Online Supporting Information Appendix

Collection of Flies for Population Survey. Flies were collected from four pairs of sympatric apple and hawthorn populations in the Midwestern U.S. during the summers of 2001 and 2002 (Table S1, Fig. S1). We also collected hawthorn flies from a site near Brazos Bend, TX in October 2005. Flies were collected from nature as larvae feeding within host fruit. Infested fruit were transported back to the lab where they were held at 21^{o} C in a temperature chamber with a 15: 9 hr. light:dark cycle. Emerging larvae leaving host fruit and forming puparia were collected on a daily basis and transferred to petri dishes containing moist vermiculite. The dishes were held for an additional 10 days at 21^{o} C before being over-wintered at 4^{o} C in a refrigerator. These conditions cause minimal selection on flies. After a 5 month overwintering period, fly pupae were removed from the cold and placed in a 21^{o} C walk-in incubator set for a 14:10 hr. light:dark cycle. Newly eclosing adults were collected on a daily basis and stored in a - 80^{o} C freezer.

Microsatellite Mapping. Evolutionary map distances for chromosomes 2 and 5 could be organized in a largely linear fashion, while those for chromosomes 1, 3, and 4 were complex, reticulate networks (Fig. S2). Despite the extensive inversion polymorphism, the microsatellites analyzed in the study appeared to be well-distributed along each of the five chromosomes surveyed. Previous crosses scored for allozymes and cDNA markers indicated overall evolutionary recombination distances of 143.2, 66.1, 50.8, and 132.5 centiMorgans for chromosomes 1-4, respectively (1) (Note: allozymes and cDNA cross data for chromosome 5 were insufficient to estimate its recombination length). In comparison, the estimates for the microsatellites in the current study generally yielded total recombination maps of 103.7, 62.7, 125.2, and 135.3 for chromosomes 1-4, respectively, and 100.4 for chromosome 5. The generally similar total recombination lengths of chromosomes were not due to us scoring larger numbers of microsatellites loci per chromosome. Indeed, fewer microsatellites were surveyed per chromosome (mean = 6.7) than in previous allozyme and cDNA analyses (mean = 11.2). The allozymes examined in the current study were NADH-diaphorase-2 (*Dia-2*), aspartate amino transferase-2 (*Aat-2*), malic enzyme (*Me*), aconitase-2 (*Acon-2*), mannose phosphate isomerase (*Mpi*), and hydroxyacid dehydrogenase (*Had*).

Outlier Analysis of Population Survey Data. We tested for F_{ST} outliers in two different ways. The first outlier analysis we used was based on the commonly employed 'fdist' method of Beaumont and Nichols (2), as implemented in LOSITAN (3). This method uses computer simulations to generate the distribution of F_{ST} values expected under neutrality, and then compares observed empirical F_{ST} values to the distribution of simulated values. Loci whose empirical F_{ST} estimates fall above the upper extreme (e.g., 95 or 99% quantile) of the simulated distribution are deemed outliers. We applied this approach by approximating the mean neutral F_{ST} in the empirical dataset using the 'mean neutral F_{ST}' and 'force mean F_{ST}' options in LOSITAN (the former removes potentially selected loci when estimating neutral F_{ST}). A total of 10,000 simulations were used in all instances. We report results from the stepwise mutation model, but very similar results were obtained using the infinite alleles model). Applying this approach to each of the four sympatric apple and hawthorn population pairs revealed three outlier loci at the 95% level (microsatellite locus P16 between apple and hawthorn populations at the Urbana, IL sympatric site, allozyme locus *Me* at Dowagiac, MI, allozyme locus *Acon-2* at the Fennville, MI site, and allozyme loci *Me* and *Acon-2* at the Grant, MI site). No outliers were detected at the 99% level for any population comparison.

The second outlier analysis we used was based on the recently developed method of Foll

and Gaggiotti (4) which directly compares models with versus without selection, thereby providing the probability that a given locus is under selection. These analyses were implemented using default settings in BayeScan. The program outputs the posterior probability of models with selection versus without selection acting on a given locus. This probability cannot be compared directly to a P-value, but is instead interpreted using the "Bayes factor", which is the $\log_{10}$ ratio of posterior model probabilities. Bayes factors >1.0, > 1.5, and > 2 are considered strong, very strong and decisive support, respectively, for the selection model over the alternative no selection model (5). The Foll and Gaggiotti method yielded results congruent with those of Beaumont and Nichols, as the same three loci (P16, *Me*, and *Acon-2*) exhibited strong evidence for outliers status (Bayes factors > 1.0) in at least one pair-wise sympatric site comparison (Fig. S3). However, only *Acon-2* exhibited strong evidence for outlier status at more than one pair-wise sympatric site comparison and 'decisive' evidence for selection (Bayes Factor > 2; Fig. S3). Additionally, linkage disequilibrium analysis for the 3 loci P16, *Me*, and *Acon-2* indicated that only two independent genomic regions defined by *Me* and *Acon*-2 (chromosome 2) and P16 (chromosome 3) exhibited outlier status.

