Heredity **63** (1989) 245–266

Received 11 April 1989

Gene frequency clines for host races of *Rhagoletis pomonella* in the midwestern United States

Jeffrey L. Feder* and
Guy L. Bush

Department of Zoology, Michigan State University, E. Lansing, Michigan 48824, U.S.A.

Speciation in many host specific parasites may be initiated in sympatry when populations shift and adapt to new hosts. The recent shift of the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae) from its native host plant hawthorn (*Crataegus* spp.) to introduced, domesticated apple (*Malus pumila*), provides a direct test of the "sympatric speciation" hypothesis by indicating whether partially reproductively isolated "host races" can evolve in the absence of geographic isolation. We report finding significant allele frequency differences for six allozymes between paired apple and hawthorn infesting populations of *R. pomonella* from across the midwestern United States. Latitudinal allele frequency clines exist among both apple and hawthorn populations, however, for a majority of the loci displaying racial differences. Inter-host genetic differentiation is therefore superimposed on clinal patterns of variation within the races such that the magnitude of host associated divergence is a function of latitude. The results indicate that host associated races can form in sympatry and implicate differences in host plant recognition and developmental timing (related to ambient temperature) as key factors restricting gene flow between apple and hawthorn populations. However, some of the same processes differentiating apple and hawthorn populations at sympatric sites also appear to be occurring within the two host races across their respective ranges. *R. pomonella* populations are therefore diverging with respect to both their host plant affiliations and local environmental conditions.

INTRODUCTION

Speciation, the rendering of one closed genetic system into two, is a fundamental process responsible for the evolution of life on earth. Unfortunately, because speciation is a relatively rare, often time consuming event, much of our understanding of the process is based on population genetics models and indirect analyses of closely related species. We therefore have very few examples directly documenting how populations become reproductively isolated from one another.

The apple maggot fly, *Rhagoletis pomonella*, is an exception, however, having a biology and natural history which makes it ideal for the empirical study of speciation. *R. pomonella* is part of an endemic, sibling species complex of true fruit flies whose members overlap broadly in their geographic distributions across North America (Bush, 1966). *Rhagoletis* larvae are internal parasites in the fruit of their hosts, with each species in the group being monophagous or oligophagous for a different set of host plants (Bush, 1966). Adult females lay their eggs directly into the host fruit which they identify by specific visual, tactile and olfactory cues (Prokopy, 1968*a*; Prokopy *et al.*, 1973, 1987, 1988; Moericke *et al.*, 1975; Fein *et al.*, 1982; Owens and Prokopy, 1986; Papaj and Prokopy, 1986). Males are attracted by the same cues and courtship and mating occurs almost exclusively on or near the fruits of the host plant (Prokopy *et al.*, 1971, 1972). Because host recognition and mate selection are directly coupled in *Rhagoletis*, variation for host preference and host associated survivorship traits can act as genetically based barriers to gene flow. These considerations led Bush (1966, 1969*a*, *b*, 1975) to propose that speciation in the *R. pomonella* group occurs sympatrically and is initiated when flies colonize and adapt to new host plants.

Verifying that sympatric host race formation can initiate speciation in the *R. pomonella* group is a three stage process. First, we must demonstrate that "host specific" traits can evolve within

* Present Address and Correspondence: Jeffrey L. Feder, Program in Ecology, Evolution and Behavior, Department of Biology, Princeton University, Princeton, New Jersey 08544, U.S.A. (609)-987-2889.

geographically continuous populations. Second, we must show that traits responsible for adapting fly populations to different host plants also form the basis for restricting inter-host gene flow. Finally, we have to determine whether host specialization is sufficient to eventually cause the complete reproductive isolation of fly populations. With regard to the first two questions, the recent shift of the species *R. pomonella* from its native host hawthorn (*Crataegus* spp.) to introduced, domestic apple (*Malus pumila*; Walsh, 1867) provides a unique ecological timeframe to determine whether host specialization is sufficient to differentiate populations in the absence of geographic barriers to gene flow. Agricultural records suggest a probable origin for the apple infesting "race" in the Hudson Valley, New York, and document the apple fly's subsequent spread across eastern North America over the last 120 years (see Illingworth, 1912; O'Kane, 1914; Bush, 1969*a*). Because domestic apples were introduced into the United States by European settlers (Chapman and Lienk, 1971), we have an upper limit of approximately 350 years for the age of the apple infesting form of *R. pomonella.*

Apple populations could have conceivably shifted from a host other than hawthorns. Besides hawthorns, however, the most likely other potential native North American hosts are several species of crab apple (*i.e.*, *M. angustifolia*, *M. diversifolia*, *M. glabrata*, *M. ioensis* and *M. coronaria*), none of which support populations of *R. pomonella* (O'Kane, 1914; Porter, 1928; Bush, 1966). Host acceptance and larval survivorship studies provide further indirect sypport for hawthorns as the original host of the apple fly. Prokopy *et al.* (1988) have shown that apple origin females prefer to oviposit into, and males reside on, hawthorn versus apple fruit (although apple origin flies do accept apple fruits significantly more often than hawthorn origin flies do; Prokopy *et al.*, 1988). In addition, survivorship is higher from egg to pupal stages for apple origin flies in hawthorn than in apple fruits (Prokopy *et al.*, 1988). These findings are consistent with *R. pomonella* having recently shifted to apples, with apple flies still retaining a degree of preference and adaptation for hawthorns as their ancestral host fruit.

In previous studies, we (Feder *et al.*, 1988) and co-workers (McPheron *et al.*, 1988*a*) have shown that sympatric apple and hawthorn populations of *R. pomonella* are genetically differentiated. Significant allele frequency differences exist for six allozyme loci between populations of the two host races collected from across the eastern United States (Feder *et al.*, 1989*a*). These six loci map to three different regions of the genome (Berlocher and Smith, 1983; Feder *et al.*, 1989*b*) and linkage disequilibrium occurs in natural populations between allozyme loci within each of these three regions (Feder *et al.*, 1988, 1989*a*). Allele frequencies for five of the six loci displaying inter-host differentiation also co-vary significantly with latitude among both apple and hawthorn populations (Feder *et al.*, 1989*a*). This suggests that north-south allele frequency clines exist for *R. pomonella.* However, the geographic analysis of eastern North America involved widely separated populations from a variety of different longitudes. Verification of clines therefore requires more detailed examination of fly populations collected along latitudinal transects.

The objective of the current study is to intensively characterize the geographic pattern of intra- and inter-host allozyme variation for *R. pomonella* in the midwestern United States. The results will let us confirm whether allele frequency clines exist for *R. pomonella* and allow an initial investigation into possible causal factors differentiating hawthorn and apple populations.

MATERIALS AND METHODS

Flies were collected from 34 different sites across the midwestern United States from 1985 to 1987 (fig. 1, table 1). The sites were organized into five north-south transects which ran from the states of Wisconsin through Illinois or from Michigan through Indiana (see fig. 1 for details). At 28 sites, both hawthorn (*C. mollis*) and feral apple (*M. pumila*) infesting populations were sampled. Host plants were sympatric (*i.e.*, separated by a distance of less than 100 metres) at 13 of the 28 paired locations (table 1). At ten of these sympatric sites apple and hawthorn trees were nearest neighbours, while at the other three sympatric sites (15, 19 and 20) flies were sampled from multiple apple and hawthorn trees within oldfields. No tree sampled at sites 15, 19 and 20 was separated from a tree of the opposite host species by a distance of more than 150 m, however. In only five instances were apple and hawthorn trees at a paired site located more than 4 km apart, with a maximum distance of 8 km at site 8 near Parkerdale, Indiana (table 1). Six of the 28 paired sites (listed as 9, 15, 19, 20, 26 and 28 in table 1) form part of the earlier allozyme study of eastern North America (these sites were designated 7, 5, 1, 2, 3 and 4, respectively, in Feder *et al.* 1989*a*). Sites 10, 13, 14 and 16 are

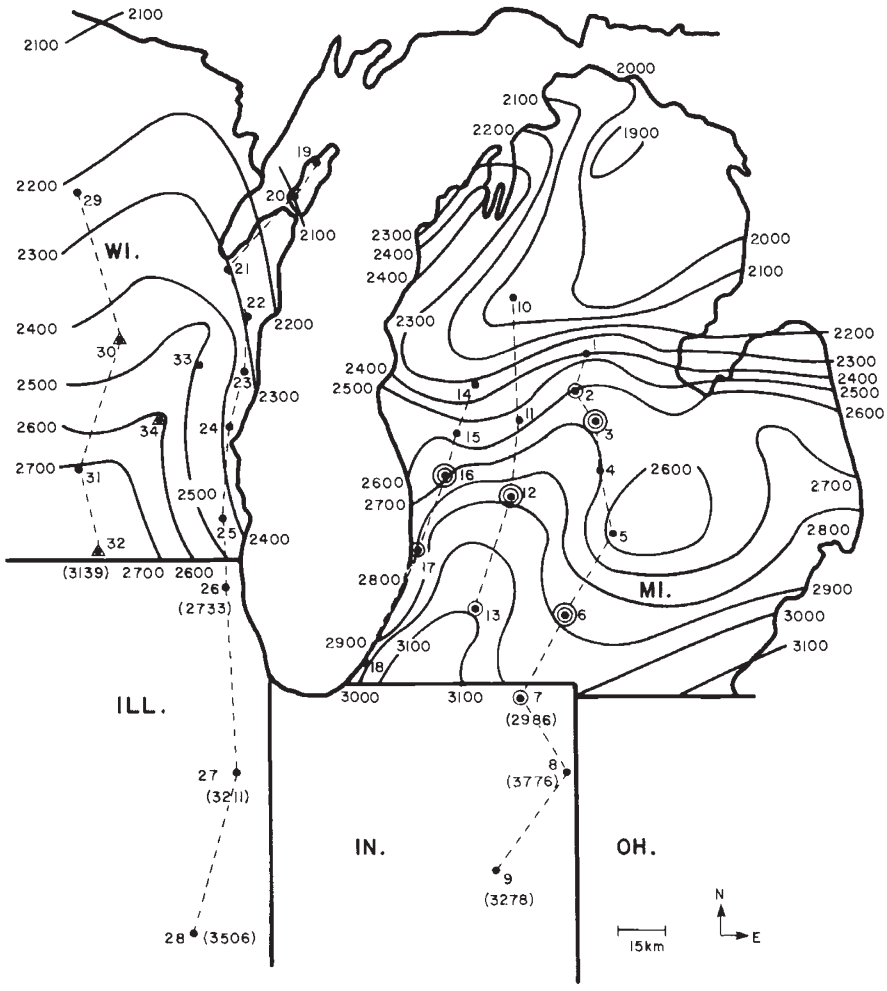

Figure 1 Collecting sites for *R. pomonella* in the midwestern United States (see table 1 for detailed descriptions of study sites). Solid lines indicate isothermal clines for growing degree days base 50°F (GDD). GDD values for Illinois and Indiana sites are given in parentheses. Dashed lines connect sites comprising each of the five latitudinal transect collected for the study. ⊙ = site where the frequency of *Acon*-2^{75} was >0·100 for either apple or hawthorn flies. ◎ = site where the frequency of *Acon*-2^{75} was >0·100 for both apple and hawthorn flies. △ = site where the frequency of *Acon*-2^{89} was >0·200 for hawthorn flies and/or >0·100 for apple flies.

the same as AA, A, H and HH in Feder *et al.* (1989*c*), where they represent part of a regional analysis of allele frequency variation in western Michigan.

