

Evidence for Inversion Polymorphism Related to Sympatric Host Race Formation in the Apple Maggot Fly, *Rhagoletis pomonella*

Jeffrey L. Feder,^{*1} Joseph B. Roethele,^{*} Kenneth Filchak,^{*} Julie Niedbalski^{*}
and Jeanne Romero-Severson^{*†}

^{*}Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556-0369 and [†]Department of Forestry and Natural Resources, Purdue University, West Lafayette, Indiana 47907-1159

Manuscript received August 24, 2001
Accepted for publication December 3, 2002

ABSTRACT

Evidence suggests that the apple maggot, *Rhagoletis pomonella* (Diptera: Tephritidae) is undergoing sympatric speciation (*i.e.*, divergence without geographic isolation) in the process of shifting and adapting to a new host plant. Prior to the introduction of cultivated apples (*Malus pumila*) in North America, *R. pomonella* infested the fruit of native hawthorns (*Crataegus* spp.). However, sometime in the mid-1800s the fly formed a sympatric race on apple. The recently derived apple-infesting race shows consistent allele frequency differences from the hawthorn host race for six allozyme loci mapping to three different chromosomes. Alleles at all six of these allozymes correlate with the timing of adult eclosion, an event dependent on the duration of the overwintering pupal diapause. This timing difference differentially adapts the univoltine fly races to an ~3- to 4-week difference in the peak fruiting times of apple and hawthorn trees, partially reproductively isolating the host races. Here, we report finding substantial gametic disequilibrium among allozyme and complementary DNA (cDNA) markers encompassing the three chromosomal regions differentiating apple and hawthorn flies. The regions of disequilibrium extend well beyond the previously characterized six allozyme loci, covering substantial portions of chromosomes 1, 2, and 3 (haploid $n = 6$ in *R. pomonella*). Moreover, significant recombination heterogeneity and variation in gene order were observed among single-pair crosses for each of the three genomic regions, implying the existence of inversion polymorphism. We therefore have evidence that genes affecting diapause traits involved in host race formation reside within large complexes of rearranged genes. We explore whether these genomic regions (inversions) constitute coadapted gene complexes and discuss the implications of our findings for sympatric speciation in *Rhagoletis*.

NEO-DARWINIAN theory posits that the genetic basis for speciation is not qualitatively different from that underlying microevolutionary change within populations. The genetic variation between and within populations is the ultimate basis for the origin of species (LEWONTIN 1997). The use of the genetics of natural populations to inform evolutionary theory traces to the collaboration and pioneering work of T. H. Dobzhansky and A. H. Sturtevant on inversion polymorphism in *Drosophila pseudoobscura* (for reviews see PROVINE 1981; KRIMBAS and POWELL 1992; POWELL 1997). At first, Dobzhansky viewed inversions as selectively neutral markers that could be used to demonstrate genetic drift in subdivided populations (DOBZHANSKY and QUEAL 1938). However, it soon became apparent that the inversions were under strong selection, which led Dobzhansky to develop his ideas concerning genetic coadaptation—*i.e.*, that nonadditive, epistatic fitness effects among loci create harmoniously interacting complexes of genes within inversions that are selectively balanced against alternate

sets in heterokaryotypes (DOBZHANSKY 1970; KRIMBAS and POWELL 1992).

Subsequent studies of genetic variation have generally focused on allele or genotype frequency distributions and single-locus approaches to understanding genetic polymorphism (although see HEDRICK *et al.* 1978; LONG *et al.* 1995; ELENA and LENSKI 1997; TEMPLETON 2000; ANDOLFATTO and PREZEWORSKI 2001; KIM and RIESEBERG 2001). However, advances in molecular techniques now make it possible to comprehensively survey the genome of almost any organism for genetic variation. The power of these methods has permitted more sophisticated tests for natural selection (HUDSON *et al.* 1987; TAJIMA 1989; McDONALD and KREITMAN 1991; FU and LI 1993) and has allowed the revisiting of questions concerning the adaptive significance of inversion polymorphism (AQUADRO *et al.* 1991; BENASSI *et al.* 1993; ROZAS and AGUADÉ 1993, 1994; WESLEY and EANES 1994; BABCOCK and ANDERSON 1996; ANDOLFATTO *et al.* 1999; CACERES *et al.* 1999; DEPAULIS *et al.* 1999; ROZAS *et al.* 1999; ANDOLFATTO and KREITMAN 2000; KOVACEVIC and SCHAEFFER 2000).

In the current study, we investigate the genetic architecture of adaptation and speciation for the apple mag-

¹Corresponding author: Department of Biological Sciences, P.O. Box 369, University of Notre Dame, Notre Dame, Indiana 46556-0369.
E-mail: feder.2@nd.edu

got fly, *Rhagoletis pomonella* (Diptera: Tephritidae). True fruit flies belonging to the *R. pomonella* sibling species complex, of which the apple maggot is a member, are at the center of a long-standing debate concerning modes of speciation (BUSH 1969a,b, 1994). Speciation in sexually reproducing animals is generally thought to require the complete geographic separation (*i.e.*, allopatry) of populations at some point in their history (MAYR 1963). But as early as the 1860s, WALSH (1864) argued that certain phytophagous insects could speciate in the absence of geographic isolation (*i.e.*, in sympatry) in the process of shifting and adapting to new host plants. *R. pomonella* shifted from utilizing the unabsconded fruit of its native host hawthorn (*Crataegus* spp.) to utilizing the introduced, domesticated apple (*Malus pumila*) sometime in the mid-1800s in the Hudson River Valley region of New York (BUSH 1966, 1969a,b). WALSH (1867) cited this shift as an example of an incipient sympatric speciation event. BUSH (1966, 1969a,b) subsequently argued that the entire complex of five (or more) sibling species comprising the *R. pomonella* group arose by sympatric speciation via host plant shifts.

Recent genetic studies have confirmed that apple- and hawthorn-infesting populations of *R. pomonella* are partially reproductively isolated "host races," the hypothesized first stage of sympatric speciation (FEDER *et al.* 1988, 1990; McPHERON *et al.* 1988). Apple and hawthorn fly races display significant allele frequency differences for six allozyme loci at sympatric field sites across the eastern United States. [The six allozymes are: *Malic Enzyme* (*Me*, 1.1.1.40), *Aconitase-2* (*Acon-2*, EC 4.2.1.3), *Mannose phosphate isomerase* (*Mpi*, 5.3.1.8), *Aspartate amino transferase-2* (*Aat-2*, EC 2.6.1.1), *NADH-Diaphorase-2* (*Dia-2*, EC 1.6.2.2), and β -*Hydroxyacid dehydrogenase* (*Had*, EC 1.1.1.30).]

Allozyme surveys have also supported Bush's contention concerning the sympatric radiation of the *R. pomonella* group (BERLOCHER *et al.* 1993; BERLOCHER 2000). Interestingly, *R. cornivora* is the sole species in the complex distinguished by fixed autapomorphic differences (BERLOCHER *et al.* 1993; BERLOCHER 2000). Only allozyme frequency differences have thus far been found to separate *R. pomonella*, *R. mendax* (blueberry maggot), *R. zephyria* (snowberry maggot), and the undescribed *Cornus florida* fly (flowering dogwood maggot; BERLOCHER *et al.* 1993; FEDER 1998; BERLOCHER 2000). These four "in group" sibling species can therefore be viewed as host races writ large, varying in degree (*i.e.*, more pronounced allele frequency differences), but not in kind, from apple and hawthorn flies (*i.e.*, "quantitative genetic" species; BERLOCHER and FEDER 2002). Despite not possessing fixed genetic differences or constituting monophyletic lineages, these species nonetheless occupy distinct adaptive host peaks. Ecological specialization is sufficient for these species to have formed and to be maintained in sympatry as recognizable genotypic clusters (MALLETT 1995) or as "evolutionary groups" (HEY 2001) in the face of potential gene flow (for fur-

ther discussion of the species/gene flow problem see FEDER 1998; WU 2001; and related commentary). An evolutionary chronology of diverged forms ranging from recently derived host races to completely isolated sibling species therefore exists in the *R. pomonella* group, consistent with the Darwinian view of speciation as a microevolutionary (adaptive) process.

Two host-associated traits have been shown to be primarily responsible for isolating *R. pomonella* group flies. First, because *Rhagoletis* adults court and mate exclusively on or near the fruit of their host plants (PROKOPY *et al.* 1971, 1972), differences in host preference behaviors translate directly into mate choice decisions and premating isolation (PROKOPY *et al.* 1988; FEDER *et al.* 1994). Mark-release recapture experiments have indicated that host-specific mating partially isolates the apple and hawthorn races of *R. pomonella*, reducing gene flow to \sim 4–6%/generation (FEDER *et al.* 1994, 1998). In comparison, host-specific mating appears to constitute a near-complete premating barrier between the sibling species *R. pomonella* and *R. mendax* (FEDER and BUSH 1989a).

Second, diapause-related traits differentially adapt *Rhagoletis* flies to variation in the fruiting times of their respective host plants. *Rhagoletis* larvae feed within the fruit of their host plants, with each taxon attacking a unique set of hosts (BUSH 1966). These plants all fruit at different times in the field season (*i.e.*, host fruit represent different temporal resource islands). Because *Rhagoletis* is univoltine and adult longevity is only 4–6 weeks in nature, the life histories of these flies must match the fruiting phenologies of their respective host plants to maximize fly fitness (SMITH 1988; FEDER *et al.* 1993, 1997a, 1998; BERLOCHER 2000). Host/insect synchrony is mediated in *R. pomonella* through variation in the depth of the overwintering pupal diapause. One consequence of this is that the different fly races and sibling species eclose and reach sexual maturity at different times of the season, resulting in allochronic mating isolation (SMITH 1988; FEDER *et al.* 1993, 1998). For example, apple fly adults eclose an average of 10 days earlier than sympatric hawthorn flies, reflecting the \sim 3-week-earlier mean fruiting phenology of apples than of haws (FEDER *et al.* 1993; FEDER 1995). We have estimated that this eclosion time difference translates into a 20–30% reduction in interbreeding between apple and hawthorn flies (FEDER *et al.* 1993, 1998). Variation in host phenology also exposes fly larvae and pupae to different environmental conditions before winter. The prewintering period is of biological significance because diapause is facultative in *R. pomonella* (DEAN and CHAPMAN 1973; BOLLER and PROKOPY 1976). Fly larvae and pupae exposed to elevated temperatures for prolonged periods will forgo an extended diapause and immediately develop into adults, with disastrous fitness consequences in the field (FEDER *et al.* 1997a). As a result, flies that infest host plants with earlier fruiting times are selected to have a more recalcitrant pupal diapause

(FEDER *et al.* 1997a,b). This life history difference can be an important postmating barrier to gene flow (FEDER *et al.* 1997a,b, 1998; FILCHAK *et al.* 2000).