It is conceivable that the comparatively low total number of microsatellite and allozyme loci we used in the outlier analyses may have affected our results, as genome scans commonly involve a hundred or more genes to set background F_{ST} levels expected for neutrality (3). The microsatellites may have been more sensitive to the detection of selection due to their presumably higher mutation rate and polymorphism levels than other types of loci, such as AFLPs and SNPs (5). Consequently, the background F_{ST} level could have been elevated because it reflected a number of loci under selection, making it difficult to detect outliers. However, even if including additional numbers of other classes of loci would have lowered the background F_{ST} level, this would have only served to heighten the number of microsatellites detected as outliers in the analysis. In this case, the outlier analysis would simply be more congruent with the bootstrap analysis and the selection experiment, both of which support the multifarious selection hypothesis by indicating selection acting on numerous genomic regions.

Statistical Analysis of Selection Experiment. Selection coefficients (s) were calculated for loci in the selection experiment by first determining the relative frequencies of genotypes in the 35 day non-diapausing and 7 day diapausing treatments. Under the assumption that limited selection was imposed by the 7 day diapausing treatment, the absolute fitness of a genotype was estimated as the frequency of the genotype in the 35-day treatment divided by its frequency in the 7 day treatment. Relative fitness values were then calculated by dividing by the highest absolute fitness value among genotype classes for a locus. The selection coefficient (s) was taken as the difference in relative fitness between homozygote genotypes, with the dominance term (h) estimated from the equation:

$W_{Aa} = 1\text{-}hs$ where, W_{Aa} = the relative fitness of the heterozygote genotype.

Marginal fitnesses (W_m) for the favored class of alleles in the selection experiment were calculated by first noting that 15% of pupae in the 35 day treatment eclosed as non-diapausing adults. We therefore considered the remaining 85% of pupae in the 35 day treatment as representing diapausing flies. Allele frequencies for these diapausing flies were estimated as:

$f_{d35} = (f_{d7} - (0.15 \times f_{nd35}))/0.8$ where,
f_{d7} = the allele frequency at a locus for diapausing flies eclosing in the 7 day treatment.
f_{nd35} = the allele frequency at a locus for non-diapausing flies in the 35 day treatment.

The relative marginal fitness was then calculated as:

$W_m = f_{d35} / f_{d7}$

Spearman rank correlations were then performed between W_m and F_{ST} values for loci using the R statistical package, with significance determined based on a Z-transformation test (6) and jackknife analysis across loci.

Patterns of linkage disequilibrium between loci, as determined by the composite method of Weir (7), were used to estimate the minimum number of independent genomic regions affected by selection in the diapause rearing experiment. Significant linkage disequilibrium was detected between nine responding microsatellites on chromosome 1 (P71, P37, P75), chromosome 2 (P26, P46, P73), and chromosome 3 (P7, P16, P69) in the rearing study and the corresponding allozymes that mapped to the same chromosomes (*Dia-2*, *Aat-2* [chr. 1]; *Me*, *Acon-2*, *Mpi* [chr. 2]; *Had* [chr. 3]; (Table S3). Significant linkage disequilibrium was not observed between any of the other 13 microsatellites which responded to selection in the rearing experiment or between any pair of loci (microsatellites or allozymes) mapping to different chromosomes. Thus, a minimum of 16 different loci or genomic regions (= 13 microsatellites in linkage equilibrium + 3 allozyme/microsatellite linked regions) dispersed throughout the genome responded to selection. Using a similar line of reasoning, we determined 1) that a minimum of 17 different genomic regions displayed host-related differentiation between apple and hawthorn flies, 13 of which overlap with regions responding to selection in the diapause rearing experiment, and 2) that a minimum of 12 different genes/genomic regions showed significant relationships with adult eclosion time, 10 of which also significantly responded in the selection experiment.

Supporting Information References

1. Feder JL, Roethele JB, Filchak K, Niedbalski J, Romero-Severson J (2003) Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* 163: 939-953.
2. Beaumont MA**,** Nichols RA **(**1996) Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B.* 263: 1619–1626.
3. T. Antao, A. Lopes, R.J. Lopes, A. Beja-Pereira, G. Luikart (2008) LOSITAN: A workbench to detect molecular adaptation based on a Fst -outlier method. *BMC Bioinformatics* 9: 323.
4. Foll M, Gaggiotti OA (2008) *Genetics* 180: 977-993.
5. Beaumont MA (2005) Adaptation and speciation: what can Fst tell us? *Trends Ecol Evol* 20: 435–440.
6. Sokal RR, Rolf FJ (1981) *Biometry 2nd edition* (Freeman and Co.,, San Francisco).
7. Weir B (1979) Inferences about linkage disequilibrium. *Biometrics* 35**:** 235-254.
8. Velez S, Taylor MS, Noor MAF, Lobo NF, Feder JL (2006) Isolation and characterization of microsatellite loci from the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae). *Molecular Ecology Notes* 6: 90-92.

