

Population structure of *Rhagoletis mendax*, the blueberry maggot

STEWART H. BERLOCHER

Department of Entomology, University of Illinois at Urbana-Champaign, 320 Morrill Hall, 505 South Goodwin Avenue, Urbana, IL 61801, U.S.A.

Rhagoletis mendax, the blueberry maggot, is a member of the *R. pomonella* species complex, well-known as a possible example of sympatric speciation via host race formation. Using electrophoretic data for 16 loci, I have determined that the population structure of *R. mendax* differs sharply from that of the well-known apple maggot fly, *R. pomonella*, as follows: (i) geographical differentiation across eastern North America is very small ($F_{ST}=0.015$ as opposed to $F_{ST}=0.220$ in *R. pomonella*), (ii) no latitudinal clines in allele frequencies exist, and (iii) consistent differentiation with respect to individual host plant and host plant species was not seen (although such differentiation could not be studied as effectively in *R. mendax* as in *R. pomonella*). *Fum* and *Had* show strong linkage disequilibrium in two northern populations, and weak but significant disequilibrium across the entire geographical range of *R. mendax*. The genetic uniformity of *R. mendax* may be a consequence of its restriction to a single host plant species over most of its range.

Keywords: allozymes, population genetics, population structure, *Rhagoletis*, sympatric speciation, Tephritidae.

Introduction

For almost three decades the fruit flies of the *Rhagoletis pomonella* species group (henceforth *pomonella* group) have been the focus of debate over sympatric speciation via host race formation, a mode of speciation in which reproductive isolation evolves, in the absence of geographical isolation, during the adaptation of parasites to novel hosts (Mayr, 1963; Bush, 1966, 1969, 1974, 1975, 1992; Futuyma & Mayer, 1980). Bush hypothesized that genetic variants with altered host selection behaviour and host survival ability can segregate out of a parasite population, producing a new population, on a new host, that is at least partially reproductively isolated from the original population on the ancestral host (namely, a host race). The existence of at least one host race formed in historical time, the apple race of *R. pomonella* (Walsh), is now well supported (Prokopy *et al.*, 1987; Feder *et al.*, 1988, 1990a,b; McPheron *et al.*, 1988b; Smith, 1988).

R. pomonella (Walsh) is the only species in the *pomonella* species complex (Bush, 1966; Berlocher *et al.* 1993; Payne & Berlocher, 1995a) that has been reported to form host races in historical time. This fact prompts the question: what makes *R. pomonella* unique? Electrophoretic and phenological studies have

revealed several features of the biology and population genetics of the species that are potentially related, to varying degrees, to its ability to colonize novel plants. *R. pomonella* possesses abundant variation, at least for allozymes (average heterozygosity over 29 loci is 0.221 ± 0.44 ; Berlocher *et al.*, 1993), populations are structured even at the level of individual host plants (McPheron *et al.*, 1988a; Feder *et al.*, 1990b), latitudinal allele frequency clines exist, with different slopes in the apple and hawthorn races (Feder & Bush, 1989), geographical differentiation over eastern North America is great ($F_{ST}=0.220$; McPheron, 1987), linkage disequilibrium of allozyme loci involved in host race differentiation exists (Feder *et al.*, 1990a), and different allozyme genotypes are associated with different diapause and eclosion phenologies (Feder *et al.*, 1993).

In this paper, I examine the *pomonella* group sibling species *R. mendax* Curran (blueberry maggot) at 16 variable allozyme loci to determine whether its population structure differs from that of *R. pomonella*. I sampled *R. mendax* sufficiently well to allow the magnitude of geographical differentiation, pattern of geographical differentiation (for example clines), and amount and degree of linkage disequilibrium to be assessed with some confidence. Provisional answers can be given to questions concerning differentiation of

R. mendax populations on individual host plants, and on different host plant species, but no data on associations between allozyme and life history phenology were obtained in this study.

Biology of *R. mendax*

The life cycle of *R. mendax* is typical of most *Rhagoletis* species (Boller & Prokopy, 1975). The insect is univoltine, with eclosion of adults in the summer as the host fruits ripen, mating on the host plants, oviposition into the fruit, larval development in a single fruit, and pupal diapause and overwintering in the soil. All *pomonella* group species have the same basic karyotype (Bush, 1966), so the linkage map of *R. mendax* should be similar to that of *R. pomonella* (Berlocher & Smith, 1983; Feder *et al.*, 1989b). *R. mendax* is restricted to eastern North America (Bush, 1966; Fig. 1). It infests almost exclusively species of the closely related genera *Vaccinium* L. (blueberries, bilberries, lingonberries, etc.) and *Gaylussacia* Humboldt, Bonpland and Kunthe (huckleberries), both in the Ericaceae (Bush, 1966). Given that in some areas *R. mendax* commonly infests plants of *Vaccinium* subgenus *Cyanococcus*, which contains all the blueberries commonly harvested for market in North America, *R. mendax* is a serious agricultural pest.

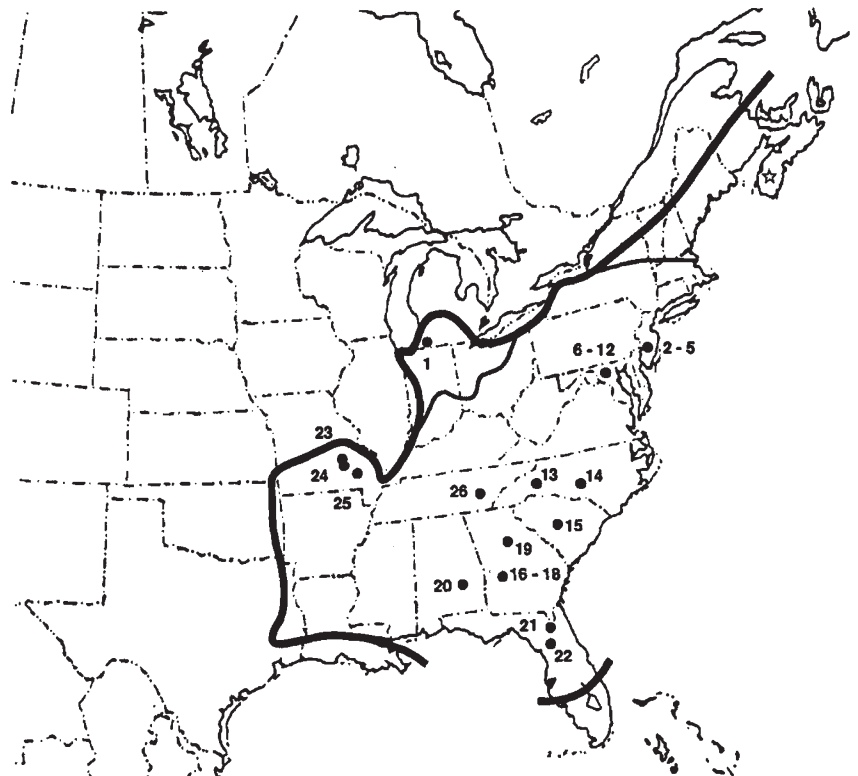
Two major steps toward a better understanding of the biology of *R. mendax* were recently made by J. Payne (Payne & Berlocher, 1995a,b). First, the populations of *pomonella* group flies infesting the fall-ripening blueberry *V. arboreum* Marshall, previously classified with *R. mendax* by Bush (1966), were found to be a new species of the *pomonella* complex (Payne & Berlocher, 1995a), and are thus not discussed here. Secondly, by far the most commonly infested host plant in much of eastern North America is a previously totally unsuspected one: *V. stamineum* L. (deerberry; Payne & Berlocher, 1995b). The discovery of this host has resulted in a substantial westward extension of the geographical range of *R. mendax* (Fig. 1; Payne & Berlocher, 1995b).