The six allozymes displaying host-related differentiation for *R. pomonella* have all been shown to correlate with the timing of adult eclosion (FEDER *et al.* 1993, 1997a,b; FILCHAK *et al.* 2000), a trait related to the depth of the pupal diapause (DANILEVSKY 1965; TAUBER and TAUBER 1976). Selection experiments in which flies were exposed to environmental conditions simulating the phenological difference between hawthorn and apple trees resulted in significant responses for these six allozymes in predicted directions (FEDER *et al.* 1997a,b; FILCHAK *et al.* 2000). Conditions emulating the earlier fruiting time for apples shifted *Me*, *Acon-2*, *Mpi*, *Aat-2*, *Dia-2*, and *Had* frequencies in surviving flies toward those typically found in the apple race, while conditions similar to those faced by flies infesting the later fruiting hawthorn favored haw race alleles. Thus, a relationship between host-related ecological specialization and population divergence has been established for Rhagoletis. Moreover, a connection has been made between genetic markers differentiating fly populations and diapause traits of adaptive significance responsible for reproductively isolating taxa.

Evidence suggests that the genetic architecture of diapause-related variation in *R. pomonella* may involve more than just first-order differences in allozyme frequencies, however. The six allozyme loci displaying allele frequency differences between the apple and hawthorn host races map to only three different regions of the *R. pomonella* genome on chromosomes 1, 2, and 3 (see Figures 1–4; BERLOCHER and SMITH 1983; FEDER *et al.* 1989; ROETHELE *et al.* 1997, 2001). Significant linkage disequilibrium has also been detected between allozyme markers within each of these three regions (FEDER *et al.* 1988, 1990). Such disequilibrium is rare in natural populations of most organisms, except when genes are associated with some type of chromosomal rearrangement (KRIMBAS and POWELL 1992; POWELL 1997). Thus, the six allozymes differentiating apple and hawthorn flies may demarcate inversion or translocation polymorphism, raising the specter of genetic coadaptation for life history traits involved in sympatric host race formation and speciation in the *R. pomonella* group.

Here, we test the inversion hypothesis for *R. pomonella* by analyzing patterns of meiotic recombination and gametic disequilibrium among complementary DNA (cDNA) and allozyme markers. Our strategy centered on documenting differences in the linear map order of loci among genetic crosses, the same classic approach first used by STURTEVANT (1926) to deduce the existence of inversions in *Drosophila*. This mapping-based approach was necessary because *R. pomonella* polytene chromosomes are currently of such poor quality that we cannot directly visualize the physical loop structures characteristic of inversion heterokaryotypes.

We report evidence that the six allozymes displaying

host-related frequency differences between the apple and hawthorn fly races reside within inversions. Moreover, inversion polymorphism subsumes a substantial portion of chromosomes 1, 2, and 3, covering perhaps one-half of the *R. pomonella* genome ($n = 6$). As a result, gametic disequilibrium is extensive in *R. pomonella*, involving many more loci than just the previously characterized six allozymes. It is therefore more accurate to characterize genetic variation between apple and hawthorn host races, and possibly among *R. pomonella* group sibling species, as involving large suites of correlated loci tied up in inversions, rather than to characterize it just by first-order differences in allozyme frequencies. We conclude by discussing the implications of our findings for sympatric host race formation and for speciation in Rhagoletis and explore the possibility that the three inverted chromosomal regions constitute coadapted gene complexes.

MATERIALS AND METHODS

Overview of research strategy: To test for chromosomal rearrangements, we first constructed an expressed sequence-tagged library for *R. pomonella* (ROETHELE *et al.* 1997, 2001). The library was then used to build an integrated linkage map of cDNA and allozyme loci for the six chromosomes comprising the *R. pomonella* genome (ROETHELE *et al.* 1997, 2001). Polymorphic markers were used to examine patterns of meiotic recombination and gametic disequilibrium from single-pair genetic crosses. Variation in the inferred linear order of genes among crosses was used to test for inversion polymorphism. The analysis was based on a set of 43 single-pair crosses performed using apple flies collected directly from nature at a field site near Grant, Michigan. We limited this study to chromosomes 1–4 because: (1) these four chromosomes presently have the most polymorphic markers and extensive genetic maps, and (2) chromosomes 1–3 are known to contain genes affecting diapause-related traits involved in host-associated adaptation and sympatric race formation for Rhagoletis (FEDER *et al.* 1988, 1997a,b; FILCHAK *et al.* 2000).

In addition, we compiled allozyme data for 66 host race populations from across the eastern United States (FEDER and BUSH 1989b; FEDER *et al.* 1990) to assess the general pattern of linkage disequilibrium throughout the range of geographic overlap between apple and haw flies. Given the magnitude and scope of nonrandom genetic associations we report below, future studies are planned to expand the population level survey of allozymes to include cDNA markers.

Building the *R. pomonella* linkage map: An earlier linkage map was constructed for *R. pomonella* based on 14 single-pair crosses involving a nondiapausing, laboratory line of *R. pomonella* (ROETHELE *et al.* 1997). The lab line was derived from a natural apple-infesting population collected near Geneva, New York, in the 1960s. At the time of the crosses, the Geneva line had undergone as many as 100 generations in the laboratory. It is therefore possible that map positions and linkage relationships for Geneva flies are not representative of wild flies.

To obtain more direct, field-based measures of recombination and gametic disequilibrium for markers on chromosomes 1–4, we performed an additional set of 43 single-pair crosses using *R. pomonella* flies collected from the wild. Parents used for the crosses were collected as larvae in infested apple fruits at a study site near Grant, Michigan, in the summer of 1995 and reared to adulthood in the laboratory. Single-pair crosses were then established and offspring reared following the same procedures discussed in FEDER *et al.* (1989).

DNA isolated from the head of a single parent or its offspring

was sufficient to score a large number of cDNA markers using PCR-based methods, leaving the thorax and abdomen for allozyme analysis. Amplifications were performed under standard PCR conditions using primer pairs generated from sequence data for cDNA clones previously identified as having a polymorphic restriction site for *AhaI*, *DdeI*, *Sau3A*, or *TaqI* (for details see ROETHELE *et al.* 1997, 2001). The amplified DNA was digested with the appropriate restriction enzyme and the resulting DNA fragments were electrophoretically resolved on 1% NuSieve + 1% agarose gels. We also scored the same flies for the following 15 different isozymes using standard horizontal starch gel electrophoresis techniques: *Acon*, *Adenylate kinase* (*Ak*, EC 2.7.4.3), *Alcohol dehydrogenase* (*ADH*, EC 1.1.1.1), *Aminoacylase* (*Acy*, EC 3.5.1.14), *Aat*, *Dia*, *Fumarase* (*Fum*, EC 4.2.1.2), *Had*, *Isocitrate dehydrogenase* (*Idh*, EC 1.1.1.42), *Malic Enzyme* (*Me*, 1.1.1.40), *Mannose phosphate isomerase* (*Mpi*, 5.3.1.8), *Peptidase* (*Pep*, 3.4.13.9), *Phosphoglucose isomerase* (*Gpi*, EC 5.3.1.9), *Phosphoglucose mutase* (*Pgm*, EC 2.7.5.1), and *Superoxide dismutase* (*Sod*, EC 1.15.1.1; for methodological details see FEDER *et al.* 1989). These 15 isozymes represent 21 different loci, 16 of which are polymorphic in *R. pomonella* on the basis of the criterion that the frequency of the most common allele is <0.95 in populations. An attempt was made to score every parent and offspring for each of the allozyme markers. However, due to resource and time constraints, not all of the crosses were analyzed for all of the cDNA loci.

Analysis of cross data: Establishing linkage relationships is straightforward in *R. pomonella* because recombination is extremely limited in males (BERLOCHER and SMITH 1983; FEDER *et al.* 1989; ROETHELE *et al.* 1997). Consequently, linked blocks of genes are inherited intact in the form of whole chromosomes from fathers, while loci on chromosomes from mothers assort randomly. Linkage group assignments established for the Geneva lab line were therefore readily tested for synteny with the new set of 43 "field" crosses by following the assortment pattern among markers in crosses involving males that were heterozygous at multiple (two or more) genes.

Single-pair crosses in this study produced 10–30 progeny. Sample sizes for individual crosses were therefore large enough to infer syntenic groups, but sometimes not to unambiguously resolve gene order for subsets of tightly linked markers. In these circumstances, sample size is usually increased by pooling data across crosses. However, this is a valid approach only if the order of markers is the same among parents and if pairwise estimates of recombination frequency do not display significant heterogeneity among crosses (STAM 1993). In our case, the raw mapping data suggested evidence for inversion polymorphism in *R. pomonella*, violating the assumptions for data pooling. The maps presented in Figures 1–4 were therefore deduced by first inferring linear gene order separately for each cross. We next compared these "multipoint" crosses to identify groups ("cliques") of crosses agreeing in gene order. We refer to the gene order supported by the largest clique of crosses as the "compatibility map" (see Figures 1–4). To evaluate the extent to which a particular cross in a clique agreed with the compatibility map, a LOD score was estimated for the cross using MAPMAKER/EXP, version 3.0 (LANDER *et al.* 1987; LINCOLN *et al.* 1993) by comparing the compatibility map order for loci to the alternative gene order having the next-highest likelihood. The remaining "incongruous" crosses disagreeing in gene order with the largest clique were used to infer the existence of inversion polymorphism. LOD score analysis was used to assess the support of incongruous crosses for alternative gene orders (rearrangements) relative to the compatibility map.

