POPULATIONS version 1.1.28 (O. Langella, 1999–2003, available free at http://bioinformatics.org/~tryphon/populations/). This distance counts the number of different alleles between multilocus genotypes. Bootstrap values were calculated by resampling loci, and are presented as percentages over 1000 replications.

Bayesian clustering and admixture analysis

To test whether our a priori definition of populations (based on the geographical location and the hosts of the sampled individuals) is consistent with genetic information, we tried a Bayesian approach using the program BAPS v. 3.2 (Corander, Marttinen & Mäntyniemi, 2005; see also http://www.abo.fi/fak/mnf/mate/jc/smack_software_eng.html). BAPS estimates posterior probabilities for the number of clusters (K) into which the multilocus genotypes sampled can be split, and determines differences between clusters based on estimated allele frequencies. It also uses prior information about the geographical sampling design. In particular, we performed a run (mixture analysis) with 219 aphid lineages pre-assigned to nine populations (ECC, ECP, NC, NP, NKT, NMT, CT, CC, and NET + NEC; see Table 1). We also performed an admixture analysis to estimate membership coefficients (Q) for all lineages in each population (i.e. the proportion of each genotype conforming to each of the defined K clusters), and also used the program STRUCTURE (Pritchard, Stephens & Donnelly, 2000) as a means of verifying K without prior information about sampling design.

Daily rhythm of sexual activity in mating females and males

Three clonal lineages of *persicae*, two from Lehonia in eastern central Greece (Mp1 and Mp2) and one from southern Italy (Mp3), and two lineages of *nicotianae* from northern Greece (Mpn1 and Mpn2) were used. All lineages originated from peach trees: those from Greece in May 2003 and the one from Italy in April 2002. After cloning they were maintained separately on excised leaves of potato *Solanum tuberosum* (L.) in small perspex boxes under long-day conditions (17 °C and 16-h light : 8-h dark), promoting parthenogenetic reproduction. Designation of their respective taxa was confirmed by multivariate morphometric analysis (Margaritopoulos *et al*., 2003, 2005).

Sexual morphs were produced by rearing the lineages for three generations under short-day conditions (17 °C and 10-h light : 14-h dark; see Margaritopoulos *et al*., 2002 for details). We determined the age at which adult mating females became sexually active by observing them at 09:00 h (onset of photophase), 13:00 h and 17:00 h every day for 14 days after adult ecdysis and recording when they started to exhibit 'signalling' behaviour, i.e. the lifting of abdomen and hind tibiae that indicated release of sex pheromone from tibial scent glands (Pettersson, 1970). The four Greek lineages were used in this experiment, examining a total of 102 mating females of *persicae* and 107 of *nicotianae*.

The daily pattern of signalling behaviour shown by mating females of the five lineages was determined by observing 10–12-day-old adult females every 2 h from 09:00 h (onset of photophase) to 19:00 h (onset of scotophase) for 1 day. A total of 140 females of *persicae* (31–64 per lineage) and 100 females of *nicotianae* (48–52 per lineage) were observed. In addition, the daily activity of males of the same five lineages was recorded, i.e. whether they were 'searching' or 'stationary'. Batches of between three and five adult males 2–5-days old were caged on young potato plants in a separate chamber from females, and their activity was recorded every 2 h from 09:00 h to 19:00 h. A total of 143 males of *persicae* (31–59 per lineage) and 104 (50–54 per lineage) of *nicotianae* were examined.

Olfactory response of males to sex pheromone

The olfactory responses of males of the four Greek lineages to sex pheromones were examined under controlled conditions in a choice chamber (16-cm long, 8-cm wide, and 8-cm high), with openings of 1 cm in diameter at opposite ends covered with carbon filters. A small fan was positioned centrally to aid air flow, and lighting was from overhead. In choice tests, one potato leaf with 10–12-day-old adult mating females of *persicae* and another with similar females of *nicotianae* were placed at opposite ends of the arena. Thirty 2–5-day-old adult males were released in the centre of the arena, testing one taxon at a time. Observations were made every hour during the 10-h photophase. Males observed near mating females or in any part of the arena outside a 2-cm-wide central zone were assumed to be responding to the odour of females. Four combinations of males and mating females of the four Greek lineages (two *persicae* and two *nicotianae*) were tested. Non-choice tests were also performed using only mating females from one lineage and by leaving the opposite side of the arena empty. Three replicate tests of each combination were conducted, observing 90 males in total for each combination.

Differences in 'signalling' behaviour of females and in male activity and olfactory response between the two taxa were analysed using the χ^2 test. When the χ^2 was significant in comparisons involving more than two lineages, then pairwise comparisons were conducted.

RESULTS

Genetic diversity within populations

A substantial intrapopulation genetic variation was observed. All five loci were polymorphic, with the number of alleles detected ranging from 7 to 11 (mean = 8.6 alleles per locus overall, range per population 4.2–6.6). Mean observed heterozygosity over all loci (Table 1)

Table 2. Single and multilocus probability tests for deviations from Hardy–Weinberg equilibrium and F_{IS} inbreeding coefficient values observed in nine populations of *Myzus persicae*. 'Crop' means field crop other than tobacco (*N*, number of clones tested; CP, cyclical parthenogens)

Population*	Host plant	*N*	myz2	myz25	M35	M37	M40	All	CP percentage
ECC	crop	11	+0.103	+0.194	+0.191	–0.111	+0.042	+0.075	19.2
			NS/NS†	NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	–
ECP	peach	26	–0.005	–0.198	+0.175	–0.013	+0.011	+0.001	100
			NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	–
NC	crop	30	–0.125	+0.105	+0.149	+0.067	–0.015	+0.042	80.8
			NS/NS	NS/NS	0.008/NS	NS/NS	NS/NS	0.02/NS	–
NP	peach	46	–0.003	+0.055	+0.075	+0.024	–0.046	+0.024	94.7
			NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	–
NKT	tobacco	21	+0.126	–0.026	+0.149	+0.067	+0.027	+0.097	68.4
			NS/NS	NS/NS	0.03/NS	NS/NS	NS/NS	0.02/NS	–
NMT	tobacco	45	–0.014	–0.028	–0.011	+0.068	+0.111	+0.26	90.7
			NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	–
NET	tobacco	11	–0.022	–0.250	–0.209	–0.163	+0.000	–0.151	0.0
			NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	NS/0.03	–
CT	tobacco	13	+0.191	+0.333	+0.336	+0.244	+0.209	+0.260	15.4
			NS/NS	0.002/NS	0.004/NS	0.05/NS	NS/NS	0.001/NS	–
CC	crop	13	+0.247	+0.442	+0.382	–0.043	+0.200	+0.245	10.0
			NS/NS	0.004/NS	0.003/NS	NS/NS	NS/NS	0.008/NS	–
NET + NEC		14	–0.049	–0.182	–0.197	–0.194	–0.040	–0.153	7.1
			NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	NS/0.01	–

*Abbreviations are defined in Table 1. †$P_{Het.Deficit}/P_{Het.Excess}$, NS = non significant.

ranged from 0.462 to 0.698, with no significant differences between populations ($P = 0.918$). There was no clear correlation between mean observed heterozygosity and proportion of cyclical parthenogenetic genotypes in each population (Pearson's correlation: $R = 0.303$, $N = 9$, $P = 0.429$).

