The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2011, **104**, 838–865. With 5 figures

Stay at home aphids: comparative spatial and seasonal metapopulation structure and dynamics of two specialist tansy aphid species studied using microsatellite markers

HUGH D. LOXDALE[1,2]*, GERHARD SCHÖFL[3], KERSTIN R. WIESNER[1], FRANKLIN N. NYABUGA[1,2], DAVID G. HECKEL[2] and WOLFGANG W. WEISSER[1]

[1]*Institute of Ecology, Friedrich Schiller University, Dornburger Strasse 159, D-07743 Jena, Germany*
[2]*Department of Entomology, Max Planck Institute for Chemical Ecology, Hans-Knöll-Strasse 8, D-07745 Jena, Germany*
[3]*Department of Cell and Molecular Biology, Leibniz Institute for Natural Product Research and Infection Biology (Hans Knöll Institute), Beutenbergstrasse 11a, D-07745 Jena, Germany*

Received 22 March 2011; revised 23 June 2011; accepted for publication 23 June 2011

Two tansy-feeding aphids, *Macrosiphoniella tanacetaria* (MA) and *Metopeurum fuscoviride* (ME), were studied at a small spatial scale in and around Jena (< 80 km^2) using polymorphic microsatellite markers. Both species were found in approximately 60% of sites formerly known to harbour the aphids, although, generally when they did occur, they occurred singly (MA ~50%; ME ~60%) and rarely together on the same plant at the same time (approximately 10%) and then usually only in the early part of the growing season. This difference may be a result of quasi-apparent competition effects elicited by ants farming ME aphids, and preferentially actively eliminating or disturbing MA aphids. In terms of population genetics, both aphids showed extreme genetic heterogeneity within a metapopulation structure, with ME more than MA (i.e. higher F_{ST} values, approximately 0.4 versus 0.15, respectively), and limited levels of interpopulation gene flow. Subpopulations often deviated from Hardy–Weinberg equilibrium and showed linkage disequilibria, as expected in animals with extended parthenogenetic reproduction, and had positive F_{IS} values for most large samples, suggesting inbreeding, and possibly philopatry, certainly in ME. Hierarchical analysis (allele range and number per locus, analysis of molecular variance and F_{ST}) strongly suggested that the plant rather than site governs the level of genetic variation. Bayesian clustering analysis revealed that both species had heterogeneous historical genetic patterning, with K (number of subgroups) in the range 3–7. Evidence is also provided from isolation-by-distance and private allele analyses indicating that, in MA, the presence of winged autumn males, absent in ME where males are wingless, influences comparative population genetic structuring, such that ME subpopulations are comparatively more inbred and genetically differentiated than MA subpopulations. Lastly, additional spatial arrangement (ALLELES-IN-SPACE) analysis showed that, in both species, certain subpopulations were genetically isolated from the remainder, probably as a result of geographical barriers, including intervening buildings and woods. As such, the biology of these tansy aphids living in semi-natural habitats is very different from many pest aphid species examined within agro-ecosystems and infesting ephemeral crops. This is because the former appear to be much more reluctant to fly and hence show contrastingly much higher levels of interpopulation divergence, even at small spatial scales as investigated in the present study. Indeed, the number of genotypic clusters found for tansy aphids using Bayesian approaches is similar to that globally for the major pest, the peach-potato aphid, *Myzus persicae*. *Biological Journal of the Linnean Society*, 2011, **104**, 838–865.

ADDITIONAL KEYWORDS: gene flow – genetic variation – historical patterns – inbreeding – migration – population divergence.

*Corresponding author. E-mail: loxdale@web.de

INTRODUCTION

I came to learn that worthwhile observations of birds and animals and insects were great in proportion to the smallness of the territory covered . . . To be a good naturalist one must be a stroller or a creeper, or better still a squatter in every sense of the word – never a traveller.

William Beebe (1877–1962) Naturalist and explorer (quoted in Gould, 2004)

An ecological scenario representing one or more assumed competing species in the same habitat, with each individual player showing metapopulation structure and dynamics (Hanski, 1999), concerns the aphids (Hemiptera: Aphididae) found on tansy, *Tanacetaria vulgaris* L. This plant is a member of the family Asteraceae (Compositae) and prefers well-drained poor soils, often in river valleys, such as the Saale valley in Jena. It usually grows as stands of 'genetically-identical' plants or genets, with each plant in turn comprising numerous stems or ramets (usually three to five). Tansy is known to be colonized by up to eight species of aphids, three species commonly: *Macrosiphoniella tanacetaria* (Kaltenbach) (MA), *Metopeurum fuscoviride* Stroyan (ME), and *Uroleucon tanaceti* (L.) (UR). The first two species are polymorphic green or brown in colour, usually green in MA, brown in ME, and the last is wine red. These species have different host-feeding preferences: MA is mainly a new shoot feeder, attacking the new flush of growth, especially in the early spring (April to early May) when its peak abundance is reached, although it may have a later secondary peak (Massonnet, 2002; Massonnet, Simon & Weisser, 2002b; Loxdale *et al*., 2011). In contrast, ME is predominantly a stem feeder, with its main peak occurring slightly later in the growing season (mid- to late-May to June), although, as the plant matures and later senesces, it moves into the (yellow) composite flower heads (Massonnet, 2002). Unlike MA, it is ant attended, usually by black garden ants, *Lasius niger* (L.) but, occasionally, by other species, including wood ants, *Formica rufa* (L.), whose guarding qualities appear to promote a larger colony size compared to the smaller guard ant species (H. D. Loxdale, pers. observ.). The third tansy aphid species, UR, peaks in mid-season (June), occurs mainly on the underside of the lower leaves (Blackman & Eastop, 2006), is not ant attended, and is very detrimental to plant health (Nowak & Komor, 2010), suggesting it not to be fully co-evolved with its host. In the case of ME, the ants actually build earth ramparts at the base of the aphid colonized plant they are tending, thereby facilitating their movements to and from the nest and, indeed, they may even move their 'cows' around physically to new parts of the same plant or perhaps to different plants, or sometimes to their nests (Keller & Gordon, 2010; M. Mehrparvar, pers. comm.). Taken together, this is likely to affect the population ecology, more especially in an insect whose populations are often derived from very limited numbers, and hence have a very low effective population size.

Besides these aspects, other aspects of the fundamental biology of the three cyclically parthenogenetic species are different. All are monoecious (i.e. inhabit the same plant throughout the growing season) and holocyclic (i.e. have a single annual autumnal/winter sexual phase after a period of asexual propagation in the spring and summer months, with perhaps up to 15 generations). The sexual phase involves the production of sexual forms [wingless sexual females (oviparae) and males] and, subsequent to the mating of these sexual forms, the laying of cold hardy overwintering eggs (Massonnet, 2002; Massonnet *et al*., 2002b; Massonnet & Weisser, 2004). However, whereas MA and UR produce winged males in autumn in response to declining temperatures and lowered day length (Dixon, 1998), ME males are wingless. This is very likely to exacerbate the effects of inbreeding in ME colonies by reducing intercolony (plant resource) aerial migration and hence gene flow.

In our studies of aphids in Jena and its surrounding landscape, MA and ME were frequently found, whereas the occurrence of UR was sporadic and confined to very local areas. MA and ME were generally much more widely spread over the region of the Saale valley. In addition, although tansy is widespread, as noted above, it is spatially separated into isolated clumps or stands of plants. These in effect act as 'islands of resource' and, as such, drive the metapopulation dynamics of these species (Massonnet, 2002; Massonnet *et al*., 2002b; Massonnet & Weisser, 2004). Because migrating aphids, when they are actively flying, are predominantly borne on the prevailing wind above their low flight speed in still air (Loxdale *et al*., 1993), being attracted to suitable plants by visual cues below the boundary layer of still air (Taylor, 1974), as well as odour cues (Pettersson, Tjallingii & Hardie, 2007), the difficulties of locating and successfully landing, feeding and reproducing on such plants are profound (Ward *et al*., 1998). This again must severely impact on the underlying population genetics of winged aphids migrating from one host plant resource (i.e. stand of plants) to another.

In the present study, we examined the occurrence of MA and ME at various collecting sites in and around Jena aiming to gain insight regarding the level of interspecific competition between them. In addition, we have used a range of polymorphic microsatellite markers to comparatively investigate the population genetics of the two most common tansy-feeding aphid species, MA and ME, at a small spatial scale. Primarily, we investigate: (1) do MA and ME inhabit the

same plants within the same stands/sites; (2) does the difference on male morphology (winged versus unwinged) fundamentally affect the population genetics, notably inbreeding; (3) does the metapopulation structure shape the population genetics and *vice versa*; (4) can we estimate how far the winged migrants fly between stands of plants and how many may do so; and (5) is there any underlying pattern of population genetic structure (i.e. historical patterning) from which we may infer the longer-term genetic turnover (genetic revolutions) of local tansy aphid populations, and does this say anything more about the spatio-temporal structure and dynamics of the two species? This dual model system offers a unique study opportunity in these respects.

MATERIAL AND METHODS

APHIDS

Life cycle

Massonnet (2002) and Massonnet *et al*. (2002b) detail the lifecycle and ecology of MA and ME. In brief, after egg hatch and the production of the fundatrix morph in the spring as new tansy plants begin to shoot (April/May), several asexual generations are produced, including winged migrants in the summer (June/July). These migrants are mostly responsible for intercolony/population gene flow and maintenance of metapopulation structure (but see also below). With the senescence of the host thereafter, populations tend to decline, often exacerbated by predator/parasitoid/pathogen pressure (Weisser, 2000; Powell & Pell, 2007; Nyabuga *et al*., 2010), especially in the non-ant attended, and hence guarded, species, MA (Weisser, 2000). In both species, but especially MA, population decline may lead to the persistence of very small number of aphids on the plant or even extinction (Massonnet, 2002; Loxdale *et al*., 2011).

Towards the end of the growing season, October and November, especially if plants are cut and resprout new foliage, a second rise in numbers occurs (Loxdale *et al*., 2011) whereupon, as a consequence of declining day length and lowered temperature (Dixon, 1998), sexual forms are produced. In MA, the winged males are able to fly between colonies and hence allow a further degree of intercolony/subpopulation gene flow, which is largely impossible in the wingless ME that can only walk between plants, especially if they fall off. In both species, males then mate with wingless females to produce eggs that are laid at the base of tansy or on foliage nearby and persist on dead plant material throughout the winter. Hence, in the autumn, only males of MA can undergo further intercolony gene flow, attracted by both plant host odours and female sex pheromones (Pickett & Glinwood, 2007). In contrast, the ME males can only mate with females within their own colonies or very close by.

Sampling

Aphids (fourth-instar nymphs and predominantly parthenogenetic winged and wingless adults) were collected from tansy genets and ramets at various sites around Jena, Germany (50°56'N, 11°35'E) in 2007, ME at 17 sites, MA at 11 sites (for site abbreviations, see Table 1; for locations, see Fig. 1). The sites were usually between a half and a few kilometres apart,

Table 1. Occupancy by the two tansy aphid species, *Macrosiphoniella tanacetaria* (MA) and *Metopeurum fuscoviride* (ME), at the various sites sampled

Site (abbreviation)	MA	ME
AA	+	+
B		+
BB	+	
D	+	+
E		+
F		+
H	*	+
IoE		+
K	+	+
L	+	+
M		+
R	+	+
S		+
T	+	+
U	+	+
V	+	+
W		+
X	+	+
Y	+	+

Sites in Jena occupied by tansy aphids: MA and ME. Note: In those cases of dual occupation of both species at the same site (grey shading), they were not normally found on the same plant together, although were occasionally (see Results).

* +, Aphids found at the site; * very few aphids found at the site (i.e. too few to analyse genetically); Sites: A, Sports stadium, Oberaue; B, Altenburger Strasse-Zeitzer Strasse; BB, TB Jena, cement factory near Kunitz; D, Burgau Park, Jena south; E, Am Flutgraben, Jena north; F, Burgau Park 'Wasserwerk'; H, opposite Hotel Best Western, Jena south; IoE, Institute of Ecology site (on left hand side in front of building facing east); K, Göschwitz bahnhof; L, Jena West Bahnhof; M, near Jena Experimental ground, Jena north; nr. Aldi, Jena north; R, Humboldt Haus, Charlottenstrasse, Jena Ost; S, Adriagrille, small meadow, Jena Ost; T, Saal bahnhof; U, Garage, Löbstedter Strasse; V, Patch near OBI, Löbstedter Strasse; W, Zeitzer Strasse; X, Porstendorf bahnhof; Y, Dornburg bahnhof.

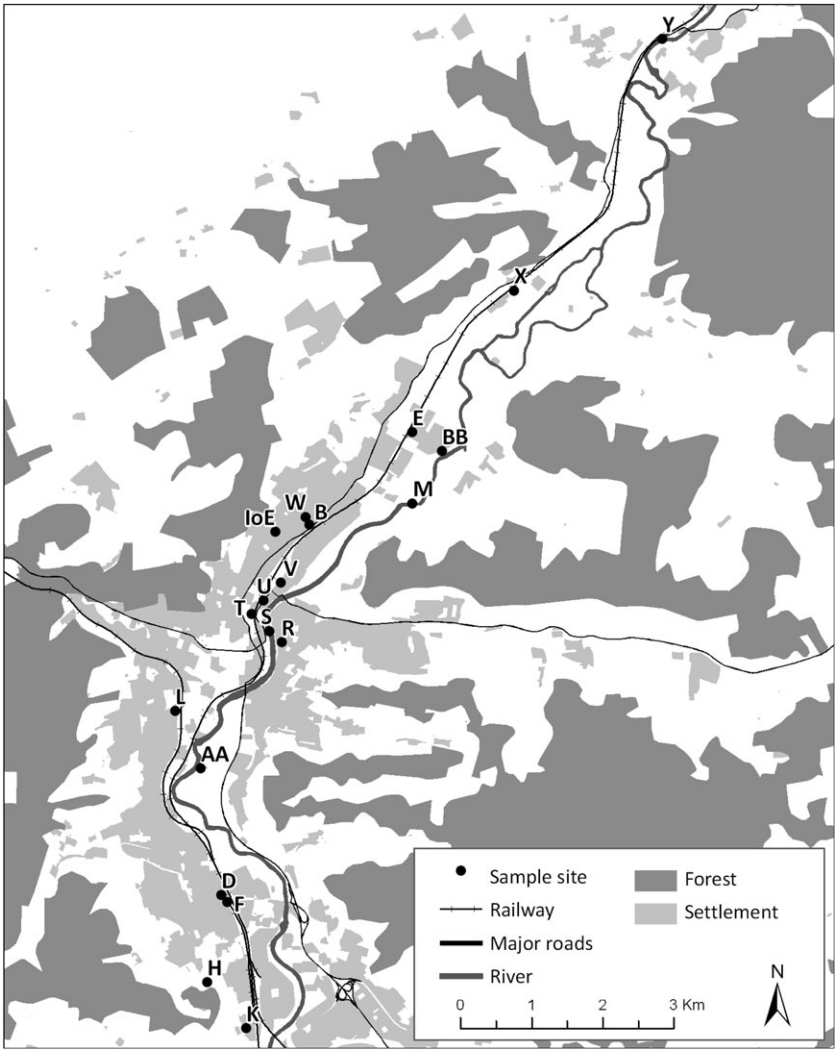

Figure 1. Map showing the collecting sites for tansy aphids in and around Jena in 2007. For site abbreviations, see Table 1.

maximally approximately 16 km (Göschwitz in the south to Dornburg in the north of Jena). For both species, most spatial samples were collected in the period April to June but, at some sites, temporal samples were also collected in July, August, September, and October (Tables 2, 3). At a given site, most plants were sampled within 30 m of one another (sometimes within a few metres), although occasionally further apart, maximally approximately 150 m. The aphids were placed in 100% ethanol in 1.0-mL Eppendorf tubes and later stored refrigerated at –20 °C until DNA extraction. To maximize genetic heterogeneity, usually only one to five aphids (maximum of eight) were randomly collected and later tested genetically per ramet. Aphids were visually checked during sorting (also in 100% ethanol) and those obviously parasitized by braconid wasp parasitoids or contaminated with entomopathogenic fungi were discarded. The longer-term temporal genetic aspects of the present samples (i.e. between seasons) in relation to earlier (i.e. pre-2007) MA and ME samples collected in Germany, including Jena, and Alsace, France, will be reported elsewhere.