The cross data were also used to test for gametic disequilibrium between linked loci in the Grant, Michigan, apple fly population ($n = 20$ –172 chromosomes scored, depending upon the markers analyzed). Gamete frequencies (linkage-phase relationships) were determined for parents on the basis of assortment patterns

TABLE 1

Sampling locations for 33 paired apple and hawthorn fly sites scored for allozymes

Site	Year	Location	Distance
1	87	Rodgers Ave, Clare, MI	3.2
2	87	MI 20, Mt. Pleasant, MI	3.0
3	87	Alger Rd., Alma, MI	s
4	87	MI 27, Eureka, MI	s
5	86	Jolly Rd., Okemos, MI	s
6	87	Mich. St. Univ., E. Lansing, MI	s
7	87	F Dr., Eckford, MI	s
8	87	Interstate 24, Parkdale, IN	8.0
9	87	Interstate 69, Gas City, IN	7.1
10	87	No. 45 Rd., Cadillac, MI	s
11	87	Grand R. Ave., Saranac, MI	4.9
12	87	Sprinkle Rd., Portage, MI	4.2
13	87	MI 20, Woodville, MI	s
14	89	112th St., Grant, MI	s
15	87	24th St., Herrington, MI	2.0
16	87	MI 89, Fennville, MI	1.5
17	87	East Rd., Three Oaks, MI	s
18	86	Hogan Farm, Ephraim, WI	s
19	85	Kuehn Farm, Carlsville, WI	s
20	87	MM Rd., Green Bay, WI	s
21	87	Gass L. Rd., Manitowoc, WI	s
22	87	EE Rd., Sheboygan, WI	s
23	87	C. Rd., Port Washington, WI	s
24	87	50th St., N. Cape, WI	s
25	87	WI Hwy. 50, Bristol, WI	s
26	86	IL 132, Waukegan, IL	3.3
27	86	Univ. IL, Urbana, IL	s
28	87	WI 29, Hatley, WI	1.6
29	87	WI 49, Berlin, WI	6.3
30	87	WI 81, Beloit, WI	2.5
31	86	Hwy. 31, Kenton, OH	5.0
32	85	Ag. Exp. Station, Geneva, NY	s
33	86	Univ. MA, Amherst, MA	1.0

Year, year sample was collected; Distance, distance (in kilometers) separating apple and hawthorn trees at paired sites; s, sympatric site where trees were <1 km apart.

observed in their offspring and used to calculate both two- and three-locus disequilibrium levels for markers on chromosomes 1–4. Standardized two-locus gametic disequilibrium values (*i.e.*, correlation coefficients = r_c) were estimated according to WEIR (1996) and tested for significance using Fisher's exact tests. Second-order disequilibrium among triads of loci was examined using the additive formulation of BENNETT (1954), which subtracts terms for first-order pairwise interactions between loci. The resulting three-locus *D* values were tested for significance by chi-square tests (WEIR 1996).

Allozyme survey of natural fly populations: To assess the extent of disequilibrium in *R. pomonella* across a wide portion of its geographic range, we compiled and analyzed allozyme data from 33 paired apple and hawthorn fly populations distributed throughout the eastern United States (see Table 1 for collecting sites; data come from FEDER and BUSH 1989b and FEDER *et al.* 1990). At the majority of the paired apple and hawthorn sites, flies were sampled from host trees <1 km apart and so were considered "sympatric." (One kilometer is well within the cruising range of *R. pomonella*; MAXWELL and PARSONS 1968; FEDER *et al.* 1994.) Flies from all 66 host race populations were scored

for *Aat-2*, *Dia-2*, *Me*, *Acon-2*, *Mpi*, and *Had*, while a subset of 22 populations were resolved for all 16 polymorphic allozymes in *R. pomonella*.

Estimation of linkage disequilibrium from genotype frequency data, as in our allozyme survey, is complicated by the fact that double heterozygotes cannot be distinguished. Nevertheless, composite disequilibrium coefficients that include components due to nonrandom associations of alleles within gametes and the nonrandom union of gametes to form zygotes can still be calculated (COCKERHAM and WEIR 1977). We calculated standardized disequilibrium coefficients (r_c) between pairs of nonallelic allozymes within apple and hawthorn populations on the basis of Burrow's composite values (COCKERHAM and WEIR 1977; WEIR 1996). The coefficients were tested for significance ($r_c > 0$) by z -transformation (ZAR 1996).

RESULTS

Genetic linkage map based on Grant, Michigan, field crosses: Six linkage groups corresponding to the haploid chromosome complement of *R. pomonella* ($n = 6$) were identified from assortment patterns in Grant, Michigan, males (Figures 1–4). Linkage group assignments were unambiguous for all 16 polymorphic allozymes and 50 cDNA markers mapped from the 43 field crosses, implying an absence of translocation polymorphism. In addition, *G-6-pdh* was mapped to chromosome 3 using redundant PCR primers constructed by SOTO-ADAMES *et al.* (1994) and flanking the second intron of the gene, and *Fum* was assigned to chromosome 3 on the basis of crosses involving *R. tabellaria* Fitch (MCIPHERON and BERLOCHER 1985). Linkage group relationships determined for the Grant, Michigan, flies were identical to those deduced for the Geneva, New York, laboratory line (ROETHELE *et al.* 1997), suggesting that synteny is conserved across the geographic range of *R. pomonella*.

Disequilibrium between allozymes in natural populations: Disequilibrium was common between linked allozymes within apple and hawthorn fly populations for each of the three genomic regions on chromosomes 1–3 displaying host-related differentiation (Figure 5). Standardized composite disequilibrium values (r_c) were significantly different from zero for 60 of 64 host race populations that could be tested between *Aat-2 100* and *Dia-2 100* (chromosome 1), 37 of 66 for *Me 100/Acon-2 95* (chromosome 2), and 17 of 22 for *Had 100/Pep-2 100* (chromosome 3; Figure 5).

In contrast, linked allozymes not displaying host-related differences were generally in equilibrium when tested in pairwise combinations both among themselves and with *Aat-2*, *Dia-2*, *Me*, *Acon-2*, *Had*, or *Pep-2*. For example, *Idh* is located 4.2 cM from *Aat-2 Dia-2* and 4.6 cM from *Dia-2* on linkage group I. (These recombination distances represent the mean values between markers averaged across all Grant apple fly crosses involving double heterozygote females.) *Idh* was in equilibrium with both *Dia-2* and *Aat-2* in all 19 host race populations tested. The same was also true in these 19 populations for comparisons between *Idh*, *Ak*, and *Pgm*.

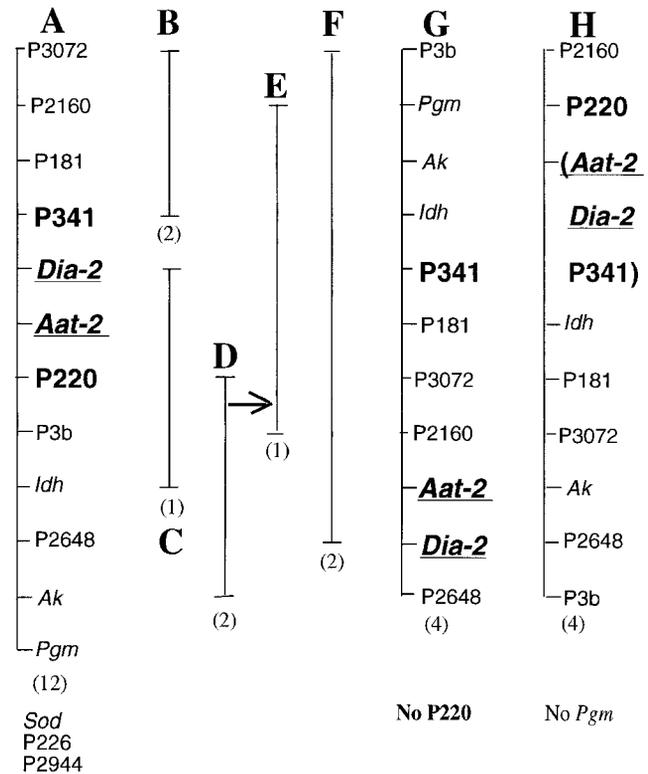


FIGURE 1.—Compatibility linkage map (A) supported by the largest clique of crosses (12) agreeing in gene order for chromosome 1. Also given are alternate hypotheses for organizing the eight remaining incompatible crosses into a series of either (B–F) five simple inversions with bracketed lines denoting regions of inverted gene order compared to the compatibility map, A, and the arrow indicating a two-step rearrangement) or (G and H) two complex inversions that require a series of multiple rearrangements to generate. Allozyme loci are italicized, while the suffix “P” indicates a cDNA locus. Underlined markers are allozymes displaying significant allele frequency differences between the apple and hawthorn host races. Boldface markers are those showing significant pairwise gametic disequilibrium in crosses. Numbers in parentheses below chromosomes are the number of crosses supporting the indicated gene order. Markers listed below the compatibility map have been assigned to chromosome 1, but their position within the chromosome could not be determined from the cross data. Map positions for markers listed below the complex inversion chromosomes could not be determined from the cliques of crosses supporting these alternate gene orders.

Unlinked allozymes on different chromosomes also tended to be in equilibrium. Only 52 of 1217 r_c values calculated between pairs of unlinked allozymes were significantly greater than zero. None of these 52 tests was significant on a table-wide basis after applying a sequential Bonferroni procedure to correct for multiple tests (RICE 1989). Genetic interactions were therefore not obvious between allozymes residing on different linkage groups within the 66 host race populations surveyed.

Disequilibrium values estimated from the apple fly crosses in which the linkage phase of markers in gametes

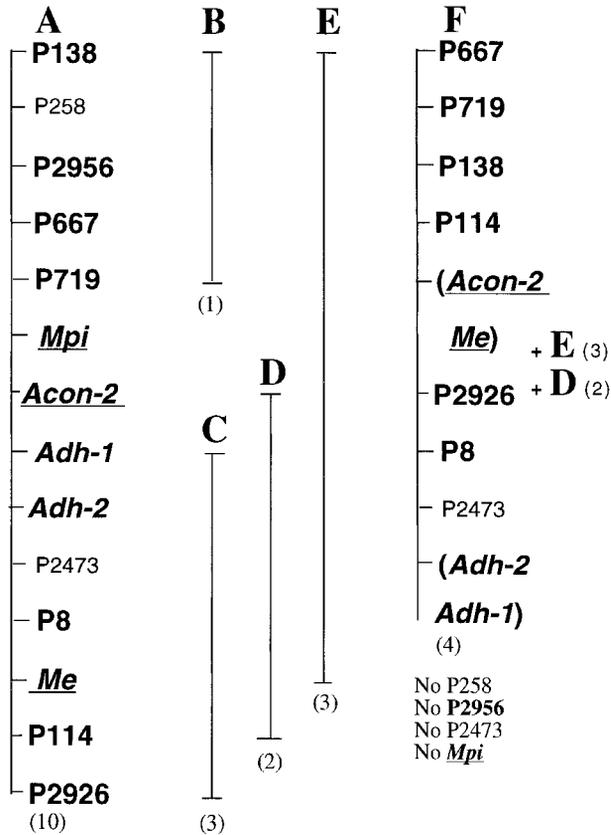


FIGURE 2.—Compatibility linkage map (A) supported by the largest clique of crosses (10) agreeing in gene order for chromosome 2. Also given are alternate hypotheses for organizing the nine remaining incompatible crosses into a series of either four simple inversions (B–E) or one complex (F) and two simple (E and D) inversions. See Figure 1 for additional details.

could be determined were similar to the composite values calculated from genotype frequencies at Grant, Michigan. Genetic cross estimates were $r_g = 0.83$ between *Aat-2* +75 and *Dia-2* 100 ($n = 168$ gametes), $r_g = 0.41$ between *Me* 100 and *Acon-2* 95 ($n = 172$ gametes), and $r_g = 0.26$ between *Had* 100 and *Pep-2* 100 ($n = 146$ gametes). These estimates were not significantly different from the composite r_c values of 0.78, 0.33, and 0.32, respectively, calculated from genotype data ($n = 416$ flies) for the apple race at Grant in 1989 ($P = 0.10$, 0.31, and 0.50 based on z -transformations; ZAR 1996). The concordance of the cross and composite disequilibrium measures suggests that gametes generally unite at random to form zygotes within the apple race at Grant. (Under random mating and in the absence of higher-order effects, composite values are equivalent to standardized gametic disequilibrium coefficients; WEIR 1996.) The random-mating hypothesis was also supported by genotype frequencies for all six allozymes in Hardy-Weinberg equilibrium at Grant (data not shown).