A few cases (7 out of 50 tests) of single-locus significant deviations from Hardy–Weinberg equilibrium indicating heterozygote deficiency were observed, mainly at loci myz25 and M35 (Table 2). Significant multilocus deviation indicating heterozygote deficiency was observed in four populations, and heterozygote excess was indicated for one population, with all five being from tobacco or other secondary crops (Table 2). Populations with heterozygote deficiency included ones with both high and low proportions of cyclical parthenogens.

Genetic differentiation between populations

The multilocus F_{ST} revealed important interregional variation (Table 3). There was no significant difference between years in the same locality or region (data not shown), thereby justifying the pooling of samples between years in the same locality to increase sample sizes. The populations from peach (ECP) and pepper (ECC) in the eastern part of the central region, where tobacco is not cultivated, did not differ significantly, but the pairwise comparisons with populations from tobacco all yielded highly significant ($P < 0.001$) and mostly high F_{ST} values. There were also significant differences between these eastern central populations and those on peaches and other crops in tobacco-growing regions, but F_{ST} values were generally lower (Table 3). Spatial population subdivision was also apparent in global (all-locus) tests for allelic differentiation (Table 3). In addition, most (102) of the 180 single-locus pairwise comparisons between regions were significant ($P < 0.05$). Among the populations from tobacco-growing regions there was a significant genetic differentiation between populations with a high percentage of cyclical parthenogens (NP, NC, NKT and NMT; see Table 1) and those consisting mostly of anholocyclic lineages (CC, CT and NET; Table 3). Notably, despite the strong regional differentiation shown by most populations, the F_{ST} value between the widely separated north-eastern and central populations was insignificant (Table 3). A similar pattern was observed when samples from all herbaceous crops in each region were pooled (data not shown).

The hierarchical AMOVA also revealed a significant F_{ST} (0.062; $P < 0.001$) over all loci and populations, with most of the variance (93.8%) being within populations. Two of the five loci (myz25 and M40) explained a high percentage of the between-population variance (14.7 and 12.9%, respectively) with F_{ST} values of 0.147

Table 3. Analysis of genetic differentiation between nine *Myzus persicae* populations (clones from the same region and crop but sampled in different years were pooled together). Below the diagonal the multilocus F_{ST} values and probabilities for the absence of differences between populations (NS, non significant, $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$) are given. Above the diagonal the probabilities for the absence of allelic differentiation between populations are given

	ECC $N = 11$	ECP $N = 26$	NC $N = 30$	NP $N = 46$	NKT $N = 21$	NMT $N = 45$	NET $N = 13$	CT $N = 13$	CC $N = 11$
ECC	–	NS	***	***	***	***	***	***	***
ECP	–0.004 NS	–	***	***	***	***	***	***	***
NC	0.069 ***	0.072 ***	–	NS	NS	*	***	**	***
NP	0.056 ***	0.051 ***	0.005 NS	–	*	***	***	***	***
NKT	0.142 ***	0.132 ***	0.010 NS	0.026 **	–	NS	**	**	***
NMT	0.157 ***	0.152 ***	0.018 **	0.040 ***	–0.012 NS	–	***	***	***
NET	0.174 ***	0.159 ***	0.071 ***	0.080 ***	0.051 **	0.059 ***	–	NS	NS
CT	0.134 ***	0.133 ***	0.030 *	0.057 ***	0.033 *	0.051 ***	0.016 NS	–	NS
CC	0.128 ***	0.127 ***	0.054 ***	0.080 ***	0.064 ***	0.077 ***	0.018 NS	–0.005 NS	–

Abbreviations for populations are defined in Table 1. N = number of clones examined.

($P < 0.001$) and 0.129 ($P < 0.001$), respectively, suggesting that there might be some linkage between these two loci and either mode of reproduction or host-plant adaptation properties.

The UPGMA tree (Fig. 2) based on DAS separated the two populations from eastern central Greece, where tobacco is not cultivated, from those collected either from tobacco or other crops and peach in tobacco-growing regions. Within the cluster of populations from tobacco-growing regions, the three populations (CC, CT, and NET) with a high percentage of anholocyclic lineages were separated from populations from north Greece consisting mainly of cyclical parthenogens. A similar pattern was observed when the samples from tobacco and other crops in each region were pooled, and also when three additional lineages from pepper from north-eastern Greece were included (tree not shown).

Bayesian clustering and admixture analyses

Of all possible combinations of 'sampling units' (samples from the same host and locality), the Bayesian analysis of population structure revealed the highest posterior probability ($P = 0.9999$) for a structure with three clusters ($K = 3$). Cluster I contained the samples from pepper (ECC) and peach (ECP) in the non-tobacco-growing eastern central region. Cluster II contained the samples from peach (NP), other crops (NC), and tobacco (NTK and NTM) in the north, with all these having a high percentage of cyclically parthenogenetic genotypes. Cluster III comprised the samples from the north-eastern region from tobacco and pepper (NET + NEC) and those from central Greece from tobacco (CTK) and pepper (CCK), i.e. all populations with a high percentage of anholocyclic genotypes.

Independent runs were also carried out for the total data set and values of K between 1 and 10 using the program STRUCTURE (Pritchard *et al*., 2000). The admixture and uncorrelated allele frequencies models were used without any prior population information. Following the pointers for choosing K provided by Pritchard *et al*. (2000) and Garnier *et al*. (2004), the best solution for K proved to be 3. There was a sharp increase of the posterior probability of the data with a K of 1–3, where a plateau was reached (for $K = 3$–7; results not shown). In such cases, the smallest K value should be taken as the correct one in order to avoid unjustified and less informative oversplitting.