Materials and methods

Sampling

All insects were obtained by rearing them from infested fruit collected in the field (Berlocher & Enquist, 1993; Payne & Berlocher, 1995b). The original intention was to obtain all fly samples as separate collections from individual host plants, but as this study progressed it became clear that adequate electrophoretic sample sizes could in most cases be

Fig. 1 Map of eastern North America with collection sites for *Rhagoletis mendax* numbered as in Table 1. The ☆ is the Nova Scotia site of Berlocher & Bush (1982). The thick line marks the approximate range limit of *R. mendax*, based on Bush (1966), Vincent & Lareau (1989), Payne & Berlocher (1995b) and Vander Kloet (1988). The thin line, which coincides with the thick line at the western and southern limits, is the range of deerberry (*Vaccinium stamineum*) from Vander Kloet (1988). Only in the northern Atlantic seaboard and midwestern USA areas does *R. mendax* occur outside the range of deerberry.



obtained only by pooling fly collections from many individual plants at a site. Three individual plant samples were made at the Fort Valley, Georgia site (Table 1). Most sites were from piedmont or coastal plain, within 3300 m of sea level. Exceptions were Powhattan Lake, N. Carolina (about 1200 m), Toccoa, Georgia (about 900 m), Taum Sauk Mountain, Missouri (about 520 m) and Savage Gulf Natural Area, Tennessee (about 520 m).

Geographical sampling was designed to obtain frequency data from the entire known range of *R. mendax*. However, because electrophoretic data for some loci have been published for Nova Scotia (Berlocher & Bush, 1982; Feder *et al.*, 1989a), collections were not made at the northernmost extreme of the range of the fly. Sites with more than one potential host plant growing together were especially sought, and at such sites collections were made from all potential

host species with ripe fruit. Plants were identified according to Vander Kloet (1988). For vouchering information see Payne & Berlocher (1995b).

Sites yielding sufficient material for electrophoresis are described in Table 1 and mapped in Fig. 1; for additional localities and locality information see Payne & Berlocher (1995b). Relevant additional details for sites are: (1) the same collection of flies was used for some data in Berlocher *et al.* (1993), across Chickaming Road from the McMarrin site of Feder *et al.* (1989a); (2–4) domestic plants in an open field, about 0.5 km from natural area (pine barrens) with three wild hosts intermixed; (6–12) three hosts growing completely intermixed; (16–18) Reade property, Peach Co., collections from seven individual plants, three of which yielded enough flies for analysis, plants of samples 16 and 17 about 25 m apart, plant of sample 18 about 500 m from the others.

Table 1 Sample information for *Ragoletis mendax*

Code	Site		Collection date	Host species	Latitude	Run as
	State	Locality (County)				
1	Michigan	Chickaming (Berrien)	7/8/90	<i>V. corymbosum</i> (D)	41°53'	L,A
2	New Jersey	Rutgers U. Cranberry and Blueberry Res. Sta. (Burlington)	31/7/91	<i>V. corymbosum</i> (D)	39°41'	L
3				<i>V. corymbosum</i>		L
4				<i>G. baccata</i>		L
5				<i>G. frondosa</i>		L
6	Maryland	Beltsville Natl. Agri. Res. Center (Prince Georges)	26/6/91	<i>G. baccata</i>	39° 4'	L
7			16/7/91	<i>V. stamineum</i>		L
8				<i>G. baccata</i>		L
9				<i>G. frondosa</i>		L
10			30/7/91	<i>V. stamineum</i>		L
11				<i>G. frondosa</i>		L
12			8/9/91	<i>V. stamineum</i>		L
13	N. Carolina	Powhattan Lake Campground (Buncombe)§	27/7/91	<i>V. stamineum</i>	35°30'	L
14		Candor (Montgomery)	28/7/91	<i>V. stamineum</i>	35°11'	L
15	S. Carolina	Clemson Univ. Sandhills Res. Sta. (Kershaw)	25/7/90	<i>V. stamineum</i>	34° 5'	L
16	Georgia	Fort Valley (Peach)	25/7/90	<i>V. stamineum</i> ‡	31°33'	A
17				<i>V. stamineum</i> ‡		A
18				<i>V. stamineum</i> ‡		A
19		Toccoa (Stevens)§	17/7/91	<i>V. corymbosum</i> (D)†	34°48'	L
20	Alabama	Tuskegee National Forest (Macon)	22/7/90	<i>V. stamineum</i>	32°30'	L
21	Florida	White Springs (Hamilton)	24/7/90	<i>V. stamineum</i>	30°33'	L
22		Ichetucknee Springs State Park (Columbia)	23/7/90	<i>V. stamineum</i>	29°55'	L
23	Missouri	Route 8 and Co. Y (Crawford)	23/7/91	<i>V. stamineum</i>	37°48'	L
24		Cortois (Washington)	23/7/91	<i>V. stamineum</i>	37°47'	L
25		Taum Sauk Mountain (Iron)§	23/7/91	<i>V. stamineum</i>	37°33'	L
26	Tennessee	Savage Gulf Natural Area (Grundy)§	4/8/91	<i>V. stamineum</i>	35°27'	L,A

D indicates a domesticated planting; A, adult; L, larva; 'code' used in Appendix.

†Rabbiteye blueberry, a geographical variant of *V. corymbosum* often classified as *V. ashei* Reade.

‡The three samples were each collected from a different individual plant.

§Sites at higher altitude than remainder; see text.

Electrophoresis

Standard single-condition horizontal starch gel electrophoresis was performed using Starch Art (Starch Art, Smithville, TX) starch. Flies from the Chickaming, Michigan site were included in every gel as controls, and *R. pomonella* from hawthorns in Urbana, IL, were included on some gels as additional controls. Gels were scored immediately after staining and were photographed for permanent record. For brevity, I equate the terms allele and electromorph.

Because nomenclature, electrophoretic conditions, staining recipes and map positions for *R. pomonella* group electrophoretic loci have been previously described in detail (Berlocher & Smith, 1983; Berlocher *et al.*, 1993; Feder *et al.*, 1993), only locus names are used here. Although as many as 29 loci have been scored from single *Rhagoletis* flies (Feder *et al.*, 1989), the data set analysed here contains 16 polymorphic loci (polymorphic = more than one allele), all expressed in both larvae and adults (Berlocher & Smith, 1983; Feder *et al.*, 1989). The loci that were excluded from analysis were either invariant in *R. mendax* or could not be reliably scored for technical reasons (the relatively small size of southern *R. mendax* was a contributing factor). Some almost monomorphic loci, such as *Dia-2* and *Ald*, were included to enable interspecific hybrids to be detected. Partial data, obtainable from the author, were obtained for some of the excluded loci.

Electrophoresis was in most cases carried out on larvae (Table 1). Both adults and larvae were electrophoresed in the large Chickaming, Michigan sample to determine whether larval and adult frequencies differed at the same site. For most sites, some flies were reserved for DNA and morphometric analysis.

Data analysis

BIOSYS-1 (Swofford & Selander, 1981) was used to compute allele frequencies and unbiased heterozygosities (Nei, 1978) from the individual genotypes. The BIOSYS-1 exact test option was used to determine whether populations were in Hardy-Weinberg Equilibrium (HWE).

Weir (1990a, p. 367) stresses the need to identify the scope of inference explicitly in analysing population differentiation. Fixed model analysis, such as contingency table analysis, permits inference only about the available samples, whereas random model analyses such as *F*-statistics permit inference about the entire species. I used contingency analysis as well as *F*-statistics, as have most past electrophoretic studies; contingency analysis permits a simple evaluation of

adequacy of sampling, independent of any particular genetic model.