The extent of disequilibrium across the genome (allozyme and cDNA markers): The field cross data from

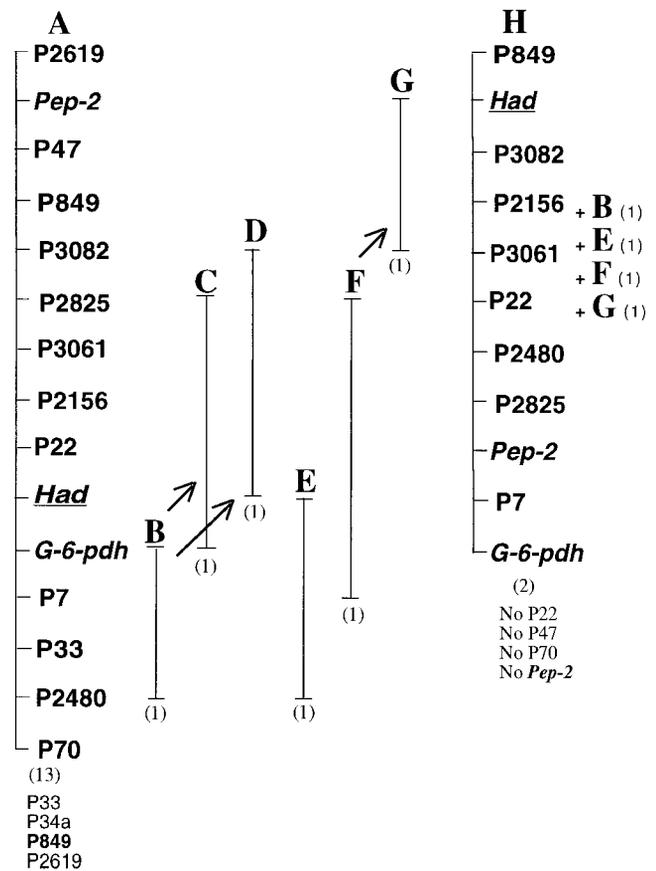


FIGURE 3.—Compatibility linkage map (A) supported by the largest clique of crosses (13) agreeing in gene order for chromosome 3. Also given are alternate hypotheses for organizing the six remaining incompatible crosses into a series of either three one-step (B, E, and F) and three two-step (C, D, and G) simple inversions or one complex (H) and four simple (B, E, F, and G) inversions. See Figure 1 for additional details.

Grant, Michigan, indicated the presence of substantial gametic disequilibrium across large portions of chromosomes 1–3 (Figures 1–3 and 6). The magnitude of pairwise, two-locus disequilibrium estimated for the apple fly population was inversely related to the mean recombination distance separating markers on chromosomes 1–3 (Figure 6). Significant pairwise disequilibrium was observed between 4 of 15 genetic markers on chromosome 1, 12 of 14 loci on chromosome 2, and 16 of 19 genes on chromosome 3 (Figures 1–3 and 6). In contrast, no significant disequilibrium was observed for any pairwise test conducted between the 10 markers on chromosome 4 (Figures 4 and 6).

Second-order disequilibrium was detected among triads of markers within all three genomic regions of chromosomes 1–3 displaying host-related differentiation. A total of 3 of 4, 21 of 146, and 100 of 241 three-locus tests were significant for chromosomes 1–3, respectively. One, 3, and 7 of these tests were significant on a table-wide basis, as determined by a sequential Bonferroni correction

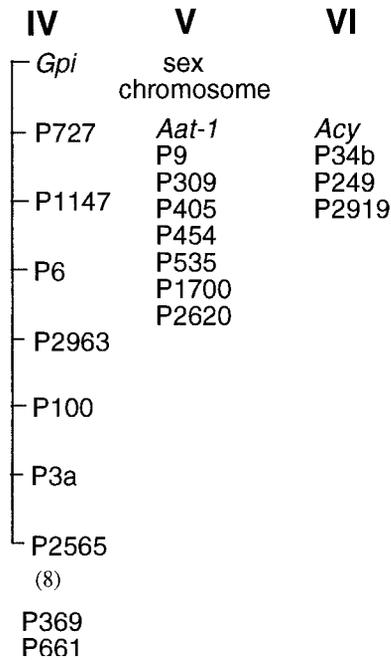


FIGURE 4.—Genetic map supported by eight crosses for chromosome 4. Also given are loci mapping to chromosomes 5 and 6 whose gene orders have yet to be determined. See Figure 1 for additional details.

procedure. Consequently, significant multilocus associations that deviate from predictions on the basis of first-order relationships between genes are present in *R. pomonella*. Out of a total of 82 higher-order three-locus disequilibrium tests, 4 were also significant among chromosome 4 markers, but none of these 4 tests was significant on a table-wide basis.

Recombination heterogeneity and variation in linear gene order: Significant recombination heterogeneity was observed among crosses for markers within each of the three chromosomal regions that displayed gametic disequilibrium (see Figure 7 for marker pairs for which we had the largest data sets). Recombination heterogeneity was not seen, however, among chromosome 4 markers (Figure 7). This suggests that exchange rates may be more uniform outside of the three genomic regions displaying host-related differentiation, but additional cross data are required to confirm this pattern.

The inferred linear order of loci in the three genomic regions displaying gametic disequilibrium on chromosomes 1–3 also varied among crosses (Figures 1–3; Tables 2–4). For chromosome 1, the largest clique of 12 crosses involving multiply heterozygous females supported the compatibility map designated by the bold-face letter “A” in Figure 1 (see Table 2 for LOD analysis). However, the remaining 8 crosses for chromosome 1 implied different gene orders designated by the letters B–H (Figure 1; Table 2). These 8 incongruous crosses could be interpreted in two mutually exclusive ways. First, they could be organized into a set of five simple and overlapping rearrangements designated by the let-

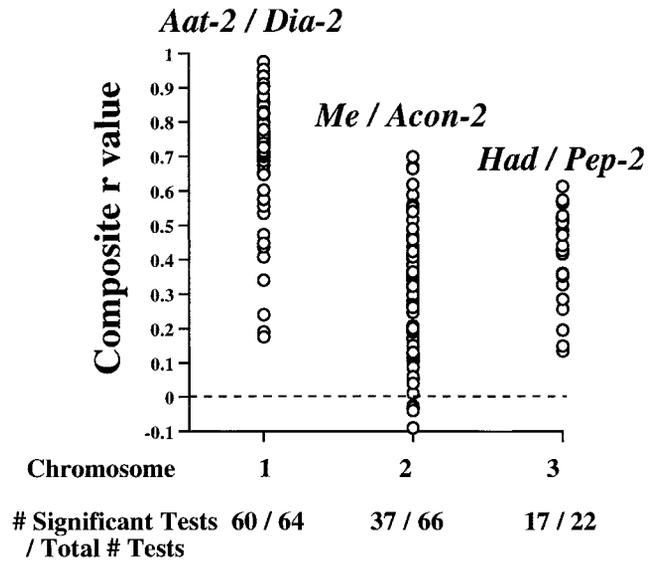


FIGURE 5.—Linkage disequilibrium within host race populations in the eastern United States. Distributions of composite disequilibrium coefficients (r_c) were calculated between pairs of linked, nonallelic allozymes within apple and hawthorn fly races on the basis of Burrow’s disequilibrium values (Δ ; COCKERHAM and WEIR 1977; WEIR 1996). Data represent a total of 66 apple and hawthorn populations collected at 33 paired sites across the eastern United States (see Table 1 for list of sites). The disequilibrium coefficients were tested for significance ($r_c \neq 0$) by z -transformation (ZAR 1996).

ters B–F in Figure 1. In this case, 7 of the 8 crosses represent four different single-inversion events, while the remaining cross is accounted for by a subsequent inversion involving one of the single-rearrangement D chromosomes (see arrow in Figure 1). Under the “simple” inversion scenario, a total of six different gene orders would therefore segregate for chromosome 1. Alternatively, the 8 incongruous crosses could be subdivided into two subcliques of 4 crosses each that differ in gene order from each other and from the compatibility map by a complex series of rearrangements (see complex inversion hypothesis in Figure 1). Under the complex inversion scenario, a total of only three different gene orders would segregate for chromosome 1 (Figure 1).

The data for chromosomes 2 and 3 were similar to those for chromosome 1. Compatibility maps were generated for both chromosomes 2 and 3, supported by a majority of crosses (chromosome 2 = 10/19 = 53%; chromosome 3 = 13/19 = 68%; Figures 2 and 3; Tables 3 and 4). The remaining nine incongruous crosses for chromosome 2 and six for chromosome 3 could be interpreted within the context of either a set of relatively simple one- and two-step inversions or a combination of complex and simple rearrangements (Figures 2 and 3; Tables 3 and 4).