A second analysis was performed excluding the anholocyclic central and north-eastern populations, so that it focused on cyclically parthenogenetic genotypes. BAPS derived two clusters with the same grouping of populations as in the first analysis. Cluster I comprised the eastern central populations, and Cluster II those in the north. The admixture analysis (Fig. 3; Table 4) showed that none of the clusters

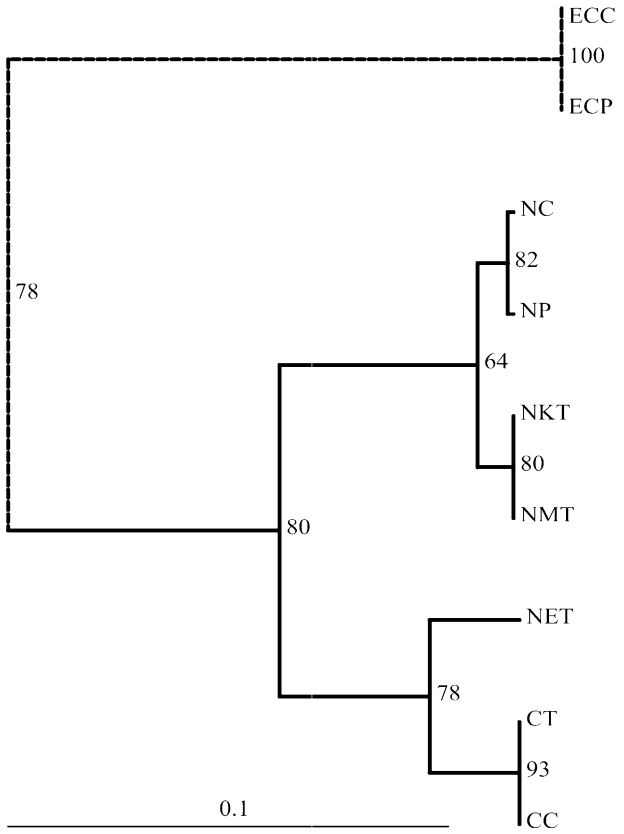

Figure 2. UPGMA dendrogram based on shared allele distances among nine *Myzus persicae* populations (clones from the same region and crop sampled in different years were pooled together). Numbers denote bootstrap percentages (from 1000 resamplings). Abbreviations are defined in Table 1.

was totally pure, and some individuals had a low membership coefficient in their original cluster. In particular the northern populations on peach and non-tobacco crops had many individuals with high cluster-I coefficients.

Daily rhythm of sexual activity in mating females and males

In both taxa the percentage of 'signalling' females increased as the adult mating females aged up to day 11–12, and then remained relatively constant (Fig. 4). However, mating females of *persicae* became sexually active more quickly than those of *nicotianae*. The daily pattern of signalling behaviour of 10–12-day-old mating females differed considerably between taxa (Fig. 5A). Females of *persicae* showed a peak of signalling behaviour at 2 h after the start of the photophase, and those of *nicotianae* showed a peak at 6 h. The differences between taxa were significant at both peaks ($\chi^2 = 17.0$, d.f. = 4, $P < 0.002$ and $\chi^2 = 27.1$, d.f. = 4, $P < 0.001$, respectively). Pairwise comparisons between lineages of the two taxa at 2 and 6 h were also all significant. The females of the two *nicotianae* lineages were also significantly more active than those of *persicae* ($\chi^2 = 10.4$, d.f. = 4, $P < 0.03$) as soon as the lights came on (Fig. 5A). In both taxa, the daily pattern of male activity was remarkably similar to that of females, although the males were isolated from them in a separate chamber (Fig. 5b). In the two *persicae* lineages from Greece, however, male activity declined more gradually after the peak compared with that of their own females. Male activity after 2 h of light was significantly higher in all three lineages of *persicae*

Table 4. Average membership coefficients (Q) for the clusters derived from the Bayesian clustering and admixture analyses, and for the populations of *Myzus persicae* examined

Analysis with all nine populations surveyed				Analysis with six populations consisting mos tly of cyclical parthenogens		
	Q_I	Q_{II}	Q_{III}		Q_I	Q_{II}
Cluster I	0.81	0.11	0.08	Cluster I	O.82	0.18
ECC	0.87	0.02	0.11	ECC	0.90	0.10
ECP	0.78	0.15	0.07	ECP	0.82	0.20
Cluster II	0.15	0.69	0.16	Cluster II	0.15	0.85
NC	0.23	0.59	0.18	NC	0.22	0.78
NP	0.23	0.65	0.12	NP	0.24	0.76
NKT	0.07	0.72	0.21	NKT	0.07	0.93
NMT	0.04	0.79	0.17	NMT	0.05	0.95
Cluster III	0.07	0.17	0.76			
NET + NEC	0.02	0.20	0.78			
CT	0.08	0.23	0.69			
CC	0.10	0.09	0.82			

Abbreviations are defined in Table 1.

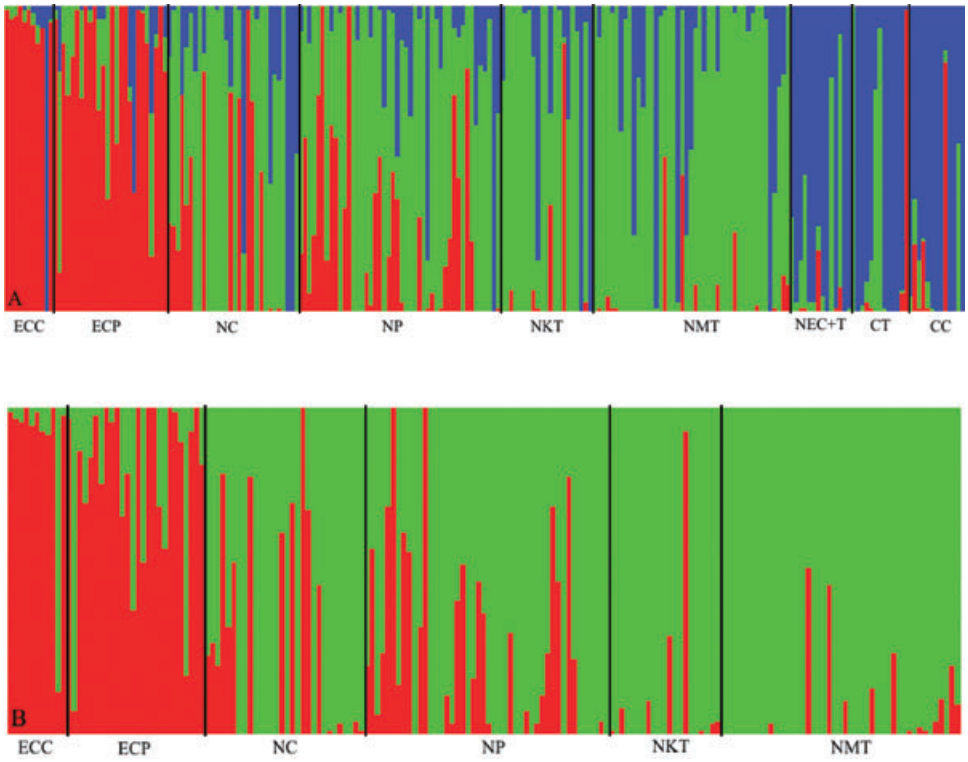

Figure 3. Admixture clustering plots of Greek *Myzus persicae* populations. (A) Nine populations; number of clusters, $K = 3$. (B) Six populations after omitting three anholocyclic populations from tobacco; $K = 2$. Each aphid lineage is represented as a vertical bar partitioned into K segments, the lengths of which are proportional to the estimated membership coefficients of the lineage in each of the K clusters. Lineages of different populations are separated by black lines. For average cluster coefficients of each population see Table 4. Abbreviations are defined in Table 1.