Table S1. *Rhagoletis pomonella* collection sites in the U.S. See Fig. S1 for map of sites. Sample sizes (n) = number of apple (A) and hawthorn (H) flies, respectively, genotyped for the 33 microsatellite loci and six allozymes (in parentheses) at a site.

Site #	City	State	Latitude	Longitude	Races	Sample size (n)
1	Grant	Michigan	43°21'00.17"N	85°53'21.98"W	A, H	46 (706), 48 (467)
2	Fennville	Michigan	42°36'06.17"N	86° 09'22.19"W	A, H	96 (30), 96 (33)
3	Dowagiac	Michigan	41°52'39.36"N	86°13'49.81"W	A, H	47 (31), 48 (32)
4	Urbana	Illinois	40°07'19.46"N	88°12'09.25"W	A, H	48 (60), 46 (60)
5	Brazos Bend	Texas	29°22'15.55"N	95°35'08.16"W	H	33 (0)

Table S2. Genetic responses in the selection experiment and eclosion study, clinal patterns of latitudinal variation, and host-related differentiation. Shown are the selective differences (Sel. Dif.) in allele frequencies for loci between flies eclosing in the 7 day diapause minus the 35 day non-diapause treatments. The letter prefix "a" denotes that diapause treatment flies had a higher frequency of the allele class more common to the apple race at the Grant, MI site, "h" denotes that these flies had a higher frequency of hawthorn race alleles, "n" denotes a higher frequency of alleles more common in northern hawthorn fly populations, and "s" denotes southern hawthorn fly populations. Estimates for selection coefficients (s) and the degree of dominance (h) for heterozygotes are also given. Eclosion Time = F-ratios for ANOVA tests for the main effect of genotype (G) or genotype x host interaction (I) on adult eclosion time. The letter prefixes "a, h, n, and s" denote whether flies possessing alleles more common to the apple race (a) or hawthorn race (h) at the Grant, MI site or southern (s) vs. northern (n) hawthorn fly populations eclosed the latest. Haw Cline and Apple Cline represent the product of the slope (b) and explained variation (r^2) for linear regressions of allele frequencies vs. latitude among hawthorn and apple fly sites. * = $P \leq 0.05$, $P \leq 0.01$, *** = $P \leq 0.001$, **** $P \leq 0.0001$, ****** $P \leq 0.000001$, as determined by Fisher's exact test for Sel. Dif. and Monte Carlo bootstrapping analysis for Haw and Apple Cline values. Host Dif. indicates whether a locus displayed genetic differentiation between apple and hawthorn fly races as determined by: 1) a significant F-test for heterogeneity between regression slopes of allele frequency vs. latitude between hawthorn and apple fly populations (denoted by "b"), 2) a significant host or a significant host x latitude interaction in the Monte Carlo bootstrap ANOVA analysis (denoted by "H" and H/L", respectively), or 3) significant allele frequency differences at individual sympatric sites (denoted by "S"). No = no significant host difference in tests. Dark bars indicate loci displaying significant pairwise linkage disequilibrium.

Chromosome	Locus	Sel. Dif.	s	h	Eclosion Time	Haw Cline	Apple Cline	Host Dif.
Chrom. 1	*Aat-2*	a n 0.1130 **	0.309	0.590	1.464	+0.0515 ******	+0.0001	H, S
	Dia-2	a n 0.0917 **	0.070	4.215	a s I 3.298 *	+0.0524 ******	+0.0015	b, H/L, S
	P71	a n 0.0833 *	0.220	0.673	1.843	+0.0245 ******	+0.0480 ******	H, S
	P37	a n 0.1202 **	0.274	0.536	1.151	+0.0469 ******	+0.0730 ******	H, S
	P75	a n 0.0795 *	0.121	1.373	h s G 4.236 *	+0.0507 ******	+0.0408 ******	No
	P4	h s 0.0090	0.069	-0.541	0.778	+0.0164 ****	+0.0187	No
	P64	a s 0.1304 *	0.099	0.550	0.238	+0.0206 ***	+0.0061	H
	P3	a s 0.0795 *	0.101	0.518	0.502	+0.0087 ***	- 0.0590 ******	b, H/L, S
Chrom. 2	P70	h s 0.1230 *	0.083	0.571	0.408	+0.0269 ******	+0.0733 ******	No
	Mpi	a s 0.0460	0.155	0.519	a s G 6.880 **	+0.0037 *	+0.0008	H/L, S
	P59	a s 0.0956	0.115	0.242	a s G 3.310 *	+0.0254 ******	+0.0119	b
	Acon-2	a s 0.1206 *	0.159	0.518	a s G 5.352 **	+0.0184 ******	+0.0014	b, H/L, S
	Me	a s 0.1747 ***	0.253	0.230	a s G 3.963 *	+0.0205 ******	+0.0089	H/L, S
	P26	a s 0.1840 ***	0.255	0.204	0.475	+0.0168 ******	- 0.0861 ******	b, H/L, S
	P46	h n 0.1497 **	0.128	0.579	1.168	+0.0333 ******	- 0.0297 **	b
	P73	a s 0.1255 *	0.109	0.657	1.104	+0.0034 *	- 0.0073	No

Table S2 (cont.)