Three different strategies were used to collect flies. At a majority of sites larvae were dissected from fallen fruit in the field and immediately frozen in liquid nitrogen (these sites are listed as L in table 1). Only late second and third instar larvae were electrophoretically analyzed from these collections. At other locations adult flies were captured directly from their respective host plants using a vacuum aspiration system and were then frozen in liquid nitrogen (these sites are designated Ad in table 1). For three hawthorn sites (4, 15, 17) and one apple site (15) both larvae and adults were collected from host trees to allow comparison of the two life history stages which represent sequential fly generations (*i.e.*, parents and their immediate offspring). At sympatric site 15, adults and larvae were collected from both apple and hawthorn host plants to determine whether inter-host genetic differentiation was consistent across different life history stages and to qualitatively gauge the extent of inter-host migration. Finally, fruits were collected from the ground beneath host trees at the five sites designated R in table 1 and the larvae infesting these fruits were reared to

Table 1 Numerical designations and descriptions of field sites. Sites were generally organized into five latitudinal transects (see fig. 1 and table 2). Latitude and growing degree days base 50°F are also given for sites

Site	Year	Location	Stage[a]	Host[b]	Distance[a]	Latit.	GDD
1	87	Rodgers Ave, Clare, MI.	L(L)	A, H	3·2	43°40′	2380
2	87	MI. 20, Mt. Pleasant, MI.	L(L)	A, H	3·0	43°35′	2660
3	87	Alger Rd., Alma, MI.	L(L)	A, H	sym	43°25′	2650
4	87	MI. 27, Eureka, MI.	L(Ad, L)	A, H	sym	43°06′	2720
5	86	Jolly Rd., Okemos, MI.	R(R)	A, H	1·0	42°41′	2570
6	87	F Dr., Eckford, MI.	L(Ad)	A, H	sym	42°09′	2880
7	87	IN. 20, Lagrange, IN.	(L)	H	—	41°38′	2986
8	87	Interstate 24, Parkerdale, IN.	L(L)	A, H	8·0	41°11′	3376
9	87	Interstate 69, Gas City, IN.	L(L)	A, H	7·1	40°50′	3278
10	87	No. 45 Rd., Cadillac, MI.	L(L)	A, H	1·0	44°11′	2060
11	87	MI. 46, Lakeview, MI.	L	A	—	43°25′	2550
12	87	Grand River Ave., Saranac, MI.	L(Ad)	A, H	4·9	42°52′	2870
13	87	Sprinkle Rd., Portage, MI.	L(L)	A, H	4·2	42°14′	3083
14	87	MI. 20, Woodville, MI.	L(Ad)	A, H	sym	43°40′	2320
15	87	112th St., Grant, MI.	Ad, L(Ad, L)	A, H	sym	43°21′	2550
16	87	24th St., Herrington, MI.	L(Ad)	A, H	2·0	43°04′	2690
17	87	MI. 89, Fennville, MI.	L(Ad, L)	A, H	1·5	42°34′	2800
18	87	East Rd., Three Oaks, MI.	Ad(Ad)	A, H	sym	41°46′	2940
19	86	Hogan Farm, Ephraim, WI.	R(R)	A, H	sym	45°07′	2060
20	85	Kuehn Farm, Carlsville, WI.	R(R)	A, H	sym	44°56′	2100
21	87	County MM Rd., Green Bay, WI.	L(L)	A, H	sym	44°28′	2310
22	87	Gass Lake Rd., Manitowoc, WI.	L(Ad)	A, H	sym	44°02′	2290
23	87	County EE Rd., Sheboygan, WI.	L(Ad)	A, H	sym	43°43′	2340
24	87	County C. Rd., Port Wash., WI.	L(Ad)	A, H	sym	43°22′	2380
25	87	50th. St., N. Cape, WI.	L(Ad)	A, H	0·5	42°46′	2450
26	86	ILL. 132, Waukegan, ILL.	R(R)	A, H	3·3	42°20′	2733
27	87	River Park, Kankakee, ILL.	(L)	H	—	41°09′	3211
28	86	Univ. ILL. Campus, Urbana, ILL.	R(R)	A, H	sym	40°05′	3506
29	87	WI. 29, Hatley, WI.	L(L)	A, H	1·6	44°52′	2230
30	87	WI. 49, Berlin, WI.	L(L)	A, H	6·3	43°54′	2450
31	87	I 90 and WI. 12, Madison, WI.	(L)	H	—	43°00′	2710
32	87	WI. 81, Beloit, WI.	L(L)	A, H	2·5	42°30′	3139
33	87	WI. 23, Greenbush, WI.	(L)	H	—	43°46′	2490
34	87	WI. 67, Iron Ridge, WI.	(L)	H	—	43°22′	2600

[a] Life history stage in which flies were collected. L = larvae, Ad = adult, R = flies samples as larvae and reared to adulthood in the laboratory. Stage designations given in parentheses are for hawthorn flies, without parentheses for apple flies.
[b] Host plants from which flies were collected at a site. A = apple, H = hawthorn.
[c] Distance (km) separating apple and hawthorn trees sampled at paired sites. sym = sympatric site where host trees were within 100 m.

adulthood in the laboratory. Laboratory rearing included chilling pupae at 4°C in a refrigerator for 5 months to simulate winter dispause conditions.

Standard horizontal starch gel electrophoretic techniques were used and are described fully elsewhere (Berlocher, 1976; Berlocher and Bush, 1982; Feder *et al.*, 1988, 1989*b*). We resolved and scored the six allozyme loci which displayed the greatest amounts of inter-host and/or latitudinal variation in the earlier survey of eastern North America (Feder *et al.*, 1989*a*). These six allozymes are: malic enzyme (*Me*), aconitase-2 (*Acon*-2), mannose phosphate isomerase (*Mpi*), NADH-diaphorase-2 (*Dia*-2), aspartate amino-transferase-2 (*Aat*-2) and hydroxyacid dehydrogenase (*Had*). A Mendelian mode of inheritance has been established for all six of these allozymes (Berlocher and Smith, 1983; Feder *et al.*, 1989*b*). Isozymes that migrated the nearest to the cathode were designated system 1, the second nearest system 2, etc. Alleles were numbered according to their relative anodal mobilities with the most common allele for each locus designated 100 and used as a standard. Electromorphs resolved for each of the six loci are given in the Appendices. Only two alleles were scored for the loci *Dia*-2 and *Me*. Consequently, only the frequencies for the rarer of the two alleles for these two loci appear in the Appendices. The same was essentially true for *Had*, although a rare third allele (Had^{97}) was occasionally scored in fly populations. However, for the sake of brevity, we only present frequencies for Had^{122} in Appendix 4. Electrophoretic alleles migrated identically in comparisons between adult and larval samples.

Genotype frequencies were examined for deviations from Hardy–Weinberg expectation by G-tests. Alleles were pooled, as required, so that all genotypic classes had expected numbers >1. The Levene correction (Spiess, 1977) was applied when sample sizes were <100 or when alleles could not be pooled to make expected numbers >1. G-contingency tests were used to test for allele frequency heterogeneity. Alleles were pooled, when necessary, to ensure that each cell in the G-contingency test had an observed number ≧5. *F*-statistics were calculated by the method of Weir and Cockerham (1984) with standard deviations estimated by jackknifing over loci or populations. Corrections for unequal sample sizes in the method of Weir and Cockerham can result in negative F_{ST} values. In these instances we report F_{ST} as 0.

First and second order linear regression analyses were done between arcsine transformed allele frequencies and either latitude or growing degree days (GDD) base 50°F. GDD is calculated by determining the number of degrees that the average temperature for a day was above 50°F and summing these daily totals throughout the course of the year (negative daily values are considered to be 0). Fifty degrees Fahrenheit is not an arbitrarily chosen temperature for calculating GDD but, rather, represents the approximate lower threshold condition under which post-pupal development occurs for *R. pomonella* (Reissig *et al.*, 1979). Fahrenheit was used as the unit of temperature in this study because weather stations in the midwestern United States do not provide GDD values in Celsius base. Unfortunately, converting GDD between the two temperature scales is not exact without daily readings from stations, a task which imposed insurmountable logistical and computational hurdles. GDD values from Wisconsin were the yearly averages from 1950–74, in Illinois from 1951–80, in Michigan from 1931–60, and in Indiana from 1975–87. Regressions were done separately for apple and hawthorn populations from Michigan/Indiana and Wisconsin/Illinois transects. At sites where both adults and larvae were collected, the allele frequencies used in the regressions were the mean between the two life history stages.

RESULTS

Hawthorn and apple populations were generally in Hardy–Weinberg equilibrium for the six polymorphic enzymes resolved in this study (see Appendices). The 16 significant departures from Hardy–Weinberg equilibrium observed out of a total of 330 tests observed for apple populations and 24 significant deviations out of 407 tests for hawthorn populations do not differ appreciably from the number of significant tests expected due to random type I error (16·5 and 20·3, respectively, for apple and hawthorn populations). None of these 40 deviations from Hardy-Weinberg equilibrium was significant at the "table wide" significance level of $P \leqq 0{\cdot}05$ using the sequential Bonferroni test (Holm, 1979; Rice, 1989). In addition, no pattern was evident in the loci, alleles or sites displaying significant departures from Hardy–Weinberg equilibrium.

Significant allele frequency differences were consistently observed for *Me*, *Acon*-2, *Mpi*, *Dia*-2, *Aat*-2, and *Had* between apple and hawthorn fly populations from across the midwestern United States (table 2; see Appendices for allele frequencies). Of the 28 paired apple and hawthorn sites, only site 4 (Eureka, Michigan) and site 32 (Beloit, Wisconsin) did not show a significant difference for at least one of the six loci analyzed. *Me* and *Acon*-2 showed significant inter-host allele frequency differences at the greatest number of paired sites (14 and 21, respectively). *Me* and *Acon*-2 are, in fact, tightly linked on chromosome II along with *Mpi* (Feder *et al.*, 1989*b*) and significant linkage disequilibrium has been found between *Me* and *Acon*-2 in natural apple and hawthorn populations (Feder *et al.*, 1988, 1989*a*). Genetic hitchhiking due to selection at a linked locus could therefore account for the observed correlation in allele frequencies between *Me* and *Acon*-2. The same may also be true for *Dia*-2 and *Aat*-2, which are separated by a map distance of 3·2 centimorgans on chromosome I (Feder *et al.*, 1989*b*). In this study, five of the seven sites which showed significant frequency differences between apple and hawthorn flies for *Aat*-2 were also significant for *Dia*-2 (table 2). High levels of linkage disequilibrium have also been observed between *Dia*-2 and *Aat*-2 (Feder *et al.*, 1988, 1989*a*).

Allele frequencies for *Me*, *Acon*-2, *Mpi*, *Dia*-2, *Aat*-2 and *Had* were similar in intra-host comparisons between adults and larvae collected from hawthorns at sites 4, 15, and 17 and from apples at site 15 (Note: Larvae dissected from host fruits at sites 4, 15 and 17 represent the offspring of adults collected from apple and hawthorn trees.). Only *Me* for hawthon flies at site 15 (Grant, Michigan) showed a significant frequency difference between adult and larval life history stages out of a total of 24 tests (Me^{100} adults = 0·593, $n = 345$; Me^{100} larvae = 0·655, $n = 467$;

Table 2 G-contingency tests for allele frequency heterogeneity between paired apple and hawthorn populations. Degrees of freedom for tests involving each locus are given in parentheses following locus abbreviations. At site 15 both adults (Ad) and larvae (L) were tested for inter-host variation

Site	*Me*(1)	*Acon*-2(2)	*Mpi*(1)	*Dia*-2(1)	*Aat*-2(3)	*Had*(1)
Transect 1						
1	***	***	***			
2	*	*	*	*		
3	**	***				
4						
5	*	**				
6		*				
8		*				
9						**
Transect 2						
10	***	***		*		***
12	***	***				
13						*
Transect 3						
14	***	***	***	*	***	
15(Ad)	***	***	***	*	*	***
15(L)	***	***	***	***	***	*
16		***		***	***	
17		***				
18	**					
Transect 4						
19	***	***		**		
20	***	***	**	***	***	**
21	*	**				
22		*				*
23					**	
24		**		*	**	
25		***				
26	***				**	
28		**				**
Transect 5						
29		**				
30	**	***				
32						

* $P \leqq 0{\cdot}05$. ** $P \leqq 0{\cdot}01$. *** $P \leqq 0{\cdot}001$. G-contingency test.

Table 3 G-contingency tests for allele frequency heterogeneity among apple and hawthorn populations along five latitudinal transects in the midwestern United States. See Table 2 for a breakdown of sites comprising each of the five transects

Host	Transect	*Me*	*Acon*-2	*Mpi*	*Dia*-2	*Aat*-2	*Had*
Apple	1		***	*			
	2	***	***				
	3	**	***		***	***	
	4	***	***	**			***
	5	**			**		***
Hawthorn	1	***	***	**			***
	2	***	***	**	***	*	***
	3	***	***	*	***	*	***
	4	***	***	***	***	***	***
	5	*	***	***			***

* $P \leqq 0{\cdot}05$. ** $P \leqq 0{\cdot}01$. *** $P \leqq 0{\cdot}001$. G-contingency test.

G-contingency test = 6·61, $P \leqq 0{\cdot}05$). The near constancy of gene frequencies observed between the two generations therefore suggests that intra-host selection is not intense between parents and their larval offspring.

The pattern and magnitude of inter-host differentiation between hawthorn and apple populations was similar for both larval and adult life history stages at site 15 (Appendices). Gene frequencies were significantly different between apple and hawthorn populations for both larvae and field captured adults for *Me*, *Acon*-2, *Mpi*, *Dia*-2, *Aat*-2, and *Had* (table 2). As mentioned before, only *Me* for hawthorns displayed significant intra-host variation at site 15 between adults and larvae. Adults collected directly from apple trees are, therefore, as genetically distinct from hawthorn adults as apple larvae are from hawthorn larvae.

Geographic variation was apparent among both apple and hawthorn populations across the midwestern United States. Substantial spatial heterogeneity was observed across each of the five north-south transects (table 3) and the magnitude of geographic variation differed between the two host races. Hawthorn flies displayed greater inter-populational differentiation than apple flies as evidenced by the 5 times higher overall F_{ST} value among hawthorn compared to apple populations (0·0626 and 0·0125, respectively; table 4). In addition, individual loci and alleles showed variable levels of geographic variation. For instance, *Me*, *Acon*, and *Had* had F_{ST} values of 0·1403, 0·0984 and 0·806, respectively, among hawthorn populations and 0·218, 0·212 and 0·183 among apple populations. In comparison, *Aat*-2 had a F_{ST} of 0·0108 for hawthorn flies and 0·0032 for apple flies across the same set of populations. Furthermore, F_{ST} among hawthorn populations for the alleles *Acon*-2^{100} and *Acon*-2^{95} (0·1033 and 0·1346, respectively; table 4) were greater than the value for the remaining pooled alleles at this locus (0·0380). A similar pattern existed for allelic variants of *Aat*-2 among hawthorn populations (table 4). The observed heterogeneity of F_{ST} values suggests that differential selection is occurring among alleles and loci, and that the effects of this selection are more pronounced for hawthorn than apple flies.