Contingency tests of allele frequency differentiation were carried out on allele counts. Instead of the usual χ^2 test, I employed a Monte Carlo simulation of an exact $R \times C$ test (MONTE CARLO $R \times C$ 2.2, Macintosh program provided by W. Engles, Department of Genetics, University of Wisconsin, Madison, WI). This eliminated the need for pooling of alleles and permitted testing involving rare alleles. However, even with exact methods or their simulations, tables with extremely low total counts of rare alleles cannot possibly refute the null hypothesis. Therefore, tables containing fewer than five copies in total of the second most common allele were not tested. MONTE CARLO $R \times C$ 2.2 estimates standard error for the probability. For most cases 10 000 replications were adequate to determine the probability of obtaining the observed data or more deviant data by chance; for these cases probabilities are given in the tables to two decimal places. However, if the probability was close to or below 0.01, 100 000 replications were carried out, which in almost all cases ensured accuracy to three decimal places. In a few cases the run was carried out to 10^6 replications, ensuring accuracies to at least four decimal places. In the event that all 10^6 random samples deviated less from expectations than did the actual sample, the probability is recorded as 0 in the tables.

In principle, contingency table analysis could be carried out in a nested analysis of the effects of individual host plants, host plant species, time of year of collection, and geography, analogous to ordinary analysis of variance. In practice, obtaining a complete balanced set of samples was impossible (see Results). The following samples permitted a test of each factor independently.

To test for individual plant effects Collections from three individual plants of *V. stamineum* sampled at the Fort Valley, Georgia site (collections 16–18) on the same day.

To test for host species effects A set of four samples (pooled from many individual plants) representing three host plant species, made on the same day at the Rutgers University Cranberry and Blueberry Research Station, New Jersey, site. One of the samples is from domesticated highbush blueberry (*V. corymbosum*) under cultivation, the others from three wild hosts, including wild *V. corymbosum*.

To test for host species effects and phenological effects A set of seven samples representing three host plant species (pooled from many individual host plants

of each species) made at four times in the summer of 1991 at the Beltsville site. Because hosts ripen at different times as the season progresses, a complete set of host-collection time combinations could not be made.

To determine whether the three samples from domesticated blueberries show any consistent genetic similarities, unweighted pair-group arithmetic mean (UPGMA) cluster analysis was carried out on the Nei unbiased genetic identity (Nei, 1978).

Before analysing the large-scale geographical differentiation among the samples, all samples at the same geographical locality, and some geographically close sites, were pooled to increase the sample sizes representing geographical areas. (Although lack of significant heterogeneity within a locality was not a criterion for pooling, I note that the two small Florida samples did not differ significantly at any locus.) Pooling produced 13 geographical samples (with the five pooled samples being sites 2-5, 6-12, 16-18, 21-22, and 23-24 in Table 1). These pooled samples were used for both the contingency and *F*-statistics analyses of large-scale geographical variation. The *F*-statistics of Cockerham & Weir (Weir, 1990a,b) were calculated using a BASIC program I wrote that accepts modified BIOSYS-1 input files. The program also computes 95 per cent confidence limits by bootstrapping over loci. The example data of Weir (1990a, p. 410) were used to confirm that the program was performing correctly. For comparison, I also calculated the *F*-statistics of Wright (1978).

To determine whether clines exist in *R. mendax*, I carried out nonparametric regression analysis of the frequency of each allele against latitude using Kendall's τ . The nonparametric approach was used because neither allele frequencies nor latitude are normally distributed. Alleles for which the highest frequency was less than 0.02 were excluded. For sites with multiple samples, I used the pooling described in the preceding paragraph. Two analyses were performed. First, all populations were included. Secondly, to remove potential altitude and interior/coastal effects, I regressed only samples from the Atlantic coastal plain and piedmont (2-5, 6-12, 14, 15, 16-18, and 21-22). For alleles of *Aat-2*, *Adh-1*, *Ald*, *Fum*, *Idh*, *Pgi* and *Pgm*, the Nova Scotia sample of Berlocher & Bush (1982) was included in the regression analysis, as the electrophoretic buffers were the same as used in this study.

As in Berlocher (1993), the analysis of linkage disequilibrium had two stages. First, contingency tests on pairwise two-locus genotype numbers were carried out. By using the Monte Carlo method described above, problems with small expected numbers were avoided. (Fisher's exact test was used on 2×2 tables.) The

criteria for selecting populations and loci for linkage disequilibrium tests were that $N \geq 30$, and that the count of the second most abundant allele at each locus was ≥ 5 .

Even given the ability to use rare alleles afforded by an exact test, large samples are needed to detect any but the most intense disequilibria (Hedrick *et al.*, 1978). Therefore, several different poolings (weighted by sample size), up to a pooling of all samples, were made to increase the ability to detect disequilibria. The danger is that artificial pooling of structured populations may create disequilibria not present in individual subpopulations (Nei & Li, 1973; see Results).

Only samples with probabilities below 0.01 (see below) in the contingency tests were carried to the second stage of measuring *D*. *D* was computed on data pooled to two alleles per locus (most common and all others) using the method of Hill (1974). As *R. mendax* populations were in HWE (see Results), the method of Hill (1974) could be used instead of that of Weir (1990b). χ^2_1 was calculated from *D* as suggested by Hill (1974). *D'*, which scales *D* between -1 and 1 (Hedrick *et al.*, 1978) was also calculated.

When multiple tests of HWE or population differentiation are carried out using contingency tables the significance level must be lowered. The nonsequential Bonferroni procedure (Weir, 1990b, pp. 109-110) was used, assuming an uncorrected critical value of 0.05, for tests of HWE, population differentiation and *D*. However, to detect any possible species-wide trends of weak linkage disequilibrium, all contingency tests with $P < 0.01$ (typical corrected *P* values were around 0.001) were carried to the second stage of computing *D*, but only those *P* values less than the corrected value are discussed.

Results

Prior to the large series of host plant rearings described in Payne & Berlocher (1995b), the impression gained from the literature on *R. mendax* (Lathrop & Nickels, 1932; Bush, 1966) was that the species infested a relatively large number of *Vaccinium* and *Gaylussacia* species across its range. This impression is erroneous: in fact, *R. mendax* is almost totally restricted to one host, *V. stamineum*, in most of its range, and utilizes more than one host species per site in only relatively small parts of its range. Thus, despite the fact that fruit collections from 10 multihost sites were made (Payne & Berlocher, 1995b), the present study, which was designed to utilize material from Payne & Berlocher (1995b), cannot answer questions about genetic differentiation on different host species very effectively; only the Rutgers Research Station and Beltsville Research Center sites are informative. Moreover, the low infesta-

tion rates outside of agricultural areas (average infestation rate is 0.7 per cent for the sites in Table 1) and multiple use of samples resulted in some small sample sizes.

Allele frequencies and average heterozygosities are in Appendix 1. Low frequency null alleles were observed at two dimeric loci, *Aat-2* and *Idh*, with the null alleles named by their relative mobility prefixed with an 'n'.

\bar{H} averaged across all 26 samples is 0.154, and average number of alleles/locus averaged across samples is 2.254. Although selection of loci for this paper was biased towards polymorphic loci, \bar{H} based on the 29 locus data set of Berlocher *et al.* (1993), in which loci were chosen simply on the basis of acceptable electrophoretic resolution, was actually slightly higher at 0.176 for the Chickaming, Michigan, site (included in both studies).