Chromosome	Locus	Sel. Dif.	s	h	Eclosion Time	Haw Cline	Apple Cline	Host Dif.
Chrom. 2	P32	a n 0.0253	0.049	-0.079	0.392	+0.0088 ******	+0.0670 ******	b
	P54	h n 0.1104 *	0.534	0.067	2.001	+0.0199 ******	- 0.0247 **	b
Chrom. 3	P69	a s 0.1680 ***	0.145	0.573	a s G 5.305 **	+0.0227 ******	+0.0370 **	S
	Had	a s 0.0649	0.091	0.219	a s I 3.138 *	+0.0517 ****	+0.0445 ****	H/L, S
	P7	a s 0.1345 **	0.152	0.389	0.302	+0.0320 ******	- 0.0215 *	b, H/L, S
	P16	a s 0.0789 *	0.061	-0.965	a s G 3.365 *	+0.0296 ******	- 0.0014	H/L, S
	P68	a s 0.0390	0.030	-1.128	a s G 3.099 *	+0.0394 ******	- 0.0027	b, H/L, S
	P80	a s 0.0649	0.091	0.219	0.247	+0.0127 ****	- 0.0312 **	b, H/L, S
	P23	a n 0.0656	0.141	-0.029	0.904	+0.0276 ******	+0.0734 ******	b
	P66	a s 0.1215 *	0.121	0.465	0.341	+0.0129 ******	+0.0052	b, H, S
Chrom. 4	P11	a s 0.1041 *	0.127	0.646	h n I 4.416 *	+0.0121 ******	- 0.0191 **	b
	P29	h n 0.0948	0.101	0.565	0.207	+0.0140 ******	- 0.0011	H, S
	P25	a s 0.1366 **	0.120	0.515	a s I 3.049 *	+0.0071 *	+0.0291 **	H, S
	P50	a s 0.1046 *	0.126	0.478	a s I 3.163 *	+0.0058 **	- 0.0303 ******	b

Table S2 (cont.)

Chromosome	Locus	Sel. Dif.	s	h	Eclosion Time	Haw Cline	Apple Cline	Host Dif.
Chrom. 4	P19	a n 0.0800 *	0.117	0.737	h n I 5.651 **	+0.0098 ****	+0.0084	No
	P12	h n 0.0108	0.039	0.051	a s G 3.957 *	+0.0064 **	- 0.0456 ******	H, S
Chrom. 5	P5	h n 0.1048 *	0.103	0.347	h n G 9.295 ***	+0.0207 ******	+0.0001	S
	P18	a s 0.0988 **	0.606	0.241	h n I 3.056 *	+0.0091 ****	- 0.0416 ******	b, H/L, S
	P9	h n 0.0050	0.020	2.491	a s G 3.635 *	+0.0065 **	+0.0001	No
	P72	h s 0.1289 **	0.151	0.683	h s I 3.303 *	+0.0049 **	+0.0151 **	No
	P55	h n 0.0628	0.060	1.072	0.483	+0.0094 ***	- 0.0361 ******	b
	P45	a s 0.0862 *	0.146	1.289	1.252	+0.0419 ******	- 0.0074	b, H/L, S
	P27	h n 0.0520	0.080	0.310	a s G 3.729 *	+0.0105 ******	- 0.0087 *	b

Table S3. Composite linkage disequilibrium values as determine by the method of Weir (7) between indicated pairs of microsatellite loci and/or allozymes. Haw Pop. = combined disequilibrium value for Urbana, IL, Dowagiac, MI , Fennville, MI and Grant, MI hawthorn fly populations estimated and tested by meta-analysis of z-transformed values for individual sites. Apple Pop. = combined disequilibrium value for Urbana, IL, Dowagiac, MI , Fennville, MI and Grant, MI apple fly populations estimated and tested by meta-analysis of z-transformed values for individual sites. Haw Diap. = composite linkage disequilibrium value for hawthorn flies from Grant, MI exposed to diapause rearing conditions of a 7 day prewinter period at 25^{o}C and 30 week overwinter period of 5^{o}C. Haw ND = composite linkage disequilibrium value for hawthorn flies from Grant, MI undergoing non-diapause development and eclosing after 35 days when reared continuously at 25^{o}C. No significant disequilibrium was detected between loci mapping to different chromosomes. * = P $\leq$ 0.05, P $\leq$ 0.01, *** = P $\leq$ 0.001, **** P$\leq$ 0.0001