Table 4 F_{ST} values (Weir and Cockerham, 1984) across the 28 paired apple and hawthorn populations analyzed in the study. R refers to the F_{ST} value calculated by pooling all alleles remaining for a locus besides those listed in the table

Locus	Allele	Apple	Hawthorn
Me	100	0·0218 ± 0·0092[a]	0·1403 ± 0·0314
Acon-2	100	0·0214 ± 0·0120	0·1033 ± 0·0305
	95	0·0312 ± 0·0175	0·1346 ± 0·0377
	R	0·0132 ± 0·0057	0·0380 ± 0·0170
	Total	0·0212 ± 0·0099	0·0984 ± 0·0257
Mpi	100	0·0123 ± 0·0060	0·0341 ± 0·0170
Dia-2	100	0·0059 ± 0·0049	0·0379 ± 0·0165
Aat-2	100	0·0023 ± 0·0033	0·0025 ± 0·0037
	75	0·0050 ± 0·0053	0·0126 ± 0·0064
	50	0·0020 ± 0·0024	0·0266 ± 0·0121
	R	0·0031 ± 0·0029	0·0061 ± 0·0036
	Total	0·0032 ± 0·0025	0·0108 ± 0·0042
Had	100	0·0183 ± 0·0063	0·0806 ± 0·0229
All loci		0·0125 ± 0·0058[b]	0·0626 ± 0·0293

Jackknife estimate of standard deviation calculated over populations[a] or over loci[b].

Patterns of geographic variation were not random across the Midwest. First and second order linear regressions between arcsine transformed allele frequencies and either latitude or growing degree days (GDD base 50°F) indicated that frequency clines exist for both apple and hawthorn flies. In only six of 52 cases did second-order polynomials significantly improve the fits of the regression lines compared to first-order linear equations (table 5). Consequently, we will confine our discussion to the results of the first order analysis. Apple flies did not display nearly the number of significant linear regressions for individual loci with either GDD or latitude as hawthorn flies did (8 compared to 22). In addition, the fit of the regression lines were better among hawthorn than apple populations as evidenced by the r^2 values in table 5. Most importantly, the absolute values of the slopes of the regression lines were also always greater for hawthorn than apple populations (table 5). The amount and pattern of genetic differentation between apple and hawthorn populations at paired sites were therefore related to latitude (or local thermal conditions, as the two factors are correlated), with inter-host divergence resulting from differences in the slopes of allele frequency clines between the two races.

Me, *Acon*-2, *Mpi*, *Aat*-2, *Dia*-2 and *Had* showed both similarities and differences in their clinal patterns of variation across the Midwest. Because allele frequency clines were more clearly defined for the hawthorn than the apple race, our discussion will initially focus on hawthorn populations along transect 1 (Michigan/Indiana) and transect 4 (Wisconsin/Indiana), the two most intensively surveyed transects in the study (Note: Site 10 was considered to be the northernmost site along transect 1 in this discussion). For transect 4, allele frequencies for *Me*^{100}, *Acon*-2^{95} and *Mpi*^{37} displayed sharp and significant drops between

Table 5 First order linear regression analyses between arcsine transformed allele frequencies and either latitude or growing degree days base 50°F (GDD). Regressions were performed separately for pooled Michigan/Indiana and Wisconsin/Illinois transects as well as for apple and hawthorn populations. r^2 = amount of genetic variation explained by the linear regression

			Apple populations				Hawthorn populations			
			Latitude		GDD[a]		Latitude		GDD[a]	
Transect	Locus	Allele	r^2	Slope ± Std. Err.	r^2	Slope ± Std. Err.	r^2	Slope ± Std. Err.	r^2	Slope ± Std. Err.
Mi./In.	*Me*	100	0·22	2·954 ± 1·439	0·37**	−1·100 ± 0·379	0·72***[t]	10·639 ± 1·717	0·83***[t]	−3·400 ± 0·395
	Acon-2	95	0·39**[t]	3·998 ± 1·304	0·37**	−1·100 ± 0·389	0·82***	11·842 ± 1·454	0·86***	−3·600 ± 0·374
	Mpi	100	0·00	−0·074 ± 1·308	0·01	0·156 ± 0·382	0·72***[t]	−5·436 ± 0·871	0·70***	1·600 ± 0·271
	Dia-2	100	0·04	0·904 ± 1·119	0·02	−0·163 ± 0·333	0·61***	4·203 ± 0·877	0·68***	−1·300 ± 0·237
	Aat-2	100	0·00	0·289 ± 1·066	0·01	−0·096 ± 0·313	0·52**	1·890 ± 0·898	0·63***	−0·622 ± 0·122
	Had	100	0·15	1·655 ± 1·006	0·26*	−0·638 ± 0·276	0·77***	7·658 ± 1·081	0·88***	−2·500 ± 0·232
	All loci[b]		0·37**	1·646 ± 0·548	0·49**	−0·555 ± 0·145	0·83***	6·945 ± 0·813	0·91***	−2·200 ± 0·174
Wi./Ill.	*Me*	100	0·19	1·455 ± 0·940	0·14	−0·406 ± 0·325	0·51**[t]	5·891 ± 1·535	0·50**[t]	−1·900 ± 0·518
	Acon-2	95	0·27	2·022 ± 1·048	0·15	−0·509 ± 0·377	0·58***	5·700 ± 1·285	0·65***	−2·000 ± 0·396
	Mpi	100	0·42*	−1·827 ± 0·680	0·46*	0·641 ± 0·218	0·50**	3·443 ± 0·918	0·60***	1·300 ± 0·273
	Dia-2	100	0·32	1·696 ± 0·785	0·23	−0·485 ± 0·278	0·46**	−3·350 ± 0·980	0·29*	−0·897 ± 0·372
	Aat-2	100	0·01	0·231 ± 0·638	0·01	−0·030 ± 0·214	0·01	0·283 ± 0·807	0·01	−0·078 ± 0·269
	Had	100	0·52**	2·892 ± 0·880	0·47*	−0·919 ± 0·308	0·80***	5·572 ± 0·746	0·81***	−1·900 ± 0·244
	All loci[b]		0·59**	1·697 ± 0·446	0·44*	−0·488 ± 0·173	0·81***	4·040 ± 0·521	0·80***	−1·300 ± 0·178

* $P \leqq 0{\cdot}05$. ** $P \leqq 0{\cdot}01$. *** $P \leqq 0{\cdot}001$. Regression coefficient (slope of first order linear regression) >0, ANOVA F-test.

[a] Slopes and standard errors for GDD regressions are in units of 10^{-2}.

[b] Mpi^{37} allele frequencies were used instead of those for Mpi^{100} in the All loci regressions.

[t] Second order polynomial significantly improved fit of regression line compared to first order equation (ANOVA).

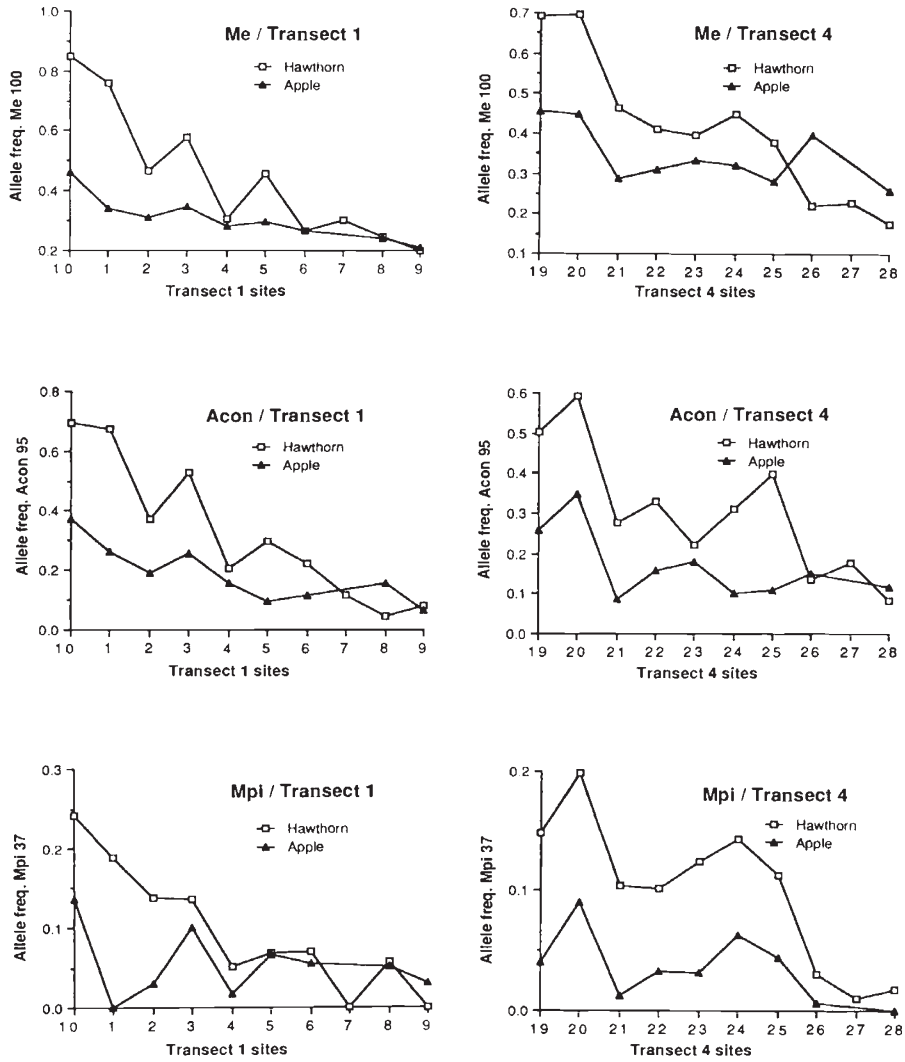

Figure 2 Allele frequencies for Me^{100}, *Acon*-2^{95} and Mpi^{37} for apple and hawthorn populations along transect 1 (Michigan/Indiana) and transect 4 (Wisconsin/Illinois). Sites are arranged by decreasing latitude from left to right along the x-axis. Site 10 is included as the northernmost site for transect 1.

hawthorn populations 20 and 21 north of Green Bay in eastern Wisconsin (fig. 2, Appendices 1, 4; G-test for Me^{100} allele frequency heterogeneity between sites 20–21 = 15·1, $P \leqq 0{\cdot}001$, 1 df [Note: all subsequent G-tests, unless specified, have 1 df]; G [*Acon*-2^{95} sites 20–21] = 27·1, $P \leqq 0{\cdot}001$; G [Mpi^{37} sites 20–21] = 4·56, $P \leqq 0{\cdot}05$). Me^{100}, *Acon*-2^{95} and Mpi^{37} frequencies leveled off and even showed somewhat of an increase among populations 21 through 24 and/or 25, before falling significantly for a second time south of site 25 near the border between Wisconsin and Illinois (fig. 2, Appendices 1, 4; G[Me^{100} sites 25–26] = 7·48, $P \leqq 0{\cdot}01$; G[*Acon*-2^{95} sites 25–26] = 23·3, $P \leqq 0{\cdot}001$; G[Mpi^{37} sites 25–26] = 6·29, $P \leqq 0{\cdot}05$). The linked loci *Me*, *Acon*-2 and *Mpi* therefore behaved in a similar fashion among transect 4 hawthorn populations. In contrast, Had^{100} frequencies did not change appreciably between hawthorn sites 20 and 21 but did fall along with Me^{100}, *Acon*-2^{95} and Mpi^{37} south of population 25 (fig. 3; G[Had^{100} sites 25–27] = 16·1, $P \leqq 0{\cdot}001$). *Dia*-2^{100} frequencies decreased steadily from hawthorn populations 20 to 24 dropping from 0·868 at site 20 to 0·425 at site 24 (fig. 3; G[*Dia*-2^{100} sites 20–24] = 16·5, $P \leqq 0{\cdot}001$). However, sites 25 and 26 formed a plateau of significantly elevated *Dia*-2^{100} frequencies (0·735 and 0·740, respectively) compared to

site 24 (G[*Dia*-2^{100} sites 24–25] = 12·2, $P \leqq 0{\cdot}001$; G[*Dia*-2^{100} sites 24–26] = 17·3, $P \leqq 0{\cdot}001$) that was more pronounced and occurred slightly further south than similar peaks for *Me*^{100}, *Acon*-2^{95}, *Mpi*^{37} and *Had*^{100} (Figs 2 and 3). Allele frequencies for *Dia*-2^{100}, like those for *Me*^{100}, *Acon*-2^{95}, *Mpi*^{37}, and *Had*^{100}, also declined sharply toward the southern end of transect 4 dipping from 0·740 at site 26 to 0·500 at site 28 (G[*Dia*-2^{100} sites 26–28] = 16·5, $P \leqq 0{\cdot}001$). *Aat*-2^{100} showed a very similar pattern of geographic variation along transect 4 as *Dia*-2^{100} (fig. 3) as would be expected due to the close proximity of *Aat*-2 and *Dia*-2 on linkage group I (Feder *et al.*, 1989*b*).