Molecular markers

Purified DNA was extracted from individual aphids (100–200 ng) using the 'salting out' procedure of Sunnucks & Hales (1996). In total, 388 MA and 590 ME were thus prepared and subsequently tested electrophoretically. Samples were stored in distilled water or in TE buffer (10 mM of Tris, 1 mM of EDTA, pH 8.0) at –80 °C before testing. Microsatellite banding profiles were obtained using the primers and protocols essentially as described by Massonnet, Leterme &

Table 2. Main genetic parameters for *Macrosiphoniella tanacetaria* collected in 2007: April/May collections and September/October collections

								Significant deviation from HWE (out of eight loci; $P < 0.05$)			
Population	N_{genets}	N_{aphids}	N_{genos}	D^*	N_a (SE)	H_O (SE)	H_E (SE)	F_{IS}	HW	HD	HE	Winged (*N*)
April/May												
AA	2	17	17	1.0	3.3 (0.65)	0.48 (0.06)	0.54 (0.05)	0.1^{NS}	0 (8)	0 (8)	0 (8)	4
BB	1	39	27	0.68	3.3 (0.55)	0.48 (0.08)	0.50 (0.05)	0.03^{NS}	5 (8)	1 (8)	2 (8)	1
D	18	80	53	0.66	4.3 (0.65)	0.51 (0.07)	0.54 (0.06)	0.06*	6 (8)	3 (8)	2 (8)	5
K	2	12	12	1.0	3.1 (0.64)	0.44 (0.09)	0.43 (0.08)	-0.01^{NS}	1 (7)	0 (7)	0 (7)	2
L	6	11	11	1.0	3.6 (0.68)	0.61 (0.08)	0.58 (0.07)	-0.06^{NS}	0 (8)	0 (8)	0 (8)	0
R	6	37	31	0.83	2.9 (0.28)	0.32 (0.08)	0.46 (0.07)	0.31***	4 (8)	4 (8)	0 (8)	0
T	9	93	44	0.47	4.6 (1.1)	0.43 (0.07)	0.52 (0.05)	0.16***	6 (8)	5 (8)	1 (8)	0
V	5	24	16	0.65	3.6 (0.56)	0.52 (0.09)	0.57 (0.05)	0.1^{NS}	5 (8)	2 (8)	0 (8)	1
Means	6.1 (1.9)	39.1 (11.0)	26.4 (5.5)	0.79 (0.07)	3.6 (0.21)	0.47 (0.03)	0.52 (0.02)	0.09 (0.04)	3.38	1.88	0.63	1.6 (0.68)
Number of tests									63	63	63	
												4.2% winged of total

Population	N_{genets}	N_{aphids}	N_{genos}	D^*	N_a (SE)	H_O (SE)	H_E (SE)	F_{IS}	HW	HD	HE	♂ (*N*)	♀ (*N*)
September/October													
BB	4	8	2	0.14	1.6 (0.17)	0.50 (0.18)	0.3 (0.09)	-0.78^{NS}	5 (5)	0 (5)	4 (5)	0	0
K	5	6	6	1.00	3.0 (0.4)	0.75 (0.09)	0.61 (0.05)	-0.25^{NS}	0 (8)	0 (8)	0 (8)	0	0
R	2	33	16	0.45	2.5 (0.25)	0.70 (0.11)	0.48 (0.05)	-0.46^{NS}	7 (8)	1 (8)	6 (8)	1	1
T	4	7	5	0.67	2.4 (0.30)	0.30 (0.1)	0.38 (0.07)	-0.21^{NS}	1 (7)	2 (7)	0 (7)	0	0
Means	3.8 (0.63)	13.5 (6.5)	7.3 (3.0)	0.56 (0.18)	2.4 (0.28)	0.56 (0.10)	0.44 (0.07)	–0.43 (0.13)	3.25	0.75	2.5	0.25 (0.25)	0.25 (0.25)
Number of tests									28	28	28		

Standard errors of means are given in parenthesis.

NS, not significant ($P > 0.05$); *$0.05 > P > 0.01$; **$0.01 > P > 0.001$; ***$P < 0.001$.

N_{genos} = number of genotypes; D^* = clonal diversity; where 0 = identical variability, and 1.0 = maximal variation for that sample; N_a = the number of alleles over all loci per population; H_O = observed heterozygosity; H_E = expected heterozygosity; F_{IS} = inbreeding coefficient; HW = Hardy-Weinberg probability ; HD = heterozygote deficiency; HE = heterozygote excess; Winged = number of winged individuals per sample.

Table 3. Main genetic parameters for *Metopeurum fuscoviride*: late April to mid May, mid–late June, and mid July to early August

									Significant deviation from HWE (out of five loci; $P < 0.05$)			
Population	N_{genets}	N_{aphids}	N_{genos}	D^*	N_a (SE)	H_O (SE)	H_E (SE)	F_{IS}	HW	HD	HE	Winged (*N*)
April/May												
D	1	5	3	0.5	1.2 (0.15)	0.08 (0.08)	0.06 (0.06)	0.27^{NS}	0 (1)	0 (1)	0 (1)	0
E	2	7	3	0.33	1.5 (0.41)	0.10 (0.06)	0.11 (0.06)	0.11^{NS}	0 (2)	0 (2)	0 (2)	0
M	2	4	4	1.0	1.7 (0.3)	0.21 (0.09)	0.24 (0.10)	0.17^{NS}	0 (3)	0 (3)	0 (3)	0
R	2	7	7	1.0	2.2 (0.28)	0.31 (0.09)	0.42 (0.09)	0.28^{NS}	2 (5)	1 (5)	0 (5)	0
S	2	8	7	0.86	1.7 (0.3)	0.27 (0.12)	0.26 (0.11)	-0.06^{NS}	0 (3)	1 (3)	0 (3)	0
T	4	7	7	1.0	2.0 (0.24)	0.24 (0.07)	0.43 (0.08)	0.46*	0 (5)	0 (5)	0 (5)	1
X	4	5	5	1.0	1.9 (0.81)	0.53 (0.14)	0.46 (0.13)	-0.17^{NS}	0 (4)	0 (4)	0 (4)	0
Means	2.4 (0.4)	7.1 (1.1)	5.14 (0.7)	0.81 (0.11)	1.7 (0.12)	0.25 (0.06)	0.28 (0.06)	0.15 (0.08)	0.29	0.29	0	0.14 (0.14)
Number of tests									23	23	23	
												2% winged of total
June												
D	5	20	7	0.32	2.3 (0.38)	0.19 (0.11)	0.34 (0.07)	0.45***	5 (5)	3 (5)	1 (5)	2
K	4	25	5	0.17	2.2 (0.28)	0.04 (0.01)	0.11 (0.04)	0.64***	3 (3)	3 (3)	0 (3)	1
U	1	20	2	0.05	1.2 (0.15)	0.11 (0.11)	0.08 (0.07)	-0.52^{NS}	1 (1)	0 (1)	1 (1)	0
V	5	40	13	0.31	2.3 (0.38)	0.26 (0.13)	0.32 (0.07)	0.18*	4 (5)	3 (5)	1 (5)	0
X	5	10	9	0.88	2.8 (0.80)	0.33 (0.13)	0.32 (0.13)	-0.03^{NS}	0 (3)	0 (3)	0 (3)	0
Means	4 (0.8)	23 (4.9)	7.2 (1.9)	0.35 (0.14)	2.2 (0.26)	0.19 (0.05)	0.23 (0.06)	0.14 (0.2)	2.6	1.8	0.6	0.6 (0.4)
Number of tests									17	17	17	
												2.6% winged of total
July/August												
B	4	84	22	0.25	3.0 (0.71)	0.13 (0.03)	0.24 (0.06)	0.48***	4 (5)	4 (5)	0 (5)	0
D	7	36	7	0.17	2.5 (0.39)	0.07 (0.02)	0.32 (0.10)	0.77***	3 (5)	3 (5)	0 (5)	0
IoE	1	12	6	0.45	1.8 (0.24)	0.24 (0.12)	0.23 (0.08)	-0.02^{NS}	1 (4)	1 (4)	0 (4)	0
K	9	88	3	0.02	1.8 (0.24)	0.01 (0.01)	0.06 (0.03)	0.82***	3 (4)	3 (4)	0 (4)	0
T	4	21	9	0.40	2.8 (0.69)	0.39 (0.13)	0.49 (0.10)	0.21*	3 (5)	2 (5)	0 (5)	0
U	1	25	1	0.00	1.0 (0.0)	0.00	0.00	1.00	0 (0)	0 (0)	0 (0)	0
V	2	17	3	0.12	1.5 (0.27)	0.04 (0.02)	0.04 (0.02)	-0.02^{NS}	0 (2)	0 (2)	0 (2)	0
X	5	101	58	0.57	3.6 (0.95)	0.24 (0.12)	0.33 (0.12)	0.26***	3 (5)	2 (5)	0 (5)	0
Means	4.1 (1.0)	48 (12.9)	13.6 (6.7)	0.23 (0.08)	2.3 (0.31)	0.14 (0.05)	0.21 (0.06)	0.43 (0.14)	2.1	1.9	0	0
Number of tests									30	30	30	
												0% winged of total

Standard errors of means are given in parenthesis.

NS, not significant ($P > 0.05$); *$0.05 > P > 0.01$; **$0.01 > P > 0.001$; ***$P < 0.001$.

N_{genos} = number of genotypes; D^* = clonal diversity; where 0 = identical variability, and 1.0 = maximal variation for that sample; *N*a = the number of alleles over all loci per population; H_O =observed heterozygosity; H_E = expected heterozygosity; F_{IS} = inbreeding coefficient; HW = Hardy-Weinberg probability ; HD = heterozygote deficiency; HE = heterozygote excess; Winged = number of winged individuals per sample.

(2001) and Massonnet *et al*., (2002a), using a 10-μL reaction mixture volume (with a final Mg^{2+} concentration of 1.5 mM) and with products detected on a Licor 4300 sequencer (Licor 6.5% polyacrylamide gels; 0.25 mm thick run on 25 × 26 cm glass plates) using fluorescently-labelled forward primers (IRD-700 and -800) and fluorescently-labelled size markers (Licor; IRD-700 and -800; 50–350 bp). *Taq* polymerase was purchased from GeneCraft or Metabion. Polymerase chain reactions (PCR) were performed using Eppendorf thermocyclers (Mastercycler®). Product size was determined visually with alleles scored by hand from electronic gel photographs. For MA, eight primers were used (Ma-1-8; Massonnet's Mt-1-8); for ME, six primers were used (Me-1-6; Massonnet's Mf-1-6) (Massonnet *et al*., 2001, 2002a).

The various genotypes identified per locus were accumulated to produce a multilocus genotype (MLG). These were assumed to be clonal in origin (i.e. predominantly the offspring of a single parthenogenetic stem mother arising from the egg stage after sexual crossing by sexual forms the previous autumn). Accordingly, approximately eight and five loci were previously found to be adequate to identify MLGs within subpopulations of MA and ME, respectively (MA: Massonnet *et al*., 2002b: fig. 6; ME: Massonnet, 2002: fig. 1). The likelihood that MLGs from a particular lineage (or multilocus lineage, MLL) originated from two or more distinct sexual reproductive events was confirmed using dedicated computer software, along with confirmation of the required number of loci, as described below.

For direct DNA sequencing, which was performed for the product of one particular ME locus allele (Me-3^{214}), the product in question was cut from the sequencing gel, eluted from the gel matrix with TE buffer, pH 8.0 and then further amplified. Thereafter, it was run again on a 6.5% sequencing polyacrylamide, cut out and purified using a Zymoclean DNA gel recovery kit. The products (control and variant allele) were identified on a 1.5% agarose gel before being sequenced on an automated Applied Biosytems 3730/XL/96 capillary DNA analyser, with the fragments being sequenced in both forward and reverse directions using the specific ME primers.

Regarding the possibility of null alleles, we experienced very few priming failures (allele drop outs) using the 14 primer pairs employed for the two aphid species, confirming previous observations made by Massonnet (2002) that null alleles are unlikely to pose a serious problem in the interpretation of the results using these particular microsatellite markers (Massonnet, 2002), Checking the present largest MA data set (spring samples) and, for comparison, the ME spring samples using MICROCHECKER version 2.2.3. (van Oosterhout *et al*., 2004) showed that, according to the simulations performed and at a confidence level of 95%, four of the eight polymorphic MA loci and two of the five polymorphic ME loci (Me-6 was found to be fixed), potentially had null alleles manifest as a homozygous excess. In the case of MA, loci Ma-3, -4, -5 and -7 had one ($N = 2$), four ($N = 5$), four ($N = 6$), and two ($N = 9$) alleles, respectively, with observed frequencies > 20% than expected (with the total number of alleles given in parenthesis), and in ME, both Me-1 and -2 showed two ($N = 2$) such alleles each. However, these excesses (even after removal of duplicate MLGs) are much more likely to be the result of the replication of particular clonal MGLs within colonies ('genetic inflation') and inbreeding as a result of the metapopulation structure of these aphids and hence a true biological phenomenon.

Statistical analysis

Conventional genetic parameters

Basic population genetic parameters were assessed for all both data sets using GENEPOP, version 4.0 (Raymond & Rousset, 1995); FSTAT, version 2.9.3.2 (Goudet, 2002), and ARLEQUIN, version 3.1 (Excoffier, Laval & Schneider, 2005), with the last specifically for analysis of molecular variance (AMOVA). These parameters included the number of alleles per locus per subpopulation (N_a) and allelic richness (R) over all loci per populations tested; and, per locus and, over all loci, observed and expected heterozygosity (H_O and H_E, respectively); the probability of deviations from Hardy–Weinberg equilibrium (HWE) using the dual null hypotheses of both heterozygote excess and deficit at individual loci, as well as over all loci; linkage disequilibrium (LD) between segregating pairs of loci; genic (allelic) and genotypic differentiation; and F-statistical parameters, F_{ST} and F_{IS}, calculated per locus and population and over all loci and populations, and F_{ST} as pairwise indices or averaged over all populations tested per species population. Hierarchical F_{ST} analysis was performed for the data collected at different spatial scales, ranging from ramet, to genet, site and total population, as well as pairwise differences (distance method) in AMOVA. Isolation-by-distance (IBD) estimates were performed by plotting the logarithm ($\log_{10}$) of population differentiation, $F_{ST}/(1 - F_{ST})$, developed by Rousset (1997), versus geographical distance in metres (expressed as $\log_{10}$ or as the natural logarithm, ln). F_{ST} values were used to compute Rousset's (1997) genetic distance, whereas IBD values were computed using the software IBDWS, version 3.15 (Jensen, Bohonak & Kelley, 2005). Using Mantel tests, IBDWS assesses the statistical significance between the genetic distance (or similarity) matrix and comparable matrix of geographical distances

(Bohonak, 2002). Population divergence, $F_{ST}/(1 - F_{ST})$, for the two aphid species was also compared using analysis of covariance in SPSS, version 15.0 (SPSS Inc.). Private allele analysis (Barton & Slatkin, 1986) was performed using GENEPOP to obtain the mean frequency of private alleles, $p(1)$, per subpopulation in an island model as well as N_m, a measure of the number of immigrants per generation between subpopulations, where N = effective population size and m = the mutation rate. Frequency analysis of MLGs was performed using SPSS, version 15.0. The parameter D^*, a simple measure of clonal diversity, was estimated by dividing the number of genotypes (G) by the total number of aphids in a particular sample (N), *sensu* Dorken & Eckert (2001) [i.e. D^*(their R) = $(G–1)/(N–1)$]. Lastly, GENCLONE, version 2.0 (Arnaud-Haond & Belkhir, 2007) was used to compute: (1) a P_{sex} value for each main sample collection for both aphid species, which provides a probability of whether or not particular MLLs derived from independent sexual events; (2) that the number of loci used for the two aphid species was adequate to assess local variability; hence, for each species, the minimum average and maximum number of discriminated MLGs was calculated per locus (1000 permutations); and (3) the frequency distribution of the pairwise number of allele differences (genetic distance) between MLGs within a population sample (Arnaud-Haond *et al*., 2005, 2007). For most of the above tests, the level of probability was assessed at $P = 0.05$, unless otherwise stated.