Adult and larval frequencies are not significantly different at the Chickaming, Michigan site. Tests of HWE indicate that *R. mendax* populations are panmictic. Out of 312 total tests, only two, both for *Acon-2* in *G. baccata* samples at the Beltsville site (6/6/93, $P=0.015$, and 16/7/93, $P=0.003$), approached or fell below the Bonferroni critical values of 0.0046 and 0.005, respectively. The deviations from HWE were in the form of heterozygote deficiencies ($f=0.213$ and 0.314, respectively), unlikely to result from inbreeding as the other loci were in HWE in these samples. A null allele could be responsible for the apparent heterozygote deficiency, as *Acon-2* has a monomeric structure and nulls would not be directly observed.

The Fort Valley, Georgia, collections did not provide evidence for population structuring arising from individual host plant effects. For the 11 tests, most probabilities were large, with the two smallest probabilities being for *Fum* at 0.02 and *Idh* at 0.008. However, for 11 loci the Bonferroni P is 0.005, so the null hypothesis for the sample as a whole is not rejected for this site.

Evidence for population structuring related to host plant species was observed at the Rutgers Research Station site for one locus, *Idh*. The P value was 0.002, below the Bonferroni value of 0.005. The P value for this locus remained significant when the geographically separate domestic blueberry sample was removed, leaving only the samples from the three intermixed wild hosts to test.

A larger number of significant tests was obtained at the Beltsville, Maryland site, although both sampling date and host species were involved. When all seven samples were tested the largest probability was 0.24 and for six of the 11 loci the probability of obtaining the results by chance alone was below the Bonferroni value of 0.005 (Table 2). However, it is not apparent

whether the differentiation arose from host species or date of collection; both poolings produced similar numbers of significant tests (Table 2).

No evidence for genetic similarity of the three samples from domestic blueberries was provided by the UPGMA analysis of Nei unbiased genetic identity. All three samples clustered distantly from one another (data not shown). No clustering of domestic blueberry samples could be produced using other distances, indicating that the choice of genetic distance was unrelated to the outcome.

Slight but significant large-scale geographical population structure was apparent in the 13 pooled samples. Seven of the test probabilities for individual loci were extremely low, so the differences are clearly not related solely to sampling error (Table 3). However, the F -statistics indicate that population differentiation is very slight in *R. mendax*. F_{ST} is only 0.015 ($F=0.040$, and

Table 2 Probabilities from contingency tests of allele frequency comparisons of seven *Rhagoletis mendax* collections at the Beltsville Agricultural Research Station

Locus	All seven samples	Pool by host (3 classes)	Pool by collection date (4 classes)
<i>Aat-2</i>	0.007	0.08	0.002
<i>Acon-2</i>	0.000	0.000	0.04
<i>Adh-1</i>	0.28	—	—
<i>Ak-2</i>	0.02	0.06	0.10
<i>Fum</i>	0.004	0.24	0.000
<i>Had</i>	0.000	0.001	0.004
<i>Idh</i>	0.000	0.03	0.000
<i>Mpi</i>	0.04	0.008	0.18
<i>Pgi</i>	0.02	0.57	0.02
<i>Pgm</i>	0.002	0.005	0.04
<i>Tpi</i>	0.001	0.01	0.00

See text for explanation of poolings. In the poolings *Adh-1* did not meet the requirements described in the text for testing.

Table 3 Contingency analysis of allele frequencies from 13 pooled samples (see text) representing the geographical range of *Rhagoletis mendax*

Locus	Probability	Locus	Probability
<i>Aat-2</i>	0.02	<i>Mpi</i>	0.0001
<i>Acon-2</i>	0	<i>Pgi</i>	0
<i>Adh-1</i>	0.0004	<i>Pgm</i>	0.09
<i>Ak-3</i>	0.09	<i>Tpi</i>	0.01
<i>Fum</i>	0	<i>Ald</i>	0.68
<i>Had</i>	0.0001	<i>Mdh-2</i>	0.91
<i>Idh</i>	0.0001		

$f = 0.025$). F_{ST} using Wright's (1978) method, with sampling variance subtracted, is very similar at 0.016. The fixed model contingency tests, and the random model F_{ST} calculation, differ with respect to statistical significance, as expected. Bootstrapping yields 95 per cent confidence limits for F_{ST} of 0.004 and 0.029, so that the null hypothesis of $F_{ST} = 0$ is barely rejected. No tests for clines in allele frequencies were significant, in either the coastal or total sample set, with or without the Nova Scotia frequencies of Berlocher & Bush (1982), which are very similar to those reported here.

Tests for autosomal linkage disequilibrium revealed only one pair of loci that displayed significant associations in both stages of the analysis. *Fum* and *Had* showed strong evidence of linkage disequilibrium in the Chickaming, Michigan, sample, in the Route 8 and Co. Y, Missouri, site and in the total pool, with very small P values (Table 4). To determine whether there was any geographical trend, all samples satisfying a slightly relaxed criterion for sample size ($N \geq 25$) were analysed for *Fum-Had* disequilibrium, and several additional poolings of samples were made (Table 4). Two of these additional poolings are of interest. When the Cortois sample is pooled with the nearby Route 8 and Co. Y sample, the value of D' remains at 1.000 whereas the χ^2 increases. Apparently, linkage disequilibrium is close to the maximum possible given the allele frequencies in this part of Missouri. The other additional analysis of interest, total pool minus the sites with large D values, was performed to determine

whether the disequilibrium seen in the total pool was generated solely by the high D sites Chickaming, Michigan, and Route 8 and Co. Y-Cortois, Missouri (pooled). The significance of the results indicates that significant *Fum-Had* disequilibrium occurs throughout the species, even if it cannot be statistically detected in most individual populations. Larger sample sizes would reveal whether disequilibria other than the *Fum*¹⁰⁰-*Had*¹⁰⁰, *Fum*¹⁵⁸-*Had*¹²⁵ association occurs.

Although disequilibrium can result from pooling of equilibrium populations with different allele frequencies (Nei & Li, 1973), this appears to be a minor contributor to the disequilibrium I observed in the pooled samples. As shown by the small F_{ST} value, allele frequency differentiation is slight in *R. mendax*. In addition, the total pooled data (with alleles pooled to two as for calculation of D) are in essentially perfect HWE for *Fum* ($\chi^2_1 = 0.24$) and only slightly out of equilibrium for *Had* ($\chi^2_1 = 4.21$). This slight degree of population structure is very unlikely to be totally responsible for the observed disequilibrium, leaving linkage disequilibrium in the strict sense as the most likely explanation.

Although interspecific hybridization in the *pomonella* group will be discussed in detail elsewhere, I note that no evidence for hybridization was observed in the 1009 flies electrophoresed, using the diagnostic loci for *pomonella* group species described in Berlocher *et al.* (1993).

Table 4 Linkage disequilibrium between *Fum* and *Had* in single samples of *Rhagoletis mendax* in which disequilibrium could potentially be detected (see text), and in various poolings of samples

Sample	N	P	D	D'	χ^2
Single samples					
Chickaming, MI	179	0	0.098	0.775	61.55
Beltsville, MD, VII/30 <i>V. stamineum</i>	59	0.068	0.017	0.389	3.56
Beltsville, MD, VIII/19 <i>V. stamineum</i>	29	0.331	0.014	0.175	0.39
Beltsville, MD, VII/16 <i>G. baccata</i>	50	0.528	0.021	0.341	1.69
Fort Valley, GA, bush 7	50	0.046	0.003	0.030	0.04
Rt. 8 and Co. Y, MO	31	0.001	0.081	1.000	17.27
Toccoa, GA	61	0.241	0.003	0.040	0.06
Pooled samples					
Rutgers Res. Sta., NJ, pool (2-5)	144	0.758	0.003	0.082	0.32
Beltsville, MD, pool (6-12)	219	0.240	-0.030	-0.225	3.81
Fort Valley, GA, pool (16-18)	108	0.001	0.024	0.299	7.83
Florida pool (21-22)	30	0.252	0.013	0.134	0.25
Rt. 8 and Co. Y-Cortois, MO, pool (23-24)	62	0	0.069	1.000	32.03
Total pool (1-26)	994	0	0.030	0.371	88.65
Pool of all except 1, 23, 24	752	0	0.013	0.198	17.99

Values in parentheses are site codes (Table 1). N is number of individual flies.