As was the case for transect 4, *Me*^{100}, *Acon*-2^{95} and *Mpi*^{37} all displayed large and highly significant frequency decreases for hawthorn populations at the northern end of transect 1 (fig. 2, Appendices 1, 4; G[*Me*^{100} sites 10-2] = 32·9, $P \leqq 0{\cdot}001$; G[*Acon*-2^{95} sites 10–2] = 17·8, $P \leqq 0{\cdot}001$; G[*Mpi*^{37} sites 10–4] = 13·3, $P \leqq 001$). However, *Me*^{100}, *Acon*-2^{95} and *Mpi*^{37} did not show as pronounced drops in frequencies among hawthorn populations at the southern end of transect 1 as they did along transect 4 (fig. 2). *Had*^{100} frequencies changed roughly equivalently along the southern portions of the two transects (fig. 3) but, unlike transect 4, *Had*^{100} frequencies fell significantly at the northern end of transect 1 (G[*Had*^{100} sites 10–2] = 13·7, $P \leqq 0{\cdot}001$). *Dia*-2^{100} and *Aat*-2^{100} clines were also different between the two transects, as both *Dia*-2^{100} and *Aat*-2^{100} displayed much more even and gradual reductions in allele frequencies along transect 1 than transect 4 (fig. 3).

"Steps" in the allele frequency clines for apple infesting populations, while not as distinct as those for hawthorn flies, were also apparent (Figs 2 and 3). For instance, *Me*^{100}, *Acon*-2^{95} and *Mpi*^{37} frequencies all dropped significantly between apple populations 20–21 (G[*Me*^{100} sites 20–21] = 15·1, $P \leqq 0{\cdot}001$; G[*Acon*-2^{95} sites 20–21] = 24·1, $P \leqq 0{\cdot}001$; G[*Mpi*^{37} sites 20–21] = 7·12, $P \leqq 0{\cdot}01$) and populations 10–2 (G[*Me*^{100} sites 10–2] = 4·36, $P \leqq 0{\cdot}05$; G[*Acon*-2^{95} sites 10–2] = 6·29, $P \leqq 0{\cdot}05$; G[*Mpi*^{37} sites 10–2] = 5·22, $P \leqq 0{\cdot}05$) just as they did for hawthorn populations. *Mpi*^{37} allele frequencies also declined significantly from 0·063 to 0·006 between apple populations 24 and 26 (G[*Mpi*^{37} sites 24–26] = 6·49, $P \leqq 0{\cdot}05$). But neither *Me*^{100} nor *Acon*-2^{95} frequencies declined markedly among apple populations along the southern end of transect 4 as they did among hawthorn populations (fig. 2). *Had*^{100} allele frequencies did drop significantly among southern apple populations along both transects 1 and 4 (G[*Had*^{100} sites 6–8] = 6·02, $P \leqq 0{\cdot}05$; G[*Had*^{100} sites 25–26] = 9·03, $P \leqq 0{\cdot}01$). However, apple population 24 did not display the same dips in allele frequencies for either *Aat*-2^{100} or *Dia*-2^{100} that hawthorn population 24 showed (fig. 3).

Allele frequencies for *Acon*-2 suggest that at least some gene flow may be occurring between local apple and hawthorn populations. Frequencies for *Acon*-2^{75} were higher for both hawthorn and apple populations in central and western Michigan than they were in any other region of the Midwest (fig. 1, Appendix 1). The same phenomenon also occurred for *Acon*-2^{89} at sites 30 and 32 in Wisconsin and Illinois (fig. 1, Appendix 1). In addition, the rare allele *Acon*-2^{73} was present only in apple and hawthorn populations along transect 4 (Appendix 1). Selection could be responsible for the elevated frequencies of certain *Acon*-2 alleles only at specific paired apple and hawthorn sites. However, *Acon*-2 allele frequencies were usually significant different between the two host races at paired sites (table 2), a result which suggests that selection pressures for *Acon*-2^{73}, *Acon*-2^{75} and *Acon*-2^{89} are probably different between hawthorn and apple populations at a given site. Inter-host gene flow is therefore the most likely cause for the increased frequencies of *Acon*-2^{75} and *Acon*-2^{89} only at certain paired sites and for the occurrence of the rare allele *Acon*-2^{73} in both host races only along transect 4. Gene flow could also account for the observation in Feder *et al.* (1989*a*) that polymorphic allozymes which do not display host associated differentiation at sympatric sites, also show low levels of intra-host geographic variation across eastern North America.

Although latitude and temperature are interrelated environmental factors, their relationship is not perfect. Several irregularities exist in the pattern of isothermal clines through Michigan and Wisconsin (fig. 1) which are informative for determining whether latitude or GDD is the most accurate predictor of gene frequencies for *R. pomonella* populations. For Michigan/Indiana transects GDD proved to be the more reliable determinant and explained approximately 10 per cent more of the genetic variation among both apple and hawthorn populations than latitude (see r^2 values for all loci in table 5). GDD and latitude accounted equally well for genetic variation among hawthorn populations in Wisconsin and Illinois (table 5) but latitude was a better predictor among apple populations (r^2 latitude = 0·59, r^2 GDD = 0·44). Our confidence in the latter result is lessened, however, by the fact that apple populations were not

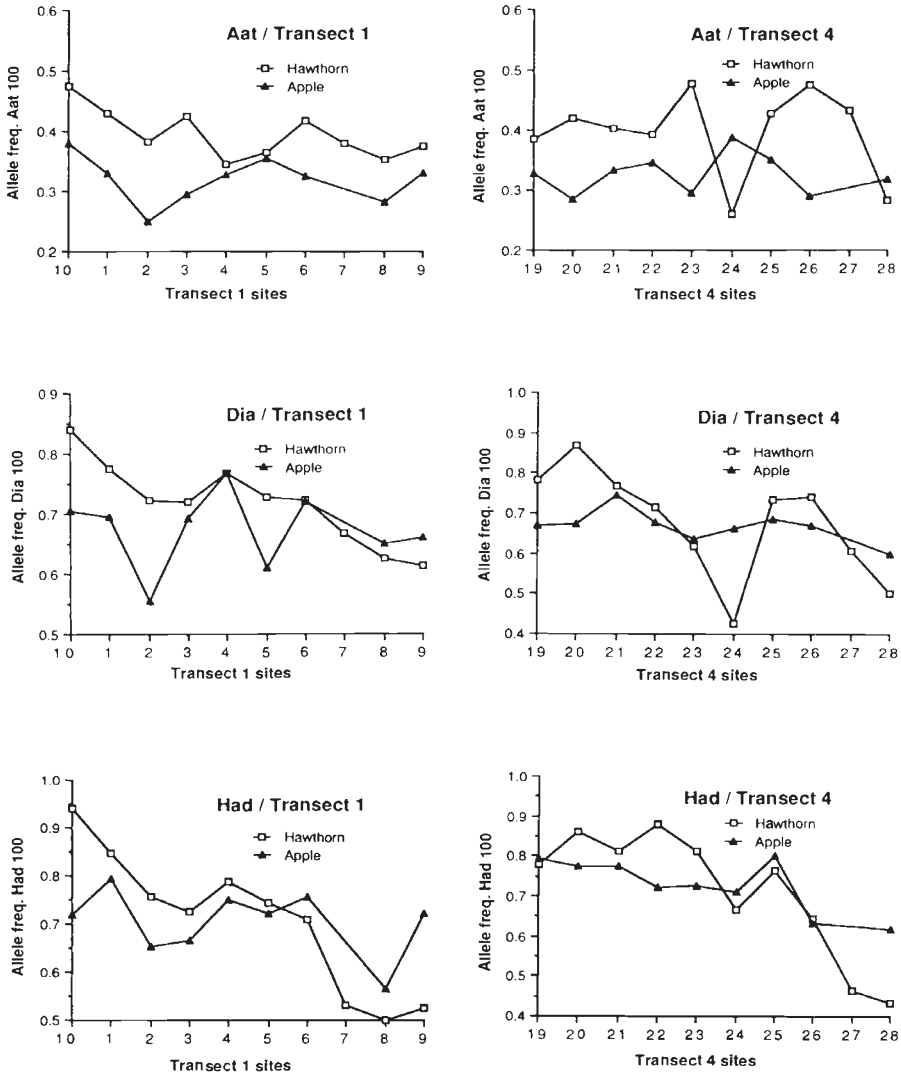

Figure 3 Allele frequencies for *Aat*-2^{100}, *Dia*-2^{100} and *Had*^{100} for apple and hawthorn populations along transect 1 (Michigan/Indiana) and transect 4 (Wisconsin/Illinois). Sites are arranged by decreasing latitude from left to right along the ×-axis. Site 10 is included as the northernmost site for transect 1.

sampled at four of the 16 collecting sites in Wisconsin and Illinois. Three of these four unpaired sites (nos. 31, 33, 34) happened to be located within areas in Wisconsin that have warmer ambient temperatures than would be expected based on their latitudes (fig. 1). The addition of apple populations from these three sites could therefore have a substantial effect on the regression between latitude and allele frequencies for apple populations in Wisconsin and Illinois.

Hawthorn populations displayed allele frequency shifts for *Me*^{100} and *Acon*-2^{95} that were, in fact, correlated with irregularities in local thermal conditions. First, site 5 (Okemos, Michigan) is located within a pocket of colder GDD in central Michigan (fig. 1). The gene frequency for *Me*^{100} was significantly higher for hawthorn flies at site 5 than at site 4 (Eureka, Michigan), which is located approximately 20 km north of site 5 but is warmer (see Appendix 1; $G[Me^{100}$ sites 4–5$] = 4{\cdot}64$, $P \leqq 0{\cdot}05$; Note: Larval and adult samples were pooled for this and all subsequent G-contingency tests involving site 4). Warmer temperature site 6 (Eckford, Michigan) is situated approximately 30 km southwest of Okemos and had a significantly lower *Me*^{100}

frequency for hawthorn flies than site 5($G[Me^{100}$ sites 5–6] = 4·01, $P \leqq 0{\cdot}05$). *Acon*-2^{95} allele frequencies displayed the same trend among hawthorn populations 4, 5 and 6 as Me^{100} (Appendix 1, fig. 2) but *Acon*-2^{95} frequencies were not significantly higher at site 5 than at sites 4 and 6. To confirm the validity of the frequency changes for Me^{100} and *Acon*-2^{95}, we analyzed an additional hawthorn population collected from the campus of Michigan State University (M.S.U.), East Lansing, Michigan in 1985 (The M.S.U. site is located 5 km northwest of site 5 and is situated in the same pocket of cooler GDD as site 5). The results from the M.S.U. site substantiated the previous findings, as Me^{100} and *Acon*-2^{95} frequencies were both significantly higher for hawthorn flies collected from M.S.U. than from sites 4 and 6 (Me^{100} M.S.U. = 0·441, $n = 144$, $G[Me^{100}$ M.S.U.-site 4] = 6·32, $P \leqq 0{\cdot}05$, $G[Me^{100}$ M.S.U.-site 6] = 4·88, $P \leqq 0{\cdot}05$; *Acon*-2^{95} M.S.U. = 0·443, $n = 124$, $G[$*Acon*-2^{95} M.S.U.-site 4] = 20·1, $P \leqq 0{\cdot}001$, $G[$*Acon*-2^{95} M.S.U.-site 6] = 12·1, $P \leqq 0{\cdot}001$). Significant frequency increases therefore occurred for both Me^{100} and *Acon*-2^{95} that corresponded with a cold bubble of GDD in central Michigan.

A second example of the relationship between local thermal conditions and Me^{100} and *Acon*-2^{95} allele frequencies involves fingers of higher GDD which extend up into Michigan around the Kalamazoo area (see site 13 in fig. 1) and in Wisconsin, northeast of the city of Madison (see site 31). Gene frequencies for Me^{100} and *Acon*-2^{95} were almost always lower for hawthorn populations within these warm areas (site 13, Portage, Michigan, and sites 33 and 34 near Greenbush and Iron Ridge, Wisconsin) compared to sites of equivalent latitude in parallel transects (sites 6, 17 and 18 in the case of site 13, sites 23 and 30 for site 33, and site 24 for site 34). The only exception to this general trend was for *Acon*-2^{95} in comparisons between sites 23 and 33, for which the frequency of the allele was lower at the colder site 23 (0·221) than at the warmer site 33 (0·306).

DISCUSSION

The results from this study strongly suggest that partially reproductively isolated host races can evolve in sympatry. However, to prove that the apple race originated via a sympatric host shift we must still show that (1) the apple infesting form of *R. pomonella* was not reproductively isolated from hawthorn flies prior to its colonization of apples and (2) that the apple race could not have evolved in allopatry from the hawthorn race and then been secondarily introduced into the northeastern United States.

All available evidence indicates that the apple race is directly descendant from a hawthorn infesting form of *R. pomonella* and was not derived from an unrecognized sibling species. For instance, experimental crosses give no indication of any postmating or ethological premating reproductive isolation between hawthorn and apple flies (Reissig and Smith, 1978; Smith, 1988). Also, no single morphological or genetic character exists which diagnostically distinguishes apple from hawthorn flies, as the host races are currently known to differ only in allele frequencies for six specific allozyme loci. In fact, because allele frequency clines are steeper among hawthorn than apple populations, hawthorn populations at northern sites in the Midwest are genetically more similar to southern apple populations than they are to southern hawthorn populations. For example, hawthorn population 19 (Door Co., Wisconsin) and hawthorn population 28 (Urbana, Illinois) are separated by a Nei genetic distance of 0·2374 ± 0·0978 (standard deviation calculated by jackknifing over loci) for the six allozymes resolved in this study. In contrast, apple population 28 and hawthorn population 19 have a Nei distance of 0·1340 ± 0·0669. Consequently, different apple populations from across the Midwest do not cluster together as a discrete genetic subdivision from hawthorn populations as might be expected if the apple race was formed from a sister taxon to the hawthorn race. Furthermore, the non-random pattern of *Acon*-2^{73}, *Acon*-2^{75} and *Acon*-2^{89} allele frequencies in Michigan and Wisconsin (fig. 1), and the similar pattern of linkage disequilibrium in apple and hawthorn populations (Feder *et al.*, 1988, 1989*a*), suggest that at least some gene flow is occurring between local apple and hawthorn populations. It is therefore very unlikely that apple and hawthorn populations of *R. pomonella* are completely reproductively isolated.