Bayesian clustering analysis of population structure

Nonhierarchical genotypic clustering of individuals independent of the sampling regime was performed using Bayesian model-based approaches. These models account for the presence of HW disequilibrium and LD among alleles at each of the marker loci by introducing population groupings that minimize deviations from equilibrium within clusters. Estimating the most likely number of clusters needed to explain the observed data is challenging and the results may be sensitive to the number of loci used, the variation at these loci, the rate of gene flow, and the number of individuals typed (Evanno, Regnaut & Goudet, 2005; Huelsenbeck & Andolfatto, 2007). We estimated the number of clusters and the assignment of individuals into clusters using two methods.

First, using an Markov chain Monte Carlo (MCMC) method implemented in STRUCTURE, version 2.2 (Pritchard, Stephens & Donnelly, 2000; Falush, Stephens & Pritchard, 2003), the most likely number of clusters for both data sets (MA2007 and ME2007) was estimated by determining the change in the marginal likelihood of the data $\Pr(X|K)$ when the numbers of clusters (K) was fixed to different values ($K = 1, 2, 3, \ldots, 12$). We used an ancestry model that allowed for admixture and correlated allele frequencies between populations. Under this model, individual MLGs were fractionally assigned to clusters using a membership coefficient (interpreted as the probability of membership or as the fraction of the genome with membership in that cluster). We ran ten replicate Markov chains with a burn-in period of 2.0×10^5 iterations followed by a sampling period of 1.0×10^6 iterations for each K. We also implemented the ΔK method *sensu* Evanno *et al*. (2005) to detect the amount of structuring beyond which a further subdivision does not substantially improve the fit of the admixture model. ΔK is the second-order rate of change of the marginal likelihood function and takes into account both the gain in posterior probabilities over a range of K-values and the variance between independent runs at given values of K.

An MLG was assigned to the cluster for which it had the highest membership coefficient. The final placement of an MLG in a cluster was based on the majority assignment across all replicate runs after 'label switching' heterogeneity had been accounted for using CLUMPP, version 1.1.1 (Jakobsson & Rosenberg, 2007).

Additionally, we calculated a stability coefficient (S_N) (Richards *et al*., 2009) across the ten independent run for each data set and each K. The stability coefficient represents the average probability of individuals i and j to be assigned to the same cluster among N runs. K-values that maximize S_N are preferred.

Second, the number of clusters and the assignment of MLGs to those clusters were estimated simultaneously using STRUCTURAMA (Huelsenbeck & Andolfatto, 2007). STRUCTURAMA implements the basic no-admixture model of STRUCTURE but additionally allows the number of populations (K) to be a random variable that follows a Dirichlet process prior. STRUCTURAMA allows the user to run analyses where the concentration parameter α of the Dirichlet process prior (which shapes the prior probability of the number of clusters) is either set to a fixed value or itself treated as a random variable drawn from a Gamma hyperprior (http://fisher.berkeley.edu/structurama/manual.html). We chose the latter approach because by appropriately parameterizing the Gamma probability distribution we impose less prior information on the number of K. For each data set, we performed four analyses consisting of a single Markov chain run for 5.0×10^6 cycles. Samples were drawn from the chain every 125th cycle. The first 20 000 of the resulting 40 000 samples were removed as burn-in before analysis. For each run, α was drawn from a differently parametrized gamma distribution with shape parameter k set to 1.8 or 2.4 and scale parameter θ set to 0.4 or 0.6. The posterior probabili-

ties of the number of populations given the data $Pr(K|X)$ were averaged across runs. For population assignment of MLGs, the mean partition (i.e. the partition of the sample that minimizes the squared distance to all partitions of the sample visited during an MCMC-run) was calculated (Huelsenbeck & Andolfatto, 2007). The final placement of an MLG in a cluster was based on the majority assignment across all replicate runs.

Geographical patterns of genetic variation

Geographical coordinates (decimal degrees) of sampling sites were determined from Google Earth (®2009) with transformation performed in the German Gauss-Krueger-Projection (Datum Potsdam, Spheroid Bessel 1841) (Transformation Service: http://cousin.de/kkisbin/trafo.tcl). The map of the sampling area (Fig. 1) was created using the database from the ESRI Data & Maps (ESRI, 2003) and the software ArcGIS.

The software ALLELES-IN-SPACE was used to visualize the spatial arrangement of genotypes within the sampled range and detect barriers to gene follow between sampled locations (Miller, 2005). Accordingly, a Delaunay triangulation (Brouns, De Wulf & Constales, 2003; Watson, 1992) was used to generate the connectivity network among collection sites. Thereafter, Monmonier's algorithm was employed to highlight geographical barriers between sampled locations, based on network edges associated with the highest rates of change between genotype and location (Manni, Guerard & Heyer, 2004).

Only populations with greater than eight individuals were used. ME subpopulations E, M, S, and Y, which fell below this number, were hence not included in the analyses (although their inclusion led to essentially the same result).

RESULTS

Co-occurrence of the two tansy aphids species

As shown in Table 1, out of a total of 19 sites surveyed, all where tansy aphids were known to occur from present or past observations of earlier studies, the majority of sites (11/19 or approximately 58%) were occupied by both tansy aphid species over the course of the growing period (overall collecting period 30 April to 8 October). However, when the number of plants (total sampled = 152) at sites showing dual occupation were examined (including all sampling dates), it was found that approximately half were occupied by either MA (74; approximately 49%) or ME (88; approximately 58%), and only 15 (approximately 10%) were occupied by both species. For such cases of dual occupation, seven involved spring (April) collections, the other eight spring and summer (April and June) collections. Hence, the two species mostly occurred singly per infested plant, whilst approximately one-tenth of plants examined were occupied by both species (in one instance, this involved a single MA as the alternative species; Table 1), although, in approximately half the cases, there was a temporal separation involved in the occupation of the same host plant.

Macrosiphoniella tanacetaria

Spatial patterns (basic parameters)

In this species, samples were collected in late April to mid May 2007 (i.e. at peak abundance; total $N = 313$) from eight main populations: AA, BB, D, K, L, R T, and V, whereas smaller numbers of samples were collected from three additional sites: U, X, and Y (Fig. 1). In addition, five sites were sampled later in the growing season that year (September to October): BB, K, R, T, and V (total $N = 54$) but predominantly from site R, which had large colonies on a few plants within a few metres of each other. For the various spring and autumn subpopulations tested, the basic collecting and genetic parameters are shown in Table 2 and/or are briefly described below.

Alleles per locus and allelic richness: Use of GENCLONE confirmed that eight dinucleotide primers were adequate to reveal the extent of genetic variability present within the local MA subpopulations studied [i.e. the maximum number of MLGs (228) was reached], as also earlier shown by Massonnet *et al*. (2002b) in their previous study of this aphid species [in the case of the closely-related grain aphid, *Sitobion avenae* (F.), see also Haack *et al*., 2000]. In addition, the distribution of pairwise differences in allele length between MLGs for the spring sample, the largest available, was seen to follow a normal distribution with a peak in frequency of approximately 42 bp in length (data not shown). This suggests that there were no undue errors in scoring, nor were there any obvious gross mutational processes occurring.

In the spring subpopulation samples, the number of alleles ranged from 2 (Ma-1 and -3) to 19 (Ma-6) (mean ± SD, 7.25 ± 5.04 SD); the mean number of alleles per locus from 1.9–7.4. The number of alleles over all loci per population, here designated as N_a, was in the range 2.9–4.6, whereas the mean ± SD overall loci and populations was 3.6 ± 1.8 (Table 2). In the smaller autumn collection, the number of alleles declined to a maximum of seven (again Ma-6), the number of alleles per locus from 1.7–3.5, whereas N_a was in the range 1.6–3.0, with mean ± SD overall loci and populations of 2.4 ± 0.79 (Table 2). In the spring

sample, mean ± SD allelic richness (R) per subpopulation sample was 2.53 ± 0.32 (minimum sample size = 11 diploid individuals), declining slightly in the autumn sample to 2.28 ± 0.56 (minimum sample size = 6 diploid individuals).

Hardy–Weinberg expectations: Assuming the total spring and autumn collections comprised, respectively, large panmictic populations at the local scale surveyed, all eight loci were found to significantly deviate from expectations ($P < 0.05$) for both data sets, except at Ma-3 in the former sample.

For the spring subpopulations, most samples (except for AA, K, and L) showed significant deviations from expectations over all loci (maximally in populations D and T; i.e. with six out of eight loci; $P < 0.05$). Removal of duplicate MLGs slightly improved the proportion of samples conforming to HWE; thus, in BB, T, and V, the number of significant deviations fell from five to four, six to five, and five to three, respectively. In the spring, four out of the eight loci (Ma-2, -4, -5, -7) showed positive mean F_{IS} values (homozygous excess; approximately 0.1) per locus over all populations, whereas, in the autumn sample, only two loci (Ma-1 and -7) showed positive values, and the remainder were negative (hence showing a homozygous deficit; –0.44). Removal of MLG (= clonal) duplicates did not affect this pattern, except that, in the spring sample, an extra locus showed positive (Ma-6), whereas, in the autumn sample, a further locus (Ma-4) showed a positive F_{IS} value in addition to Ma-1 and -7. The mean values for observed and expected heterozygosity over all population samples were rather similar ($H_O = 0.47$; $H_E = 0.52$) with most subpopulations showing an excess of homozygotes (6/8), and with a slightly positive mean F_{IS} value (mean = 0.09) (Table 2). In the autumn sample, mean observed and expected heterozygosity changed slightly ($H_O = 0.56$; $H_E = 0.44$) such that F_{IS} values were all negative, reflecting a small but nonsignificant heterozygote excess overall (Table 2). The largest sample collected at site R (33 individuals) was seen to significantly deviate from HWE at 7/8 loci, except at Ma-7. Removal of duplicates reduced this to 5/8 loci, with Ma-1, -7, and -8 now nonsignificant.

LD: For the spring collection, LD values for each locus pair across all subpopulations were all highly significant (i.e. in disequilibrium; 28 pairwise comparisons; Fisher's method; $P < 0.001$); for the autumn samples, 13/28 (approximately 46%) pairwise comparisons were highly significantly different.

Genic and genotypic differentiation: Except at Ma-3, in terms of genic differences, all spring subpopulations were significantly different at the eight loci tested, as they were globally for each population pair across all loci (28 comparisons; Fisher's method; $P < 0.05$). Similarly, all samples were significantly different for genotypic differences, both for all population comparisons and globally (i.e. each population pair across all loci; Fisher's method). With the autumn samples, with the exception of Ma-2 and -3, all were significantly different for genic differences at the eight loci tested, as they were globally. Genotypically, all the autumn samples were significant at the eight loci, and across population pairs also, with one exception (5/6 comparisons; 83%). Hence, the subpopulations tested were generally highly differentiated genetically at the local spatial scale investigated.

F_{ST}: In the spring subpopulation collection, comprising the largest collection, F_{ST} values were in the range 0.067–0.351 (mean ± SE, 0.15 ± 0.014 over all loci and populations; 28 pairwise comparisons) and all pairwise comparisons were significant at $P < 0.05$.

AMOVA: For the spring samples, pairwise differences (distance method) calculated in AMOVA revealed that 13.12% of variation occurred among populations [d.f. = 7; sum of squares (SOS) = 174.74; variance components (VC) = 0.314], 9.83% among individuals within populations (d.f. = 305; SOS = 706.92; VC = 0.235); and 77% within individuals (d.f. = 313; SOS = 578.00; VC = 1.846): hence, most of the variance was accounted for within individuals themselves, showing that populations were highly heterogeneous, more especially at the level of individual genotypes. Even so, 13% of the variance was accounted for 'among populations', showing that interpopulation gene flow was not negligible. Overall, the number of winged individuals was small, representing only 4.2% of the total sample (Table 2).

In the autumn sample, pairwise differences calculated for four subpopulations, revealed most variation to again reside within individuals (110.35%; d.f. = 54; SOS = 134.5; VC = 2.49) rather than among populations (21%; d.f. = 3; SOS = 32.0.; VC = 0.46).