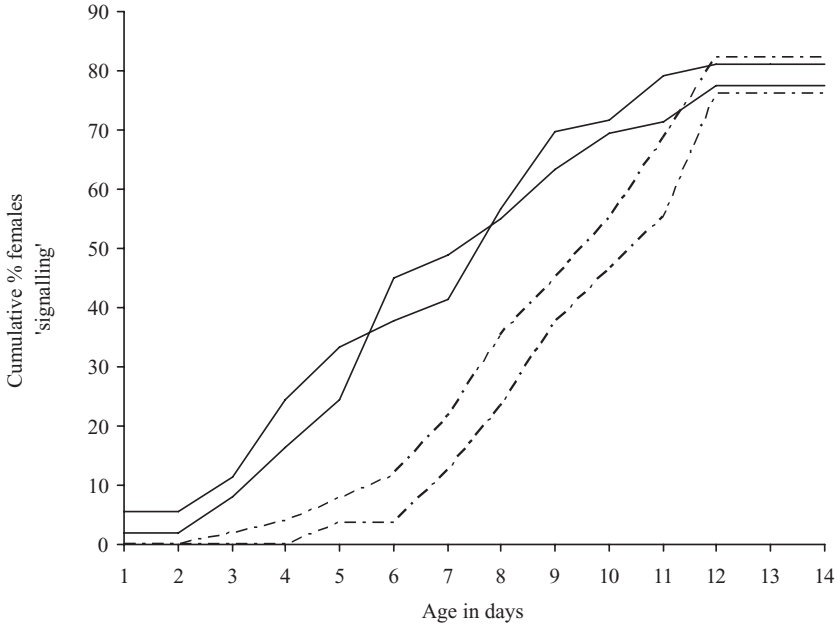

Figure 4. Cumulative percentage (%) of sexual females of two clonal lineages of *persicae* (solid lines) from eastern central Greece and two of *nicotianae* (dashed lines) from northern Greece 'signalling', in relation to their age.

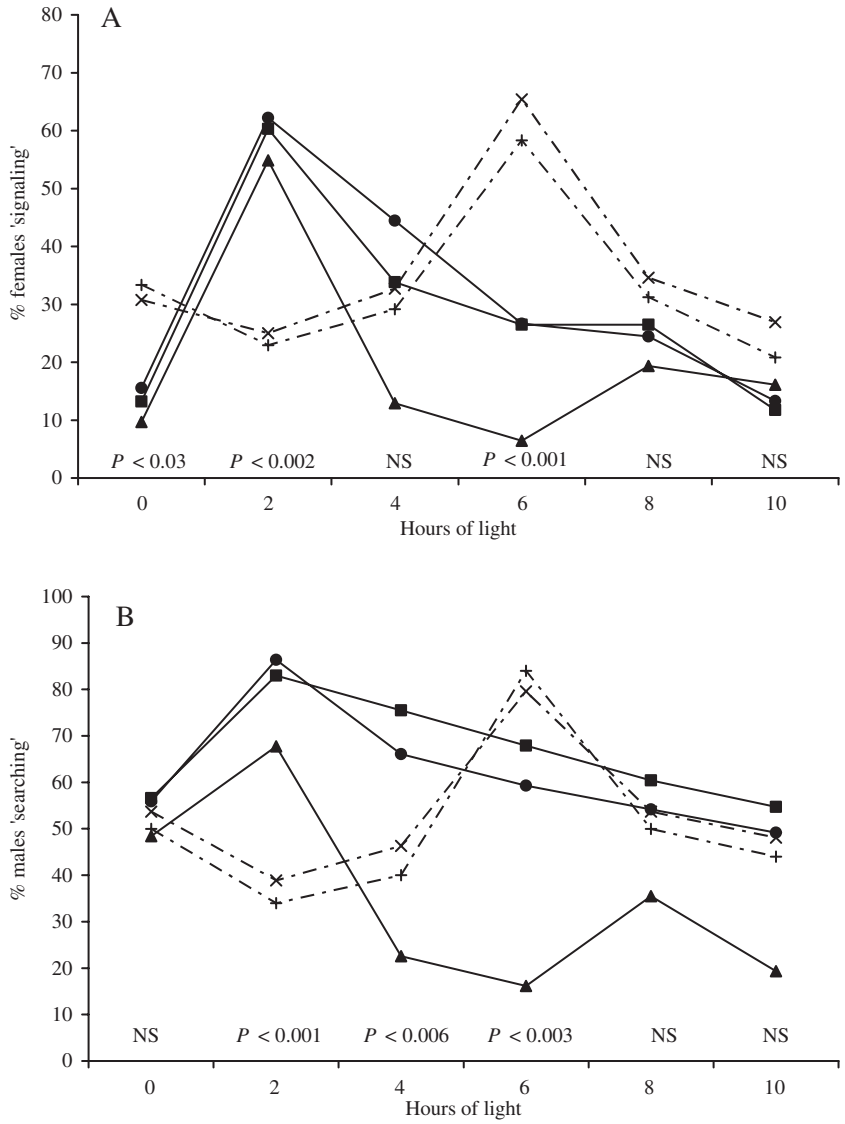

Figure 5. Percentage of (A) mating females 'signalling' and (B) males 'searching' during the day, in *persicae* (two lineages from eastern central Greece, ■ and ●; one from southern Italy, ▲) and *nicotianae* (two lineages from northern Greece, + and ×).

than in the two *nicotianae* ($\chi^2 = 20.5$, d = 4, $P < 0.001$; pairwise comparisons between lineages of the two taxa were also all significant). Significant differences were also observed at 4 h ($\chi^2 = 14.4$, d.f. = 4, $P < 0.006$) and 6 h ($\chi^2 = 16.3$, d.f. = 4, $P < 0.003$), with the two *nicotianae* lineages both peaking at 6 h, but only the Italian *persicae* males were significantly less active than *nicotianae* males at these times, because of the prolonged period of male activity in the two *persicae* lineages from Greece.

Olfactory response of males to sex pheromone

In non-choice tests, where the responses of males were observed in the presence only of mating females of their own or of the other taxon, males were attracted to the pheromones of both subspecies, but the percentage attracted by females of the wrong subspecies was significantly lower in two of the four combinations (Table 5). In most choice tests, males were attracted similarly by pheromones from their own and the other subspecies at the time of peak signalling of the females (Table 6). Only males of *nicotianae* lineage Mpn1 were attracted significantly more by females of their own subspecies than by females of the other subspecies ($\chi^2 = 14.4$, d.f. = 1, $P < 0.001$). In general, the responses of males under the confined conditions of the choice chamber overrode the innate daily rhythm of activity that they showed in isolation. Some evi-

Table 5. Percentage of male *persicae* and *nicotianae* responding to the odour of mating females of their own and the other taxon in non-choice tests

	% of *persicae* males responding		% of *nicotianae* males responding	
At 2 h after onset of photophase*	83.3 (Mp1♂–Mp1♀)	80.0 (Mp2♂–Mp2♀)	70.0 (Mpn1♂–Mp1♀)	56.7 (Mpn2♂–Mp2♀)
At 6 h after onset of photophase†	63.3 (Mp1♂–Mpn1♀) $\chi_1^2 = 9.2$, $P < 0.002$	83.3 (Mp2♂–Mpn2♀) $\chi_1^2 = 0.33$, $P = 0.563$	76.7 (Mpn1♂–Mpn1♀) $\chi_1^2 = 1.0$, $P = 0.312$	83.3 (Mpn2♂–Mpn2♀) $\chi_1^2 = 19.9$, $P < 0.001$

*Peak of signalling of *persicae* females. †Peak of signalling of *nicotianae* females.