It is still possible, of course, that different races of *R. pomonella* exist on different species of hawthorns and that the apple race was derived from just one of these "hawthorn races". Almost 100 different species of *Crataegus* comprising 19 different species groups may be endemic to North America (Fernald, 1950; Correll and Johnston, 1970; Muniyamma and Phipps, 1985). However, the taxonomic status of a majority of these *Crataegus* species is questionable as hybridization, polyploidy and apomixis is common in the genus

(Phipps, 1983, 1984; Muniyamma and Phipps, 1984, 1985). Many of the trees that we identified as *C. mollis* in this study may therefore have been of hybrid origin. Nevertheless, *R. pomonella* has been reported to attack only a restricted set of *Crataegus* "species", with infestations confirmed for only 14 endemic hawthorns (Bush, 1966; Wasbauer, 1972; Berlocher, 1976; McPheron, 1987; McPheron *et al.*, 1988*b*). In the northeastern United States, only hawthorn species which share certain phenotypic traits with *C. mollis* in the *Brainerdianae*, *Coccinnea*, *Macracantha*, *Pruinosa*, *Punctata* and *Tenuifoliae* series have been reported to be parasitized by the fly (O'Kane, 1914; Bush, 1966; Wasbauer, 1972). It therefore appears that *R. pomonella* utilizes only a small number of potential hawthorn hosts in the Northeast, preferentially attacking hawthorns that possess a similar fruiting phenology and fruit characteristics; A trend which is not conducive to the formation of different hawthorn races. In addition, electrophoretic analysis of *R. pomonella* populations infesting *C. punctata*, *C. brachyacantha*, *C. douglasii* and *C. monogyna* from the southern and western United States has given no evidence for genetically differentiated hawthorn races (Berlocher, 1976; McPheron, 1987; McPheron *et al.*, 1988*b*). Although further sampling of *R. pomonella* at sites with different sympatric hawthorn species is needed to completely discount the possibility of host specific hawthorn races, all available data suggest that hawthorn populations of the fly in the United States and Canada represent a single race which displays extensive latitudinal variation.

Several points also make it unlikely that the apple race formed in allopatry from the hawthorn race and was subsequently introduced into the Hudson Valley region of New York. First, genetic bottlenecks due to reductions in population size frequently occur when a species colonizes a new area. Reduced amounts of genetic variation have, in fact, been observed for newly introduced populations of *R. pomonella* and *R. completa* in the western United States (McPheron *et al.*, 1988*b*; Berlocher, 1984). However, apple populations in the eastern United States have as much allelic diversity and genetic heterozygosity as eastern hawthorn populations (Feder *et al.*, 1989*a*). There is, therefore, no genetic evidence indicating that the apple race was introduced into the New England area.

Second, the *R. pomonella* species group is endemic to North America (Bush, 1966), thereby ruling out the possibility that the apple race originated in Europe and was brought to the United States along with cultivated apples. Consequently, if the apple race formed in allopatry then it must have originated somewhere in North America. The range of domestic apples is, however, contained entirely within that of hawthorns in North America. Therefore, even if *R. pomonella* did not originally shift onto apples in the Hudson Valley, the apple race could not have been geographically isolated from hawthorn infesting populations at the time of its inception.

Third, B. D. Walsh (1867), who first described *R. pomonella*, noted at the time of the initial discovery of the apple race that although *R. pomonella* "exists both in the East and in the West, it attacks the cultivated apple only in a certain limited region, even in the East, for according to Dr Trimble (New York Semi-weekly Tribune, July 19, 1867) this new and formidable enemy of the apple is found in the Hudson-river valley, but has not yet reached New Jersey". Therefore, up until 1867, no apple infesting population of *R. pomonella* had been reported outside of the New England area. We now know, however, that an isolated population of *R. pomonella* does infest both hawthorns and domestic apples in the highlands of Mexico (Bush, 1966). But Mexican flies are unique in that they possess a hyaline spot at the base of the apical band on their wings which makes them morphologically distinguishable from all other taxa in the *R. pomonella* group (Bush, 1966). The apple race in the United States could not, therefore, have been derived from the Mexican population.

Although the Mexican population does not clarify the issue of the geographic origin of the apple race, it does lend further indirect support for hawthorns as the original host for the apple race. In Mexico, *R. pomonella* has independently shifted from hawthorns to domestic apples within historical times (apples were introduced into Mexico from Spain in 1522; Standley, 1922). However, Mexico does not have any endemic species of crab apple (Standley, 1922; J. Beaman, personal communication). Therefore, in Mexico at least, *R. pomonella* shifted directly from hawthorns to domestic apples.

At least two major factors are responsible for the observed genetic differences between hawthorn and apple populations. First, differential selection occurs between flies infesting apples and hawthorns and appears to be related in some way with ambient temperature (see below for further discussion). Second, the consistency of inter-host genetic differenation between larvae and adults flies at site 15 suggests that adult flies do not migrate randomly between hawthorn and apple

trees but, instead, tend to attack the same species of host tree they infested as larvae. Although we did not examine the pattern of microgeographic genetic variation among trees in this study, results from Feder *et al.* (1988, 1989*c*, manuscript in preparation) show that intra-host genetic variation among trees for larvae and adults at site 15 is almost an order of magnitude less than that of inter-host divergence. These data do not rule out the possibility that adults stay on their "natal trees" and, consequently, that selection alone is responsible for inter-host differentiation. However, mark and recapture studies suggest that adults continually move among different host trees within orchards and oldfields (Phipps and Dirks, 1933; Bourne *et al.*, 1934; Neilson, 1971; Buriff, 1973; Roitberg, 1982), with only about 10 per cent of marked flies remaining on "release trees" (Maxwell, 1968; Reissig, 1977). Observations of *R. pomonella* dispersing over distances of at least 1·6 km in the field (Maxwell and Parsons, 1968) also testify to the vagility of the fly. Allele frequency differences between hawthorn and apple adults at site 15 therefore indicate that at least some host fidelity exists for *R. pomonella.* Because courtship and mating in *R. pomonella* occur almost exclusively on or near the fruits of the host plant (Prokopy *et al.*, 1971, 1972), differential host utilization by hawthorn and apple flies produces a system of positive assortative mating which helps maintain genetic polymorphism between the races. Genetically based differences in host preference and conditioning in adults (Prokopy *et al.*, 1982, 1986; Papaj and Prokopy, 1986) are the two most likely causes for host fidelity, as larval conditioning has not been found for *Rhagoletis* flies (Prokopy *et al.*, 1982, 1988).

Ambient temperature has always been regarded as an important environmental factor influencing host plant-parasite interactions for *Rhagoletis* (Prokopy, 1968*b*; Reissig *et al.*, 1979). *R. pomonella* flies are, in general, univoltine across their range and adult life expectancy has been estimated at from 3 to 6 weeks in the field (see review by Boller and Prokopy, 1976). Adults must therefore eclose at times closely synchronized with the fruiting phenologies of their host plants. "Temporal windows" for infesting apples and hawthorns are different, however, with the fruit of domestic apple varieties favored by *R. pomonella* generally ripening and falling approximately one month earlier than fruit of native hawthorn species. In addition, flies must develop as larvae and pupae at rates matching local environmental conditions. If development proceeds too quickly flies may not diapause and emerge as a second generation in the fall when suitable host fruit is no longer available. Alternatively, if development occurs too slowly then flies run the risk of freezing to death during the first frost. Ambient temperature does, of course, correlate with latitude and is a major determinant of the phenology of fruit maturation. Hawthorn and apple flies may therefore have development rates differentially adapting them to local thermal conditions and the phenologies of their host plants. The temperature hypothesis is contingent, of course, on the six allozymes resolved in this study being related in some way to developmental timing in *R. pomonella.* Preliminary studies indicate that such a relationship does, in fact, exist as allele frequencies for *Me*, *Acon*-2 and *Had* correlate with the timing of diapause termination and adult eclosion for hawthorn flies (Feder *et al.*, manuscript in preparation)

It is not uncommon for allozyme surveys to uncover one or more loci whose alleles display a clinal pattern of variation correlated with some aspect of the physical environment (Johnson, 1971; Clegg and Allard, 1972; Christiansen and Frydenberg, 1974; Koehn *et al.*, 1976; see review by Hedrick *et al.*, 1976). Latitudinal, seasonal and altitudinal variation in temperature have been implicated as causal factors for a number of allozyme clines including those for the freshwater fish, *Catostomus clarkii* (Koehn, 1969); the marine killifish, *Fundulus heteroclitus* (Mitton and Koehn, 1975); the fathead minnow, *Pimephales promelas* (Merritt, 1972); the fruit fly, *D. melanogaster* (Berger, 1971; Vigue and Johnson 1973; Johnson and Schaffer, 1973; Miller, Pearcy and Berger, 1975); and the butterfly, *Colias meadii* (Johnson, 1976). What is unusual about *R. pomonella* is that ambient temperature apparently has different consequences for conspecific populations infesting different host plants. Host plant associations therefore add an additional level of environmental heterogeneity for *R. pomonella* which helps maintain increased amounts of intraspecific genetic variation, a result consistent with predictions of niche theory (Levene, 1953; Maynard Smith, 1966, 1970; Levins, 1968, Christiansen and Feldman, 1975; Felsenstein, 1976).

Allele frequency clines for hawthorn and apple flies are not completely discordant, however, and display some broad similarities which may have important evolutionary implications. The three areas where the largest allele frequency changes occur for apple and/or hawthorn populations are located in or near major ecological transition zones in the Midwest. Sites 1–4 in Michigan and 20–21

in Wisconsin are situated at the boundary between deciduous and boreal forest habitats. Likewise, sites 25-26 in Illinois/Wisconsin lie where the central prairie meets deciduous forest. The ecological consequences of these zones are significant and mark cutoff points in the distribution of numerous plant species. Changes in voltinism and/or host plant associations occur across these regions for several different phytophagous insects including the European corn-borer, (Showers, 1981) and the swallowtail butterfly, *Papillio glaucus* (for review see Scriber and Hainze, 1987). The finding of a similar pattern for *R. pomonella* supports the hypothesis that these zones have a general impact on the evolution of life history traits for phytophagous insects.

The observed frequency clines in *R. pomonella* could, of course, be due to secondary contact rather than selection. However, for several reasons, discussed fully in Feder *et al.* (1989*a*), the secondary contact hypothesis is unlikely. Perhaps most importantly, the finding that frequencies for Me^{100} and $Acon\text{-}2^{95}$ among hawthorn populations covary with irregularities in local thermal conditions argues for differential selection as the cause for allele frequency clines in *R. pomonella.*

A crucial remaining question is whether apple and hawthorn host races represent incipient species. A host race is a parasitic population which, by adapting to a preferred host, has become partially reproductively isolated from other conspecific populations specialized on alternative hosts (Diehl and Bush, 1984). Host races may therefore be at various stages of divergence ranging from populations which almost freely interbreed to those that rarely exchange genes (Bush, 1969*a*). Based on available data, it is difficult to judge whether apple and hawthorn flies have reached or will ever reach the evolutionary point of no return when they should be considered distinct species. In essence, the results of this and other studies (Feder *et al.*, 1988, 1989*a*; McPheron *et al.*, 1988*a*; Smith *et al.*, 1988) have shown that host specific traits can evolve within geographically continuous populations. The next key question then is whether these host associated adaptations are sufficient to eventually cause the complete reproductive isolation of populations. Indirect support that they are has come from additional population genetic studies involving sibling species in the *R. pomonella* complex. For instance, *R. mendax* is a closely related sympatric sibling species to *R. pomonella* which infests blueberries (*Vaccinium* spp.) and huckleberries (*Gaylussacia* spp.) in the eastern United States and Canada (Bush, 1966). Although field captured *R. pomonella* and *R. mendax* adults will hybridize in the laboratory and produce fertile and viable F1 progeny (J. Frey, personal communication; Feder and Bush, 1989), co-occurring populations of the sibling species are genetically distinct in nature and do not interbreed (Berlocher and Bush, 1982; Feder *et al.*, 1989*d*). Genetic analysis of reproductively mature adults collected from interdigitated blueberry bushes and apple trees indicate that *R. mendax* and *R. pomonella* maintain complete host fidelity in the field (Feder and Bush, 1989). Differential host plant recognition is therefore a very effective pre-mating isolating barrier between at least two sympatrically distributed sibling species in the *R. pomonella* group.

The allele frequency clines indicate that both hawthorn and apple races are not evolving uniformly across the Midwest, however. Primary frequency clines have been implicated in speciation (Fisher, 1930; Murray, 1972; Endler, 1977) and it is possible that *R. pomonella* populations are "isolated by distance" as a consequence of being differentially adpated to local thermal conditions. Both hawthorn and apple races do display sharp allele frequency "steps" across environmental transition zones in the Midwest (figs 2 and 3). In addition, hawthorn populations at the northern and southern ends of *R. pomonella's* range are essentially fixed for alternative alleles at the loci *Had*, *Dia*-2 and *Me* (see Appendices and McPheron, 1987). *R. pomonella* populations at different latitudes may therefore be "genetically incompatible" with one another even though they are inter-connected by a series of populations exchanging genes. It will be interesting to see whether allele frequency clines for *R. pomonella* can become steep enough to result in "hybrid zones" given the fly's apparent ability to disperse over relatively long distances.