In contrast to the results for chromosomes 1–3, we found no disagreement in gene order among the eight

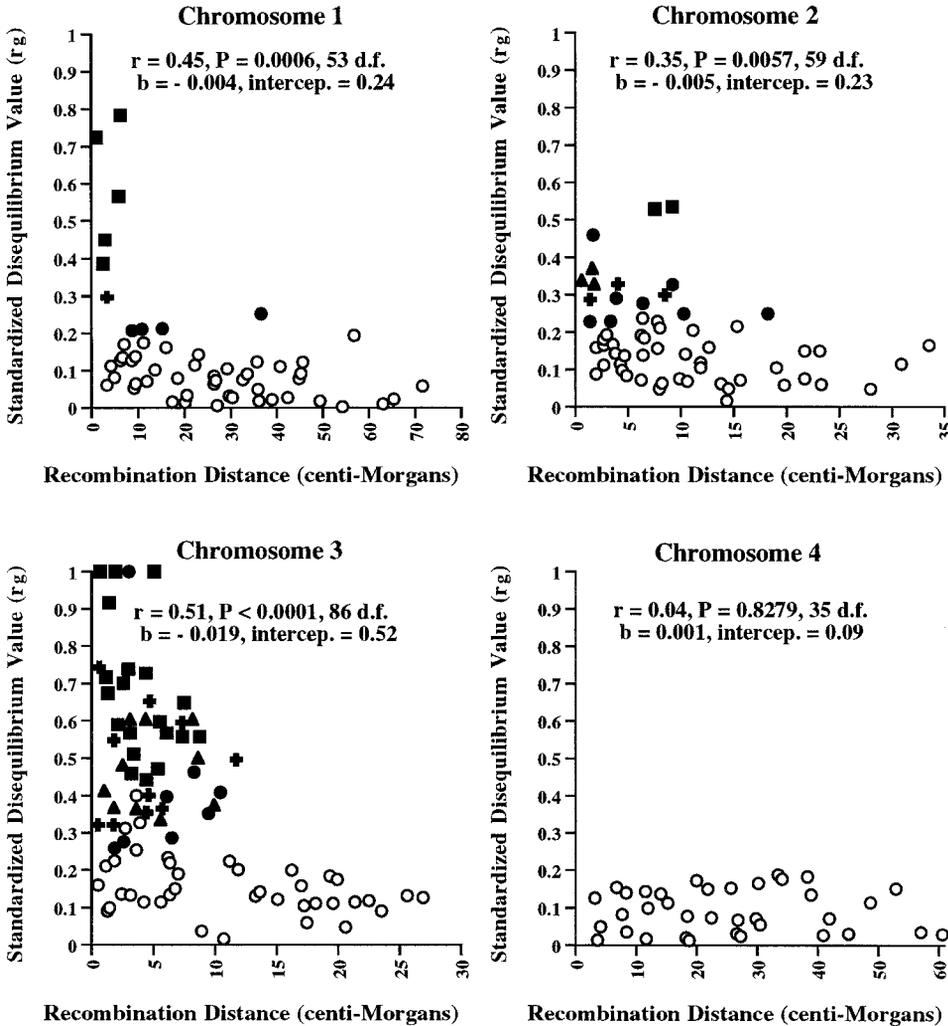


FIGURE 6.—Pairwise gametic disequilibrium coefficients (r_g) between linked genetic markers on chromosomes 1–4 plotted against their recombination map distances. Data were derived from 43 single-pair crosses between apple flies from Grant, Michigan, with map distances between markers representing the mean value among crosses. Significance of disequilibrium values ($r_g \neq 0$) was determined by chi-square tests (open circle, $P > 0.05$; solid circle, $P \leq 0.05$; solid cross, $P \leq 0.01$; solid triangle, $P \leq 0.001$; solid square, $P \leq 0.0001$). Also given are the linear regression equations between disequilibrium coefficients and recombination distances for each chromosome.

multipoint female crosses for chromosome 4 (Figure 4, Table 5).

DISCUSSION

Evidence for inversion polymorphism: Our data indicate that all six allozyme loci currently known to display host-associated differentiation in *R. pomonella* reside within three different inverted regions of the fly's genome. *Aat-2*, *Dia-2*, *Me*, *Acon-2*, *Mpi*, and *Had* are located within segments of chromosomes 1–3 that show highly significant and extensive gametic disequilibrium. Such disequilibrium is rare in natural populations of most organisms, except when genes are associated with inversions or some other type of chromosomal rearrangement (*e.g.*, translocations; KRIMBAS and POWELL 1992; POWELL 1997). In *R. pomonella*, gametic disequilibrium exists for large numbers of markers on chromosomes 1–3 and is present between allozymes at a majority of host race populations surveyed in the eastern United States. Moreover, the sign and magnitude of disequilibrium values among linked allozymes is similar to that between apple and hawthorn fly populations across sites. This pattern is inconsistent with the disequilibrium

being generated by stochastic processes (genetic drift) in small, local demes (see LEWONTIN 1974). Rather, it supports the existence of shared rearrangements between the host races. We also observed significant recombination heterogeneity and variation in the linear order of genes among single-pair crosses for loci on chromosomes 1–3. Recombination heterogeneity and differences in linear gene order are expected among crosses in which parents are alternately homozygous and heterozygous for chromosomal inversions. We can rule out translocations because chromosome assignments were invariant for loci among crosses. Genetic differentiation between apple and hawthorn host races, and probably among *R. pomonella* sibling species, is therefore best characterized by large suites of correlated loci tied up in inversions, rather than just by first-order variation in allozyme frequencies.

In contrast, the homogeneity of recombination rates, concordance of gene order, and lack of gametic disequilibrium for chromosome 4 markers (Figures 4, 6, and 7) imply an absence of inversion polymorphism. Additional crosses and markers are needed to investigate chromosomes 5 and 6 to determine if the same holds true for these two chromosomes. The current cDNA

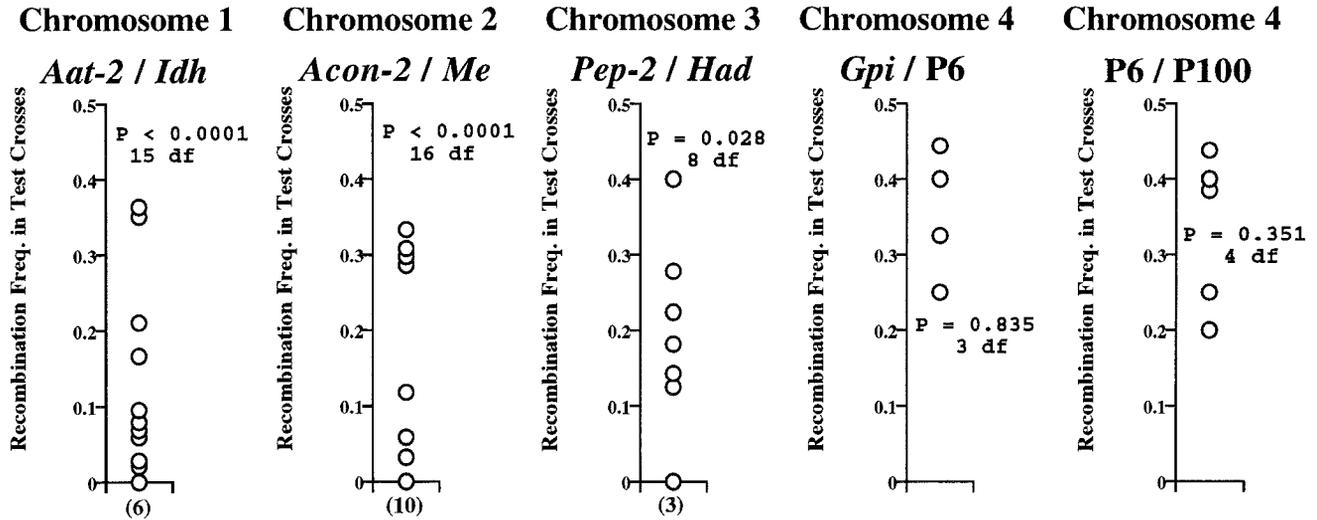


FIGURE 7.—Recombination distance heterogeneity among crosses between indicated pairs of genetic markers on chromosomes 1–4 (10–70 offspring scored/cross). Numbers in parentheses are the number of crosses in which no recombination was detected between markers. Significance was determined by Fisher’s exact tests.

clones in female parents were not sufficiently polymorphic to yield a large number of multipoint crosses for assessing gene order for chromosomes 5 and 6.

Unfortunately, polytene chromosome preparations are currently of such poor quality in *R. pomonella* that we cannot cytologically confirm the existence or pinpoint the location of loop-like structures indicative of inversions. This also precludes us from determining whether the rearrangements represent pericentric or paracentric inversions or differ from one another by a simple, as opposed to a complex, sequence of evolutionary steps. Inversion polymorphism is not unprecedented for Tephritid flies, however, and has been physically documented for *Procecidochares utilis* on the basis of polytene chromosome spreads (BUSH and TAYLOR 1969). Fluorescence *in situ* hybridization of probes to meiotic chromosomes or other more refined cytogenetic techniques might provide physical verification of inversions for *R. pomonella*.

Genetic coadaptation? The existence of several blocks of linked, nonrandomly associated genes in *R. pomonella* affecting diapause traits raises the issue of whether the inversions represent coadapted gene complexes. Here, we restrict our definition of coadaptation to Dobzhansky’s use of the term as nonadditive, epistatic fitness interactions among genes producing a state of gametic disequilibrium enhancing mean population fitness. However, while it is clear that the rearrangements in *R. pomonella* are under host-dependent selection related to diapause, and that extensive gametic disequilibrium exists within and/or surrounding the rearrangements, these observations alone are not sufficient to verify coadaptation. For example, it is still possible that each of the three inverted regions of the genome contains only a single locus under balancing- or frequency-dependent selection, with genetic hitchhiking of linked, neutral

genes being responsible for the disequilibrium. Differential, additive selection on multiple loci that adapt flies to alternative plants, coupled with interhost migration, could also generate a permanent state of gametic disequilibrium between linked genes (FELSENSTEIN 1965; NEI and LI 1973; PROUT 1973; LI and NEI 1974; FELDMAN and CHRISTIANSEN 1975; SLATKIN 1975).

Examining the hitchhiking hypothesis: Until such time as more sophisticated genetic constructs and techniques are available for experimentally dissecting the *Rhagoletis* genome, the question of whether hitchhiking explains the observed disequilibrium necessitates a theoretical approach. In this context, the key issue is whether recombination and gene conversion have had sufficient time to restore random genetic associations among loci to discount an historical cause for the observed disequilibrium. General analytical solutions (approximations) to this question have been derived for the case of a two-locus epistatic system by ISHII and CHARLESWORTH (1977) and for a more general deterministic model by NEI and LI (1980). As discussed by ISHII and CHARLESWORTH (1977, p. 100),

unless a neutral locus is very closely linked to one of the selected loci involved in the maintenance of the inversion polymorphism, the half-life of the decay of an association between the neutral locus and the inversion is of the order of the reciprocal of the rate of double crossing over in heterokaryotypes. It therefore seems likely that experimental estimates of the rate of exchange of alleles at an allozyme locus between two mutually inverted gene arrangements will provide a good estimate of the rate of decay for the neutral (hitchhiking) hypothesis.