Table 6. Mean percentages of males of *persicae* and *nicotianae* that responded to the odour of mating females of their own subspecies and the other subspecies in choice tests

	% of *persicae* males responding				% of *nicotianae* males responding			
	Mp1♂		Mp2♂		Mpn1♂		Mpn2♂	
	Mpn1♀	Mp1♀	Mpn2♀	Mp2♀	Mpn1♀	Mp1♀	Mpn2♀	Mp2♀
At 2 h after the onset of photophase*	10.0	70.0	6.7	70.0	17.0	66.7	10	57.0
At 6 h after the onset of photophase†	63.3	3.3	63.3	3.3	90.0	10.0	66.7	3.3
Mean percentage at off-peak times	20.4a	48.9b	22.2a	48.9b	54.4a	30.4b	42.2a	24b

*Peak of the signalling of *Myzus persicae females*. †Peak of the signalling of *nicotianae* females. Percentages followed by a different letter (a or b) differed significantly by a χ^2 test ($P < 0.05$).

dence was obtained that males discriminate between the pheromones of their own and the other taxon by excluding observations at the time of peak signalling by females of both taxa (i.e. including only data obtained when presumably similar numbers of mating females of both taxa were releasing pheromone). A significantly higher mean percentage of males then responded to the females of their own taxon rather than the other taxon (Table 6).

DISCUSSION

The present study has revealed that populations of *M. persicae* have a broadly heterogeneous genetic structure in mainland Greece, as evidenced by the high allelic differentiation and the mostly significant F_{ST} among populations. The observed spatial genetic structuring can be attributed to two main factors: host specialization and variation in life cycle category.

Genetic differentiation between populations with different host specialization

Samples from eastern central Greece, a region where tobacco is not cultivated, showed a strong genetic differentiation from populations in tobacco-growing regions, with F_{ST} values generally almost twofold higher than those obtained for any other comparisons between populations, irrespective of their mode of reproduction. This corroborated the results of previous morphometric studies (Margaritopoulos *et al*., 2000, 2003), which found that tobacco-adapted populations (*M. persicae nicotianae*) differ genomically from those of generalist populations (*M. persicae s.s*.).

Supporting evidence that the two taxa differ genomically is provided by the Bayesian analysis, which derived three clusters: cluster I corresponding to *persicae* lineages, and clusters II and III corresponding to *nicotianae* lineages. The *nicotianae* lineages were split into two clusters according to their life cycle category, with cluster III comprising populations with predominantly anholocyclic genotypes. The admixture plot (Fig. 3) shows the presence of some pure cluster-III (anholocyclic *nicotianae*) genotypes in all populations except in those on peach. When the analysis was performed without the central and north-eastern populations, so that it focused on cyclically parthenogenetic genotypes, this resulted in a two-cluster structure (cluster I = *persicae*; cluster II = *nicotianae*).

Populations on peach and non-tobacco crops within the northern tobacco-growing region (NP and NC) showed least conformity with the defined *K* clusters, containing in particular many genotypes with high cluster-I membership coefficients, which are likely to

be *persicae* rather than *nicotianae* (including some with pure cluster-I genotypes). The lack of purity of many genotypes in these and other populations could be at least partly caused by hybridization between taxa, and also between anholocyclic and cyclically parthenogenetic genotypes, both of which are certainly known to occur. However, it may simply be a reflection of the close genetic relationship between the two taxa and/or the limitations of the analysis.

Most of the genetic differences between tobacco and other herbaceous crops in the same region were not significant. This could be explained by the fact that tobacco aphids have the ability to colonize other plants within the vast host range of *M. persicae* (Semtner, Tilson & Dara, 1998; Nikolakakis *et al.*, 2003). All of the common (anholocyclic) genotypes found on tobacco in Greece and southern Italy were also found on other host plants in tobacco-growing regions (Blackman *et al.*, 2007). Tobacco aphids colonize the growing points and young leaves of tobacco, and can build up large dense colonies, whereas on other crops, such as potato, pepper, sugar beet *Beta vulgaris* L., and brassicas, individual *M. persicae* are found dispersed sparsely on the older leaves. The economic importance of this species on most crops results from its efficiency as a vector of plant viruses (Blackman & Eastop, 2000). In tobacco-growing regions, tobacco-adapted aphids will therefore massively predominate, and are likely to be collected not only on tobacco but also on other crops. In northern Greece, where tobacco and peaches are grown in close proximity, the winged migrants going to and from peach will also predominantly be tobacco aphids. Host-associated trade-offs that reduce the fitness of tobacco aphids on other crops, along with selection against cross-host winged migrants and their subsequent generations (Nikolakakis *et al.*, 2003; Margaritopoulos *et al.*, 2005), could explain the absence of tobacco-adapted genotypes from regions where tobacco is not grown.

Genetic differentiation between populations with different reproductive strategy

Apart from the differentiation between populations of the two taxa, genetic structuring (significant allelic differentiation and moderate but significant F_{ST} values, mostly between 0.05 and 0.08) was observed between the samples from the tobacco-growing regions, with a high percentage of cyclical parthenogenetic genotypes (northern Greece) and of those (from central and north-eastern Greece) consisting mostly of anholocyclic clones. These two sample groups also formed two separate clusters in the Bayesian analysis. In several aphid species the mode of reproduction has been shown to be responsible for the genetic structuring of populations. Some of these studies (e.g. Simon *et al.*, 1999; Delmotte *et al.*, 2002; Guillemaud *et al.*, 2003) found that the distribution of anholocyclic genotypes is affected by the winter temperature (the 'climatic' selection hypothesis), as predicted by models (Rispe & Pierre, 1998; Rispe *et al.*, 1998). In mainland Greece, however, the occurrence of anholocylic genotypes is related to the absence of the primary host, peach (Margaritopoulos *et al.*, 2002; Blackman *et al.*, 2007).