Whether or not *R. pomonella* host races are "cohesive evolutionary units" is not central to the issue of sympatric speciation, however. The important question is whether host recognition behaviours are constant across the range of a race. For example, if flies infesting hawthorns continue to prefer hawthorns when experimentally transplanted into different portions of their distribution, then host shifts ensure the divergence of populations regardless of whether selection pressures vary from locality to locality within races. The results from site 15 confirm that some host fidelity exists for *R. pomonella* in at least one portion of the fly's range. However, apple and hawthorn populations at sites 4 and 32 showed no significant genetic differentiation (table 2). Is the lack of genetic diver-

gence between apple and hawthorn populations at these two sites due to high levels of inter-host gene flow? Or, alternatively, is host fidelity still strong at sites 4 and 32 but selection pressures for *Acon*-2, *Mpi*, *Dia*-2, *Aat*-2 and *Had* similar for apple and hawthorn flies? The shapes of the allele frequency clines among hawthorn and apple populations suggest that the latter possibility is true. However, detailed studies measuring inter-host gene flow across the range of *R. pomonella* are needed to resolve this point and establish the extent to which apple and hawthorn races are discrete populations on separate evolutionary pathways.

In conclusion, *R. pomonella* is rapidly becoming a model organism for the study of speciation. Genetic differences exist between recently formed host races of the species and the geographic pattern of allozyme variation suggests that differences in host plant utilization and host associated development (in some way related to ambient temperature) are responsible for reducing gene flow between apple and hawthorn races. Our next task is to characterize exactly how these factors operate at the population level and to translate their mode of action to the speciation process. In this regard, the ability to hybridize several different *R. pomonella* species in the laboratory will help us elucidate the genetic basis for host associated traits, such as differential host recognition, which appear to play a major role in reproductively isolating taxa in the *R. pomonella* group.

Acknowledgments We would like to thank S. H. Berlocher, B. A. McPheron, S. M. Williams and two anonymous reviewers for reading and making useful comments on an earlier draft of this manuscript. C. A. Chilcote, J. E. Zablotney, M. D. B. Liebold, A. Wertheimer, T. A. Hunt, S. Jamrog, S. Hess, S. Bourbeau, A. Lake, D. Swiderski, S. H. Berlocher, B. A. McPheron, D. C. Smith, B. Angus MacMillan and J. Millett assisted in various aspects of fly collecting, fly rearing and gel running. B. S. Weir gave valuable advice on the calculation of F-statistics. We are grateful to F. Nurnberger, D. Clark, W. M. Wendland and K. Scheeringa for providing us with weather station data from the states of Michigan, Wisconsin, Illinois and Indiana, respectively. S. Tolin provided much appreciated secretarial support and K. Cebra drew fig. 1.

REFERENCES

BERLOCHER, S. H. 1976. The genetics of speciation in *Rhagoletis* (Diptera: Tephritidae). Ph.D. Thesis, University of Texas at Austin.

BERLOCHER, S. H. 1984. Genetic changes coinciding with the colonization of California by the walnut husk fly, *Rhagoletis completa*. *Evolution*, *38*, 906-918.

BERLOCHER, S. H. AND BUSH, G. L. 1982. An electrophoretic analysis of *Rhagoletis* (Diptera: Tephritidae) phylogeny. *Syst. Zool.*, *31*, 136-155.

BERLOCHER, S. H. AND SMITH, D. C. 1983. Segregation and mapping of allozymes of the apple maggot fly. *J. Hered.*, *74*, 337-340.

BERGER, E. M. 1971. A temporal survey of allelic variation in natural and laboratory populations of *Drosophila melanogaster*. *Genetics*, *67*, 121-136.

BOLLER, E. F. AND PROKOPY, R. J. 1976. Bionomics and management of *Rhagoletis*. *Ann. Rev. Ent.*, *21*, 223-246.

BOURNE, A. I., THIES, W. H. AND SHAW, F. R. 1934. Some observations on long distance dispersal of apple maggot flies. *J. Econ. Ent.*, *27*, 352-355.

BURIFF, C. R. 1973. Recapture of released adult apple maggot flies on sticky-board traps. *Environ. Entomol.*, *2*, 757-758.

BUSH, G. L. 1966. The taxonomy, cytology and evolution of the genus *Rhagoletis* in North America (Diptera: Tephritidae). Museum of Comparative Zoology, Cambridge, Mass., U.S.A.

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

BUSH, G. L. 1969*b*. Mating behavior, host specificity, and the ecological significance of sibling species in frugivorous flies of the genus *Rhagoletis* (Diptera: Tephritidae). *Am. Nat.*, *103*, 669-672.

BUSH, G. L. 1975. Modes of animal speciation. *Ann. Rev. Ecol. Syst.*, *6*, 339-364.

CHAPMAN, P. J. AND LIENK, S. E. 1971. Tortricid *Fauna of Apple in New York*, Cornell Univ., Ithaca, New York. 122pp.

CHRISTIANSEN, F. B. AND FRYDENBERG, O. 1974. Geographic patterns of four polymorphisms in *Zoarces viviporus* as evidence of selection. *Genetics*, *77*, 765-770.

CHRISTIANSEN, F. B. AND FELDMAN, M. W. 1975. Subdivided populations: A review of the one- and two-locus deterministic theory. *Theor. Pop. Biol.*, *7*, 13-38.

CLEGG, M. T. AND ALLARD, R. W. 1972. Patterns of genetic differentiation in the slender wild oat species, *Avena barbata*. *PNAS*, *69*, 1820-1824.

CORRELL, D. S. AND JOHNSTON, M. C. 1970. *Manual of the Vascular Plants of Texas*. Texas Research Foundation, Renner, Texas, 1881pp.

DIEHL, S. R. AND BUSH, G. L. 1984. An evolutionary and applied perspective of insect biotypes. *Ann. Rev. Entomol.*, *29*, 471-504.

ENDLER, J. A. 1977. *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, New Jersey.

FEDER, J. L. AND BUSH, G. L. Differential host utilization is an effective premating barrier between sibling species in the *Rhagoletis pomonella* group. *Evolution* (submitted).

FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1988. Genetic differentiation between sympatric host races of the apple maggot fly, *Rhagoletis pomonella*. *Nature*, *336*, 61-64.

FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1989*a*. Geographic pattern of genetic differentiation between host associated populations of *Rhagoletis pomonella* (Diptera: Tephritidae) in the eastern United States and Canada. *Evolution* (In press).

FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1989*b*. Inheritance and linkage relationships of allozymes in the apple maggot fly. *J. Hered. 80*, 277-283.

FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1989*c*. Regional, local and microgeographic allele frequency variation between host faces of *Rhagoletis pomonella* in western Michigan. *Evolution* (In press).

FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1989*d*. Are the apple maggot, *Rhagoletis pomonella* and the blueberry maggot, *R. mendax*, distinct species? Implications for sympatric speciation. *Entomol. Exp. Appl.* (In press).

FEIN, B. L., REISSIG, W. H. AND ROELOFS, W. L. 1982. Identification of apple volatiles attractive to the apple maggot, *Rhagoletis pomonella*. *J. Chem. Ecol.*, *8*, 1473-1487.

FELSENSTEIN, J. 1976. The theoretical population genetics of variable selection and migration. *Ann. Rev. Genet.*, *10*, 253-280.

FERNALD, M. L. 1950. *Gray's Manual of Botany*, 8th edn. American Book Company, New York, 1632pp.

FISHER, R. A. F. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press. Oxford, England.

HEDRICK, P. W., GINEVAN, M. E. AND EWING, E. P. 1976. Genetic polymorphism in heterogeneous environments. *Ann. Rev. Ecol. Syst.*, *7*, 1-32.

HOLM, S. 1979. A simple sequential rejective multiple test procedure. *Scan. J. Stat.*, *6*, 65-70.

ILLINGWORTH, J. F. 1912. A study of the biology of the apple maggot (*Rhagoletis pomonella*), together with an investigation of methods of control. *Cornell Univ. Agr. Expt. Stat. Bull.*, *324*, 129-187.

JOHNSON, F. M. AND SCHAFFER, H. E. 1973. Isozyme variability in species of the genus *Drosophila*. VII. Genotype-environmental relationships in populations of *Drosophila melanogaster* from the eastern United States. *Biochem. Genet.*, *10*, 149-163.

JOHNSON, G. B. 1976. Polymorphism and predictability of the alpha-glycerophosphate dehydrogenase locus in *Colias* butterflies. *Biochem. Genet.*, *14*, 403-425.

JOHNSON, M. S. 1971. Adaptive lactate dehydrogenase variation in the crested blenny, *Anoplorchus*. *Heredity*, *27*, 205-226.

KOEHN, R. K. 1969. Esterase heterogeneity: Dynamics of a polymorphism. *Science*, *163*, 943-944.

KOEHN, R. K., MILKMAN, R. AND MITTON, J. B. 1976. Population genetics of marine pelecypods. IV. Selection, migration and genetic differentiation in the blue mussel, *Mytilus edulis*. *Evolution*, *30*, 2-32.

LEVENE, H. 1953. Genetic equilibrium when more than one ecological niche is available. *Am. Nat.*, *87*, 331-333.

LEVINS, R. 1968. *Evolution in Changing Environments*. Princeton University Press, Princeton, New Jersey.

MAXWELL, C. W. 1968. Apple maggot adult dispersion in a New Brunswick apple orchard. *J. Econ. Ent.*, *61*, 103-106.

MAXWELL, C. W. AND PARSONS, E. C. 1968. The recapture of marked apple maggot adults in several orchards from one release point. *J. Econ. Ent.*, *61*, 1157-1159.

MAYNARD SMITH, J. 1966. Sympatric speciation. *Am. Nat.*, *100*, 637-650.

MAYNARD SMITH, J. 1970. Genetic polymorphism in a varied environment. *Am. Nat.*, *104*, 487-490.

MCPHERON, B. A. 1987. The population genetics of the colonization of the western United States by the apple maggot, *Rhagoletis pomonella* (Walsh) (Diptera: Tephritidae). Ph.D. Dissertation, Univ. of Illinois, Urbana, Illinois.

MCPHERON, B. A., SMITH, D. C. AND BERLOCHER, S. H. 1988a. Genetic differences between host races of the apple maggot fly. *Nature*, 336, 64-66.

MCPHERON, B. A., JORGENSEN, C. D. AND BERLOCHER, S. H. 1988*b*. Low genetic variability in a Utah cherry-infesting population of the apple maggot, *Rhagoletis pomonella*. *Entomol. Exp. Appl.*, *46*, 155-160.

MERRITT, R. B. 1972. Geographic distribution and enzymatic properties of lactate dehydrogenase allozymes in the fat head minnow, *Pimephales promelas*. *Am. Nat.*, *196*, 173-184.

MILLER, S., PEARCY, R. W. AND BERGER, E. 1975. Polymorphism at the alpha-glycerophosphate dehydrogenase locus in *Drosophila melanogaster*. I. Properties of adult allozymes. *Biochem. Genet.*, *13*, 175-188.

MITTON, J. B. AND KOEHN, R. K. 1975. Genetic organization and adaptive response of allozymes to ecological variables in *fundulus heteroclitus*. *Genetics*, *79*, 97-111.

MOERICKE, V., PROKOPY, R. J., BERLOCHER, S. AND BUSH, G. L. 1975. Visual stimuli eliciting attraction of *Rhagoletis pomonella* flies to trees. *Entomol. Exp. Appl.*, *18*, 497-507.

MUNIYAMMA, M. AND PHIPPS, J. B. 1984. Studies in *Crataegus*. XI. Further cytological evidence for the occurrence of apomixis in North American hawthorns. *Can. J. Bot.*, *62*, 2316-2324.

MUNIYAMMA, M. AND PHIPPS, J. B. 1985. Studies in *Crataegus*. XII. Cytological evidence for sexuality in some diploid and tetraploid species of North American hawthorns. *Can. J. Bot.*, *63*, 1319-1324.

MURRAY, J. 1972. *Genetic Diversity and Natural Selection*. Oliver and Boyd. Edinburgh, Scotland.

NEILSON, W. T. A. 1971. Dispersal studies of a natural population of apple maggot adults. *J. Econ. Ent. 64* (3), 648-653.

O'KANE, W. C. 1914. The apple maggot. *New Hampshire Expt. Sta. Bull.*, No. 171, 120pp.

OWENS, E. D. AND PROKOPY, R. J. 1986. Relationship between reflectance spectra of host plant surfaces and visual detection of host fruit by *Rhagoletis pomonella* flies. *Physiol. Entomol.*, *11*, 297-307.

PAPAJ, D. R. AND PROKOPY, R. J. 1986. Phytochemical basis of learning in *Rhagoletis pomonella* and other herbivorous insects. *J. Chem. Ecol.*, *12*, 1125-1143.

PHIPPS, C. R. AND DIRKS, C. O. 1933. Dispersal of the apple maggot fly—1932 studies. *J. Econ. Ent.*, *5*, 576-582.

PHIPPS, J. B. 1983. Biogeographic, Taxonomic, and Cladistic relationships between East Asiatic and North American *Crataegus*. *Ann. Missouri Bot. Gard.*, *70*, 667-700.