Unfortunately, data on rates of gene flux in heterokaryotypes do not exist for *Rhagoletis* and are relatively sparse even for *Drosophila* species (ISHII and CHARLESWORTH 1977; NEI and LI 1980; NAVARRO *et al.* 2000). Neverthe-

TABLE 2
LOD scores for the largest clique of compatible crosses on chromosomes 1–4

Cross no.	Chromosome 1	Chromosome 2	Chromosome 3	Chromosome 4
C15		3.22	0.25	
C16			×	0.22
C17	×	×	4.08	0.20
C20			×	
C23	0.88	1.88	8.28	
C25	0.60			
C26	×			
C32			0.77	
C33			0.48	0.81
C34			0.35	
C37	0.21	×	5.00	
NC1	0.54	×	×	
NC22	0.37	×		
NC32	×		0.18	
NC37		6.22		
NC40	×	0.49		0.54
NC44	1.02	0.32	0.41	
NC51	1.34	×	0.33	
NC58	×	×	1.15	
NC94			×	
NC97	0.81	1.30		1.30
A29	0.37	0.34		0.68
A45			×	0.17
A87		×		
A100	0.87			
A105	×	×	0.35	
A109		1.67		
A114	×			
A121		×		
A124		0.55	2.61	
A125	×	0.71		2.12
A127	0.32			
A246	0.37		×	

Scores were calculated as \log_{10} (likelihood for the compatibility map order for a cross/likelihood for second-most likely gene order for the cross). ×, cross in which the maximum-likelihood gene order was different from the compatibility map. No data shown, cross in which female was not heterozygous for three or more loci on a given chromosome (*i.e.*, not a multipoint cross). The compatibility map order for each chromosome is given as order “A” in Figures 1–4.

less, it has been suggested that double-exchange/gene conversion rates between inversions on the order of 10^{-4} – 10^{-5} /meiosis are probably typical in *Drosophila* (ISHII and CHARLESWORTH 1977; NEI and LI 1980). If these rates also generally apply to heterokaryotypic *Rhagoletis* females, then the decay half-life for gametic disequilibrium will range from $\sim 20,000$ to $200,000$ generations (years) for neutral markers within inversions. (Recall that *R. pomonella* is univoltine, having one generation per year, and that recombination is all but nonexistent in males, thereby halving the overall gene flux rate between inversions.)

Elsewhere, we present nucleotide sequence data for cDNA loci and mtDNA implying that inversion polymorphism for *R. pomonella* chromosomes 1–3 has been segregating for anywhere from 0.84 to 1.39 million years (J. L.

FEDER, unpublished results). Unless gene flux rates between heterokaryotypes is routinely on the order of $\sim 5 \times 10^{-7}$ or less in *Rhagoletis* females, then the estimated ages of the inversions make it difficult to account for the high levels of observed disequilibrium in the absence of some form of nonadditive (epistatic) selection [*e.g.*, r_g between *Dia-2/Aat-2* (chromosome 1) = 0.83, $P < 10^{-6}$, $n = 168$ gametes; *P2956/Aconitase-2* (chromosome 2) = 0.63, $P < 5 \times 10^{-4}$, $n = 46$; *P7/Had* (chromosome 3) = 1.0, $P < 10^{-6}$, $n = 96$]. Moreover, 0.84 million years greatly predates the ~ 150 -year-old origin of the apple race. Inversion polymorphism therefore appears to have been segregating in *R. pomonella* long before it was seized upon by host-associated selection during the shift to apple, discounting interhost migration as the primary cause for gametic disequilibrium. Finally, the large num-

TABLE 3
LOD scores for cross disagreeing with the compatibility map for chromosome 1

Cross no.	Order							
	A	B	C	D	E	F	G	H
C17	0.39	0.39	0.39	1.41	1.41	ML	ML	1.41
C26	5.11	ML	5.11	5.11	5.11	5.11	ML	5.11
NC32	3.74	8.96	5.36	9.35	7.24	ML	8.96	ML
NC40	1.32	ML	1.32	2.98	4.18	1.92	ML	1.20
NC58	0.86	0.86	3.48	ML	0.61	9.57	8.71	ML
A105	3.91	3.91	10.21	ML	6.28	0.47	1.24	ML
A114	1.15	1.15	0.85	2.12	ML	1.15	ML	3.83
A125	0.41	0.41	ML	0.41	0.41	0.41	0.41	ML

A, compatibility gene order; B–H, alternative gene orders shown in Figure 1. ML, maximum-likelihood gene order(s) for a cross. LOD scores were calculated as \log_{10} (likelihood for the ML gene order for a cross/likelihood for the alternative gene order designated by the column heading).

ber of significant multi-locus associations within chromosomes deviating from linear predictions on the basis of pairwise, first-order disequilibrium values between loci is also suggestive of possible higher-order fitness interactions among genes. Our current data therefore imply that linked blocks of genes within inversions on *Rhagoletis* chromosomes 1–3 behave as coadapted complexes. Further empirical and theoretical work is needed to verify this point, however, and we found no convincing evidence from our disequilibrium survey of natural populations supporting epistatic interactions between different chromosomes.

Are inversions pivotal to host race formation? Although our results indicate that three different inverted regions of the *R. pomonella* genome encode important diapause traits involved in sympatric host race formation and speciation, we caution that this does not mean that all host-associated genes will be found to reside within rearrangements. Indeed, given the magnitude and extent of gametic disequilibrium that exists in *R. pomonella*, surveys of natural populations will inevitably be biased toward detecting markers and traits associated with inversions.

To demonstrate, assume that half of the physical length of the *R. pomonella* genome is tied up in inversions and the other half is not. Moreover, assume that nine loci under host-related selection are located within the inversions, while an additional 20 genes are scattered through 200 cM of the remaining genome. A total of 16 polymorphic markers are scored to detect host-associated differentiation (the number of allozymes resolved for *R. pomonella*), 6 of which happen to reside within the inversions and 10 of which do not. If a neutral marker outside of an inversion must map to within 0.1 cM of a selected locus to be in strong enough linkage disequilibrium to display host-related differentiation via hitchhiking, then the probability is high ($\sim 81.7\% = [1 - (20 \times 2 \times 0.1 \text{ cM}/200 \text{ cM})]^{10}$) that none of the

10 markers outside the inversions will show a difference between the host races. In contrast, all 6 markers within the inversions will display host-related differences.

Detailed quantitative trait locus mapping studies of host-related traits are therefore needed to complement the population surveys and cytogenetics of *R. pomonella* before we can claim an exclusive role for inversions in sympatric host race formation and speciation. Obviously, inversion polymorphism is an important genetic consideration for host shifts, but inversions may not be the complete story.

Are host performance and preference genes linked?

The presence of extensive gametic disequilibrium in *R. pomonella* raises the intriguing possibility of genetic linkage between host-plant recognition and host perfor-

TABLE 4
LOD scores for crosses disagreeing with the compatibility map for chromosome 2

Cross no.	Order					
	A	B	C	D	E	F
C17	3.18	ML	3.18	3.50	3.18	ML
C37	0.94	1.72	0.94	ML	0.94	0.94
NC1	1.92	2.63	6.26	ML	4.94	1.50
NC22	1.22	1.22	ML	2.44	1.22	ML
NC51	2.76	2.76	0.91	2.76	ML	2.76
NC58	1.08	1.08	7.40	4.13	ML	7.40
A87	7.15	5.81	7.15	7.15	ML	7.15
A105	2.19	2.19	ML	6.01	0.89	ML
A121	3.03	5.00	ML	3.03	5.00	ML

A, compatibility gene order; B–F, alternative gene orders shown in Figure 2. ML, maximum-likelihood gene order(s) for a cross. LOD scores were calculated as \log_{10} (likelihood for the ML gene order for a cross/likelihood for the alternative gene order designated by the column heading).

TABLE 5
LOD scores for crosses disagreeing with the compatibility map for chromosome 1

Cross no.	Order							
	A	B	C	D	E	F	G	H
C16	3.61	1.73	ML	2.61	3.61	3.61	3.61	ML
C20	5.44	ML	5.44	5.44	6.14	6.14	6.14	5.44
NC1	12.52	12.52	12.52	12.52	14.58	14.58	ML	12.52
NC94	>15	>15	>15	>15	ML	>15	>15	>15
A45	1.84	1.84	1.84	ML	0.81	3.44	3.44	ML
A246	2.55	2.55	2.55	2.55	2.55	ML	2.55	2.55

A, compatibility gene order; B–H, alternative gene orders shown in Figure 3. ML, maximum-likelihood gene order(s) for a cross. LOD scores were calculated as \log_{10} (likelihood for the ML gene order for a cross/likelihood for the alternative gene order designated by the column heading).

mance (survivorship) traits. Resolution of this question is important because several models of sympatric speciation are predicated on such an association (*e.g.*, MAYNARD-SMITH 1966; BUSH 1969a,b; FELSENSTEIN 1981). Indeed, host acceptance and performance traits have recently been found to be linked in the pea aphid, *Acyrtosiphon pisum* (HAWTHORNE and VIA 2001). As we mentioned above, the three genomic regions displaying disequilibrium in *R. pomonella* all correlate with diapause traits adapting flies to differences in host fruiting phenology (FILCHAK *et al.* 2000). The inversions therefore contain genes affecting host-related performance. Genetically based differences in host preference have also been shown to be important prezygotic barriers to gene flow between the apple and hawthorn host races, as well as among related sibling species in the *R. pomonella* complex (PROKOPY *et al.* 1988; FEDER *et al.* 1994, 1998). It therefore is important to determine whether genes affecting host recognition map to the same genomic regions as diapause traits. If so, then a major tenet of sympatric host race formation would be verified.