Clonal structure: In the spring subpopulations, overall, 56% of MLGs in the total population occurred as 'uniques' (i.e. occurred only once in the population sample), with the remaining 44% as multiple or repeat copies (i.e. 2, 3, 4, etc.), reaching a maximum of 15 copies (Cl. 162, representing 5% of the total) (Table 4A). In the autumn sample, the proportion of uniques declined to 36% of the total, whereas the largest multiple MLG comprised ten copies (Cl. 237, representing 19% of the total) (Table 4A) (Loxdale, Massonnet & Weisser, 2010). In the spring sample, the proportion of multilocus genotypes found per population, D^*, was in the range 0.47–1.0 (mean 0.79)

Table 4. Frequency distribution of multilocus genotypes

A) Early and late season collections of *Macrosiphoniella tanacetaria*

Clone number	Population	Season	Frequency
19	BB	APR–MAY	2 (i.e. twice)
27	BB	APR–MAY	2
35	BB	APR–MAY	3
36	BB	APR–MAY	2
38	BB	APR–MAY	2
41	BB	APR–MAY	2
42	BB	APR–MAY	3
43	BB	APR–MAY	3
50	D	APR–MAY	2
52	D	APR–MAY	4
58	D	APR–MAY	2
60	D	APR–MAY	2
67	D	APR–MAY	5
71	D	APR–MAY	3
72	D	APR–MAY	6
73	D	APR–MAY	2
75	D	APR–MAY	2
79	D	APR–MAY	2
80	D	APR–MAY	2
85	D	APR–MAY	3
87	D	APR–MAY	2
88	D	APR–MAY	4
131	R	APR–MAY	2
133	R	APR–MAY	2
135	R	APR–MAY	2
146	R	APR–MAY	4
154	T	APR–MAY	6
157	T	APR–MAY	3
158	T	APR–MAY	3
160	T	APR–MAY	2
162	T	APR–MAY	15
167	T	APR–MAY	4
168	T	APR–MAY	2
173	T	APR–MAY	3
184	T	APR–MAY	2
188	T	APR–MAY	4
189	T	APR–MAY	4
191	T	APR–MAY	2
194	T	APR–MAY	6
195	T	APR–MAY	5
200	V	APR–MAY	2
204	V	APR–MAY	8
226	BB	SEPT–OCT	7
237	R	SEPT–OCT	10
239	R	SEPT–OCT	4
240	R	SEPT–OCT	2
243	R	SEPT–OCT	2
247	R	SEPT–OCT	4
249	R	SEPT–OCT	2
256	T	SEPT–OCT	3

B) Early, middle and late season collections of *Metopeurum fuscoviride*

Clone number	First population	Season	Frequency	Second population	Season	Frequency	Third population	Season	Frequency
2	D	JUNE	3 (i.e. three times)						
3	V	JUNE	3						
4	V	JUNE	3						
8	T	JULY–AUG	2						
10	T	JULY–AUG	2						
12	U	JUNE	14	T	JUNE	2			

Table 4. *Continued*

B) Early, middle and late season collections of *Metopeurum fuscoviride*

Clone number	First population	Season	Frequency	Second population	Season	Frequency	Third population	Season	Frequency
17	U	JUNE	6	T	JUNE	3			
25	T	JULY–AUG	3						
28	R	APR–MAY	4						
31	U	JULY–AUG	25						
40	L	JUNE	2						
42	D	JULY–AUG	3						
45	B	JULY–AUG	2						
48	L	JUNE	2						
51	B	JULY–AUG	3						
59	V	JUNE	4						
62	K	JULY–AUG	3						
69	X	JULY–AUG	3						
71	X	JULY–AUG	2						
72	X	JULY–AUG	2						
73	X	JULY–AUG	2						
76	X	JULY–AUG	10						
83	X	JULY–AUG	2						
87	X	JULY–AUG	2						
95	M	APR–MAY	1	T	JULY–AUG	1			
98	X	JUNE	2						
101	X	JULY–AUG	2						
105	X	JUNE	1	T	JULY–AUG	1			
107	D	JULY–AUG	2						
109	X	JULY–AUG	2						
110	X	JULY–AUG	2						
113	B	JUNE	7	X	JUNE	1	X	JULY–AUG	6*
120	X	JULY–AUG	2						
122	X	APR–MAY	1	T	JULY–AUG	1			
124	W	APR–MAY	1	X	JULY–AUG	9			
131	X	JULY–AUG	5						
136	X	JULY–AUG	4						
138	X	JULY–AUG	2						
145	B	JULY–AUG	3						
147	X	JULY–AUG	2						
149	T	JULY–AUG	1	B	JULY–AUG	1			
156	IoE	JUNE	2	IoE	JULY–AUG	4*			
161	T	JULY–AUG	3						
163	I0E	JUNE	1	IoE	JULY–AUG	1*			
166	IoE	JUNE	2						
170	X	JULY–AUG	2						
178	B	JULY–AUG	3						
240	R	APR–MAY	3						
241	R	APR–MAY	2						
256	S	APR–MAY	2						
262	D	APR–MAY	2						
263	D	APR–MAY	2						
181	D	JUNE	5	K	JULY–AUG	10			
				D	JULY–AUG	13			
191	L	JUNE	2	B	JULY–AUG	8			
195	S	APR–MAY	1	T	APR–MAY	1			
				K	JUNE	18			
				W	JUNE	1			
				X	JUNE	1	X	JULY–AUG	1*
				B	JULY–AUG	48			
				D	JULY–AUG	5			
				K	JULY–AUG	75			
				V	JULY–AUG	15			
216	E	APR–MAY	2	D	JUNE	7			
				K	JUNE	1			
				V	JUNE	21			
				X	JULY–AUG	1			

Table 4. *Continued*

B) Early, middle and late season collections of *Metopeurum fuscoviride*

Clone number	First population	Season	Frequency	Second population	Season	Frequency	Third population	Season	Frequency
230	V	JUNE	1	AA	APR–MAY	1			
				E	APR–MAY	4			
				B	JUNE	1			
				K	JUNE	4			
				Y	JUNE	2			
245	T	JULY–AUG	4						
250	D	JULY–AUG	2						
252	IoE	JULY–AUG	4						
255	T	JULY–AUG	4						
259	B	JULY–AUG	2						
264	D	APR–MAY	1	D	JULY–AUG	1			
266	T	JUNE	2						
273	D	JULY–AUG	10						
276	F	JUNE	5						
278	D	JUNE	2						

*Temporal sample. APR–MAY = late April to mid-May; JUNE = mid–late June; JULY–AUG = mid July to early August; SEPT–OCT = early September to mid-October. Frequency = frequency of occurrence (once, twice, three times, etc.).

and with three out of eight samples with a value of 1.0 (i.e. all unique genotypes). In the autumn sample, D^* was in the range 0.14–1.0 (mean 0.56), with only one sample of four tested having a value of 1.00 (Table 2). These results show that, in most samples in the two sampling periods, subpopulations predominantly comprised unique genotypes and with relatively smaller proportions of repeat MLG copies (clones). There were only three examples where D^* was < 0.5 (i.e. T for the early sample, and BB and R in the later sample) (Table 2). Despite the disparity in the number of samples tested (less later on), the fact that D^* declines from a mean of approximately 0.8 in the spring to 0.6 later in the growing season (Table 2) reveals the subpopulations to have become slightly more clonal with time (Loxdale *et al.*, 2010). All the various MLGs were seen to be unique to the sites they were collected from, with no overlap between sites.

When the entire large spring sample, comprising 228 MLGs ($N = 333$ individuals), was tested for evidence that particular MLLs had originated from two or more sexual events, P_{sex} values were seen to range from 1.9×10^{-50} to 5×10^{-3}, hence strongly supporting the notion that such lineages derived from a single sexual foundress. A similar conclusion was drawn for the autumn collection (29 MLGs and 55 individuals; P_{sex} values 9.8×10^{-22} to 1×10^{-2}).

Hierarchical partitioning: For hierarchical partitioning of the variances obtained, only the spring samples, the largest available, were considered. These were analyzed at both the level of the genet and ramet. In the case of genets (sample size range 4–47), the number of alleles per locus ranged from 2 (Ma-1 and -3), to 17 (Ma-6) (mean ± SD, 6.88 ± 4.46); the mean number of alleles per locus, N_a for genets 2.24 ± 0.77, whereas, for ramets (sample size 3–5), the number of alleles per locus ranged from 2 (Ma-1 and -3), to 14 (Ma-6) (mean ± SD, 6.13 ± 3.66); the mean number of alleles per locus, N_a 2.06 ± 0.65. Hence, the range of alleles was little changed from that found at the site level, although the mean allele number per locus was reduced from 3.6 in subpopulations to approximately 2.0–2.4 for genets and ramets.

On examining F_{ST} at the scale of genet-ramet and ramet alone (and accepting that the method is very approximate because sample size is so reduced, especially for ramets), values were seen to increase from a mean of 0.129 over all loci and populations to 0.298 for genets and ramets combined and 0.316 for ramets only. In terms of significance, 661/780 (84.74%) and 864/1035 (83.48%) pairwise comparisons were significant for the two cases, respectively. Thus, the number of significant cases declined on reducing sample size. F_{IS} was negative for both ramets and genets combined and ramets only (i.e. –0.133 and –0.126, respectively), hence showing a heterozygote excess. Pairwise differences for the highly subdivided samples were much as with the larger site-based samples, with most of the variance being found to reside 'at the individual level' [79.53% (d.f. = 311; SOS = 586.0; VC = 1.88) and 77.1% (d.f. = 207; SOS = 374.0; VC = 1.81) for genet and ramet combined and ramets alone but with higher 'among population' values [i.e. 29.8% (d.f. = 39; SOS = 469.9; VC = 0.71) and 32%. (d.f. = 45;

SOS = 362.6; VC = 0.74), respectively]. These results show that, with increasing subdivision of the resource, and hence sample size from aphids infesting clumps of plants at given sites, to genets and finally to ramets (sample size 3–5), the mean number of alleles decreased, whereas the level of genetic differentiation of colonies increased, with a concomitant decrease in levels of estimated gene flow, as expected.

On partitioning the number of microsatellite MLG repeats (i.e. > 1.0) versus the number of ramets and genets on which they were sampled, not unexpectedly, the greater the number of MLGs of a given lineage, the more frequently these were found, on average, on a range of ramets or genets in the same stand of tansy plants. Thus, for ramets, MLGs on the same ramet occurred at a mean ± SD of 2.4 ± 0.91 times (number of examples (single ramets) = 16; range 2–5 MLG repeats); on different ramets, including the natal one on the same genet, at a mean ± SD of 3.48 ± 2.58 times (number of examples = 66; range 2–8 ramets; 2–15 MLG repeats); and on different genets, including the natal one, at a mean ± SD of 4.3 ± 2.71 times (number of examples = 22; range 2–4 genets; 2–10 MLG repeats). None of the MA MLGs was found to occur either at sites other than the one sampled, nor at a different sampling period other than that sampled (Table 4A). From the available data, it is clear that the spread of MLGs between ramets on the same plant, let alone genets in the same stand, is apparently extremely limited in MA.

To further demonstrate the extreme heterogeneity of MA colonies at very small spatial scales, the most variable locus sampled in the study, Ma-6, were additionally examined for genets from four spring subpopulations (D, R, T, and V) (Table 5). Many of these genets were, per site, only a few metres apart, yet the distributions of even the most abundant alleles is (barring sampling effects in relation to the small sample sizes) clearly very variable and reflects the isolated genetic nature of the colonies studied.

Aphid number versus allele/genotype correlations: There were clear linear correlations on plotting for both data sets (spring and autumn) combined the number of aphids versus mean number of alleles ($y = 0.021x + 2.547$, $R^2 = 0.535$) and versus number of genotypes ($y = 0.522x + 4.023$, $R^2 = 0.896$), as well as between the mean number of alleles versus number of genotypes ($y = 0.039x + 2.395$, $R^2 = 0.586$)(data not shown). The linearity of the first two correlations implies that the graphs are a long way from reaching the asymptote of the expected exponential relationship (Loxdale & Macdonald, 2004; Arnaud-Haond *et al*., 2007), in turn suggesting that the population samples tested reflect but a very small proportion of a huge population diversity locally.

METOPEURUM FUSCOVIRIDE

Spatial patterns (basic parameters)

In 2007, three spatio-temporal series of subpopulations were sampled: late April to mid May (spring) at sites D, E, M, R, S, T, U, W, X, Y; mid-late June (summer) at six main sites B, D, K, U, V, X and with smaller numbers of sample from seven additional sites, F, H, IoE, L, T, W,Y, where the colony size was small and it was therefore not possible to get obtain large samples for genetic analysis; and in mid July to early August (late summer) where generally larger samples were obtained at sites, B, D, IoE, K, T, U, V, X (Table 3). Me-6, although previously found to be polymorphic in some populations in France and Germany (Massonnet, 2002), was monomorphic in the present local study and hence excluded from further analysis. In addition, some populations showed a very large allele at Me-3 (Me-3^{214}), three times larger than the original sequenced microsatellite, a $(TG)_{20}$ repeat (Massonnet *et al*., 2002a). Direct sequencing of this larger product showed it to be an amplified version of the same microsatellite [i.e. $(TG)_{60}$], with identical flanking regions. The spatial distribution of aphids bearing the Me-3^{214} allele are shown in Table 6.

Alleles per locus and allelic richness: Use of GENCLONE confirmed that five polymorphic microsatellites of a total of six (two tri- and four dinucleotides) were adequate to reveal the extent of genetic variability within the local ME populations studied; i.e. the maximum number of MLGs (April/May, $N = 37$; June, $N = 43$; July/August, $N = 103$) was reached, as previously shown by Massonnet (2002) in a previous study of this aphid species. However, in the present study, Me-6 $(TGG)_{10}$, was invariant in the local Jena subpopulations screened (see below), whereas it had been found to be polymorphic by Massonnet (2002) in an earlier, very much larger spatial study of the aphid. Use of the GENCLONE also revealed that the distribution of pairwise number of allele differences between MLGs was in ME seen to follow a tri-phasic pattern with peaks some 80 bp apart at approximately 30, 110 and 190 bp (all seasonal samples pooled; data not shown).

For the 20 subpopulations tested in the spring, early and late summer, the basic collecting and genetic parameters are shown in Table 3. The number of alleles ranged from 1 (Me-6, fixed) to 13 (Me-3) (mean ± SD, 3.94 ± 3.65); the mean number of alleles per locus per population, N_a, from 1 (fixation, subpopulation U; July/August) to 3.6 (mean ± SD over all loci and populations, 2.05 ± 0.67). Mean ± SD allelic richness (R) for the three seasonal collections was 1.79 ± 0.51 (April/May) (minimum sample size = 4 diploid individuals); 2.00 ± 0.62 (June) (minimum

Table 5. Distribution of MA-6 alleles in samples of *Macrosiphoniella tanacetaria* collected from genets in local tansy stands

Population/ allele	189	194	196	199	203	205	207	208	209	213	217	219	221	223	232	249	253	2*N**
D1	0	0	0	0.5	0	0	0.125	0	0	0	0	0	0.375	0	0	0	0	8
D2	0	0	0	0.357	0	0	0.071	0	0	0	0	0	0.571	0	0	0	0	14
D8	0	0	0	0.25	0.25	0	0.125	0	0	0	0.375	0	0	0	0	0	0	8
D11	0	0	0	0	0	0	0.5	0	0	0	0	0	0.5	0	0	0	0	8
D12	0	0	0	0	0	0	0.5	0	0	0	0.5	0	0	0	0	0	0	10
D13	0	0	0	0.571	0.143	0	0	0	0	0	0	0	0.286	0	0	0	0	14
D14	0	0	0	0	0.375	0	0.125	0	0	0	0	0	0.5	0	0	0	0	8
D15	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	10
D16	0	0	0	0	0	0	0.125	0	0	0.375	0	0	0.5	0	0	0	0	8
D18	0	0	0	0.5	0	0	0	0	0	0.25	0	0	0.25	0	0	0	0	8
D19	0	0	0	0	0.5	0	0	0	0	0	0.167	0	0.333	0	0	0	0	18
D42	0	0	0	0.5	0	0	0.125	0	0	0	0	0	0.375	0	0	0	0	8
D63	0	0	0	0.5	0	0	0.25	0	0	0	0.125	0	0.125	0	0	0	0	8
D64	0	0	0	0.5	0	0	0.25	0	0	0	0.125	0	0.125	0	0	0	0	8
D74	0	0	0	0.6	0.1	0	0.1	0	0	0	0.1	0	0	0	0	0	0.1	10
D75	0	0	0	0.6	0.1	0	0.1	0	0	0	0.1	0	0	0	0	0	0.1	10
D105	0	0	0	0	0.4	0	0	0	0	0	0.6	0	0	0	0	0	0	10
R3	0	0	0	0.167	0	0	0	0	0	0	0.167	0	0.667	0	0	0	0	6
R4	0	0	0	0.571	0	0	0	0	0	0	0.095	0	0.333	0	0	0	0	42
R5	0	0	0	0.5	0	0	0	0	0	0	0.125	0	0.375	0	0	0	0	16
R8	0	0	0	0.375	0	0	0	0	0	0	0.25	0	0.375	0	0	0	0	8
T2	0	0	0	0	0.5	0	0	0	0	0	0	0	0.5	0	0	0	0	18
T3	0	0	0	0	0.5	0	0	0	0	0	0.167	0	0.333	0	0	0	0	6
T5	0	0	0	0	0.227	0	0	0	0	0	0.182	0	0.455	0.136	0	0	0	22
T8	0	0	0	0.521	0.011	0	0.372	0	0.021	0	0	0	0.032	0	0.043	0	0	94
T9	0	0	0	0.5	0	0	0.5	0	0	0	0	0	0	0	0	0	0	20
T11	0	0	0	0.125	0.375	0	0	0	0	0	0	0	0.5	0	0	0	0	8
T12	0	0.1	0.1	0	0.1	0	0	0	0	0	0.6	0	0.1	0	0	0	0	10
V1	0	0.1	0	0	0.5	0	0	0	0	0	0.1	0	0.3	0	0	0	0	10
V2	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0	0	0	20
V3	0	0	0.125	0	0	0.375	0	0	0	0	0	0	0.5	0	0	0	0	8
V5	0	0	0.625	0	0	0.125	0	0	0	0	0.25	0	0	0	0	0	0	8

*2*N* = diploid number of genes, where *N* = sample size.