Discussion

The population structure of *R. mendax* is clearly different from that of *R. pomonella*. Differences concerning large-scale geographical differentiation are especially striking. *R. mendax* is almost uniform across its range, with little differentiation ($F_{ST}=0.015$) and no clinal structure whereas *R. pomonella* is quite structured, with very significant differentiation ($F_{ST}=0.220$; McPheron, 1987) and strong clinal structure (Feder & Bush, 1989). The genetic uniformity of *R. mendax* is from a lack of variation; based on the 29 locus data set of Berlocher *et al.* (1993), the expected heterozygosities \pm SE (0.221 ± 0.044 in *R. pomonella*, 0.176 ± 0.038 in *R. mendax*) and average number of alleles (2.7 ± 0.3 in *R. pomonella* and 2.3 ± 0.2 in *R. mendax*) do not differ significantly between the two species.

The small amount of geographical variation observed in *R. mendax* agrees with previous studies with more limited sampling. Berlocher & Bush (1982) analysed three *R. mendax* samples (Florida, Michigan, Nova Scotia) for 15 loci and reported a very small average Nei unbiased distance of 0.011 (F -statistics were not calculated). Feder *et al.* (1989a) studied three Michigan and one Nova Scotia populations for 16 variable loci (partially different from the set used here), and observed $F_{ST}=0.040 \pm 0.013$, which is probably not significantly greater than the value reported here (Feder & Bush (1989) report a variance of F_{ST} from bootstrapping instead of 95 per cent confidence limits).

With respect to linkage disequilibrium, differences between *R. pomonella* and *R. mendax* are also apparent. In both species significant disequilibrium occurs between only a small fraction of loci, in agreement with other studies (Langly, 1977). In *R. mendax* disequilibrium was observed only between *Fum* and *Had* whereas in *R. pomonella* significant pairwise disequilibria involving seven loci organized into three linkage groups (linkage group I, *Dia-2/Aat-2*; linkage group II, *Me/Acon-2/Mpi*; linkage group III, *Had/Pep-2*; Feder *et al.*, 1990a) were seen in both hawthorn and apple races. However, meaningful comparisons between the species are difficult to make, as only one locus pair (*Acon-2/Mpi*) could be tested in both species (*Fum* is monomorphic in *R. pomonella*, *Dia-2* is effectively monomorphic in *R. mendax* and *Me* and *Pep-2* were not studied here). For *Acon-2* and *Mpi*, at which all common alleles are shared between the two species, no disequilibrium was seen in this study, and only sporadic disequilibrium in *R. pomonella* (Feder *et al.*, 1990a). I note that the *R. mendax* disequilibrium data imply that *Fum* and *Had* are linked, as does the fact that these two loci are linked in *R. tabellaria* (McPheron & Berlocher, 1985).

Differences between *R. mendax* and *R. pomonella* involving differentiation of populations on different individual host plants and different host plant species cannot be completely characterized in this study; the unexpected low infestation rate and exclusive use of only one host species in much of the range of *R. mendax* compromised the original sampling design. The available data offer no evidence for differentiation of *R. mendax* populations with respect to individual host plants. Some evidence was found for differentiation associated with both host plant species and phenology, although it is not clear which has the greatest effect. Moreover, the pattern of involvement of different loci in differentiation in *R. mendax* is very inconsistent. In *R. pomonella*, the same set of six loci (*Me*, *Acon-2*, *Mpi*, *Dia-2*, *Aat-2*, *Had*), which are associated into three linkage groups, are consistently differentiated between apple and hawthorn races (McPheron *et al.*, 1988b; Feder *et al.*, 1990a) whereas seven other polymorphic loci are differentiated only sporadically. Five of these 'host-plant' loci, four with sufficient variation to test (*Acon-2*, *Mpi*, *Aat-2* and *Had*), were studied in *R. mendax*, but these loci are not consistently differentiated at the two sites where I sampled from multiple host species. At the Rutgers site the only locus showing significant differentiation is *Idh*, which has been intensively studied in *R. pomonella* and is rarely differentiated along host-plant lines. At the Beltsville site two of the 'host-plant' loci of *R. pomonella* (*Acon-2*, and *Had*) are differentiated among the seven host-plant/collection date samples, but *Idh* and *Pgm*, characteristic 'non-host-plant' loci in *R. pomonella*, are also differentiated. Overall, the inconsistent involvement of loci in population differentiation with respect to host plant species does not support the existence of host races or other host-related substructuring in *R. mendax*, although further study is clearly required. One possibility that needs investigation is that population densities are so low in *R. mendax* that only a few sibling groups are being sampled, although the excellent fit to HWE in almost all samples/loci argues against this explanation.

An explanation for the overall geographical genetic uniformity of *R. mendax* may lie in its host plant ecology. The restriction of the species to a single host, *V. stamineum*, across the great majority of its range may make for a very uniform selective environment. If allozymes are under selection, as now seems likely for at least some allozymes (for example Chambers, 1988), then the host-specificity of *R. mendax* may be the primary force promoting genetic uniformity. Another very host-specific species (the yet-to-be-named 'sparkleberry fly') in the *pomonella* group also displays very little geographical variation (Payne & Berlocher, 1995a). On the other hand, other aspects of the popu-

lation biology of *R. mendax* would seemingly promote geographical differentiation caused by drift. The insect is totally absent from the many regions of eastern North America that lack the acid soils needed by its *Vaccinium* host plants, and even in areas with suitable hosts populations may be scattered and have low population densities (Payne & Berlocher, 1995b). Study of neutral or near-neutral molecular markers in *R. mendax* may reveal a different pattern of differentiation than that seen at allozyme loci.

Acknowledgements

This project would never have been possible without the assistance of S. Lyons, J. A. Payne, W. S. Sheppard, C. Anelli, J. Ballington, J. L. Feder, B. A. McPheron, T. Anderson, M. G. Anderson, P. Lyrene and G. L. Bush (Cmdr., Ret.). NSF 89-06217 provided funding.