A final irony: Through the *Genetics of Natural Population* series and his influential book *Genetics and the Origin of Species*, Theodosius Dobzhansky began to build a “unified science of population biology out of the elements of ecology and population genetics” (LEWONTIN 1997). Dobzhansky’s research scheme started with the characterization of standing genetic variation and its modulation in natural populations. Within this context, the process of speciation reduced to understanding how this polymorphism became converted into interspecific genetic differences in geographically isolated populations. The refinement of genetic techniques to resolve polymorphism at the nucleotide level has since resulted in the verification of the action of balancing and directional and purifying selection on particular genes (see reviews by KREITMAN and AKASHI 1995; WAYNE and SIMONSEN 1998). However, studies of population divergence following Dobzhansky’s lead have generally failed to link specific genetic changes to how organisms come

to occupy alternative adaptive peaks (acquire new ways of making a living) and become reproductively isolated during the speciation process (LEWONTIN 1997; but see LEE *et al.* 1995; METZ and PALUMBI 1996; TING *et al.* 1998).

In contrast, the apple maggot has become a model system for ecological speciation in sympatry. Due to its importance as an economic pest, much is known about the biology and natural history of the fly, as well as related *Rhagoletis* species (DEAN and CHAPMAN 1973; BOLLER and PROKOPY 1976). From an evolutionary perspective, traits involved in host recognition and diapause have been shown to differentially adapt *R. pomonella* flies to their respective host plants (FEDER *et al.* 1994, 1997a,b, 1998; FEDER 1998; FILCHAK *et al.* 2000). In addition, the relationship of this ecological specialization to reproductive isolation has been established (FEDER 1995, 1998).

Here, we show that three regions of the *R. pomonella* genome associated with host-related adaptation and population divergence are subsumed by inversion polymorphism (either overlapping sets of simple inversions or complex series of rearrangements). Moreover, epistatic interactions affecting fitness may exist between genes within these inversions. Thus, a convergence appears to be emerging between the genetic architecture of divergence between *R. pomonella* and *D. pseudoobscura*. For example, NOOR *et al.* (2001a,b) have recently presented evidence that regions of the *D. pseudoobscura* and *D. persimilis* genome associated with male hybrid sterility and female species mating preferences primarily involve two fixed inversions separating the flies. Moreover, NOOR *et al.* (2001b) and RIESEBERG (2001) have proposed theoretical models whereby inversions that differentiate hybridizing taxa can disproportionately contribute to the persistence of taxa and their continued divergence in sympatry in the face of gene flow. It is therefore ironic that the *Rhagoletis* model for sympatric speciation could eventually contribute some of the most compelling evidence fulfilling Dobzhansky’s grand vi-

sion for the study of speciation and supporting his theory of genetic coadaptation.

The authors thank the following individuals for their assistance, moral support, and/or conversational input: Stewart Berlocher, Andrew Berry, Guy Bush, Drew Denker, Scott Freeman, Marty Kreitman, Kristin Lewis, Bruce McPheron, Martin Taylor, Joseph O'Tousa, William L. Perry, Dave Prokrym, Jim Smith, Uwe Stolz, the USDA/APHIS/PPQ facility at Niles, Michigan, and four anonymous reviewers. We also thank Paul Lewis and Dmitri Zaykin for their help in perfecting a computer program written by J.L.F. to test for higher-order linkage disequilibrium. This research was supported, in part, by grants from the National Science Foundation, the USDA, and the 21st Century Fund of the state of Indiana to J.L.F. and is dedicated in form and spirit to T. Dobzhansky and to A. H. Sturtevant, as well as to the late Tom Wood.

LITERATURE CITED

- ANDOLFATTO, P., and M. KREITMAN, 2000 Molecular variation at the *In(2L)t* proximal breakpoint site in natural populations of *Drosophila melanogaster* and *D. simulans*. *Genetics* **154**: 1681–1691.
- ANDOLFATTO, P., and M. PREZEWSKI, 2001 Regions of lower crossing over harbor more rare variants in African populations of *Drosophila melanogaster*. *Genetics* **158**: 657–665.
- ANDOLFATTO, P., J. D. WALL and M. KREITMAN, 1999 Unusual haplotype structure at the proximal breakpoint of *In(2L)t* proximal breakpoint in a natural population of *Drosophila melanogaster*. *Genetics* **153**: 1297–1311.
- AQUADRO, C. F., A. L. WEAVER, S. W. SCHAFFER and W. W. ANDERSON, 1991 Molecular evolution of inversions in *Drosophila pseudoobscura*: the amylase gene region. *Proc. Natl. Acad. Sci. USA* **88**: 305–309.
- BABCOCK, C. S., and W. W. ANDERSON, 1996 Molecular evolution of the sex-ratio inversion complex in *Drosophila pseudoobscura*: analysis of the *Esterase-5* gene region. *Mol. Biol. Evol.* **13**: 297–308.
- BENASSI, V., S. AULARD, S. MAZEAU and M. VEUILLE, 1993 Molecular variation of *Adh* and *P6* genes in an African population of *Drosophila melanogaster* and its relation to chromosomal inversions. *Genetics* **134**: 789–799.
- BENNETT, J. H., 1954 On the theory of random mating. *Ann. Eugen.* **18**: 311–317.
- BERLOCHER, S. H., 2000 Radiation and divergence in the *Rhagoletis pomonella* species group: inferences from allozymes. *Evolution* **54**: 543–557.
- BERLOCHER, S. H., and J. L. FEDER, 2002 Sympatric speciation in phytophagous insects: Moving beyond controversy? *Annu. Rev. Entomol.* **47**: 773–815.
- BERLOCHER, S. H., and D. C. SMITH, 1983 Segregation and mapping of allozymes of the apple maggot fly. *J. Hered.* **74**: 337–340.
- BERLOCHER, S. H., B. A. MCPHERON, J. L. FEDER and G. L. BUSH, 1993 A revised phylogeny of the *Rhagoletis pomonella* (Diptera: Tephritidae) sibling species group. *Ann. Entomol. Soc.* **86**: 716–727.
- BOLLER, E. F., and R. J. PROKOPY, 1976 Bionomics and management of *Rhagoletis*. *Annu. 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., 1969a Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera: Tephritidae). *Evolution* **23**: 237–251.
- BUSH, G. L., 1969b 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., 1994 Sympatric speciation in animals: new wine in old bottles. *Trends Ecol. Evol.* **9**: 285–288.
- BUSH, G. L., and S. C. TAYLOR, 1969 The cytogenetics of *Procecidochares*. I. The mitotic and polytene chromosomes of the Pamakani fly, *P. utilis* (Tephritidae: Diptera). *Caryologia* **22**: 311–322.
- CACERES, M., J. M. RANZ, A. BARBADILLA, M. LONG and A. RUIZ, 1999 Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**: 415–418.
- COCKERHAM, C. C., and B. S. WEIR, 1977 Digenic descent measures for finite populations. *Genet. Res.* **30**: 121–147.
- DANILEVSKY, A. S., 1965 *Photoperiodism and Seasonal Development of Insects*. Oliver & Boyd, London.
- DEAN, R. W., and P. J. CHAPMAN, 1973 *Bionomics of the Apple Maggot in Eastern New York*, no. 3. Search Agricultural Entomology, Geneva, NY.
- DEPAULIS, F., L. BRAZIER and M. VEUILLE, 1999 Selective sweep at the *Drosophila melanogaster* *Suppressor of Hairless* locus and its association with the *In(2L)t* inversion polymorphism. *Genetics* **152**: 1017–1024.
- DOBZHANSKY, T., 1970 *Genetics of the Evolutionary Process*. Columbia University Press, New York.
- DOBZHANSKY, T., and M. L. QUEAL, 1938 Genetics of natural populations. II. Genetic variation in natural populations of *Drosophila pseudoobscura* inhabiting isolated mountain ranges. *Genetics* **23**: 463–484.
- ELENA, S. F., and R. E. LENSKI, 1997 Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**: 395–398.
- FEDER, J. L., 1995 The effects of parasitoids on sympatric host races of the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae). *Ecology* **76**: 801–813.
- FEDER, J. L., 1998 The apple maggot fly, *Rhagoletis pomonella*: Flies in the face of conventional wisdom about speciation?, pp. 130–144 in *Endless Forms: Species and Speciation*, edited by D. HOWARD and S. H. BERLOCHER. Oxford University Press, Oxford.
- FEDER, J. L., and G. L. BUSH, 1989a A field test of differential host plant usage between two sibling species *Rhagoletis pomonella* fruit flies (Diptera: Tephritidae) and its consequences for sympatric models of speciation. *Evolution* **43**: 1813–1819.
- FEDER, J. L., and G. L. BUSH, 1989b Gene frequency clines for host races of *Rhagoletis pomonella* (Diptera: Tephritidae) in the Midwestern United States. *Heredity* **63**: 245–266.
- FEDER, J. L., C. A. CHILCOTE and G. L. BUSH, 1988 Genetic differentiation between sympatric host races of *Rhagoletis pomonella*. *Nature* **336**: 61–64.
- FEDER, J. L., C. A. CHILCOTE and G. L. BUSH, 1989 Inheritance and linkage relationships of allozymes in the apple maggot fly. *J. Hered.* **80**: 277–283.
- FEDER, J. L., C. A. CHILCOTE and G. L. BUSH, 1990 The geographic pattern of genetic differentiation between host associated populations of *Rhagoletis pomonella* (Diptera: Tephritidae) in the eastern United States and Canada. *Evolution* **44**: 570–594.
- FEDER, J. L., T. A. HUNT and G. L. BUSH, 1993 The effects of climate, host plant phenology and host fidelity on the genetics of apple and hawthorn infesting races of *Rhagoletis pomonella*. *Entomol. Exp. Appl.* **69**: 117–135.
- FEDER, J. L., S. OPP, B. WLAZLO, K. REYNOLDS, W. GO *et al.*, 1994 Host fidelity is an effective pre-mating barrier between sympatric races of the apple maggot fly. *Proc. Natl. Acad. Sci. USA* **91** (17): 7990–7994.
- FEDER, J. L., J. B. ROETHELE, B. WLAZLO and S. H. BERLOCHER, 1997a Selective maintenance of allozyme differences between sympatric host races of the apple maggot fly. *Proc. Natl. Acad. Sci. USA* **94**: 11417–11421.
- FEDER, J. L., U. STOLZ, K. M. LEWIS, W. PERRY, J. B. ROETHELE *et al.*, 1997b The effects of winter length on the genetics of apple and hawthorn races of *Rhagoletis pomonella* (Diptera: Tephritidae). *Evolution* **51**: 1862–1876.
- FEDER, J. L., S. H. BERLOCHER and S. B. OPP, 1998 Sympatric host race formation and speciation in *Rhagoletis* (Diptera: Tephritidae): a tale of two species for Charles D., pp. 408–441 in *Genetic Structure in Natural Insect Populations: Effects of Host Plants and Life History*, edited by S. MOPPER and S. STRAUSS. Chapman & Hall, New York.
- FELDMAN, M. W., and F. B. CHRISTIANSEN, 1975 The effect of population subdivision on two loci without selection. *Genet. Res.* **24**: 151–162.
- FELSENSTEIN, J., 1965 The effect of linkage on directional selection. *Genetics* **52**: 349–363.
- FELSENSTEIN, J., 1981 Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35**: 124–138.
- FILCHAK, K. E., J. B. ROETHELE and J. L. FEDER, 2000 Natural selec-