Table 6. The distribution of Me-3^{214} alleles within *Metopeurum fuscoviride* populations sampled in 2007

Season and year/Site	Late April to Mid May 2007		Mid to late June 2007		Mid July to early August 2007	
	Allele frequency	*N*	Allele frequency	*N*	Allele frequency	*N*
B	0.000	0	0.000	0	0.780	84
D	0.000	5	0.750	20	0.569	36
E	0.857	7	0.000	0	0.000	0
K	0.000	0	0.920	25	0.966	83
S	0.562	8	0.000	0	0.000	0
T	0.357	7	0.000	0	0.262	21
U	0.000	0	0.000	20	0.000	25
V	0.000	0	0.775	40	0.941	17
X	0.167	6	0.200	10	0.302	101

N, sample size.

sample size = 10 diploid individuals) and 1.95 ± 0.63 (July/August) (minimum sample size = 12 diploid individuals).

Hardy–Weinberg expectations: On performing HW tests on the three population sets as if they comprised, respectively, one larger panmictic population each, all five usable polymorphic loci were found to significantly deviate from expectations [except at Me-5; June (summer) collection], and globally over all loci and populations. With one exception (Me-5; June collection), F_{IS} values were all positive. Removal of duplicate MLGs did not change the general picture, except at Me-5, where April/May populations were also seen not to significantly depart from HWE.

For all three seasonal samples, observed and expected heterozygosity were low (H_O means ranged from 0.14–0.25; H_E means from 0.21–0.28). These values are lower than for MA, respectively. Although the samples sizes are different between seasonal collections, there is a trend of decreasing H_O (and H_E) with time, presumably reflecting the increase of clonal copies as the growing season progressed and probably clonal competition/selection. In terms of HWE, in the subpopulations of larger sample size (June and July/August), most tests showed significant deviations from expectations, declining in the later sample [i.e.13/17 (76%) and 17/30 (57%), respectively]. Removal of clonal copies was not possible as a result of the high incidence of clonal copies in all three ME data sets (as seen by the low D^* values; see below). Mean F_{IS} values per locus over all populations were mostly positive (with one exception, i.e. 14/15 tests), rising from approximately 0.19 in spring, to 0.33 in summer and 0.39 in late summer, signifying increasing homozygosity with time, probably as a result of clonal replication (Table 3).

LD: For the spring collection, most LD values for each locus pair across all subpopulations were highly significant [i.e. 7/10 (70%) pairwise comparisons; Fisher's method; $P < 0.001$]; for the summer sample, 5/10 (50%) samples were; and, for the late summer samples, all were.

Genic and genotypic differentiation: All populations showed significant differences ($P < 0.05$) in gene and genotype frequencies, as they did globally over all population pairs across all loci (Fisher's method) for both the midsummer and late summer collections, respectively, with one exception in the spring sample (1/21 genic and genotypic comparisons; approximately 5%) and one in the latter sample (1/28 genotypic comparisons; approximately 3.6%). As with MA, the populations are clearly highly divergent genetically.

F_{ST}: F_{ST} values were in the range 0.090–0.833 (mean ± SE, 0.377 ± 0.04 over all loci and populations; 21 pairwise comparisons) in the spring sample; 0.058–0.875 (mean ± SE, 0.403 ± 0.97 over all loci and populations; ten pairwise comparisons) in the summer sample; and 0.031–0.976 (mean ± SE, 0.400 ± 0.0.05 over all loci and populations; 28 pairwise comparisons) in late summer. All pairwise comparisons were significant ($P < 0.05$) for all sets of seasonal samples, barring one exception in the last.

AMOVA: Pairwise differences (distance method) calculated in AMOVA revealed that in the spring samples, 36.32% variation occurred among populations (d.f. = 6; SOS = 51.96; VC = 0.54); 11.93% among individuals within populations (d.f. = 43; SOS = 48.37; VC = 0.177) and 51.75% within individuals (d.f. = 50; SOS = 38.5; VC = 0.77); in the summer samples, 48.3% variation occurred among populations (d.f. = 4;

SOS = 20.1; VC = 0.664), 11.81% among individuals within populations (d.f. = 110; SOS = 95.9; VC = 0.162), and 39.87% within individuals (d.f. = 115, SOS = 63.0; VC = 0.547). The late summer samples showed a similar trend, with 39.4%, 24.4%, and 36.2% for the three categories, respectively.

Clonal structure: In the ME 2007 spring sample, uniques comprised 60% of the total, whereas, of the remaining multiple copies, the largest repeat only comprised 15% of the total (i.e. 2×4 repeats for clones 28 and 230 each). In the summer sample, uniques comprised 18% of the total; the largest repeat comprising 21 repeats was 14% (Cl. 216). Lastly, in the late summer sample, uniques again comprised 18% of the total, whereas the largest clonal repeat comprised 20% (Cl. 195, 75 repeats) (Table 4B), hence being slightly in excess of the unique category, and comprising the only time this was found in any of the population samples used in the present study for either species (Loxdale *et al*., 2010). Early in the season, D^* was in the range 0.33–1.0, and with five out of seven samples showing > 0.86 (mean 0.81), and hence a majority of unique MLGs. By June, values were in the range 0.05–0.88 in seven samples surveyed, with a mean of 0.35, whereas, in the late season (July/August), values were in the range 0.00–0.57 (mean 0.23) in eight samples, so that both later sampling periods show a clear decline in the proportion of unique MLGs with a concomitant increase in clonal copies, as seen with MA above (Table 2).

Hierarchical partitioning: For ME, the largest collection was made in late summer (July/August), so that only this was analyzed for hierarchical trends. In the case of genets (sample size range 4–63), the number of alleles per locus ranged from 1 (Me-6, fixed) to 13 (Me-3) (mean ± SD, 4.5 ± 4.27 over all five loci and eight subpopulations); the mean ± SD number of alleles per locus, N_a, for genets was 1.48 ± 0.42 (31 subsamples), whereas, for ramets (sample size 3–8), the number of alleles and range was the same as for genets, although the mean ± SD number of alleles per locus was 1.39 ± 0.43 (81 subsamples). As a further example of this decreasing trend, and in relation to Me-3, which had the most alleles of any of the Me loci tested (13 maximum), the genet samples from the July/August sampling period still showed this overall maximum number of alleles over all subpopulations; however, the mean ± SD number of alleles per locus decreased from 4.38 ± 2.45 to 2.13 ± 0.71 (31 samples), decreasing further still at the ramet scale [i.e.1.88 ± 1.41 (81 subsamples). In the subpopulations, seven of eight of these (approximately 88%) had the Me-3^{214} allele, four of them (i.e. 50%) as the dominant allele ($f > 0.569$–0.966); in the genet subsamples, 26/31 (83.9%) had this allele, 18/31 (58%) as the dominant allele ($f > 0.4$–1.0), whereas, in the ramet subsamples, 67/81 (approximately 83%) had the allele, and, in 40 (approximately 50%), it was dominant ($f > 0.4$–1.0).

In terms of genetic variance, F_{ST} values were seen to rise from approximately 0.4 in the subpopulation samples to 0.51 for genets and ramets combined (465 pairwise comparisons) and 0.47 (3321 pairwise comparisons) for ramets alone. The significance levels were 85.4% for genets–ramets (397/465 comparisons) and 71% (2359/4421 comparisons), respectively. This progressive decline in the level of significance from 100% in the subpopulations tested to 85% (genets–ramets) and then 71% (ramets) undoubtedly reflects the inaccuracy of F_{ST} measured at the much smaller sample sizes appertaining to these very small spatial scales, as well as the fact that such small samples are more likely to become similar as alleles are lost in the process of sample size reduction (sampling effect).

At the genet–ramet scale, pairwise differences were increased proportionality to 53.2% (genet–ramet) (d.f. = 30; SOS 411.3; VC 0.56) and 57.7% (ramet) (d.f. = 81; SOS 483.1; VC 0.59), respectively, for 'among population' differences. Not unexpectedly, compared with the subpopulations with a 'among population' value of approximately 40%, these much higher values reflect an increasing level of genetic differentiation at these smaller spatial scales.

As with the hierarchical partitioning of the MA samples, on partitioning the number of ME MLG repeats versus the number of ramets and genets sampled, it was again found that the greater the number of MLGs per lineage the more such MLGs were, on average, found over a range of ramets or genets in the same stand. However, unlike MA, in ME, MLGs were also detected also on different stands of tansy within the sampling region (Table 4B). Thus, for ramets, MLGs on the same ramet occurred at a mean ± SD of 2.54 ± 0.86 times [number of examples (single ramets) = 21; range 2–5 MLG repeats]; on different ramets (including the natal one) on the same genet, at a mean ± SD of 5.64 ± 6.8 times (number of examples tested = 30; range 2–6 ramets; 2–25 MLG repeats); and on different genets (including the natal one), at a mean ± SD of 19.6 ± 23.25 times (number of examples tested = 29; range 2–8 genets; 2–75 MLG repeats). Some ME clones, even rare ones, were found at different sites in the same season (e.g. Cl. 12 and 17) or at different sites at different seasons (e.g. Cl. 95, 105, 122), whereas some of the more abundant clones showed both trends (e.g. Cl. 195, 216) (Table 4B). Occasionally, some clones, even rare ones (e.g. cl. 156, 163), were found to occur at the same site in different seasons (Table 4B). As with MA, the spread of ME MLGs between ramets on the same plant, let alone

genets in the same stand, was generally very limited. However, unlike the former species, in ME, the spread of the more abundant clones was sometimes considerable in terms of the small overall spatial scale examined (e.g. Cl. 195, site K versus X; i.e. approximately 11 km apart out of a total sampling distance of approximately 16 km). Hence, it here appears that, although movement between ramets and genets is generally limited, ME aphids had spread comparatively further than MA individuals.

Aphid number versus allele/genotype correlations: As with MA, there were clear, albeit generally weaker, linear correlations on plotting the data for ME (here spring, summer and autumn data sets combined) as the number of aphids versus the mean number of alleles ($y = 0.012x + 1.71$, $R^2 = 0.304$) and the number of genotypes ($y = 0.291x + 1.17$, $R^2 = 0.476$); and the mean number of alleles versus the number of genotypes ($y = 6.628x - 6.57$, $R^2 = 0.628$) (data not shown). However, the same general conclusions can be drawn as for the former species; namely that, in terms of the number of aphids versus the number of alleles and genotypes, the population sampled is far from the asymptote.

IBD for the two aphid species

On plotting measures of population divergence [logarithm of Rousset's, (1997) $F_{ST}/(1 - F_{ST})$] versus log geographical distance (metres), no obvious relationship is apparent in either aphid species studied (Fig. 2A). This is confirmed statistically ($P > 0.05$) upon performing a Mantel tests for matrix correlation between genetic distance and the log of geographical distance for the two tansy aphids species: MA 2007: $Z = -77.2792$, $r = 0.1703$, intercept = -3.783 ± 0.578 (SE); slope = 0.8634 ± 0.167 (SE); $N = 28$: one-sided $P = 0.251$ from 1000 randomizations; for ME 2007 (summer): $Z = -9.5601$, $r = -0.1941$, intercept = 5.250 ± 1.921 (SE); slope = -1.524 ± 0.528 (SE); $N = 10$: one-sided $P = 0.629$ from 1000 randomizations; and for ME 2007 (later summer) $Z = -23.1847$, $r = -0.3187$, intercept = 4.754 ± 1.932 (SE); slope = -1.451 ± 0.270 (SE); $N = 28$: one-sided $P = 0.922$ from 1000 randomizations. Interestingly, when the natural log (ln) values of population divergence [Rousset's $F_{ST}/(1 - F_{ST})$] are compared using analysis of covariance (SPSS), the ME 2007 samples are seen to be clearly significantly higher than the MA 2007 samples (Fig. 2B; Table 7); hence, the MA subpopulations are significantly less divergent genetically than the ME samples.

Private allele analysis

This analysis gave, for the largest directly comparable data sets available (April to June) the following results (with N_m being corrected for sample size):

MA April/May:
mean sample size = 39.19; mean frequency of private allele, $p(1) = 0.04$; $N_m = 3.08$

ME April/May:
mean sample size = 7.14; mean frequency of private allele, $p(1) = 0.25$; $N_m = 0.27$

ME June:
mean sample size = 23.0; mean frequency of private allele, $p(1) = 0.09$; $N_m = 0.84$

Bayesian clustering analysis of population structure

For MA, average log-likelihoods across ten replicate STRUCTURE runs reached a first plateau at $K = 4$ (Fig. 3A; empirical and simulation evidence suggests that a biologically meaningful number of K may be

Table 7. Analysis of covariance of Rousset's (1997) measure of genetic divergence as a function of tansy aphid species (*Macrosiphoniella tanacetaria* vs. *Metopeurum fuscoviride*). Factor: species, covariate: distance, $F_{ST}/(1 - F_{ST})$ log-transformed for analysis. Dependent variable: ln $F_{ST}/(1 - F_{ST})$

	Tests of between group effects				
Source of variation	Sum of squares	d.f.	Mean squares	F	Significance
Corrected model	23.313*	2	11.656	11.679	0.000
Constant term	3.674	1	3.674	3.681	0.060
ln distance	0.352	1	0.352	0.353	0.555
Species	21.515	1	21.515	21.556	0.000
Error	53.897	54	0.998		
Total	164.198	57			
Corrected total variation	77.209	56			

*$R^2 = 0.302$ (corrected $R^2 = 0.276$). Highly significant difference among species/sampling dates; geographical distance not significant.