References

- BERLOCHER, S. H. 1993. Gametic disequilibrium between allozyme loci and sex chromosomes in the genus *Rhagoletis*. *J. Hered.*, **84**, 431–437.
- 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 ENQUIST, M. 1993. Distribution and host plants of the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae) in Texas. *J. Kans. Entomol. Soc.*, **66**, 145–166.
- BERLOCHER, S. H. AND SMITH, D. C. 1983. Segregation and mapping of allozymes of the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae). *J. Hered.*, **74**, 337–340.
- BERLOCHER, S. H., MCPHERON, B. A., FEDER, J. L. AND BUSH, G. L. 1993. Genetic differentiation at allozyme loci in the *Rhagoletis pomonella* (Diptera: Tephritidae) species complex. *Ann. Entomol. Soc. Am.*, **86**, 716–727.
- BOLLER, E. F. AND PROKOPY, R. J. 1975. Bionomics and management of *Rhagoletis*. *Ann. Rev. Entomol.*, **21**, 223–246.
- BUSH, G. L. 1966. The taxonomy, cytology, and evolution of the genus *Rhagoletis* in North America (Diptera: Tephritidae). *Bull. Mus. Comp. Zool.*, **134**, 431–562.
- BUSH, G. L. 1969. Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera: Tephritidae). *Evolution*, **23**, 237–251.
- BUSH, G. L. 1974. The mechanisms of sympatric host race formation in the true fruit flies (Tephritidae). In: White, M. J. D. (ed.) *Genetic Mechanisms of Speciation in Insects*, pp. 2–23. Australia and New Zealand Book Co., Sydney.
- BUSH, G. L. 1975. Sympatric speciation in phytophagous parasitic insects. In: Price, P. W. (ed.) *Evolutionary Strategies of Parasitic Insects and Mites*, pp. 187–206. Plenum Press, New York.
- BUSH, G. L. 1992. Host race formation and sympatric speciation in *Rhagoletis* fruit flies (Diptera: Tephritidae). *Psyche*, **99**, 335–357.
- CHAMBERS, G. K. 1988. The *Drosophila* alcohol dehydrogenase gene–enzyme system. *Adv. Genet.*, **25**, 481–488.
- FEDER, J. L. AND BUSH, G. L. 1989. Gene frequency clines in host races of *Rhagoletis pomonella* (Diptera: Tephritidae) in the midwestern United States. *Heredity*, **63**, 245–266.
- FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1988. Genetic differentiation between sympatric host races of *Rhagoletis pomonella*. *Nature*, **336**, 61–64.
- FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1989a. Are the apple maggot (*Rhagoletis pomonella*) and the blueberry maggot (*R. mendax*) distinct species? Implications for sympatric speciation. *Entomologia exp appl.*, **51**, 113–123.
- FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1989b. 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. 1990a. The geographic pattern of genetic differentiation between host-associated populations of *Rhagoletis pomonella* (Diptera: Tephritidae) in the eastern United States. *Evolution*, **44**, 570–594.
- FEDER, J. L., CHILCOTE, C. A. AND BUSH, G. L. 1990b. Regional, local, and microgeographical allele frequency variation between apple and hawthorn populations of *Rhagoletis pomonella* in western Michigan. *Evolution*, **44**, 595–608.
- FEDER, J. L., HUNT, T. A., AND BUSH, G. L. 1993. The effect of climate, host plant phenology, and host fidelity on the genetics of apple and hawthorn-infesting races of *Rhagoletis pomonella*. *Entomologia exp. appl.*, **69**, 117–135.
- FUTUYMA, D. J. AND MAYER, G. C. 1980. Non-allopatric speciation in animals. *Syst. Zool.*, **29**, 254–271.
- HEDRICK, P., JAIN, S. AND HOLDEN, L. 1978. Multilocus systems in evolution. *Evol. Biol.*, **11**, 104–154.
- HILL W. G. 1974. Estimation of linkage disequilibrium in randomly mating populations. *Heredity*, **33**, 229–239.
- LANGLEY, C. 1977. Nonrandom associations between allozymes in natural populations of *Drosophila melanogaster*. In: Christiansen, F. B. and Fenchel, T. M. (eds) *Measuring Selection in Natural Populations*, pp. 265–273. Springer-Verlag, Berlin.
- LATHROP, F. H. AND NICKELS, C. B. 1932. The biology and control of the blueberry maggot in Washington County, Maine. *U. S. Dept. Agr. Tech. Bull.*, no. 275.
- MAYR, E. 1963. *Animal Species and Evolution*. Belknap Press, Cambridge, MA.
- 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. Thesis, University of Illinois, Urbana.
- MCPHERON, B. A. AND BERLOCHER, S. H. 1985. Segregation and linkage of allozymes of *Rhagoletis tabellaria*. *J. Hered.*, **76**, 218–219.
- MCPHERON, B. A., SMITH, D. C. AND BERLOCHER, S. H. 1988a. Microgeographic genetic variation in the apple maggot, *Rhagoletis pomonella*. *Genetics*, **119**, 445–451.
- MCPHERON, B. A., SMITH, D. C. AND BERLOCHER, S. H. 1988b. Genetic differences between host races of the apple maggot fly. *Nature*, **336**, 64–66.

- NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- NEI, M. AND LI, A. 1973. Linkage disequilibrium in subdivided populations. *Genetics*, **75**, 213–219.
- PAYNE, J. A. AND BERLOCHER, S. H. 1995a. Phenological and electrophoretic evidence for a new blueberry-infesting species in the *Rhagoletis pomonella* (Diptera:Tephritidae) sibling species complex. *Entomologia exp. appl.*, (in press).
- PAYNE, J. A. AND BERLOCHER, S. H. 1995b. Distribution and host plants of the blueberry maggot, *Rhagoletis mendax* (Diptera: Tephritidae) in southeastern North America. *J. Kans. Entomol. Soc.*, (in press).
- PROKOPY, R. J., DIEHL, S. R. AND COOLEY, S. S. 1987. Behavioral evidence for host races in *Rhagoletis pomonella* flies. *Oecologia*, **76**, 138–147.
- SMITH, D. C. 1988. Heritable divergence of *Rhagoletis pomonella* host races by seasonal asynchrony. *Nature*, **336**, 66–68.
- SWOFFORD, D. L. AND SELANDER, R. B. 1981. BIOSYS-1: A Fortran program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *J. Hered.*, **72**, 281–283.
- VANDER KLOET, S. P. 1988. *The genus Vaccinium in North America*. Research Branch, Agriculture Canada, Publication 1828.
- VINCENT, C. AND LAREAU, M. J. 1989. Update on the distribution of the blueberry maggot, *Rhagoletis mendax* (Diptera: Tephritidae), in Canada. *Acta Hort.*, **241**, 333–337.
- WIER, B. S. 1990a. Intraspecific variation. In: Hillis, D. M. and Moritz, C. (eds) *Molecular Systematics*, pp. 398–410. Sinauer Associates, Sunderland, MA.
- WIER, B. S. 1990b. *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.
- WRIGHT, S. 1978. *Evolution and the Genetics of Populations*, vol. 4, *Variability in and among Natural Populations*. University of Chicago Press, Chicago.

Appendix 1 Allele frequencies in *Rhagoletis mendax*

Locus	Population code												
	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>Aat-2</i>													
(N)	197	20	52	35	37	26	14	50	12	58	25	30	39
-16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
01	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
21	0.244	0.125	0.337	0.271	0.230	0.173	0.429	0.340	0.208	0.198	0.200	0.217	0.269
n21	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.000
25	0.000	0.000	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000	0.000	0.033	0.000
50	0.670	0.875	0.596	0.657	0.676	0.712	0.429	0.640	0.750	0.759	0.720	0.633	0.705
n50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.000	0.017	0.000
59	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
75	0.069	0.000	0.067	0.071	0.095	0.038	0.071	0.000	0.042	0.026	0.080	0.067	0.026
100	0.018	0.000	0.000	0.000	0.000	0.077	0.036	0.020	0.000	0.009	0.000	0.017	0.000
<i>Acon-2</i>													
(N)	180	19	52	35	37	24	19	49	12	59	25	30	39
73	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
75	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
89	0.125	0.000	0.106	0.057	0.189	0.188	0.184	0.102	0.042	0.220	0.040	0.150	0.026
95	0.164	0.132	0.058	0.071	0.041	0.125	0.053	0.163	0.000	0.144	0.120	0.183	0.154
100	0.556	0.789	0.817	0.829	0.743	0.688	0.711	0.673	0.750	0.517	0.820	0.650	0.692
106	0.147	0.079	0.019	0.029	0.027	0.000	0.053	0.061	0.125	0.119	0.020	0.017	0.128
114	0.003	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.083	0.000	0.000	0.000	0.000
<i>Adh-1</i>													
(N)	194	19	52	35	36	25	19	49	9	59	25	29	38
35	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.020	0.000	0.000
-38	0.005	0.026	0.029	0.029	0.000	0.020	0.026	0.020	0.000	0.000	0.000	0.000	0.013
-46	0.000	0.000	0.038	0.100	0.042	0.000	0.026	0.010	0.000	0.000	0.040	0.000	0.053
-100	0.995	0.974	0.933	0.857	0.944	0.980	0.947	0.969	1.000	1.000	0.940	1.000	0.921
-140	0.000	0.000	0.000	0.014	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013
<i>Ak-3</i>													
(N)	184	20	52	35	37	25	19	50	12	59	25	30	39
68	0.000	0.000	0.000	0.000	0.000	0.000	0.026	0.000	0.000	0.025	0.000	0.000	0.000
83	0.019	0.150	0.048	0.014	0.027	0.000	0.000	0.010	0.000	0.034	0.020	0.067	0.000