Chromosome	**Loci**	**Haw Pops.**	**Apple Pops.**	**Haw Diap.**	**Haw ND**
Chrom. 1	P71/P37	+0.502 ****	+0.707****	+0.413 ***	+0.568 ****
	P71/P75	+0.642 ****	+0.801****	+0.564 ****	+0.686 ****
	P37/P75	+0.543 ****	+0.569****	+0.451 ****	+0.685 ****
	P71/*Dia-2*	---	---	+0.709 ****	+0.522 ****
	P71/ *Aat-2*	---	---	+0.541 ****	+0.435 ****
	P37/*Dia-2*	---	---	+0.486 ****	+0.532 ****
	P37/ *Aat-2*	---	---	+0.348 **	+0.529 ****
	P75/*Dia-2*	---	---	+0.757 ****	+0.766 ****
	P75/ *Aat-2*	---	---	+0.658 ****	+0.622 ****
	Dia-2/ *Aat-2*	---	---	+0.692 ****	+0.574 ****

Table S3 (cont.)

Chromosome	Loci	Haw Pops.	Apple Pops.	Haw Diap.	Haw ND
Chrom. 2	P59/P73	+0.233 **	- 0.077	+0.007	+0.105
	P26/P73	+0.037	+0.025	+0.177	+0.210 *
	P59/*Me*	---	---	+0.226 *	+0.105
	P59/ *Acon-2*	---	---	+0.259 *	+0.135
	P26/*Me*	---	---	+0.277 **	+0.121
	P46/*Me*	---	---	+0.226 *	+0.121
	P46/ *Acon-2*	---	---	+0.273 *	+0.017
	P73/*Me*	---	---	+0.236 *	- 0.100
	P73/*Mpi*	---	---	+0.257 *	- 0.048
	Me/*Acon-2*	---	---	+0.435 ****	+0.575 ****
	Me/*Mpi*	---	---	+0.181	+0.221 **
	Acon-2/*Mpi*	---	---	+0.134	+0.285 ***
Chrom. 3	P69/P7	+0.108	+0.220 *	+0.068	- 0.004
	P7/P16	+0.195 **	+0.372 ****	+0.145	+0.309 ***
	P7/P68	+0.340 ****	+0.356 ****	- 0.036	- 0.013
	P7/P80	+0.111	+0.306 ****	+0.235 *	- 0.019
	P7/P23	+0.158 *	+0.156 *	- 0.030	+0.031
	P16/P68	+0.256 **	+0.409 ****	+0.401 ***	+0.106
	P16/P23	- 0.091	+0.248 **	+0.243	+0.112
	P68/P80	+0.272 ****	+0.282 ***	+0.045	- 0.002
	P68/P23	+0.315 ****	+0.277 ****	+0.327 *	- 0.027

Table S3 (cont.)

Chromosome	Loci	Haw Pops.	Apple Pops.	Haw Diap.	Haw ND
Chrom. 3	P80/P23	+0.237 **	+0.138 *	+0.167	+0.059
	P7/ *Had*	---	---	+0.379 ***	+0.285 ***
	P16/ *Had*	---	---	+0.405 ***	+0.415 ****
	P68/ *Had*	---	---	+0.504 ****	+0.014
	P80/ *Had*	---	---	+0.465 ****	+0.076
	P23/ *Had*	---	---	+0.287 *	+0.195 **
Chrom. 4	P29/P11	+0.146 *	+0.051	- 0.081	- 0.064
	P29/P25	- 0.197 **	+0.010	- 0.103	- 0.228 *
Chrom. 5	P18/P9	- 0.039	- 0.289 ***	- 0.300 **	- 0.222 *
	P72/P55	+0.149 *	+0.076	+0.232 *	- 0.045

Table S4. Linear regressions of microsatellite and allozyme genotypes vs. adult eclosion times for apple and hawthorn flies in the 7 day prewinter, 5 month overwinter treatment in the eclosion study. Given are the explained variation (r^2) and slopes (b = days difference in eclosion time for substitution of a northern allele in the genotype of a locus) for the regression equations for the 20 loci displaying significant relationships with eclosion time, as determined by ANOVA (see Table S2). Dark bars indicate loci displaying significant pairwise linkage disequilibrium.

Chromosome	Locus	r2 Apple	b Apple	r2 Haw	b Haw
Chrom. 1	*Dia-2*	0.002	+0.53	0.024	- 2.63
	P75	0.045	- 3.48	0.012	- 1.72
Chrom. 2	*Mpi*	0.129	- 5.29	0.001	- 0.40
	P59	0.039	- 2.12	0.018	- 1.21
	Acon-2	0.025	- 2.00	0.081	- 2.83
	Me	0.039	- 4.84	0.061	- 2.60
Chrom. 3	P69	0.077	- 3.17	0.023	- 1.35
	Had	0.016	+1.80	0.058	- 3.18
	P16	0.009	- 1.24	0.020	- 1.55
	P68	0.002	- 0.62	0.086	- 3.02
Chrom. 4	P11	0.049	- 2.90	0.045	+2.08
	P25	0.008	+0.90	0.013	- 0.97
	P50	0.016	+1.56	0.050	- 2.21
	P19	0.093	- 6.28	0.062	+2.85
	P12	0.054	- 2.92	0.006	- 0.91

Table S4 (cont.)