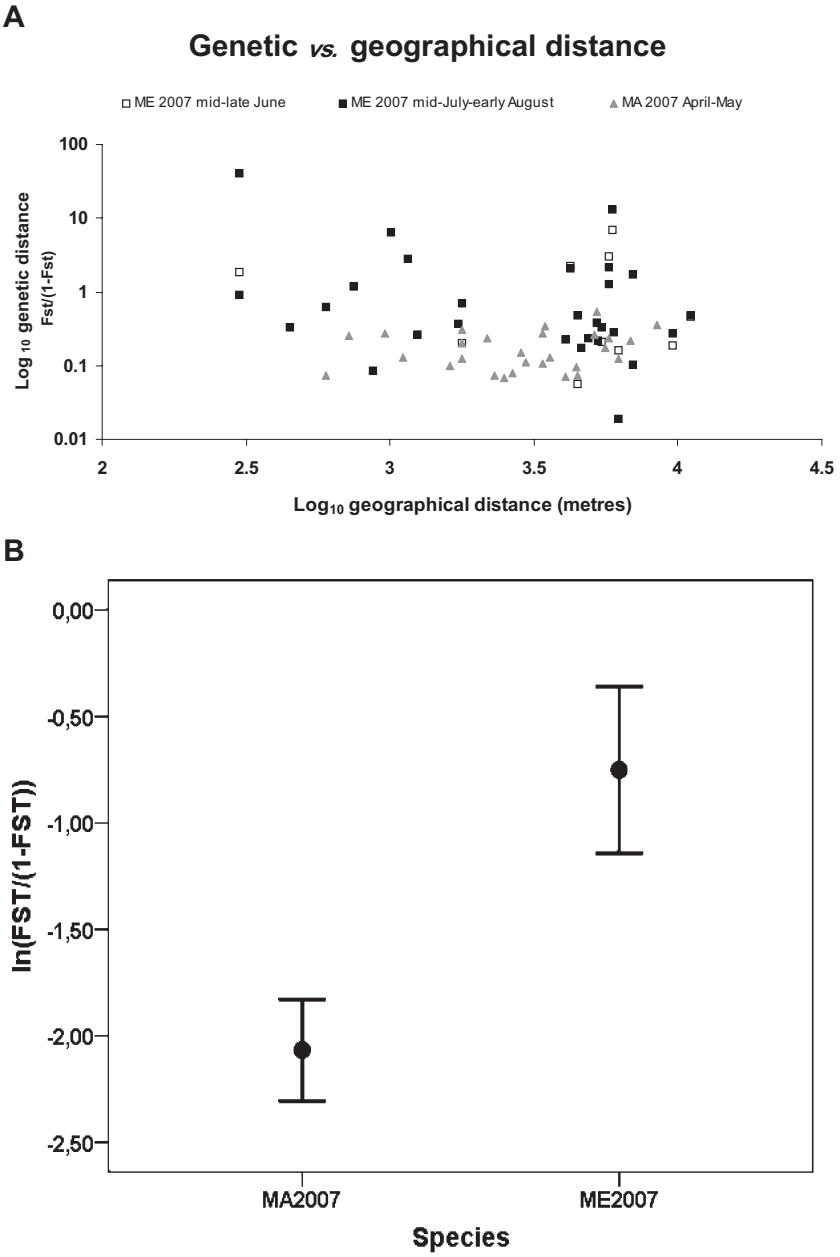

Figure 2. A, isolation by distance (IBD). The graph shows $\log_{10}$ genetic distance [as Fst/(1 – Fst)] versus $\log_{10}$ geographical distance (metres). For *Macrosiphoniella tanacetaria* (MA), only the larger spring samples were tested, whereas, for *Metopeurum fuscoviride* (ME), only the two larger summer (mid-late June) and late summer (mid July to early August) samples were used. B, relative 'migratoryness' of tansy aphids as measured after analysis of covariance. Plotted as $\ln[F_{ST}(1 - F_{ST})]$ versus MA and ME.

indicated by a declining rate of increase in $\Pr(X \mid K)$ as K increases rather than by the absolute maximum likelihood; Pritchard *et al*. 2000; Evanno *et al*. 2005). This estimate was matched by high coefficients of stability of individual assignment to clusters, S_N, until $K = 4$ (Table 8) and the highest posterior probability for $K = 4$ using STRUCTURAMA (Fig. 3A). Evanno's ΔK peaked at $K = 2$, but declined sharply only after $K = 4$ (Fig. 3A).

Individual assignments to clusters were largely concordant when K was fixed to 3 (using STRUCTURE) or treated as a random variable (using STRUCTURAMA) (Fig. 4A). Although, in the latter case, individuals were assigned to four different clusters, one minor cluster

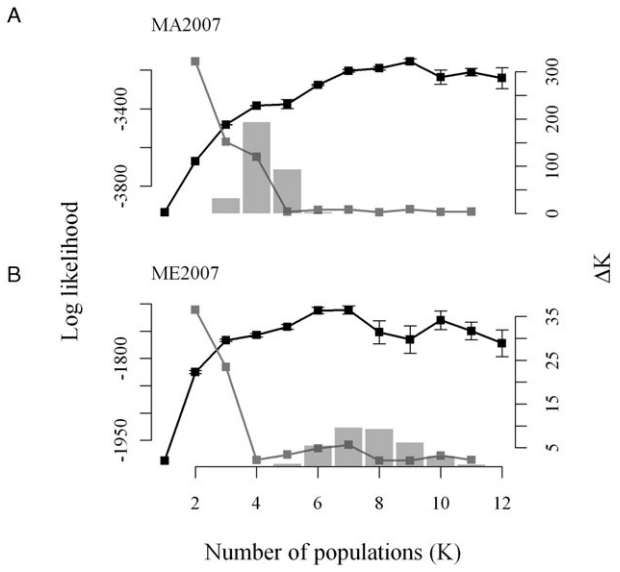

Figure 3. Estimates of the likely number of genetically homogeneous *Macrosiphoniella tanacetaria* (MA) (A) and *Metopeurum fuscoviride* (ME) (B) populations. Black squares show the marginal log likelihoods of the data $\Pr(X|K)$ when the number of populations (K) is fixed to different values averaged over six STRUCTURE runs. The grey squares denote ΔK, an *ad hoc* indicator of the uppermost hierarchical level of structure detected, based on the rate of change in $\Pr(X|K)$ between successive K-values. Grey bars denote the posterior probability distributions $\Pr(K|X)$ for the number of populations averaged over four STRUCTURAMA runs where K is treated as a random variable.

Table 8. Stability coefficients (S_N) for cluster assignment across ten independent STRUCTURE runs

	Macrosiphoniella tanacetaria	*Metopeuruma fuscoviride*
$K = 2$	**> 0.99**	**> 0.99**
$K = 3$	**> 0.99**	**0.94**
$K = 4$	**> 0.99**	0.63
$K = 5$	0.96	0.58
$K = 6$	0.92	0.54
$K = 7$	0.95	0.42
$K = 8$	0.79	0.37

Bold face = highest coefficients of stability.

defined by STRUCTURAMA (the five 'yellow' MLGs in Fig. 4A) did not match the fourth cluster when K was fixed to 4 using STRUCTURE. Both approaches placed the majority of individuals from sampling sites AA, BB, T, U, and V into one genetic group and the majority of individuals from sampling sites D and R into another. These groupings do not suggest any overt geographical pattern (Fig. 1). The third major group (the 'green' MLGs in Fig. 4A) was formed by all the individuals sampled during autumn irrespective of the location of the sampling sites. A few individuals sampled during early summer were assigned to the same group (Fig. 4A).

For ME, log-likelihoods derived from STRUCTURE plateaud at $K = 3$ (Fig. 3B), the S_N-values dropped after $K = 3$ (Table 8), whereas the posterior probabilities of the number of clusters derived from STRUCTURAMA were highest for $K = 7$ and 8 (Fig. 3B). Evanno's ΔK peaked at $K = 2$ and 3 (Fig. 3B). Thus, STRUCTURE does not support any substructuring beyond $K = 3$, whereas STRUCTURAMA assigns individuals to seven different clusters. Between these two methods, the assignments did not agree (Fig. 4B).

Spatial arrangement of genotypes

The Delaunay triangulation was based on midpoints of each in the triangles, confirmed with the Monmonier's algorithm analysis. This showed that with MA (spring samples only), the subpopulation BB (site near Kunitz), was genetically differentiated from the seven other subpopulations sampled (Fig. 5A). With the three pooled seasonal ME subpopulations, populations U and V were seen to be genetically differentiated (i.e. show restricted interpopulation gene flow) in terms of the remaining urban subpopulations sampled (B, D, IoE, K, and R) and the more distant suburban subpopulation X (Porstendorf) (Fig. 5B).

DISCUSSION

The present results confirm the metapopulation structure and dynamics of these two specialist tansy-feeding aphid species (Weisser, 2000; Weisser & Härri, 2005). At a purely ecological level, the data show that approximately 60% of sites harbouring tansy plants and previously known (Massonnet, 2002) to be colonized by one or more species of tansy-feeding aphids in the immediate area of Jena, were similarly colonized in the present study. However, when the data was further analyzed at the plant level per site, it was found that approximately 50% were infested with MA only, approximately 60% by ME only, and approximately 10% by both species together, mostly early in the growing season. From this, it appears that the two species preferentially infest tansy plants singly rather than together. There was also some evidence that MA tended to colonize tansy plants slightly earlier than ME, whereas, sometimes, the former species showed a late season (around September) population increase (Massonnet, 2002; Loxdale *et al.*, 2011). It may well be that the guard ants,

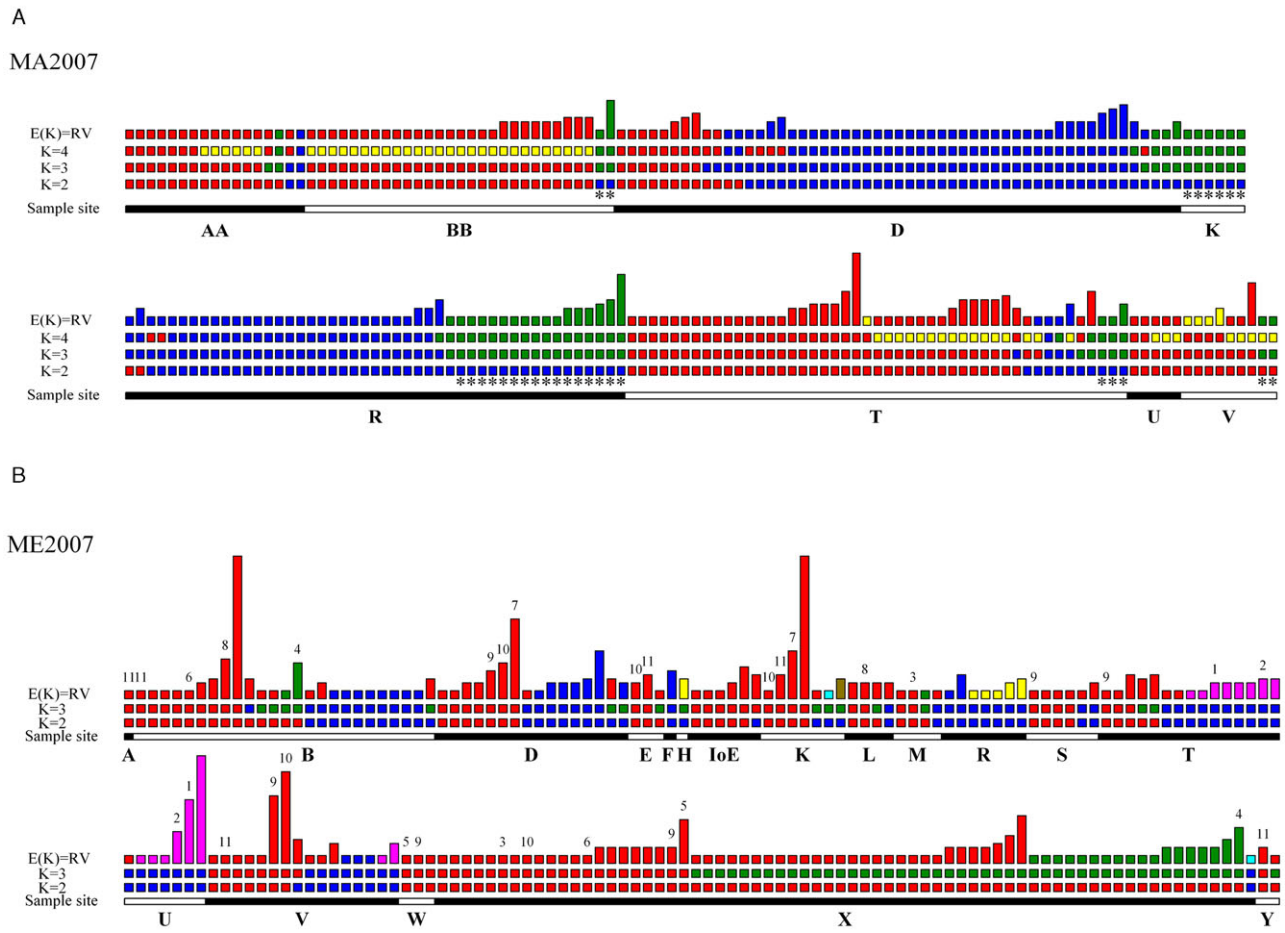

Figure 4. Assignment of multilocus genotypes (boxes) to different genetic clusters (indicated by colour). Analyses were performed either in STRUCTURE with the number of populations fixed ($K = 2$, 3, or 4) or in STRUCTURAMA 'with the number of populations treated as a random variable [$E(K) = \mathrm{RV}$]. A, analyses of the *Macrosiphoniella tanacetaria* (MA) sample. Letters indicate sampling sites (Fig. 1). Asterisks beneath boxes indicate multilocus genotypes (MLGs) sampled during autumn. The height of the boxes in the topmost row is proportional to the number of individuals exhibiting a certain MLG. B, analyses of the *Metopeurum fuscoviride* (ME) samples. Letters indicate sampling sites (Fig. 1). The height of the boxes in the top most row is proportional to the number of individuals exhibiting a certain MLG. Numbers above boxes indicate identical MLGs sampled from different sites (See *Biological Journal of the Linnean Society* online for colour version of this figure).

L. niger and the larger, more aggressive *F. rufa*, the latter occurring especially at the northernmost sites sampled, Porstendorf and Dornburg (sites X and Y, respectively; Fig. 1), may significantly contribute to the observed spatial structuring of the two species by actively killing and/or disturbing MA The behaviour of the two aphid species is different: the ant attended *M. fuscoviride* is calm and is not easily induced to fall off the plant, whereas the non-ant attended (and hence unguarded) *M. tanacetaria* is decidedly more nervous and readily responds to disturbance by falling.

The late increase of MA, if and when it comes towards autumn when sexual forms are produced, may arise because the aphid remains in low numbers in the flower heads and may thus escape detection, especially in plants not occupied by ME and thus (usually) ants. Furthermore, the apparent temporal separation of the species, namely the generally earlier population rise of MA (mainly on shoots) and later rise of ME (mainly a stem feeder), is a result of the fact that the ants kill or drive off (by disturbing them so that they fall off the plant) the former species as the season progresses. The competition between the two species (Table 1) is itself then a kind of 'apparent competition'; the ants may predate both species, especially in times of scare food resources (as well as farming ME in the more prosperous times of summer, including perhaps even carrying the aphids to their nests and between ramets and genets) but preferentially eliminate MA aphids.

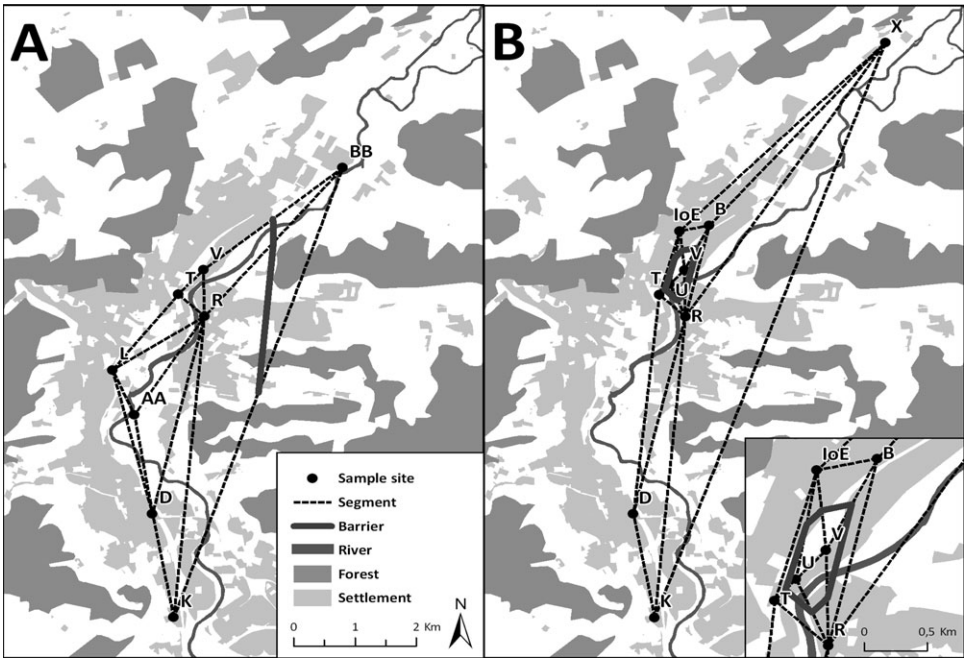

Figure 5. Geographical locations of *Macrosiphoniella tanacetaria* (MA) (A) and *Metopeurum fuscoviride* (ME) (B) samples The lines between locations represent the connectivity network created by Delaunay triangulation. Monmonier's algorithm was used to detect a genetic barrier, which separate in (A) the BB population and in (B) the V and U populations.