Appendix 1 *Continued*

Locus	Population code												
	1	2	3	4	5	6	7	8	9	10	11	12	13
87	0.000	0.000	0.000	0.000	0.000	0.000	0.026	0.010	0.000	0.008	0.000	0.033	0.000
100	0.978	0.850	0.952	0.986	0.973	1.000	0.947	0.980	1.000	0.932	0.980	0.900	1.000
111	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
126	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Ald</i>													
(N)	182	20	52	35	37	25	19	50	12	59	25	30	39
100	0.986	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.987
162	0.014	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013
217	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Dia-1</i>													
(N)	181	20	52	35	36	26	19	50	12	59	25	30	39
100	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
150	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Dia-2</i>													
(N)	173	20	52	35	37	26	19	50	12	58	25	29	39
70	0.000	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
85	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Fum</i>													
(N)	197	20	52	35	37	25	19	50	12	59	25	29	39
-50	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.038
100	0.264	0.275	0.115	0.086	0.095	0.320	0.105	0.080	0.042	0.119	0.100	0.207	0.026
158	0.734	0.725	0.885	0.914	0.905	0.680	0.895	0.920	0.958	0.881	0.900	0.793	0.936
<i>Had</i>													
(N)	183	20	52	35	37	26	19	50	12	59	25	30	39
72	0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	0.156	0.075	0.029	0.043	0.054	0.038	0.053	0.220	0.042	0.051	0.060	0.100	0.064
125	0.817	0.925	0.971	0.957	0.946	0.942	0.947	0.770	0.958	0.949	0.940	0.900	0.936
150	0.005	0.000	0.000	0.000	0.000	0.019	0.000	0.010	0.000	0.000	0.000	0.000	0.000
<i>Idh</i>													
(N)	185	20	52	35	37	26	19	50	12	59	25	29	39
82	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
93	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.042	0.000	0.000	0.000
100	0.554	0.625	0.635	0.486	0.392	0.692	0.579	0.480	0.625	0.551	0.560	0.517	0.654
113	0.000	0.000	0.000	0.029	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
118	0.414	0.325	0.356	0.486	0.554	0.192	0.395	0.500	0.333	0.339	0.340	0.483	0.346
n120	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.000	0.000	0.000	0.000	0.000
128	0.016	0.050	0.010	0.000	0.054	0.115	0.026	0.010	0.042	0.068	0.100	0.000	0.000
<i>Mdh-1</i>													
(N)	183	20	52	35	37	26	19	50	12	59	25	30	39
-73	0.000	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-100	1.000	1.000	1.000	1.000	0.986	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.974
-108	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.026
<i>Mdh-2</i>													
(N)	184	20	52	35	37	26	19	50	12	59	25	30	39
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	0.008	0.000	0.010	0.000	0.000	0.000	0.026	0.000	0.000	0.000	0.000	0.000	0.000
67	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	0.989	1.000	0.971	1.000	1.000	1.000	0.947	0.990	1.000	0.983	1.000	1.000	0.987
147	0.003	0.000	0.019	0.000	0.000	0.000	0.026	0.010	0.000	0.017	0.000	0.000	0.013
<i>Mpi</i>													
(N)	196	20	52	35	37	26	17	49	12	58	25	30	39
63	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
70	0.046	0.000	0.019	0.029	0.027	0.173	0.088	0.061	0.000	0.095	0.000	0.050	0.154

Appendix 1 *Continued*

Locus	Population code												
	1	2	3	4	5	6	7	8	9	10	11	12	13
85	0.046	0.025	0.019	0.014	0.027	0.000	0.029	0.020	0.042	0.026	0.040	0.017	0.038
100	0.880	0.975	0.962	0.957	0.932	0.827	0.882	0.878	0.958	0.862	0.960	0.933	0.782
105	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.026
109	0.026	0.000	0.000	0.000	0.014	0.000	0.000	0.041	0.000	0.017	0.000	0.000	0.000
112	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Pgi</i>													
(<i>N</i>)	199	20	52	35	37	25	19	50	12	59	25	30	39
80	0.005	0.000	0.010	0.014	0.014	0.000	0.000	0.040	0.000	0.042	0.000	0.000	0.026
100	0.839	0.900	0.827	0.700	0.797	0.900	0.974	0.890	0.833	0.932	0.960	0.850	0.949
130	0.148	0.100	0.154	0.286	0.189	0.100	0.026	0.040	0.167	0.025	0.040	0.150	0.026
145	0.008	0.000	0.010	0.000	0.000	0.000	0.000	0.010	0.000	0.000	0.000	0.000	0.000
<i>Pgm</i>													
(<i>N</i>)	185	20	52	35	37	25	19	50	12	59	25	30	39
75	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
80	0.014	0.000	0.010	0.014	0.027	0.000	0.000	0.020	0.000	0.017	0.000	0.033	0.051
92	0.092	0.025	0.087	0.100	0.068	0.060	0.053	0.070	0.125	0.169	0.000	0.183	0.013
95	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	0.832	0.975	0.885	0.814	0.784	0.940	0.921	0.860	0.792	0.746	0.880	0.750	0.859
111	0.046	0.000	0.010	0.000	0.081	0.000	0.026	0.040	0.000	0.017	0.020	0.033	0.038
118	0.014	0.000	0.010	0.043	0.027	0.000	0.000	0.010	0.042	0.051	0.060	0.000	0.026
125	0.003	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.042	0.000	0.040	0.000	0.013
<i>Tpi</i>													
(<i>N</i>)	152	6	51	35	23	14	19	50	12	55	25	28	29
85	0.000	0.000	0.010	0.000	0.000	0.000	0.026	0.000	0.000	0.000	0.000	0.000	0.000
90	0.000	0.000	0.000	0.000	0.000	0.179	0.000	0.020	0.000	0.000	0.000	0.018	0.000
100	0.993	0.917	0.980	1.000	0.957	0.821	0.974	0.980	1.000	0.991	0.980	0.982	0.966
112	0.000	0.083	0.010	0.000	0.043	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
120	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.020	0.000	0.034
\bar{H}	0.191	0.142	0.149	0.160	0.167	0.172	0.160	0.169	0.133	0.171	0.130	0.184	0.157

Locus	Population code												
	14	15	16	17	18	19	20	21	22	23	24	25	26
<i>Aat-2</i>													
(<i>N</i>)	30	36	20	30	55	61	31	23	7	29	30	20	30
-16	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000
01	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000	0.000	0.000
21	0.183	0.278	0.250	0.333	0.300	0.246	0.274	0.261	0.286	0.155	0.150	0.300	0.217
n21	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.017
25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
50	0.717	0.625	0.675	0.617	0.655	0.672	0.661	0.717	0.643	0.741	0.850	0.650	0.750
n50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.000	0.000
59	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
75	0.067	0.056	0.075	0.033	0.036	0.049	0.065	0.000	0.071	0.086	0.000	0.050	0.017
100	0.033	0.042	0.000	0.017	0.009	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Acon-2</i>													
(<i>N</i>)	29	37	20	29	54	61	31	22	7	31	31	16	30
73	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.000	0.032	0.000	0.000	0.017
75	0.000	0.027	0.000	0.000	0.009	0.000	0.016	0.000	0.000	0.048	0.000	0.000	0.017
89	0.259	0.081	0.025	0.000	0.046	0.008	0.000	0.023	0.000	0.081	0.081	0.000	0.017
95	0.069	0.135	0.125	0.121	0.185	0.287	0.081	0.091	0.071	0.097	0.145	0.156	0.217