Chromosome	Locus	r2 Apple	b Apple	r2 Haw	b Haw
Chrom. 5	P5	0.010	+0.99	0.022	+1.27
	P18	0.034	- 2.81	0.048	+3.30
	P9	0.014	- 1.49	0.072	- 3.03
	P72	0.006	+0.88	0.036	- 1.89
	P27	0.026	- 1.97	0.016	- 1.29

Table S5. List of *Rhagoletis pomonella* microsatellite primer pairs (designated by "P + number") developed by Velez *et al*. (8).

Genbank accession No.	Primer sequences (5' to 3')	Expected allele size	Repeat motif	Genbank accession No.	Primer sequence (5' to 3')	Expected allele size	Repeat motif
P1) AY734885	GGAAACGACATCCGGTAAAA ACGGGCTCACAAACGAAATA	248	$[T_3(GT)_n]_n$	P20) AY734904	TGCGTCTATTTTTCCATCACTG CGCGATTGAACACAACAATC	299	$(GT)_n$
P2) AY734886	TCCACTCAAATACGGCAACA AGAGATCCCGGTGTCGTTC	228	$(GT)_n$	P21) AY734905	CAAGTGCGTGTCTGCGTAAG CCATTGCCAATTTGAATCAC	229	$(GT)_nT_3$
P3) AY734887	TCCACTCAAATACGGCAACA GCAGCCGATCTTTTCGTCTA	188	$(GT)_n$	P22) AY734906	GGGGCAATTGACACTTCCTA TTCGCCTGGTAACAAAAACC	226	$(GT)_n$
P4) AY734888	GCAAGCGAGTCGTAATCACA CCCTCATCATTGTGGTCCTC	185	$(GT)_n(G_3T_3)_n$	P23) AY734907	AAACTGCCTTGCCTGTCATT GCACTTTGTCGTTGATGCAC	210	$(GT)_n$
P5) AY734889	GAGCAGCAGAGGAAAAAGGA TGCACTGGTGTATTCCAAGG	227	$G_3(GT)_nG_3$	P24) AY734908	TCAACACACGCACAACCTTT TGGTTGCTTTTTGCATTTTG	248	$(GT)_nA(GT)_n$
P6) AY734890	AGTCAGAGTGCGGCAAAAGT CGGTAGACCTCAGGCTGATAG	162	$(GT)_n$	P25) AY734909	ATGACATTCGCTACGGGGTA TCTCGGAGAGTGGCAGTTTT	223	$(GT)_n$
P7) AY734891	CATTGGCAACGCTAGTTCAA GCGCTGAAACCATGAAAAAT	236	$(GT)_n$	P26) AY734910	GCCGTCGGACTGTTATGAAT TGCAATGTCAACAGCAATGA	249	$(GT)_n$
P8) AY734892	GCGATTTTTGAAAGCGAAAG ATTCACTAGTCCGCGGTTTC	220	$T_3(GT)_nT_3$	P27) AY734911	TTCTCACATTTTCGCGTTTG CTGGCCAATGCATAAATCCT	158	$(GT)_n$
P9) AY734893	CGGCAGGTAAATGACCAAAA GCAATGACCGTTGGCTATTA	159	$(GT)_n$	P28) AY734912	ACTGGCTGATGATGATGCAA TTATGTTGCATTCGGAGCAC	226	$(G3T)_n(GT)_n$
P10) AY734894	TAACAGCAGCCGTAATGCAG CCATTAACCAACCGCCATAC	175	$(GT)_n$	P29) AY734913	TCCATGTGTGCCAGAACATT GACGTTATTTCGCTCGGTTG	203	$(GT)_n$
P11) AY734895	ATGCAGCCATGACTGAGATG TGGAAAGTAATTTCACAAAGGCTA	304	$(GT)_n$	P30) AY734914	CGATCGCAGCAAAGTACAGG GATATGTGCACCAATGGGACT	151	$(GT)_n$
P12) AY734896	GGGTGTTCATGGTAGTTGTAGAT ACTAGTAAAGGAAAGGCGCAAT	250	$T_3(GT)_nG_3(GT)_n$	P31) AY734915	TGGGTGCCACTTCCTTTAAC ATCAAAAACGCGCAAAAACT	162	$(GT)_n$
P13) AY734897	GACATCAACTGGTGGTACGC ACCAGCCACCGATCATATTT	368	$(GT)_n$	P32) AY734916	CAGTGCCAAGTGAAGCGTAA TGACACTGTCTGGCATTTCC	247	$(GT)_n$
P14) AY734898	TTGCTAGCAGAAGCACTTCAA CCTGTGTGCAAGGACTTCAA	215	$(GT)_n$	P33) AY734917	ATAATGCCAACGCAAGGAAC TTTCGGAAAGGTGGAAAACA	192	$(GT)_n$
P15) AY734899	CGCGAGAATTTAGTTGAGCA TGCCAAGAAGTGTTGTTTCC	286	$(GT)_n$	P34) AY734918	AGGAAACCAAAAAGCAAGCA TACACGCAGAGGCTATGTGG	239	$(GT)_n$
P16) AY734900	CGCTTTAGATTTTCGCTACACA ACGCAGTGCCAAATCTTCTT	314	$(GT)_n$	P35) AY734919	CCGCTACAGTGCAATACACAA ACGGCTAATAGTTTTCACACG	220	$(GT)_n$
P17) AY734901	TTCGAAACCGTTTGTTACCTCT CGCTATTGGAGGCAATGAAT	258	$(GT)_n$	P36) AY734920	TACCACTTTTCGCCGATTTT AAACAGCCGGATTCAATGTC	246	$(GT)_n$
P18) AY734902	CCCAATGTCCCGTAAACTTC TTCACTCAATGCCCATTTCA	291	$(GT)_n$	P37) AY734921	CAACAGCGCGACTTAGTGAA TGGCTTCCACCTTTGTTTTT	214	$(GT)_n$
P19) AY734903	TCTGCCTTTGCTTCTCCATT TGGTGAAAAATTCCAGATCACA	343	$(GT)_n$	P38) AY734922	TTGGACGGACAAACATGAAA TTGCTAGCAGAAGCCTGTCA	234	$(GT)_n$