The observation that MA and ME are usually spatially separate as a result of a competition-like effect(s) induced by ants may be a prime reason for the maintenance of the dual metapopulation structuring of these two tansy aphid species. In this respect, it is also of interest that both aphids have preferred feeding sites on individual plants (as previously noted), again showing that they generally remain separate even at this microspatial scale, perhaps also as a result of some kind of competition for the direct resource of phloem (i.e. resource partitioning) (Denno, Mcclure & Ott, 1995; Stewart, 1996). However, there may also be plant-related anatomical reasons for this separation, reflected in morphological adaptations–specializations, both of the plant and aphids.

Our molecular results also confirm the findings of earlier studies using polymorphic microsatellite markers (Massonnet, 2002; Massonnet *et al*., 2002b; Massonnet & Weisser, 2004; Loxdale *et al*., 2010, 2011), demonstrating that populations of the tansy-feeding aphid species, MA and ME, show a clear metapopulation structure as seen from the spatio-temporal patterns of allele and genotype frequencies. Values for allelic variation and observed and expected heterozygosity for the two species are largely comparable with the results obtained in these earlier studies [although, in ME; Massonnet (2002) found the locus Me-6 to be polymorphic in subpopulations from Germany and Alsace, France, with a maximum of six alleles versus the single allele found in the present study] and many subpopulations deviated significantly from HWE, emphasizing their clonal nature and the effect of abiotic/biotic and stochastic effects on genotype frequencies (Massonnet, 2002; Massonnet *et al*., 2002b; Massonnet & Weisser, 2004). The high overall F_{ST} values (MA = 0.15; ME ~ 0.4), are much greater than found in highly migratory aphids such as *S. avenae* (F.) (i.e. << 0.05; Llewellyn *et al*., 2003) and the mostly positive F_{IS} values obtained in both tansy-feeding species (Tables 2, 3; except the small late sample of MA, which was negative, i.e. heterozygote excess) suggests that both are inbred with little interpopulation movement. At the same time, the fact that pairwise comparisons of allelic variance often showed linkage disequilibrium in both species is expected for organisms that indulge in long periods of parthenogenetic reproduction during the spring and summer months (Massonnet, 2002; Massonnet *et al*., 2002b).

Further support for the notion that subpopulations of the two tansy aphids are genetically isolated, even over the small spatial scale sampled, is given by the fact that genic (allelic) and genotypic tests of subpopulation divergence are all significant. Furthermore, the finding that the MA colonies tested as sampled from site D (Burgau Park) using the highly

polymorphic locus Ma-6 were seen to be highly differentiated (Table 5), even between tansy genets a few metres apart, whilst sometimes samples from ramets on the same plant were clearly different (data not shown), emphasizes the extreme population genetic heterogeneity.

Doubtless, these findings relate to the original foundation events of individual colonies by one or a few winged asexual females during the summer flight period of the aphids (April/July; Massonnet, 2002; Massonnet *et al*., 2002b; Loxdale *et al*., 2011), a differentiation that is maintained to some large extent during the subsequent course of the summer (Loxdale *et al*., 2011). This is either because of the unwillingness of the aphids in terms of the urge and/or ability to migrate between plants of their own or of other nearby stands, or a result of their general inability to successively locate and land on these during migratory flights (see below). Our results clearly provide evidence for a further case of limited inter subpopulation migration and hence gene flow in aphids (i.e. two tansy feeding species). As such, the population genetic patterns of the species are very different from assumed long distance migrants such as the predominantly anholocyclic grain aphid, *S. avenae* (Llewellyn *et al*., 2003), and much more similar to the patterns found in its predominantly holocyclic sister species the blackberry-grain aphid, *Sitobion fragariae* (Walker) (Loxdale & Brookes, 1990), as well as the holocyclic damson-hop aphid, *Phorodon humuli* (Schrank) (Loxdale *et al*., 1998), only more so (i.e. even much more locally heterogeneous) (Loxdale & Lushai, 2007).

The discovery of the large allele in ME (Me-3^{214}), three times the size of the original cloned dinucleotide $(TG)_{20}$ microsatellite (Massonnet, 2002; Massonnet *et al*., 2002a), is very interesting. [Massonnet's original cloned ME locus-3 allele was Me-3^{134}; the commonest alleles in the 2002 study of aphids comprising a large spatial study were from Breisac, Alsace, France (Me-3^{134}, $f = 0.80$) and Munster, Germany (Me-3^{146}, $f = 0.86$)]. It is apparently unique in aphids at such hypervariable loci, although similar length variation is known in the intergenic spacer (IGS) regions of the ribosomal DNA (rDNA) cistron of the peach-potato aphid, *Myzus persicae* (Sulzer) (Fenton *et al*., 2003, 2005). As with such ribosomal DNA regions, the microsatellite size expansion could be the result of unequal crossing over at meiosis, perhaps as a result of multiple rounds in different sexual generations of the same original mutated lineage [Fenton *et al*., 2003; in relation to insecticide selection of rDNA IGS bands in the Greenbug aphid, *Schizaphis graminum* (Rondani) (Hemiptera: Aphididae), see also Shufran, Mayo & Crease, 2003]. The carboxylesterase-4 (E4) gene in *M. persicae*, which confers cross resistance to pesticides, is also amplified in the resistant strains but, here, the whole E4 gene is involved (= amplicon) (Field & Blackman, 2003). Alternatively, the large Me-3 allele may have arisen from a direct mutation event such as insertion in the genome. This view is supported by the observation that the distribution of pairwise distances among individuals in terms of microsatellite allelic divergence showed a tri-phasic pattern (see Results). Because the aforementioned large Me-3 allele was not seen in the ME populations sampled by Massonnet in Germany and Alsace, France in 1999 (Massonnet, 2002), this event has apparently occurred in the local population subsequent to that period, as was first noted by us in preliminary samples collected around Jena in summer 2006 (data not shown) and later in 2007, for the present study (Table 6). The fact that individuals bearing this large allele are only found in some subpopulations sampled, and not others, suggests that the lineage bearing this allele has not yet had time to spread throughout the entire population.

In terms of clonal structure, subpopulations of both MA and ME are dominated by unique MLGs or low copy (repeat) number MLGs, with larger copy number MLGs only appearing later in the growing season (e.g. ME cl. 195 with 143 copies spread over four sites: B, D, K, and V, in July/August) (Table 4B). The increase of clonality with time is also shown with the decline of average D^* values: in MA, from 0.79 in April/May to 0.56 in September/October (Table 2) and, for ME, from 0.81 in April/May to approximately 0.23 later on (Table 3). These findings are explored in more detail elsewhere (Loxdale *et al*., 2010), and essentially show that, at the time of peak flight abundance (June/July), populations of both tansy aphid species, presumably also including aerial samples, are not dominated by particular clones because these are apparently continually being eliminated by intense clonal selection/competition.

With regard to hierarchical examination of the data sets for the two species, in MA, allele range was seen to decline from a mean of approximately 7.3 to 6.9 and 6.1 for site, genet, and ramet, respectively, and this involved a loss of approximately five alleles in total (from 19 to 14), whereas the average number of alleles per locus declined from approximately 3.6, 2.2, and 2.1, respectively. In ME, the allele range was seen not to change from a mean of approximately 4.5 (1–13 alleles total) between samples at the site, genet, and ramet level, although the mean number of alleles per locus declined from approximately 4.5, to 1.5 and 1.4 for the three sampling levels, respectively. These results strongly suggest that the major change is between site and genet/ramet, with little difference between genet and ramet. This means that the genet, rather than site, is the level of population genetic variance in both species, and such a result further

emphasizes the fine-grain level of heterogeneity present within the metapopulation of these aphids, and hence the very restricted levels of interpopulation gene flow.

AMOVA showed that, in MA, the larger, earlier sample possessed less 'among population' variance (approximately 13%), than the smaller autumn samples (approximately 20%). The higher value means that there is slightly less interpopulation gene flow in these late season samples compared to the early ones, although the sample sizes are very different. There were few sexual forms recorded in the later sample (indeed, only one winged male and one ovipara; Table 2). Hierarchical partitioning revealed that the 'among population' variance increased to approximately the same value for both genet and ramet (30%), showing the genet to be the unit of genetic diversity; in effect, a green island(s) of plant resource in an otherwise unsuitable habitat for the aphids.

In ME, 'among population' variance ranged at the site level (June and July/August) from approximately 40–50% to 12–24% 'among individuals within populations' and 36–40% 'within individuals' (see Results). First, compared with the MA subpopulations, it is immediately clear than ME populations are much more differentiated at the 'among population' variance level. That the 'among population' variance level declines from approximately 50% to 40% between the two sampling periods is support for a declining level of aerial movement (gene flow) as the season progresses. Unlike MA subpopulations, most variation is found among populations, with the 'within individuals' category coming next. In all three ME seasonal populations, the proportion of winged individuals was negligible (ranging from 0% to 2.6% overall). In terms of hierarchical changes, 'among population' variance increases at the genet-ramet scale to approximately 53% and at the ramet scale to approximately 58%. These data also support the view that, in ME too, the genet is the unit of population variance.

F_{ST} as a function of hierarchy was seen to increase in MA from approximately 0.15 to 0.3 for both genet–ramet and ramet alone. Meanwhile, the level of significance for pairwise comparisons fell from 100% to approximately 85% in both cases. In ME, F_{ST} increased from approximately 0.4 to 0.5 in both cases, whereas the level of significance of pairwise comparisons fell from to approximately 100% to 85% and 70% in genet–ramets and ramets, respectively. These data again show that the level of genetic differentiation of colonies is at the genet rather than at the site level.

The spread of genotypes among plants showed contrasting patterns in the two species. In MA, interestingly, no MLGs were found beyond the site and season from which they were collected (Table 4A). In contrast, even some rare ME genotypes showed spatio- and or spatio-temporal changes in distribution, whereas abundant ME clones such as cl. 195 were distributed over a number or plants and sites (maximum 4) and seasons (Table 4B). This apparently paradoxical situation may perhaps be explained by assuming that ME have a slightly different flight behaviour than MA (ratio of attraction to white light: green targets; Hardie, 1993) leading to comparatively more local landings and hence colony formation than MA. Either way, both species are very rare as winged forms in 12.2-m high suction traps samples run by the Rothamsted Insect Survey (Harrington, Hullé & Plantegenest, 2007) over many years (see Loxdale *et al*., 2010).

There was a failure to find any significant IBD relationship [$F_{ST}/(1 - F$ST)] in MA (Massonnet & Weisser, 2004) at even large spatial scales, and this was only the case above approximately 450 km. For ME, the data were preliminary and consisted of seven geographical populations [five French (Alsace) and two German]; there was no significant IBD trend (Rousset's 1997 method) in the Alsace data collected locally (< 50 km apart), although there was between pairwise assessments of the Alsace and German populations (between approximately 354 and 624 km. apart) (Massonnet, 2002: fig. 2). In the present study, similalrly, there was no significant relationship between measures of population genetic divergence and geographical distance for either MA and ME over the area sampled (maximum approximately 16 km north–south; Fig. 2A).

The fact that IBD was nonsignificant in the two cases tested (MA and ME) suggests that, despite the problems of successfully locating plant hosts, enough winged migrants do ultimately succeed and by so doing, offset the differentiating effects of selection and genetic drift. Perhaps as earlier proposed by Wright (1990; see also Wang, 2004), one or a very few individuals only are necessary to achieve this. The fact that, on replotting the IBD data after analysis of covariance, MA aphids subpopulations are significantly less differentiated than those of ME (both years; Fig. 2B; Table 7) reveals a fundamental difference in the biology/behaviour of the two species, probably related to the fact that MA has winged males, whereas ME does not (also an additional reason for the reduced F_{ST} in MA versus ME and the generally higher positive F_{IS} values in ME (average 0.24) versus the lower positive value for the large early MA collection, i.e. 0.09] (Tables 2, 3). We consider that this is the first time that such a direct comparative result has been shown in aphids feeding on the same plant host but with differing ecologies and lifecycle. It may be that such asymmetric dispersal patterns are a

contributory factor in the maintenance of the dual tansy aphid metapopulation structure (Salomon, Connolly & Bode, 2010).

The private allele analysis for the MA and ME subpopulation samples collected early in the season (April to June), for which direct comparative data exists, generally supports the notion of the former species being generally more migratory than the latter.

Bayesian analysis of the two species revealed that both were heterogeneous in terms of structure (Fig. 4). Thus, the relatively well-supported analysis for MA with eight polymorphic loci suggests that individuals may be grouped into two to four clusters (Fig. 3A), whereas, for the less well-supported ME with only five polymorphic loci out of six tested, the number of supported clusters was in the range two and eight (Fig. 3B), depending on the estimation method used. Moreover, for ME, there was only little concordance between the population subgroupings as estimated by STRUCTURE and STRUCTURAMA, respectively (Fig. 4B). This makes the biological relevance of the ME estimates rather difficult to assess. Thus, although ME may have a more heterogeneous population genetic patterning than MA as a result of being somewhat less aerially mobile than MA (Fig. 2B), the larger *K*-values in the case of ME are only supported by STRUCTURAMA. It has been suggested, however, that STRUCTURAMA could overestimate the number of biologically relevant subgroupings compared to STRUCTURE (Groot *et al*., 2011). Considering only the more conservative ΔK criterion as an indicator of the uppermost hierarchical level of structure supported by the data (Evanno *et al*., 2005), MA and ME hardly differ in the number of genetic clusters supported by the data (MA: 2–4; ME: 2–3).

The *K*-values derived in the present study for tansy aphids may be compared with that shown after STRUCTURE analysis for *M. persicae* (Fenton *et al*., 2010), a moderate-distance migrant (Wilson *et al*., 2002; Guillemaud, Mieuzet & Simon, 2003; Loxdale & Lushai, 2007), using six polymorphic microsatellite markers (Kasprowicz *et al*., 2008). In that study, 215 globally distributed sexual and asexual lineages could be optimally subdivided into three genetic clusters corresponding to a European, Australasian, and a Tobacco-specialized subgroup (Fenton *et al*., 2010: fig. 2). *Myzus persicae* is a global pest of great economic significance (Blackman & Eastop, 2000), which has been transported round the world in historical times by human agency (e.g. on the 'eyes' of sprouting potatoes), including to Australia, and may almost be said to comprise a global population (Wilson *et al*., 2002; Loxdale & Lushai, 2007). Thus, the tansy aphid populations as investigated at an extremely local scale in the present study appear to be as heterogeneous as samples of *M. persicae* collected around the world.