Heterozygote deficiencies were common in most populations irrespective of their predominant reproductive strategy. Previous studies have found populations of aphids with cyclical parthenogenesis to have an excess of homozygotes (Simon *et al.*, 1999; Delmotte *et al.*, 2002; Guillemaud *et al.*, 2003; Vorburger *et al.*, 2003), although Wilson *et al.* (2002) found generally no deviation from the Hardy–Weinberg equilibrium in sexual populations from peach in Australia. Various explanations have been proposed for homozygote excess, such as null alleles, Wahlund effect, inbreeding, and selection. No evidence of null alleles was found at the five loci used in the present study. This situation could simply reflect a Wahlund effect of sampling from distinct gene pools in the same population, as revealed by the Bayesian analysis. For instance, almost all samples contained lineages that were assigned to Cluster III. Further research is needed to clarify whether inbreeding or selection are involved. The finding that anholocyclic lineages were generally more heterozygote deficient than cyclical parthenogens is contrary to expectation, as parthenogenetic lineages are generally expected to accumulate heterozygosity with time by mutation (Birky, 1996; Simon *et al.*, 1999; Delmotte *et al.*, 2002). Many of the anholocyclic lineages in central Greece may have been generated recently by inheritance of the anholocyclic trait through the sexual phase, as suggested by a previous study (Margaritopoulos *et al.*, 2002), and by the results of F_{ST} and Bayesian analyses reported here. Such 'contagious parthenogenesis' has been demonstrated for *Sitobion avenae* (Simon *et al.*, 1999) and *Rhopalosiphum padi* (Delmotte *et al.*, 2001). Only the north-eastern anholocyclic populations at Xanthi showed heterozygote excess, suggesting that they included more ancient lineages.

Migration and selection

The F_{ST} values between populations (Table 3), as well as the estimated overall value (0.062), are among the highest obtained for aphids using microsatellites, and suggest strong genetic differentiation even over small distances (Lehonia and Karditsa are less than 100 km apart, see Fig. 1). For *S. avenae,* Simon *et al.* (1999) obtained an average F_{ST} value of 0.032 in France, and

in the UK, Llewellyn *et al*. (2003) reported most values lower than 0.05. For *R. padi*, Delmotte *et al*. (2002) reported 0.022 and 0.032 for anholocyclic and cyclically parthenogenetic genotypes, respectively. The authors suggested that genetic homogeneity over a large geographical scale results from the high migratory capabilities of these aphids. However, *M. persicae* seems to differ in this respect, as shown also by previous studies in Australia (Wilson *et al*., 2002: F_{ST} = 0.058–0.202, with an average 0.087) and France (Guillemaud *et al*., 2003: F_{ST} up to 0.17–0.21 in some cases). This could mean that long-distance migration is not so common in *M. persicae*, or that it has a low success rate (Loxdale *et al*., 1993). Nevertheless, Vorburger *et al*. (2003) found substantial migration among Australian populations, and suggested that local selection of genotypes investing in different modes of reproduction might explain the differences in regional distribution. The occurrence of common genotypes in various regions of mainland Greece, some of which have also been found in southern Italy (Blackman *et al*., 2007), shows that some long distance migration does occur, but the genetic stability of populations in the same locality from year to year suggests that most of the population does not move far. Local selection for specific genotypes with respect to host plant or to mode of reproduction seems likely to be the main reason for genetic differentiation between populations of *persicae* and *nicotianae*, or between *nicotianae* populations investing in different reproduction strategies. Supporting evidence for local selection is the fact that the anholocyclic ('common') genotypes sampled from pepper in the eastern central region were not found anywhere in the tobacco-growing regions surveyed (Blackman *et al*., 2007).

Host specialization and the evolution of reproductive isolation

The present study has clearly revealed genetic differentiation and reduced gene flow between two host-related groups within *M. persicae*, providing further justification for their designation as host-related subspecific taxa (Drès & Mallet, 2001). This and previous papers (Nikolakakis *et al.*, 2003; Blackman *et al*., 2007) have shown that both cyclically parthenogenetic and anholocyclic genotypes of *M. persicae* exhibit host specialization. This conflicts somewhat with theories that anholocylic genotypes may either evolve ecological specialization more easily than their sexual counterparts (Hales *et al*., 1997) or be selected to have 'general purpose genotypes' (Lynch, 1984).

How do the two taxa retain their host-associated differences in regions where both have the ability to reproduce sexually on peach, and where there is undoubtedly some degree of interbreeding? In closely related aphid taxa that mate on the same host, two cases have been reported where interbreeding is restricted by divergence in the mate recognition system (Guldemond & Dixon, 1994; Thieme & Dixon, 1996). In the present study we found differences between the two taxa in the daily pattern of sexual activity of both males and mating females. Similar results were obtained for more than one genotype per taxon, and the same pattern was observed in *persicae* lineages from both Greece and Italy. The timing of male activity matched that of mating females, but male activity was more prolonged in the two Greek lineages of *persicae*, which may increase the probability of gene flow between the two taxa. This has also been observed in *Cryptomyzus* (Guldemond & Dixon, 1994) and in other insect species, where it has been attributed to high competition for mates among males (Thornhill & Alcock, 1983). The divergence of sexual activity rhythms did not differ as much as they did in closely related *Cryptomyzus* species (Guldemond & Dixon, 1994), but was comparable with that reported between subspecies of *A. fabae* (Thieme & Dixon, 1996). This and other evidence discussed previously indicates that reproductive isolation between *nicotianae* and *persicae* is not complete, and that we are currently observing the evolution of this phenomenon. This may explain why males are attracted, although somewhat differentially, by the sexual pheromone released by the mating females of both their own subspecies and other subspecies. However, specificity of sex pheromones varies between species. For instance, males of *S. avenae* respond to the sex pheromone released by both conspecific mating females and those of *S. fragariae* (Lilley & Hardie, 1996). It seems that the subspecies are evolving a mechanism for host-independent assortative mating, which complements the observed divergent ecological selection. This could be a case of reinforcement (Drès & Mallet, 2001; Schluter, 2001; Turelli, Barton & Coyne, 2001), but confirmation of this would require evidence of *persicae* × *nicotianae* hybrid dysfunction.

It seems most likely that the combination of properties involved in adaptation to tobacco arose originally in a genetically recombining population, with the tobacco-adapted genome quickly achieving dominance in a tobacco-growing area. Then it could have spread to more tropical regions (probably in Southeast Asia where it was first reported as a tobacco pest; de Jong, 1929), where it would have become further established as large permanently parthenogenetic populations genetically isolated from those of *M. persicae s.s*., before reaching other parts of the world. The two taxa may have had a period of geographical isolation, but not enough for the completion of reproductive isolation. Tobacco was first cultivated in Greece in about 1600 (Pouqueville, 1820–21), but

the appearance of tobacco aphid as a pest in Greece dates only from the late 1970s (Tobacco Institute of Greece, pers. comm.), and it is only in the late 1980s that it was first observed to have a sexual phase on peach. It is unclear whether behavioural mechanisms reducing gene flow between the two subspecies were already developed when cyclically parthenogenetic populations came into contact again in Greece, or whether this has happened subsequently. Examination of the sexual activity rhythms of more lineages of both taxa from other countries, along with hybridization experiments, should aid our understanding of the evolution of these isolating mechanisms. Linkage analysis and mapping of the genes affecting parameters of host performance and preference of *M. persicae nicotianae* could give further insights into the origins of this subspecies, as well as clues to its possible future course of evolution.