Table S5 (cont.)

Genbank accession No.	Primer sequences (5' to 3')	Expected allele size	Repeat motif	Genbank accession No.	Primer sequence (5' to 3')	Expected allele size	Repeat motif
P39) AY734923	GCGAAAATGTGGTCGTAGGT ACAGTGCGGCTGACACATAG	217	$(GT)_n$	P58) AY734942	ATGTACCTATTAATGGAGTTTTTCG TGCCGTTATACTCAGTGCATGT	172	$(GT)_nG_3(GT)_n$
P40) AY734924	CGAGCAGGTGATGATAATGC CCGAAATTTGAGCCCCTTA	234	$(GT)_n$	P59) AY734943	CGCGTCCAACTAAGAAGTCG CACTCTTCCGTTGCTTGTCA	214	$(GT)_n$
P41) AY734925	TGCTGCTGCTGTTCTGTTCT TTGTGTGTTCGCTCAAAGAAA	185	$(GT)_n$	P60) AY734944	TACAACCTAGGCAGCCCAAC GTCTGGTTTGGCGATCACTT	152	$(GT)_n$
P42) AY734926	ACAGCAACAGCAGCAACAAC AACAGCCATGCCAAAAGAGT	168	$(GT)_n$	P61) AY734945	TCTTTGGCGCTTTATTCGTT CAGCGCTGCAATAAAAGTCA	221	$(GT)_n[G_3(GT)_2]_n$
P43) AY734927	TTGTGTACGGCGCTGAGTTA TCTGCGCTACCCAAACTACG	151	$(GT)_n$	P62) AY734946	TGAAGAACGCCCTATCTCAAA TGCTGATACAAAACGCCCTA	150	$(GT)_n$
P44) AY734928	AGAACGCCTCGTAAAAAGCA GGCTTTTTGTTGCTCCATTC	156	$(GT)_n$	P63) AY734947	CAAGACGCCTACGTGTCAGA GCCACTCAAATGCAGACATC	195	$(GT)_n$
P45) AY734929	GCGCAAACTCCTCAAACTCT GTGTCTGGCGATAGCATTCA	178	$(GT)_n$	P64) AY734948	CCTCCTTTTGACTGCCTTTT GCCACTAGTGCACTTGACCTC	158	$(GT)_n$
P46) AY734930	GCGCATTTCTCCATTCATTT GCGGTAATTGTGCGTATGTG	246	$(GT)_n$	P65) AY734949	CCTTTTGTAAAATCACGTTTTCAT CGGAACAAATGAAGAAAGCA	131	$(GT)_n$
P47) AY734931	TGGTCAGAGAACACTTGCTGA TGTGCAAGGACTTCAATGGT	183	$(GT)_n$	P66) AY734950	GCAAACCATTTTCCACGAAT CGAAGCATGAATGCAACAAC	235	$(GT)_n$
P48) AY734932	TTACCCACAATCTGGCATCA AACTCCAGTGAATTATGGTTGGA	158	$(GT)_n$	P67) AY734951	TGCTGCATAATGCTCCTCAG TTCACCACCCGCAATACATA	186	$(GT)_n$
P49) AY734933	TATTTTCGGTTCGGACTGCT GGCATTGACTGCATTTCTCA	166	$(GT)_n$	P68) AY734952	CTTGCCATTGTCGACACCTA GTGGCGGACAATTTTACTGG	164	$(GT)_n$
P50) AY734934	GTGCAACCAGTGAGCAGTGT TCTGACTGGCCCGTATTTGT	160	$(GT)_n$	P69) AY734953	CGACCAATAACAAAGGTAGAAGG AGTTAGCGCTTGTGGATGCT	181	$(GT)_n$
P51) AY734935	TGCTGGGTGTCAACAACTTT TGAAACAATAAAAAGCACTACGTGA	162	$(GT)_n$	P70) AY734954	CAGCCTGCCAACACCATT GCAACGCCTTCAAATTCATC	200	$(GT)_n$
P52) AY734936	ATAGTTGGACGCGCTTGACT AAGAAATGTTGGGACGGTTG	244	$(GT)_n$	P71) AY734955	CGCAAGCACTTTTTGAACTG CTGCTGAATTGGCAGCATAA	181	$(GT)_n$
P53) AY734937	ATTTGTGCGCTAGTGTGCTG GCGGTGCAAAATAAACACAA	160	$(GT)_n$	P72) AY734956	CATTGCGATTTTCCACACAC TGTGCGTTAGAGTGCTGCTT	194	$(GT)_n$
P54) AY734938	TGTGCTAAATTACCCAAAAGCA GCGTCATTCAGTCAACCAAA	225	$(GT)_n$	P73) AY734957	TTTTCTCGTCTACTCGTGTTAGTTAAT AAAATGCACTTTGTAAATAGTCACTCA	158	$(GT)_n$
P55) AY734939	GGCTATTGAAATCACGGCTA AGCGAAATTGGCGTAAACAA	167	$(GT)_n$	P74) AY734958	ACAGCCGCTCTCTGACTCTC AAGCGCTTGTAGGGGAATATAG	216	$(GT)_n$
P56) AY734940	AGGTTCAATGCCAATGTCGT GCATGCAGGGCGTATCTAAA	249	$(GT)_n$	P75) AY734959	GCCGACTGTCGATTCTCTTG GGCAGTGATGACGAGAAACA	206	$(GT)_n$
P57) AY734941	TTTTTCAAATTGCGGTTTCC TTGTCGATGATTTATTGCAAATG	218	$(GT)G_3(GT)_n$	P76) AY734960	TTGAGTTTCAACGCCATTTG GGGGCGTACTGAATGAGATG	215	$(GT)_n$