PHIPPS, J. B. 1984. Problems of hybridity in the cladistics of *Crataegus*. In Grant, W. F., (ed.) *Plant Biosystematics* Academic Press, Toronto, Canada, pp. 417-438.

PORTER, B. A. 1928. The Apple Maggot. *U.S. Dept. Agr. Tech. Bull. 66*, 48pp.

PROKOPY, R. J. 1968*a*. Visual responses of apple maggot flies, *Rhagoletis pomonella*: Orchard studies. *Entomol. Exp. Appl.*, *11*, 403-422.

PROKOPY, R. J. 1968*b*. Influence of photoperiod, temperature and food on the initiation of diapause in the apple maggot. *Canad. Ent.*, *100*, 318-329.

PROKOPY, R. J., BENNETT, E. W. AND BUSH, G. L. 1971. Mating behavior in *Rhagoletis pomonella* (Diptera: Tephritidae). I. Site of assembly. *Canad. Ent.*, *103*, 1405-1409.

PROKOPY, R. J., BENNETT, E. W. AND BUSH, G. L. 1972. Mating behavior in *Rhagoletis pomonella* (Diptera: Tephritidae). II. Temporal organization. *Canad. Ent.*, *104*, 97-104.

PROKOPY, R. J., MOERICKE, V. AND BUSH, G. L. 1973. Attraction of apple maggot flies to odor of apples. *Envir. Entomol.*, *2*, 743-749.

PROKOPY, R. J., AVERILL, A. L., COOLEY, S. S. AND ROITBERG, C. A. 1982. Associative learning in egglaying site selection by apple maggot flies. *Science*, *218*, 76-77.

PROKOPY, R. J., PAPAJ, D. R., COOLEY, S. S. AND KALLET, C. 1986. On the nature of learning in oviposition site acceptance by apple maggot flies. *Anim. Behav.*, *34*, 98-107.

PROKOPY, R. J., ALUJA, M. AND GREEN, T. A. 1987. Dynamics of host odor and visual stimulus interaction in host finding behavior of apple maggot flies. In Labeyrie, V., Fabres, G. and Lachaise, D. (eds) *Insects-Plants*, Junk, Netherlands, pp. 161-166.

PROKOPY, R. J., DIEHL, S. R. AND COOLEY, S. S. 1988. Behavioral evidence for host races in *Rhagoletis pomonella* flies. *Oecologia, 76*, 138–147.

RESISSIG, W. H. 1977. Capture of released apple maggot flies, *Rhagoletis pomonella* (Diptera: Tephritidae), on sticky panels. *Can. Ent., 109*, 631–634.

REISSIG, W. H. AND SMITH, D. C. 1978. Bionomics of *Rhagoletis pomonella* in *Crataegus*. *Ann. Ent. Soc. Amer., 71*, 155–159.

REISSIG, W. H., BARNARD, J., WEIRES, R. W., GLASS, E. H. AND DEAN, R. W. 1979. Prediction of apple maggot fly emergence from thermal unit accumulation. *Environ. Entomol., 8*, 51–54.

RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution, 43*, 223–225.

ROITBERG, B. D., VAN LENTEREN, J. C., VAN ALPHEN, J. J. M., GALIS, F. AND PROKOPY, R. J. 1982. Foraging behavior of *Rhagoletis pomonella*, a parasite of hawthorn (*Crataegus*), in nature. *J. Anim. Ecol., 51*, 307–325.

SCRIBER, J. M. AND HAINZE, J. H. 1987. Geographic invasion and abundance as facilitated by differential host plant utilization abilities. In Barbosa, P. and Shultz, J. C. (eds) *Insect Outbreaks: Ecology and Evolutionary Processes*, Academic Press, New York, pp. 433–468.

SHOWERS, W. B. 1981. Geographic variation of the diapause response in the European corn borer. In Denno, R. F. and Dingle, H. (eds) *Insect Life History Patterns*, Springer-Verlag, Berlin and New York, pp. 97–111.

SMITH, D. C. 1988. Heritable divergence of *Rhagoletis pomonella* host races by seasonal asynchrony. *Nature, 336*, 66–67.

SPIESS, E. B. 1977. *Genes in Populations*. Wiley and Sons, New York. 780pp.

STANDLEY, P. C. 1922. *Trees and Shrubs of Mexico*. Smithsonian Press, Washington, D.C., 1721pp.

VIGUE, C. L. AND JOHNSON, F. M. 1973. Isozyme variability in speces of the genus Drosophila. VI. Frequency-property-environment relationships of allelic *alcohol dehydrogenase* in *D. melanogaster*. *Biochem. Genet., 9*, 213–227.

WALSH, B. J. 1867. The apple-worm and the apple maggot. *J. Hort., 2*, 338–343.

WASBAUER, M. S. 1972. An annotated host catalog of the fruit flies of America north of Mexico (Diptera: Tephritidae). Bureau of Entomol., Dept. of Agric., Sacramento, Calif. Occasional papers no. 19.

WEIR, B. S. AND COCKERHAM, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution, 38*, 1358–1370.

Appendix 1 Allele frequencies for *Me* and *Acon*-2 for hawthorn (H) and apple (A) populations of *R. pomonella* across the mid-western United States. At sites 4, 15 and 17 where both larvae and adults were collected, (L) refers to larvae and (Ad) to adults. Sites are arranged within the five transects according to latitude from north to south

	Me						*Acon*-2																	
Site		N	80		N	80		N	73	75	89	95	100	106	114		N	73	75	89	95	100	106	114
Transect 1 (Mi./In.)																								
1	H	46	0·239	A	50	0·660	H	51		0·029	0·059	0·676	0·157	0·059	0·020	A	50		0·020	0·050	0·260	0·510	0·160	
2	H	47	0·532	A	40	0·687	H	47		0·011	0·011	0·372	0·479	0·128		A	39		0·102	0·026	0·192	0·538	0·128	0·013
3	H	38	0·421	A	48	0·656	H	38		0·132	0·026	0·526	0·276	0·040		A	47		0·245	0·053	0·255	0·372	0·074	
4(Ad)	H	30	0·667	A	—	—	H	30			0·017	0·200	0·650	0·133		A	—	—	—	—	—	—	—	—
4(L)	H	30	0·717	A	30	0·717	H	28				0·214**	0·696*	0·072	0·018	A	29		0·017		0·155	0·638	0·190	
5	H	44	0·545	A	45	0·702	H	44		0·079	0·023	0·295	0·500	0·079	0·023	A	41		0·024	0·012	0·098	0·646	0·220	
6	H	36	0·736	A	49	0·735	H	36		0·361	0·069	0·2	0·278	0·069		A	47		0·223	0·032	0·1	0·511	0·117	
7	H	30	0·700	A	—	—	H	30		0·133	0·050	0·117	0·617	0·050	0·033	A	—	—	—	—	—	—	—	—
8	H	35	0·757	A	30	0·760	H	35		0·014	0·029	0·043	0·743	0·1		A	29				0·155	0·552	0·293	
9	H	40	0·800*	A	52	0·789	H	51			0·029	0·078	0·794	0·078	0·020	A	52		0·010		0·067	0·788	0·125	0·010
Transect 2 (Mi.)																								
10	H	50	0·150	A	39	0·538	H	39			0·051	0·692	0·244	0·013		A	39		0·064	0·089	0·372	0·372	0·103	
11	H	—	—	A	41	0·537	H	—	—	—	—	—	—	—	—	A	41		0·024	0·049	0·256	0·524	0·146	
12	H	40	0·587	A	45	0·837	H	40		0·138	0·037	0·312	0·425	0·088		A	45		0·189	0·089	0·089	0·511	0·122	
13	H	48	0·771	A	34	0·706	H	48		0·135	0·115	0·167	0·552	0·021	0·010	A	34		0·088	0·044	0·147	0·544	0·162	0·015
Transect 3 (Mi.)																								
14	H	42	0·226	A	50	0·480	H	42		0·012	0·107	0·619	0·250	0·012		A	50		0·100	0·060	0·160**	0·500	0·180	
15(Ad)	H	345	0·407	A	357	0·620	H	317		0·047	0·062	0·435	0·377	0·077	0·002	A	357		0·083	0·040	0·164	0·542	0·165	0·006
15(L)	H	467	0·345	A	706	0·633	H	429		0·026	0·062	0·479	0·346	0·078	0·009	A	665		0·064	0·029	0·194	0·560	0·148	0·005
16	H	33	0·561	A	36	0·708	H	32		0·172	0·016	0·391	0·344	0·078		A	36		0·157	0·043	0·100	0·543	0·157	
17(Ad)	H	25	0·660	A	—	—	H	36		0·153	0·028	0·361	0·333	0·111	0·014	A	—	—	—	—	—	—	—	—
17(L)	H	33	0·637	A	30	0·634	H	32		0·047	0·016	0·406	0·469	0·062		A	30		0·054	0·036	0·054	0·820	0·036	
18	H	32	0·750	A	31	0·500	H	32		0·094	0·062	0·250	0·532	0·062		A	31		0·097	0·048	0·177*	0·436	0·242	
Transect 4 (Wi./Ill.)																								
19	H	110	0·309	A	106	0·542	H	111		0·054	0·122	0·504	0·266	0·045	0·009	A	106	0·005	0·014	0·024	0·259	0·514	0·175	0·009
20	H	156	0·304*	A	133	0·553	H	158	0·003	0·009	0·070	0·592	0·256	0·066	0·003	A	132	0·004		0·045	0·348	0·436	0·163	0·004
21	H	43	0·535	A	42	0·714	H	43		0·023	0·081	0·279*	0·511	0·105		A	40	0·025	0·025	0·075	0·088	0·562	0·213	0·012
22	H	39	0·590	A	45	0·689	H	38		0·026	0·053	0·329	0·487	0·079	0·026	A	45		0·022	0·067	0·156	0·589	0·144	0·022
23	H	43	0·605	A	51	0·666	H	43		0·058	0·046	0·221*	0·523	0·151		A	50		0·030	0·070	0·180	0·610	0·110	
24	H	29	0·552	A	50	0·680**	H	29		0·018	0·086	0·310	0·500	0·086		A	49		0·010	0·061	0·102	0·673	0·154	
25	H	49	0·623	A	80	0·719*	H	49		0·010	0·143	0·398	0·408	0·041		A	79		0·032	0·070	0·108	0·595	0·190	0·006
26	H	84	0·780	A	82	0·604	H	82			0·061	0·134	0·683	0·116	0·006	A	76	0·013	0·020	0·033	0·151	0·586	0·171	0·026
27	H	53	0·774	A	—	—	H	51		0·029	0·167	0·176*	0·490	0·127	0·010	A	—	—	—	—	—	—	—	—
28	H	60	0·825	A	60	0·742	H	60	0·017	0·017	0·125	0·083	0·642	0·092	0·025	A	60	0·092	0·017	0·099	0·117	0·475	0·183	0·017
Transect 5 (Wi.)																								
29	H	32	0·470	A	42	0·595	H	32			0·125	0·484	0·344	0·047		A	42			0·062	0·238	0·538	0·150	0·012
30	H	44	0·568	A	44	0·795	H	44		0·068	0·239	0·420	0·193	0·011		A	44		0·093	0·116	0·140	0·593	0·058	
31	H	29	0·930	A	—	—	H	29		0·017	0·103	0·345	0·448	0·086		A	—	—	—	—	—	—	—	—
32	H	45	0·667	A	44	0·636	H	45		0·011	0·200	0·089	0·489	0·211		A	44		0·013	0·105	0·184	0·539	0·158	
Misc. (Wi.)																								
33	H	20	0·750	A	—	—	H	18			0·110	0·306	0·528	0·056		A	—	—	—	—	—	—	—	—
34	H	37	0·815	A	—	—	H	36			0·236	0·125	0·528	0·111		A	—	—	—	—	—	—	—	—

* $P \leqq 0{\cdot}05$. ** $P \leqq 0{\cdot}01$. *** $\leqq 0{\cdot}001$. G-test for significant deviation from Hardy–Weinberg equilibrium.