ACKNOWLEDGEMENTS

This work was supported by the Commission of the European Communities Tobacco Information and Research Fund, project 96/T/18 'Management of the insect pests and viruses of tobacco using ecologically compatible technologies'. We are grateful to Alex Wilson, who did some preliminary analyses of the microsatellite data and commented on an earlier version of the manuscript, and to Dimitra Zagakou and Sotiris Sotiriou for help on behavioural experiments.

REFERENCES

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

Birky CW. 1996. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* **144:** 427–437.

Blackman RL. 1987. Morphological discrimination of a tobacco feeding form from Myzus persicae (Sulzer) (Hemiptera: Aphididae), and a key to New World Myzus (Nectarosiphon) species. *Bulletin of Entomological Research* **77:** 713–730.

Blackman RL. 1990. Specificity in aphid/plant genetic interactions, with particular attention to the role of the alate colonizer. In: Campbell RK, Eikenbary RD, eds. *Aphid–Plant Genotype Interactions*. New York: Elsevier, 251–274.

Blackman RL, Eastop VF. 2000. *Aphids on the World's Crops An Identification and Formation Guide*, 2nd edn. London: John Wiley and Sons.

Blackman RL, Malarky G, Margaritopoulos JT, Tsitsipis JA. 2007. Distribution of common genotypes of *Myzus persicae* in Greece, in relation to life cycle and host plant. *Bulletin of Entomological Research* **97:** 253–263.

Brain CK. 1942. The tobacco aphid. *Rhodesian Agricultural Journal* **39:** 241–243.

Chakraborty R, Jin L. 1993. A unified approach to the study of hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ, eds. *DNA Fingerprinting: State of the Science.* Basel, Switzerland: Birkhauser, 153–175.

Claridge MF. 1988. Species concepts and speciation in parasites. In: Hawksworth DL, ed. *Prospects in Systematics*. Oxford: Systematics Association: Clarendon Press, 92–111.

Clements KM, Sorenson CE, Wiegmann BM, Neese PA, Roe RM. 2000a. Genetic, biochemical and behavioral uniformity among populations of *Myzus nicotianae and Myzus persicae*. *Entomologia Experimentalis et Applicata* **95:** 269–281.

Clements KM, Wiegmann BM, Sorenson CE, Smith CF, Neese PA, Roe RM. 2000b. Genetic variation in the *Myzus persicae* complex (Homoptera: Aphididae): evidence for a single species. *Annals of the Entomological Society of America* **93:** 31–46.

Corander J, Marttinen P, Mäntyniemi S. 2005. Bayesian Identification of Stock Mixtures from Molecular Marker Data. *Fishery Bulletin* **104:** 550–558.

De Jong JK. 1929. Enkele resultaten von het onderzoek naar de biologie von de tabaksluis *Myzus persicae* Sulzer. *Bulletin Van Het Deli Proefstation Te Medan* **28:** 1–36.

Delmotte F, Leterme N, Bonhomme J, Rispe C, Simon JC. 2001. Multiple routes to asexuality in an aphid species. *Proceedings of the Royal Society Series B* **268:** 1–9.

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. *Molecular Ecology* **11:** 711–723.

Douglas AE. 1997. Provenance, experience and host plant affiliation in the polyphagous aphid, *Aphis fabae*. *Entomologia Experimentalis et Applicata* **83:** 161–170.

Drès M, Mallet J. 2001. Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of Royal Society of London Series B* **357:** 471–492.

Eastop VF. 1973. Deductions from the present day host plants of aphids and related insects. In: van Emdem HF, ed. *Insect Plant Relationships. Symposium of Royal Entomological Society of London* 6: 157–178.

Eastop VF, Blackman RL. 2005. Some new synonyms in Aphididae (Hemiptera). *Zootaxa* **1089:** 1–36.

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

Feder JL, Roethele JB, Filchak K, Niedbalski J, Romero-Severson J. 2003. Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* **163:** 939–953.

Field LM, Javed N, Stribley MF, Devonshire AL. 1994. The peach-potato aphid *Myzus persicae* and the tobacco aphid *Myzus nicotianae* have the same esterase-based mechanisms of insecticide resistance. *Insect Molecular Biology* **3:** 143–148.

Futuyama DJ, Mayer GC. 1980. Non-allopatric speciation in animals. *Systematic Zoology* **29:** 254–271.

Garnier S, Alibert P, Audiot P, Prieur B, Rasplus J-Y. 2004. Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, *Carabus solieri*. *Molecular Ecology* **13:** 1883–1897.

Guillemaud T, Mieuzet L, Simon J-C. 2003. Spatial and temporal genetic variability in French populations of the peach-potato aphid, *Myzus persicae*. *Heredity* **91:** 143–152.

Guldemond JA. 1990. Evolutionary genetics of the aphid *Cryptomyzus,* with a preliminary analysis of the inheritance of host preference, reproductive performance and host-alternation. *Entomologia Experimentalis et Applicata* **57:** 65–76.

Guldemond JA, Dixon AFG. 1994. Specificity and daily cycle of release of sex pheromones in aphids: a case of reinforcement? *Biological Journal of the Linnean Society* **52:** 287–303.

Guldemond JA, Mackenzie A. 1994. Sympatric speciation in aphids. I. Host race formation by escape from gene flow. In: Leather SR, Watt AD, Mills NJ, Walters KFA, eds. *Individuals, Populations and Patterns in Ecology*. Andover, UK: Intercept, 367–378.

Guo SW, Thompson EA. 1992. Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics* **48:** 361–372.

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. *Molecular Ecology* **9:** 2055–2066.

Hales DF, Tomiuk J, Wöhrmann K, Sunnucks P. 1997. Evolutionary and genetic aspects of aphid biology: a review. *European Journal of Entomology* **94:** 1–55.

Lilley R, Hardie J. 1996. Cereal aphid responses to sex pheromones and host-plant odours in the laboratory. *Physiological Entomology* **21:** 304–308.

Llewellyn KS, Loxdale HD, Harrington R, Brookes CP, Clark SJ, Sunnucks P. 2003. Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites. *Molecular Ecology* **12:** 21–34.

Loxdale HD, Hardie J, Halbert S, Foottit Royal, Kidd NAC, Carter CI. 1993. The relative importance of short- and long-range movement of flying aphids. *Biology Reviews* **68:** 291–311.

Lynch M. 1984. Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *Quarterly Review of Biology* **59:** 257–290.

Mackenzie A. 1996. A trade-off for host plant utilization in the black bean aphid, *Aphis fabae*. *Evolution* **50:** 155–162.

Mackenzie A, Guldemond JA. 1994. Sympatric speciation in aphids. II. Host race formation in the face of gene flow. In: Leather SR, Watt AD, Mills NJ, Walters KFA, eds. *Individuals, Populations and Patterns in Ecology* Andover, UK: Intercept, 379–395.

Margaritopoulos JT, Blackman RL, Tsitsipis JA, Sannino L. 2003. Coexistence of different host-adapted forms of *Myzus persicae* in the region of Caserta in South Italy. *Bulletin of Entomological Research* **93:** 131–135.