The local heterogeneity found in MA and ME using Bayesian approaches probably reflects a historical pattern of distribution produced over several growing seasons. Thus, despite the view (as outlined above) that there are apparently sufficient winged migrants to offset the effects of differentiating forces in terms of a lack of IBD relationship, over the course of a single or even several growing seasons, there clearly has been insufficient interpopulation gene flow to homogenize MA and ME population structuring in the area sampled in and around Jena. Rather, the populations of tansy aphids are genetically differentiated at small spatial scales within a growing season and apparently persist so in subsequent seasons, as also found earlier in the case of *S. fragariae* at small spatial scales (Loxdale & Brookes, 1990: fig. 3). In addition to the normal problems tansy aphids face in locating other stands of tansy to colonize within the metapopulation, such genetic differentiation is also probably maintained and indeed enhanced by the geographical barriers (e.g. buildings) between some of the collecting sites within the town of Jena itself, which may hamper interpopulation gene flow, as emphasized by the spatial analysis (ALLELES-IN-SPACE) (Fig. 5). However, it should be stated that ME subpopulation U was collected from one large plant and ME subpopulation V from a few nearby plants, approximately 500 metres away (Fig. 5B) and both were very inbred; thus, this is likely responsible for some of the differentiation seen between these and other ME subpopulations.

In conclusion, from the findings obtained in the present study, the two species of tansy aphids MA and ME are largely spatially isolated within the metapopulation structure and both are extremely immobile, certainly compared to migratory pest species such as *S. avenae*, *M. persicae*, and the holocyclic bird cherry-oat aphid, *Rhopalosiphum padi* (L.), common at peak flight times in 12.2-m high suction trap samples (Woiwod *et al*., 1988), and with generally low F_{ST} values (*M. persicae* < 0.05: Wilson *et al*., 2002; Guillemaud *et al*., 2003; *R. padi*, sexual populations, $F_{ST} = 0.026$; asexual populations = 0.169: Delmotte *et al*., 2002). In terms of local genetic heterogeneity, the tansy aphid population genetic structure is more similar to species of low to middle range flight urge and/or ability, such as *S. fragariae* and *P. humuli* (Loxdale & Lushai, 2007), although they are probably less mobile even than these species as the data from the present study strongly suggest (i.e. the winged forms appear to have difficulty even moving over distances of a few kilometres, although some individuals may be carried higher by prevailing winds

and hence travel longer distances). That historical patterns of genetic structuring locally are heterogeneous, comprising between three to seven genotypic groupings, argues that colonies have remained isolated on their plant resource patches for several growing seasons at least, may be longer, a likelihood promoted by the metapopulation structure and the behaviour of the aphids themselves. This involves weak powers of aerial dispersal with concomitant inbreeding (undoubtedly involving mating of sexuals derived from clone mates, especially in ME and also involving the laying of overwintering sexual eggs on the same of nearby host plant stands), and perhaps even with a degree of philopatry, actively selected for because of the widespread, heterogeneous distribution of the plant patches themselves.

Considering that aphids are passively borne on the wind (when actively flying), it must be extremely difficult for such specialist aphids to successfully locate and land on the requisite host plant species, a scenario that is even difficult for very common pest species such as *R. padi* seeking, in autumn, their primary woody overwintering hosts (Ward *et al*., 1998). With MA and ME, such difficulties of inter-colony migration and hence gene flow are apparently very difficult even at the small spatial scale, as studied in Jena in the present study. Lastly, there was no evidence in the present study of any dominant clones (i.e. superclones) in the samples tested, which represents another difference between *M. persicae* and other pest species that live in agro-ecosystems comprising large areas of intensively cultivated monocultures.

ACKNOWLEDGEMENTS

We thank Sylvia Creutzburg for her assistance with running PCRs; Dr Kathrin Henniges-Janssen for help with DNA purification and allele size determination; Norma Nitschke for statistical advice; Christina Braun and Osama Mustafa for cartographic assistance; Dr Tobias Kaiser for advice on certain molecular analyses; Mohsen Mehrparvar for enlightening discussions on aphid ecology; and last, but no means least, Dr Nicola von Mende-Loxdale for help both in collecting tansy aphids and for valuable advice on matters relating to PCR. We also wish to thank the three anonymous reviewers for their most helpful comments on the manuscript, which have greatly improved the paper. H.D.L. performed the experimental work during the tenure of an EU Marie-Curie Senior Research Fellowship. F.N.N. thanks the Deutscher Akademischer Austausch Dienst (DAAD) for financial support and K.R.W thanks the German Ministry for Education and Research through the BIOLOG program (project DIVA-Jena 01LC0613A2).

REFERENCES

Arnaud-Haond S, Alberto F, Teixeira S, Procaccini G, Serrão EA, Duarte CM. 2005. Assessing genetic diversity in clonal organisms: low diversity or low resolution? Combining power and cost efficiency in selecting markers. *Journal of Heredity* **96:** 434–440.

Arnaud-Haond S, Belkhir K. 2007. GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes* **7:** 15–17.

Arnaud-Haond S, Duarte CM, Alberto F, Serrão EA. 2007. Standardizing methods to address clonality in population studies. *Molecular Ecology* **16:** 5115–5139.

Barton NH, Slatkin M. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* **56:** 409–415.

Blackman RL, Eastop VF. 2000. *Aphids on the world's crops: an identification and information guide*, 2nd edn. Chichester: John Wiley & Sons Ltd.

Blackman RL, Eastop VF. 2006. *Aphids on the world's herbaceous plants and shrubs. Volume 2: the aphids*. Chichester: John Wiley & Sons Ltd.

Bohonak AJ. 2002. IBD (isolation by distance): a program for analyses of isolation by distance. *Journal of Heredity* **93:** 153–154.

Brouns G, De Wulf A, Constales D. 2003. Delaunay triangulation algorithms useful for multibeam echosounding. *Journal of Surveying Engineering-Asce* **129:** 79–84.

Delmotte F, Leterme N, Gauthier JP, 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.

Denno RF, Mcclure MS, Ott JR. 1995. Interspecific interactions in phytophagous insects – competition reexamined and resurrected. *Annual Review of Entomology* **40:** 297–331.

Dixon AFG. 1998. *Aphid ecology*, 2nd edn. London: Chapman & Hall.

Dorken ME, Eckert CG. 2001. Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *Journal of Ecology* **89:** 339–350.

ESRI. 2003. Data & maps. Redlands, CA: ESRI.

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

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

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164:** 1567–1587.

Fenton B, Malloch G, Navajas M, Hillier J, Birch ANE. 2003. Clonal composition of the peach-potato aphid *Myzus*

persicae (Homoptera: Aphididae) in France and Scotland: comparative analysis with IGS fingerprinting and microsatellite markers. *Annals of Applied Biology* **142:** 255–267.

Fenton B, Malloch G, Woodford JAT, Foster SP, Anstead J, Denholm I, King L, Pickup J. 2005. The attack of the clones: tracking the movement of insecticide resistant peach-potato aphids *Myzus persicae* (Hemiptera: Aphididae). *Bulletin of Entomological Research* **95:** 483–494.

Fenton B, Margaritopoulos JT, Malloch G, Foster SP. 2010. Micro-evolutionary change in relation to insecticide resistance in the peach-potato aphid, *Myzus persicae*. *Ecological Entomology* **35** (Suppl. 1)**:** 131–146.

Field LM, Blackman RL. 2003. Insecticide resistance in the aphid *Myzus persicae* (Sulzer): chromosome location and epigenetic effects on esterase gene expression in clonal lineages. *Biological Journal of the Linnean Society* **79:** 107–113.

Goudet J. 2002. FSTAT, a program to estimate and test gene diversities and fixation indices. 2.9.3.2 edn. Available at: http://www2.unil.ch/popgen/softwares/fstat.htm

Gould CG. 2004. *The remarkable life of William Beebe, explorer and naturalist*. 194. Washington, DC: Island Press, Shearwater Books.

Groot AT, Classen A, Inglis O, Blanco CA, López JJ, Téran Vargas AT, Schal C, Heckel DG, Schöfl G. 2011. Genetic differentiation across North America in the generalist moth *Heliothis viriscens* and the specialist *H. subflexa*. *Molecular Ecology* **20:** 2676–2692.

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.

Haack L, Simon JC, Gauthier JP, Plantegenest M, Dedryver CA. 2000. Evidence for predominant clones in a cyclically parthenogenetic organism provided by combined demographic and genetic analyses. *Molecular Ecology* **9:** 2055–2066.

Hanski I. 1999. *Metapopulation ecology*. Oxford: Oxford University Press.

Hardie J. 1993. Flight behaviour in migrating insects. *Journal of Agricultural Entomology* **10:** 239–245.

Harrington R, Hullé M, Plantegenest M. 2007. Monitoring and forecasting. In: van Emden HF, Harrington R, eds. *Aphids as crop pests*. Wallingford: CABI, 515–536.

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

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

Jensen JL, Bohonak AJ, Kelley ST. 2005. Isolation by distance, web service. *BMC Genetics* **6:** 13.

Kasprowicz L, Malloch G, Pickup J, Fenton B. 2008. Spatial and temporal dynamics of *Myzus persicae* clones in fields and suction traps. *Agricultural and Forest Entomology* **10:** 91–100.

Keller L, Gordon E. 2010. *The lives of ants*. Oxford: Oxford University Press.

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, Brookes CP. 1990. Genetic stability within and restricted migration (gene flow) between local populations of the blackberry-grain aphid *Sitobion fragariae* in south-east England. *Journal of Animal Ecology* **59:** 495–512.

Loxdale HD, Brookes CP, Wynne IR, Clark SJ. 1998. Genetic variability within and between English populations of the damson-hop aphid, *Phorodon humuli* (Hemiptera: Aphididae), with special reference to esterases associated with insecticide resistance. *Bulletin of Entomological Research* **88:** 513–526.

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

Loxdale HD, Lushai G. 2007. Population genetic issues: the unfolding story revealed using molecular markers. In: van Emden HF, Harrington R, eds. *Aphids as crop pests*. Wallingford: CABI, 31–67.

Loxdale HD, Macdonald C. 2004. Tracking parasitoids at the farmland field scale using microsatellite markers. In: Werner D, ed. *Biological resources & migration*. Berlin: Springer-Verlag, 107–125.

Loxdale HD, Massonnet B, Schöfl G, Weisser WW. 2011. Evidence for a quiet revolution: seasonal variation in colonies of the specialist tansy aphid, *Macrosiphoniella tanacetaria* (Kaltenbach) (Homoptera, Aphididae) studied using microsatellite markers. *Bulletin of Entomological Research* **101:** 221–239.

Loxdale HD, Massonnet B, Weisser WW. 2010. Why are there so few aphid clones? *Bulletin of Entomological Research* **100:** 613–622.

Manni F, Guerard E, Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Human Biology* **76:** 173–190.

Massonnet B. 2002. *Metapopulation structure and population genetics of monophagous aphids on tansy* (Tanacetum vulgare L.). PhD thesis, University of Basel, Switzerland.

Massonnet B, Leterme N, Simon J-C, Weisser WW. 2001. Characterization of microsatellite loci in the aphid species *Macrosiphoniella tanacetaria* (Homoptera, Aphididae). *Molecular Ecology Notes* **1:** 14–15.

Massonnet B, Leterme N, Simon J-C, Weisser WW. 2002a. Characterization of microsatellite loci in the aphid species *Metopeurum fuscoviride* (Homoptera, Aphididae). *Molecular Ecology Notes* **2:** 127–129.

Massonnet B, Simon J-C, Weisser WW. 2002b. Metapopulation structure of the specialised herbivore *Macrosiphoniella tanacetaria* (Homoptera, Aphididae). *Molecular Ecology* **11:** 2511–2521.

Massonnet B, Weisser WW. 2004. Patterns of genetic

differentiation between populations of the specialised herbivore *Macrosiphoniella tanacetaria* (Homoptera: Aphididae). *Heredity* **93:** 577–584.

Miller MP. 2005. Alleles-In-Space: computer software for the joint analysis of inter-individual spatial and genetic information. *Journal of Heredity* **96:** 722–724.

Nowak H, Komor E. 2010. How aphids decide what is good for them: experiments to test aphid feeding behaviour on *Tanacetum vulgare* (L.) using different nitrogen regimes. *Oecologia* **163:** 973–984.

Nyabuga FN, Loxdale HD, Heckel DG, Weisser WW. 2010. Spatial population dynamics of a specialist aphid parasitoid, *Lysiphlebus hirticornis* Mackauer (Hymenoptera: Braconidae: Aphidiinae): evidence for philopatry and restricted dispersal. *Heredity* **105:** 433–442.

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

Pettersson J, Tjallingii WF, Hardie J. 2007. Host-plant selection and feeding. In: van Emden HF, Harrington R, eds. *Aphids as crop pests*. Wallingford: CABI, 87–113.

Pickett JA, Glinwood RT. 2007. Chemical ecology. In: van Emden HF, Harrington R, eds. *Aphids as crop pests*. Wallingford: CABI, 235–260.

Powell W, Pell JK. 2007. Biological control. Chemical ecology. In: van Emden HF, Harrington R, eds. *Aphids as crop pests*. Wallingford: CABI, 469–513.

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

Raymond M, Rousset F. 1995. Genepop (version 4.0) – population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86:** 248–249.

Richards CM, Volk GM, Reilley AA, Henk AD, Lockwood DR, Reeves PA, Forsline PL. 2009. Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species of domesticated apple. *Tree Genetics & Genomes* **5:** 339–347.

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

Salomon Y, Connolly SR, Bode L. 2010. Effects of asymmetric dispersal on the coexistence of competing species. *Ecology Letters* **13:** 432–441.

Shufran KA, Mayo ZB, Crease TJ. 2003. Genetic changes within an aphid clone: homogenization of rDNA intergenic spacers after insecticide selection. *Biological Journal of the Linnean Society* **79:** 101–105.

Stewart AJA. 1996. Interspecific competition reinstated as an important force structuring insect herbivore communities. *Trends in Ecology and Evolution* **11:** 233–234.

Sunnucks P, Hales DF. 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I–II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Molecular Biology and Evolution* **13:** 510–524.

Taylor LR. 1974. Insect migration, flight periodicity and boundary layer. *Journal of Animal Ecology* **43:** 225–238.

Wang J. 2004. Application of the one-migrant-per-generation rule to conservation and management. *Conservation Biology* **18:** 332–343.

Ward SA, Leather SR, Pickup J, Harrington R. 1998. Mortality during dispersal and the cost of host specificity in parasites: how many aphids find hosts? *Journal of Animal Ecology* **67:** 763–773.

Watson DF, ed. 1992. *Contouring: a guide to the analysis and display of spatial data*. 321. New York, NY: Pergamon Press.

Weisser WW. 2000. Metapopulation dynamics in an aphid-parasitoid system. *Entomologia Experimentalis et Applicata* **97:** 83–92.

Weisser WW, Härri SA. 2005. Colonisations and extinctions at multiple spatial scales: a case study in monophagous herbivores. *Annales Zoologici Fennici* **42:** 363–377.

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

Woiwod I, Tatchell GM, Dupuch MJ, Macaulay EDM, Parker SJ, Riley AM, Taylor MS. 1988. Rothamsted Insect Survey: nineteenth annual summary: suction traps 1987. *Annual Report of Rothamsted Experimental Station, 1987, Part 2*, pp. 195–229.

Wright S. 1990. Evolution in Mendelian populations (reprinted from *Genetics* **16:** 97–159, 1931). *Bulletin of Mathematical Biology* **52:** 241–295.

ARCHIVED DATA

Data deposited at Dryad: doi:10.5061/dryad.bs325