Appendix 1 *Continued*

Locus	Population code												
	14	15	16	17	18	19	20	21	22	23	24	25	26
100	0.569	0.703	0.700	0.638	0.648	0.541	0.839	0.864	0.857	0.500	0.710	0.781	0.533
106	0.103	0.054	0.150	0.241	0.111	0.164	0.065	0.000	0.071	0.242	0.065	0.063	0.200
114	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Adh-1</i>													
(N)	30	30	20	30	44	59	31	22	7	31	31	20	30
35	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.016	0.000	0.000	0.000
-38	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-46	0.033	0.017	0.000	0.000	0.023	0.008	0.032	0.045	0.000	0.000	0.016	0.050	0.033
-100	0.950	0.983	1.000	1.000	0.955	0.983	0.968	0.955	1.000	0.968	0.984	0.950	0.967
-140	0.017	0.000	0.000	0.000	0.023	0.000	0.000	0.000	0.000	0.016	0.000	0.000	0.000
<i>Ak-3</i>													
(N)	30	37	20	30	55	61	31	22	7	31	31	20	30
68	0.017	0.000	0.000	0.017	0.009	0.008	0.032	0.000	0.000	0.000	0.000	0.000	0.000
83	0.033	0.027	0.075	0.000	0.027	0.025	0.032	0.000	0.000	0.016	0.032	0.025	0.033
87	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	0.950	0.973	0.925	0.983	0.964	0.951	0.935	1.000	1.000	0.984	0.935	0.975	0.967
111	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
126	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.000	0.000
<i>Ald</i>													
(N)	30	36	20	30	55	61	31	23	6	31	31	20	30
100	1.000	0.986	1.000	1.000	0.982	0.975	0.984	1.000	1.000	1.000	1.000	1.000	1.000
162	0.000	0.000	0.000	0.000	0.018	0.025	0.016	0.000	0.000	0.000	0.000	0.000	0.000
217	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Dia-1</i>													
(N)	30	30	20	30	55	61	31	22	7	31	31	20	30
100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
150	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Dia-2</i>													
(N)	30	25	20	30	46	61	31	22	7	31	31	20	30
70	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
85	1.000	1.000	1.000	1.000	1.000	1.000	0.968	1.000	1.000	1.000	1.000	1.000	1.000
100	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.000	0.000	0.000	0.000	0.000	0.000
<i>Fum</i>													
(N)	30	37	20	30	50	61	31	23	7	31	31	20	30
-50	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000	0.000	0.000	0.000	0.000
100	0.117	0.068	0.000	0.083	0.120	0.123	0.065	0.304	0.071	0.097	0.065	0.150	0.083
158	0.883	0.932	1.000	0.917	0.880	0.877	0.919	0.696	0.929	0.903	0.935	0.850	0.917
<i>Had</i>													
(N)	30	37	20	30	55	61	31	22	7	31	31	20	30
72	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	0.050	0.041	0.075	0.067	0.136	0.082	0.000	0.068	0.286	0.161	0.129	0.075	0.067
125	0.950	0.959	0.925	0.933	0.855	0.918	0.984	0.932	0.714	0.839	0.871	0.925	0.933
150	0.000	0.000	0.000	0.000	0.009	0.000	0.016	0.000	0.000	0.000	0.000	0.000	0.000
<i>Idh</i>													
(N)	30	37	20	30	55	62	31	22	7	31	31	20	30
82	0.000	0.014	0.000	0.000	0.000	0.008	0.016	0.000	0.071	0.000	0.000	0.050	0.000
93	0.000	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.033
100	0.467	0.662	0.350	0.600	0.609	0.532	0.645	0.568	0.429	0.694	0.500	0.375	0.633
113	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
118	0.533	0.324	0.650	0.383	0.373	0.379	0.339	0.409	0.500	0.306	0.500	0.550	0.317
n120	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
128	0.000	0.000	0.000	0.000	0.018	0.081	0.000	0.023	0.000	0.000	0.000	0.025	0.017

Appendix 1 *Continued*

Locus	Population code												
	14	15	16	17	18	19	20	21	22	23	24	25	26
<i>Mdh-1</i>													
(N)	30	37	20	30	55	61	31	22	7	31	31	20	30
-73	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-108	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Mdh-2</i>													
(N)	30	37	20	30	55	61	24	22	7	31	31	20	30
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.000	0.000	0.000	0.000	0.000
20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000
67	0.000	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	1.000	1.000	1.000	1.000	0.982	1.000	1.000	0.977	1.000	1.000	0.984	1.000	1.000
147	0.000	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Mpi</i>													
(N)	30	37	20	30	55	61	29	22	7	31	31	20	30
63	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
70	0.150	0.054	0.150	0.083	0.082	0.082	0.052	0.068	0.071	0.194	0.032	0.025	0.017
85	0.000	0.054	0.075	0.000	0.045	0.025	0.034	0.091	0.000	0.016	0.016	0.050	0.017
100	0.850	0.865	0.775	0.917	0.873	0.885	0.914	0.841	0.929	0.742	0.952	0.925	0.967
105	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
109	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.048	0.000	0.000	0.000
112	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Pgi</i>													
(N)	30	37	20	30	55	61	31	22	7	31	31	20	30
80	0.000	0.000	0.000	0.000	0.018	0.025	0.000	0.045	0.000	0.048	0.016	0.000	0.033
100	0.967	1.000	1.000	0.950	0.900	0.934	0.968	0.909	1.000	0.935	0.984	1.000	0.917
130	0.033	0.000	0.000	0.050	0.082	0.041	0.032	0.045	0.000	0.016	0.000	0.000	0.050
145	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Pgm</i>													
(N)	30	37	20	30	55	61	31	22	7	31	31	20	30
75	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
80	0.067	0.000	0.025	0.033	0.045	0.033	0.016	0.000	0.000	0.000	0.048	0.000	0.000
92	0.033	0.054	0.150	0.033	0.055	0.098	0.065	0.091	0.071	0.065	0.081	0.050	0.133
95	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000
100	0.867	0.865	0.775	0.917	0.891	0.803	0.903	0.818	0.929	0.919	0.823	0.925	0.783
111	0.033	0.054	0.025	0.017	0.009	0.049	0.016	0.023	0.000	0.016	0.032	0.025	0.050
118	0.000	0.014	0.000	0.000	0.000	0.008	0.000	0.068	0.000	0.000	0.000	0.000	0.033
125	0.000	0.000	0.025	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Tpi</i>													
(N)	26	30	20	30	55	45	17	15	7	19	22	7	24
85	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.068	0.000	0.000
90	0.000	0.000	0.000	0.000	0.009	0.011	0.000	0.000	0.000	0.000	0.045	0.000	0.000
100	1.000	1.000	0.975	1.000	0.973	0.978	1.000	1.000	1.000	1.000	0.886	1.000	1.000
112	0.000	0.000	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
120	0.000	0.000	0.000	0.000	0.018	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000
\bar{H}	0.155	0.135	0.149	0.134	0.170	0.173	0.126	0.158	0.134	0.164	0.143	0.134	0.150