Margaritopoulos JT, Mamuris Z, Tsitsipis JA. 1998. Attempted discrimination of *Myzus persicae and Myzus nicotianae* (Homoptera: Aphididae) by random amplified polymorphic DNA polymerase chain reaction technique. *Annals of the Entomological Society of America* **91:** 602–607.

Margaritopoulos JT, Shigehara T, Takada H, Blackman RL. 2007. Host-related morphological variation within *Myzus persicae* group (Homoptera: Aphididae) from Japan. *Applied Entomology and Zoology* **42:** 329–335.

Margaritopoulos JT, Tsitsipis JA, Goundoudaki S, Blackman RL. 2002. Life cycle variation of *Myzus persicae* (Sulzer) (Hemiptera: Aphididae) in Greece. *Bulletin of Entomological Research* **92:** 309–320.

Margaritopoulos JT, Tsitsipis JA, Zintzaras E, Blackman RL. 2000. Host-correlated morphological variation of *Myzus persicae* (Hemiptera: Aphididae) populations in Greece. *Bulletin of Entomological Research* **90:** 233–244.

Margaritopoulos JT, Tsourapas C, Tzortzi M, Kanavaki OM, Tsitsipis JA. 2005. Host selection by winged colonisers within the *Myzus persicae* group: a contribution towards understanding host specialization. *Ecological Entomology* **30:** 406–418.

Müller FP. 1985. Genetic and evolutionary aspects of host choice in phytophagous insects, especially aphids. *Biologisches Zentralblatt* **104:** 225–237.

Nikolakakis NN, Margaritopoulos JT, Tsitsipis JA. 2003. Performance of *Myzus persicae* (Hemiptera: Aphididae) clones on different host-plants and their host preference. *Bulletin of Entomological Research* **93:** 235–242.

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

Pettersson J. 1970. An aphid sex attractant. Biological studies. *Entomologia Scandinavica* **1:** 63–73.

Pouqueville F-CHL. 1820–21. *Voyage Dans la Grêce Comprenant la Description Ancienne et Moderne de l'Epire*. Paris: F. Didot.

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

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

Raymond M, Rousset F. 1995b. An exact test for population differentiation. *Evolution* **49:** 1280–1283.

Raymond B, Searle JB, Douglas AE. 2001. On the processes shaping reproductive isolation in aphids of the *Aphis fabae* (Scop.) complex (Aphididae: Homoptera). *Biological Journal of the Linnean Society* **74:** 205–215.

Rispe C, Pierre JS. 1998. Coexistence between cyclical parthenogens, obligate parthenogens, and intermediates in a fluctuating environment. *Journal of Theoretical Biology* **195:** 97–110.

Rispe C, Pierre JS, Simon JC, Gouyon PH. 1998. Models of sexual and asexual coexistence in aphids based on constraints. *Journal of Evolutionary Biology* **11:** 685–701.

Schluter D. 2001. Ecology and the origin of species. *Trends in Ecology and Evolution* **16:** 372–380.

Schneider S, Roessli D, Excoffier L. 2000. *Arelquin. A software for population genetic data analysis, Version 2.0.* Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.

Semtner PJ, Tilson WM, Dara SK. 1998. Performance of the tobacco aphid (Homoptera: Aphididae) on various host plants. *Journal of Entomological Science* **32:** 180–195.

Simon J-C, Baumann S, Sunnucks P, Hebert PDN, Pierre J-C, Le Gallic J-F, Dedryver CA. 1999. Reproductive mode and population genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Molecular Ecology* **8:** 531–545.

Simon JC, Rispe C, Sunnucks P. 2002. Ecology and evolution of sex in aphids. *Trends in Ecology and Evolution* **17:** 34–39.

Sloane MA, Sunnucks P, Wilson ACC, Hales DF. 2001. Satellite isolation, linkage group identification and determination of recombination frequency in the peach-potato aphid, *Myzus persicae* (Sulzer) (Hemiptera: Aphididae). *Genetical Research Cambridge* **77:** 251–260.

Sunnucks P, De Barro PJ, Lushai G, Maclean N, Hales DF. 1997. Genetic structure of an aphid studied using microsatellites: cyclic parthenogenesis, differentiated lineages and host specialization. *Molecular Ecology* **6:** 1059–1073.

Thieme T, Dixon AFG. 1996. Mate recognition in the *Aphis fabae* complex: daily rhythm of release and specificity of sex pheromones. *Entomologia Experimentalis et Applicata* **79:** 85–89.

Thornhill R, Alcock J. 1983. *The Evolution of Insect Mating Systems*. Cambridge, MA: Harvard University Press.

Turelli M, Barton NH, Coyne JA. 2001. Theory and speciation. *Trends in Ecology and Evolution* **16:** 330–343.

Vanlerberghe-Masutti F, Chavigny P. 1998. Host-based genetic differentiation in the aphid *Aphis gossypii* Glover, evidenced from RAPD fingerprints. *Molecular Ecology* **7:** 905–914.

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. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology and Evolution* **16:** 381–390.

Via S. 2002. The ecological genetics of speciation. *American Naturalulist* **159:** 1–7.

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

Via S, Hawthorne DJ. 2002. The genetic architecture of ecological specialization: correlated gene effects on host use and habitat choice in pea aphids. *American Naturalist* **159:** 76–88.

Vorburger C, Lancaster M, Sunnucks P. 2003. Environmentally related patterns of reproductive modes in the aphid *Myzus persicae* and the predominance of two 'superclones' in Victoria, Australia. *Molecular Ecology* **12:** 3493–3504.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.

Wilson ACC, Sunnucks P, Blackman RL, Hales DF. 2002. Microsatellite variation in cyclically parthyenogenetic populations of *Myzus persicae* in south-eastern Australia. *Heredity* **88:** 258–266.

Wu CI. 2001. The genic view of the process of speciation. *Journal of Evolutionary Biology* **14:** 851–865.

SUPPLEMENTARY MATERIAL

The following material is available for this article online:

Figure S1. UPGMA dendrogram based on shared allele distances among *Myzus persicae* populations. Numbers denote bootstrap percentages (from 1000 resamplings). Samples from tobacco and other crops in each tobacco growing region were pooled (C + T). Samples from different years in each region and crop were also pooled. Abbreviations are defined in Table 1.

Table S1. Allele frequencies, observed (H_O), and expected heterozygosity (H_{EXP}) in the nine populations of *Myzus persicae* examined (N = number of clones examined).

Table S2. Analysis of genetic differentiation between *Myzus persicae* populations. Below the diagonal the multilocus F_{ST} values and probabilities for absence of differences between populations (NS = non-significant, $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$) are given.

Table S3. Analysis of genetic differentiation between Greek mainland *Myzus persicae* populations. Below the diagonal the multilocus F_{ST} values and probabilities for absence of differences between populations (NS = non significant, $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$) are given. Above the diagonal probabilities for absence of allelic differentiation between populations are given.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1095-8312.2007.00828.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other then missing material) should be directed to the corresponding author for the article.