- tion and sympatric divergence in the apple maggot, *Rhagoletis pomonella*. *Nature* **407**: 739–742.
- FU, Y. X., and W. H. LI, 1993 Statistical test of neutrality of mutations. *Genetics* **133**: 693–709.
- HAWTHORNE, D. J., and S. VIA, 2001 Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* **412**: 904–907.
- HEDRICK, P. W., S. JAIN and L. HOLDEN, 1978 Multilocus systems in evolution. *Evol. Biol.* **11**: 101–184.
- HEY, J., 2001 The mind of the species problem. *Trends Ecol. Evol.* **16**: 326–329.
- HUDSON, R. R., M. KREITMAN and M. AGUADÉ, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153–159.
- ISHII, K., and B. CHARLESWORTH, 1977 Associations between allozyme loci and gene arrangements due to hitch-hiking effects of new inversions. *Genet. Res.* **30**: 93–106.
- KIM, S.-C., and L. RIESEBERG, 2001 The contribution of epistasis to species differences in annual flowers. *Mol. Ecol.* **10**: 683–690.
- KOVACEVIC, M., and S. W. SCHAEFFER, 2000 Molecular population genetics of X-linked genes in *Drosophila pseudoobscura*. *Genetics* **156**: 155–172.
- KREITMAN, M., and H. AKASHI, 1995 Molecular evidence for natural selection. *Annu. Rev. Ecol. Syst.* **26**: 403–422.
- KRIMBAS, C. B., and J. R. POWELL, 1992 Introduction, pp. 1–52 in *Drosophila Inversion Polymorphism*, edited by C. B. KIMBRAS and J. R. POWELL. CRC Press, Boca Raton, FL.
- LANDER, E. S., P. GREEN, J. ABRAHAMSON, A. BARLOW, M. J. DALY *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**: 174–181.
- LEE, Y.-H., T. OTA and V. D. VACQUIER, 1995 Positive selection is a general phenomena in the evolution of abalone sperm lysin. *Mol. Biol. Evol.* **12**: 231–238.
- LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- LEWONTIN, R. C., 1997 Dobzhansky's *Genetics and the Origin of Species*: Is it still relevant? *Genetics* **147**: 351–355.
- LI, W.-H., and M. NEI, 1974 Stable linkage disequilibrium without epistasis in subdivided populations. *Theor. Popul. Biol.* **6**: 173–183.
- LINCOLN, S. E., M. J. DALY and E. S. LANDER, 1993 *Constructing Genetic Linkage Maps with MAPMAKER/EXP Version 3.0: A Tutorial and Reference Manual*, Ed. 3. Whitehead Institute for Biomedical Research Technical Report, Whitehead Institute, Cambridge, MA.
- LONG, A. D., S. L. MULLANY, L. A. REID, J. D. FRY, C. H. LANGLEY *et al.*, 1995 High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. *Genetics* **139**: 1273–1291.
- MALLET, J., 1995 A species definition for the modern synthesis. *Trends Ecol. Evol.* **10**: 294–299.
- MAXWELL, C. W., and E. C. PARSONS, 1968 The recapture of marked apple maggot adults in several orchards from one release point. *J. Econ. Entomol.* **61**: 1157–1159.
- MAYNARD-SMITH, J., 1966 Sympatric speciation. *Am. Nat.* **100**: 637–650.
- MAYR, E., 1963 *Animal Species and Evolution*. Harvard University Press, Cambridge, MA.
- METZ, E. C., and S. R. PALUMBI, 1996 Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Mol. Biol. Evol.* **13**: 397–406.
- MCDONALD, J. H., and M. KREITMAN, 1991 Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* **351**: 652–654.
- MCPHERON, B. A., and S. H. BERLOCHER, 1985 Segregation and linkage of allozymes of *Rhagoletis tabellaria*. *J. Hered.* **76**: 218–219.
- MCPHERON, B. A., D. C. SMITH and S. H. BERLOCHER, 1988 Genetic differences between *Rhagoletis pomonella* host races. *Nature* **336**: 64–66.
- NAVARRO, A., A. BARBADILLA and A. RUIZ, 2000 Effect of inversion polymorphism on the neutral nucleotide variability of linked chromosomal regions in *Drosophila*. *Genetics* **155**: 685–698.
- NEI, M., and W.-H. LI, 1973 Linkage disequilibrium in subdivided populations. *Genetics* **75**: 213–219.
- NEI, M., and W.-H. LI, 1980 Non-random association between electrophores and inversion chromosomes in finite populations. *Genet. Res.* **35**: 65–83.
- NOOR, M. A. F., K. L. GRAMS, L. A. BERTUCCI, Y. ALMENDAREZ, J. REILAND *et al.*, 2001a The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. *Evolution* **55**: 512–521.
- NOOR, M. A. F., K. L. GRAMS, L. A. BERTUCCI and J. REILAND 2001b Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* **98**: 12084–12088.
- POWELL, J. R., 1997 *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. Oxford University Press, Oxford.
- PROKOPY, R. J., E. W. BENNETT and G. L. BUSH, 1971 Mating behavior in *Rhagoletis pomonella* (Diptera: Tephritidae). I. Site of assembly. *Can. Entomol.* **103**: 1405–1409.
- PROKOPY, R. J., E. W. BENNETT and G. L. BUSH, 1972 Mating behavior in *Rhagoletis pomonella*. II. Temporal organization. *Can. Entomol.* **104**: 97–104.
- PROKOPY, R. J., S. R. DIEHL and S. S. COOLEY, 1988 Behavioral evidence for host races in *Rhagoletis pomonella* flies. *Oecologia* **76**: 138–147.
- PROUT, T., 1973 Appendix to MITTON, J. B., and R. C. KOEHN: population genetics of marine pelecypods. III. Epistasis between functionally related isoenzymes in *Mytilus edulus*. *Genetics* **73**: 487–496.
- PROVINE, W. B., 1981 Origins of the genetics of natural populations series, pp. 5–83 in *Dobzhansky's Genetics of Natural Populations*, Vols. I–XLIII, edited by R. C. LEWONTIN, J. A. MOORE, W. B. PROVINE and B. WALLACE. Columbia University Press, New York.
- RICE, W. R., 1989 Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
- RIESEBERG, L. H., 2001 Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **16**: 351–358.
- ROETHELE, J. B., J. L. FEDER, S. H. BERLOCHER, M. E. KREITMAN and D. LASHKARI, 1997 Toward the construction of a molecular genetic linkage map for the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae): a comparison of alternative strategies. *Ann. Entomol. Soc. Am.* **90**: 470–479.
- ROETHELE, J. B., J. ROMERO-SEVERSON and J. L. FEDER, 2001 Evidence for broad-scale conservation of linkage map relationships between *Rhagoletis pomonella* (Diptera: Tephritidae) and *Drosophila melanogaster* (Diptera: Drosophilidae). *Ann. Entomol. Soc. Am.* **94**: 936–947.
- ROZAS, J., and M. AGUADÉ, 1993 Transfer of genetic information in the *rp49* region of *Drosophila subobscura* between different gene arrangements. *Proc. Natl. Acad. Sci. USA* **90**: 8083–8087.
- ROZAS, J., and M. AGUADÉ, 1994 Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **91**: 11517–11521.
- ROZAS, J., C. SEGARRA, G. RIBO and M. AGUADE, 1999 Molecular population genetics of the *rp49* region in different chromosomal inversions of *Drosophila subobscura*. *Genetics* **151**: 189–202.
- SLATKIN, M., 1975 Gene flow and selection in a two-locus system. *Genetics* **81**: 787–802.
- SMITH, D. C., 1988 Heritable divergence of *Rhagoletis pomonella* host races by seasonal asynchrony. *Nature* **336**: 66–67.
- SOTO-ADAMES, F. N., H. R. ROBERTSON and S. H. BERLOCHER, 1994 Phylogenetic utility of partial DNA sequences of *G-6-pdh* at different taxonomic levels in Hexapoda with emphasis on Diptera. *Ann. Entomol. Soc. Am.* **87**: 723–736.
- STAM, P., 1993 Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. *Plant J.* **3**: 739–744.
- STURTEVANT, A. H., 1926 A crossover reducer in *Drosophila melanogaster* due to inversion of a section of the third chromosome. *Biol. Zent. Bl.* **46**: 697–702.
- TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- TAUBER, M. J., and C. A. TAUBER, 1976 Insect seasonality: diapause maintenance, termination, and postdiapause development. *Annu. Rev. Entomol.* **21**: 81–107.
- TEMPLETON, A. R., 2000 Epistasis and complex traits, pp. 41–57 in *Epistasis and the Evolutionary Process*, edited by J. B. WOLF, E. D. BRODIE III and M. J. WADE. Oxford University Press, New York.
- TING, C.-T., S.-C. TSAUR, M.-L. WU and C.-I WU, 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* **282**: 1501–1504.
- TURNER, J. R. G., 1977 Butterfly mimicry: the general evolution of

- an adaptation, pp. 163–206 in *Evolutionary Biology*, edited by M. K. HECHT, W. C. STEERE and B. WALLACE. Plenum Press, New York.
- WALSH, B. J., 1864 On phytophagous varieties and phytophagous species. *Proc. Entomol. Soc. Phila.* **3**: 403–430.
- WALSH, B. J., 1867 The apple-worm and the apple maggot. *J. Hort.* **2**: 338–343.
- WAYNE, M. L., and K. L. SIMONSEN, 1998 Statistical tests of neutrality in the age of weak selection. *Trends Ecol. Evol.* **13**: 236–240.
- WEIR, B. S., 1996 *Genetic Data Analysis II*. Sinauer Associates, Sunderland, MA.
- WESLEY, C. S., and W. F. EANES, 1994 Isolation and the analysis of the breakpoint sequences of chromosomal inversion In(3L) Payne in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **91**: 3132–3136.
- WU, C.-I., 2001 The genic view of the process of speciation. *J. Evol. Biol.* **14**: 851–865.
- ZAR, J. H., 1996 *Biostatistical Analysis*, Ed. 3. Prentice-Hall, Upper Saddle River, NJ.

Communicating editor: C.-I Wu