Appendix 2 Allele frequencies for *Aat*-2 for hawthorn (H) populations of *R. pomonella* across the mid-western United States. At sites 4, 15 and 17 were both larvae and adults were collected, (L) refers to larvae and (Ad) to adults. Sites are arranged within the five transects according to latitude from north to south

			Aat-2											
Site	Host	N	01	21	32	40	50	59	75	84	100	112	123	130
Transect 1 (Mi./In.)														
1	H	43		0·070			0·151		0·337		0·430		0·012	
2	H	47		0·074			0·223		0·245		0·383		0·064	
3	H	34		0·044			0·235*		0·250		0·426		0·044	
4(Ad)	H	30		0·100			0·134	0·017	0·433		0·266		0·050	
4(L)	H	27	0·018	0·111			0·167		0·259		0·426		0·018	
5	H	37		0·148			0·230		0·230		0·365		0·027	
6	H	36	0·028	0·069	0·028		0·222	0·042	0·139		0·417		0·042	0·014
7	H	25		0·140			0·140	0·020	0·280		0·380		0·040	
8	H	34		0·206			0·235*		0·176		0·353		0·029	
9	H	40		0·200			0·200	0·038	0·175		0·375			0·013
Transect 2 (Mi.)														
10	H	38		0·066			0·118		0·276		0·513		0·026	
11	H	—	—	—	—	—	—	—	—	—	—	—	—	—
12	H	34		0·118			0·162		0·265		0·412		0·044	
13	H	47	0·011	0·106			0·309	0·011	0·213		0·351			
Transect 3 (Mi.)														
14	H	42		0·059			0·167		0·274		0·476		0·024	
15(Ad)	H	271	0·006	0·103	0·006		0·138	0·013	0·275		0·415		0·039	0·006
15(L)	H	414	0·001	0·091	0·002	0·001	0·127		0·271		0·461		0·040	0·006
16	H	33		0·061			0·076		0·303		0·545		0·015	
17(Ad)	H	36		0·153	0·014		0·167		0·181		0·389		0·097	
17(L)	H	27		0·167	0·037		0·148		0·241		0·389		0·018	
18	H	32		0·140			0·234		0·250		0·359		0·016	
Transect 4 (Wi./Ill.)														
19	H	92	0·005	0·077	0·005		0·141		0·353		0·386		0·027	0·005
20	H	156		0·083	0·003		0·080	0·003	0·353		0·420	0·003	0·048	0·006
21	H	41		0·098	0·012		0·134	0·012	0·280	0·012	0·402		0·049	
22	H	37		0·176			0·176		0·216*		0·392		0·027	0·013
23	H	43	0·023	0·128			0·221	0·012	0·116		0·477		0·012	0·012
24	H	23		0·130			0·456		0·152		0·261*			
25	H	48	0·010	0·135			0·198		0·198*		0·427		0·031	
26	H	80		0·106			0·156	0·006	0·213		0·475		0·044	
27	H	52		0·183			0·183*		0·192		0·433		0·009	
28	H	60	0·017	0·142			0·317		0·200		0·283		0·025	0·017
Transect 5 (Wi.)														
29	H	28		0·089			0·161	0·018	0·339		0·339		0·054	
30	H	44	0·011	0·136			0·136	0·011	0·307		0·375		0·023	
31	H	18		0·167			0·167		0·083		0·528		0·056	
32	H	43		0·105			0·221		0·232		0·407		0·035	
Misc. (Wi.)														
33	H	20		0·075			0·175		0·225		0·475		0·050	'
34	H	33		0·121			0·152		0·242		0·394*		0·091	

* $P \leqq 0{\cdot}05$. ** $P \leqq 0{\cdot}01$. *** $P \leqq 0{\cdot}001$. G-test for deviation from Hardy–Weinberg equilibrium.

Appendix 3 Allele frequencies for *Aat*-2 for apple (A) populations of *R. pomonella* across the mid-western United States. At sites 4, 15 and 17 where both larvae and adults were collected, (L) refers to larvae and (Ad) to adults. Sites are arranged within transects according to latitude from north to south

			Aat-2											
Site	Host	N	01	21	32	40	50	59	75	84	100	112	123	130
Transect 1 (Mi./In.)														
1	A	50		0·140			0·160		0·330		0·330		0·040	
2	A	34		0·191*			0·236*		0·279	0·015	0·250		0·029	
3	A	44		0·170		0·011	0·273		0·227		0·295*		0·023	
4(Ad)	A	—	—	—	—	—	—	—	—	—	—	—	—	—
4(L)	A	29		0·069			0·259		0·310		0·328		0·017	0·017
5	A	41		0·220			0·134		0·268		0·354		0·024	
6	A	46		0·130			0·196	0·011	0·272**	0·011	0·326		0·054	
7	A	—	—	—	—	—	—	—	—	—	—	—	—	—
8	A	30		0·150			0·283		0·283		0·283			
9	A	50		0·090			0·220	0·010	0·300	0·010	0·330		0·040	
Transect 2 (Mi.)														
10	A	38		0·132			0·237*	0·013	0·263		0·381		0·013	
11	A	36	0·014	0·097			0·153		0·292		0·403		0·042	
12	A	42		0·119			0·190		0·381		0·274		0·036	
13	A	34		0·118			0·190		0·221		0·368		0·103	
Transect 3 (Mi.)														
14	A	47		0·181	0·032		0·149	0·011	0·447		0·181			
15(Ad)	A	297	0·002	0·163	0·003		0·175		0·300		0·337		0·020	
15(L)	A	662	0·002	0·150	0·002	0·001	0·186*	0·001	0·285**		0·344*		0·026	0·002
16	A	33		0·288			0·242		0·121		0·319			
17(Ad)	A	—	—	—	—	—	—	—	—	—	—	—	—	—
17(L)	A	25		0·160			0·300		0·140		0·400			
18	A	31	0·016	0·161			0·306		0·242		0·194		0·065	0·016
Transect 4 (Wi./Ill.)														
19	A	104		0·149	0·010		0·168		0·332		0·327		0·010	0·005
20	A	133	0·004	0·180	0·008		0·173	0·041	0·274		0·286		0·026	0·008
21	A	39		0·115	0·026		0·141	0·013	0·372		0·333			
22	A	45		0·167			0·200*		0·278		0·344		0·011	
23	A	51		0·176			0·167		0·304		0·294		0·038	0·020
24	A	49		0·153	0·020		0·173	0·010	0·245		0·388		0·010	
25	A	76		0·184			0·151		0·270		0·349		0·046	
26	A	71		0·190	0·007		0·176		0·296		0·289	0·014	0·028	
27	A	—	—	—	—	—	—	—	—	—	—	—	—	—
28	A	60		0·167			0·233		0·258		0·317		0·008	0·017
Transect 5 (Wi.)														
29	A	41		0·085			0·122	0·016	0·293		0·439		0·024	
30	A	42		0·119			0·202	0·012	0·286	0·012	0·369			
31	A	—	—	—	—	—	—	—	—	—	—	—	—	—
32	A	33		0·167	0·015		0·182		0·197		0·409		0·021	
Misc. (Wi.)														
33	A	—	—	—	—	—	—	—	—	—	—	—	—	—
34	A	—	—	—	—	—	—	—	—	—	—	—	—	—

* $P \leqq 0{\cdot}05$. ** $P \leqq 0{\cdot}01$. *** $P \leqq 0{\cdot}001$. G-test for deviation from Hardy-Weinberg equilibrium.

Appendix 4 Allele frequencies for *Mpi*, *Had* and *Dia*-2 for hawthorn (H) and apple (A) populations of *R. pomonella* across the mid-western United States. At sites 4, 15 and 17 where both adults and larvae were collected, (L) refers to larvae and (Ad) to adults. Sites are arranged within transects according to latitude from north to south

			Mpi							*Had*								*Dia*-2					
Site		N	31	37	70	100	125		N	31	37	70	100	125		N	122		N	122		N	85		N	85
Transect 1 (Mi./In.)																									
1	H	37	0·013	0·189	0·095	0·703		A	34	0·029		0·015	0·941	0·015	H	49	0·153	A	51	0·206	H	49	0·224*	A	51	0·304
2	H	47		0·138	0·043	0·819		A	33	0·030	0·030		0·940		H	47	0·245	A	39	0·346	H	47	0·277**	A	28	0·446*
3	H	37	0·040	0·136	0·040	0·784		A	35	0·086	0·100	0·029	0·786		H	38	0·276	A	48	0·333	H	32	0·281**	A	44	0·307
4(Ad)	H	30		0·050	0·067	0·883		A	—	—	—	—	—	—	H	29	0·241	A	—	—	H	30	0·233*	A	—	—
4(L)	H	30		0·050	0·067	0·883		A	30		0·017	0·033	0·950		H	30	0·183	A	30	0·250**	H	28	0·232	A	30	0·233
5	H	45		0·067	0·044	0·867	0·022	A	38	0·013	0·066	0·026	0·895		H	45	0·256	A	45	0·278	H	44	0·273	A	45	0·389
6	H	36	0·028	0·069	0·042	0·861		A	28	0·107	0·054		0·839		H	36	0·292	A	49	0·245	H	36	0·278	A	48	0·281
7	H	31			0·016	0·968	0·016	A	—	—	—	—	—	—	H	31	0·468	A	—	—	H	27	0·333	A	—	—
8	H	35	0·014	0·057	0·029	0·900		A	30		0·050		0·950		H	35	0·500	A	30	0·433	H	36	0·375	A	30	0·350
9	H	40	0·025		0·050	0·925		A	50	0·060	0·030	0·050	0·860		H	40	0·475	A	52	0·279	H	40	0·388*	A	41	0·341
Transect 2 (Mi.)																									
10	H	29	0·035	0·241	0·035	0·672	0·017	A	33	0·015	0·136	0·091	0·758		H	50	0·060	A	39	0·282	H	47	0·160	A	39	0·295
11	H	—	—	—	—	—	—	A	42	0·012	0·095		0·893		H	—	—	A	42	0·238	H	—	—	A	38	0·329
12	H	40	0·025	0·050	0·088	0·838		A	45	0·022	0·067	0·056	0·833	0·022	H	40	0·250	A	45	0·300	H	40	0·275	A	45	0·267
13	H	30		0·050	0·033	0·900	0·017	A	34	0·015	0·089	0·059	0·823	0·015		48	0·458*	A	33	0·288	H	47	0·404	A	31	0·274
Transect 3 (Mi.)																									
14	H	42	0·012	0·214	0·048	0·714	0·012	A	47	0·032	0·042	0·011	0·915		H	42	0·107	A	50	0·180	H	31	0·145*	A	47	0·319
15(Ad)	H	308	0·013	0·112	0·047	0·818	0·010	A	353	0·013	0·045	0·054	0·885	0·004	H	340	0·154	A	356	0·226	H	313	0·244	A	307	0·298
15(L)	H	417	0·035	0·096	0·047	0·818***	0·005	A	660	0·011	0·049	0·046	0·892		H	464	0·170	A	688	0·203	H	407	0·206	A	672	0·310
16	H	28	0·036	0·143	0·036	0·786		A	34		0·059	0·074	0·867		H	33	0·303	A	36	0·292	H	33	0·136	A	35	0·386
17(Ad)	H	31	0·039	0·013	0·066	0·882		A	—	—	—	—	—	—	H	29	0·362	A	—	—	H	38	0·302	A	—	—
17(L)	H	38	0·016	0·064	0·081	0·839		A	27	0·019	0·037	0·074	0·870		H	33	0·318	A	30	0·183	H	29	0·327	A	30	0·467
18	H	32	0·016	0·078	0·016	0·890		A	31		0·065	0·080	0·855		H	32	0·313	A	31	0·322	H	32	0·359	A	31	0·484
Transect 4 (Wi./Ill.)																									
19	H	104	0·024	0·149	0·034	0·788	0·005	A	99	0·020	0·040	0·076	0·854	0·010	H	109	0·220	A	106	0·208	H	108	0·218	A	105	0·329
20	H	158		0·199	0·092	0·668	0·041	A	132	0·004	0·091	0·079	0·789	0·038	H	157	0·137	A	133	0·226	H	155	0·132	A	132	0·326
21	H	43	0·012	0·105	0·047	0·802	0·023	A	38	0·013	0·013	0·066	0·908		H	43	0·186	A	42	0·226	H	41	0·232	A	41	0·256
22	H	39		0·102	0·026	0·872		A	45	0·022	0·033	0·033	0·912		H	25	0·120	A	45	0·278	H	37	0·284	A	45	0·322
23	H	40		0·125	0·087	0·788		A	31	0·032	0·032	0·032	0·903		H	43	0·186	A	51	0·275	H	43	0·384	A	51	0·363
24	H	21		0·143	0·023	0·833		A	40	0·012	0·063	0·050	0·875		H	15	0·333	A	47	0·287	H	29	0·575	A	49	0·337
25	H	40	0·025	0·113		0·862		A	80	0·044	0·044	0·025	0·888		H	34	0·235	A	65	0·200	H	34	0·265	A	65	0·315
26	H	83	0·042	0·030	0·048	0·880		A	79	0·025	0·006	0·070	0·899		H	84	0·357	A	83	0·367	H	77	0·260	A	74	0·331*
27	H	53	0·028	0·010	0·019	0·943		A	—	—	—	—	—	—	H	52	0·538	A	—	—	H	43	0·395	A	—	—
28	H	60	0·017	0·017	0·033	0·933		A	45	0·022		0·044	0·933		H	60	0·567	A	60	0·383	H	58	0·500	A	60	0·400
Transect 5 (Wi.)																									
29	H	32		0·062	0·062	0·860	0·016	A	41	0·037	0·037	0·037	0·853	0·037	H	32	0·125	A	42	0·083	H	32	0·156	A	42	0·143
30	H	42	0·012	0·083	0·107	0·762	0·036	A	40	0·025	0·050	0·113	0·812		H	44	0·295*	A	44	0·250	H	39	0·295	A	42	0·321
31	H	29			0·034	0·966		A	—	—	—	—	—	—	H	29	0·345	A	—	—	H	27	0·296	A	—	—
32	H	45			0·044	0·956		A	34	0·029		0·029	0·942		H	43	0·372	A	44	0·341	H	38	0·250	A	37	0·365
Misc. (Wi.)																									
33	H	20	0·075	0·025	0·050	0·825	0·025	A	—	—	—	—	—	—	H	20	0·250	A	—	—	H	20	0·225	A	—	—
34	H	33	0·076	0·030	0·015	0·879		A	—	—	—	—	—	—	H	36	0·333	A	—	—	H	35	0·314*	A	—	—

* $P \leqq 0{\cdot}05$ ** $P \leqq 0{\cdot}01$ *** $P \leqq 0{\cdot}001$ G-test for significant deviation from Hardy-Weinberg equilibrium.