Table S5 (cont.)

Genbank accession No.	Primer sequences (5' to 3')	Expected allele size	Repeat motif	Genbank accession No.	Primer sequence (5' to 3')	Expected allele size	Repeat motif
P77) AY734961	GGCTTTTGTTATCATCCAGCA CAACAGCAGCACATCCACTC	198	$(GT)_n$	P80) AY734964	GGACAGTTGTGGTTGCTGAA TCCTTTGCAATGTTATGGTAATTG	205	$(GT)_n$
P78) AY734962	CGCTTGTTGTGCAATTCTGT TGCCAGTTGCTCAGAGAAGA	236	$(GT)_n$	P81) AY734965	GAAGAACAGGGGGCTAATCC TTTTTGCTTCCTCGGGTTTT	226	$(GT)_n$
P79) AY734963	TTTTGGATCACCATAATGTGC GATGCCAAACAAATCTGCTG	198	$(GT)_n$				

Supporting Information Figure Legends

Fig. S1. *Rhagoletis pomonella* collecting sites genotyped in the study. Also shown are the ranges of the apple and hawthorn fly races in the eastern Unites States. A = apple, H = hawthorn population. See Table S1 for site descriptions.

Fig. S2. Linkage relationships of microsatellites on chromosomes 1-5 depicted as networks of evolutionary map distances (in centiMorgans) between loci based on data from seven test crosses. Widths of bolded lines demarcate relative degree to which microsatellites show linkage disequilibrium in natural populations and the selection experiment. Asterisks in panel B) indicate that P59, P26, P46, and P73 also display linkage disequilibrium, but not in all pairwise combinations. The locations of allozymes are not shown because they were not mapped in the crosses. However, the allozymes *Dia-2* and *Aat-2* (chrom. 1); *Me*, *Acon-2*, and *Mpi* (chrom. 2); and *Had* (chrom. 3) were all in linkage disequilibrium with certain of the microsatellites on the same chromosome displaying disequilibrium (see Table S3 for all significant pairwise composite disequilibrium values among microsatellites and allozymes). Numbered microsatellites in parentheses below networks give sets of loci showing no recombination in a test cross, implying an inversion polymorphism in the female parent.

Fig. S3. Results of outlier analyses using Bayescan (4), where separate analyses were run for each of the four sympatric apple and hawthorn population pairs ('comparison'). Three loci (P16, *Me* and *Acon-2*) were detected as statistical outliers (Bayes Factor > 1 being indicative of strong support for a model with selection on a particular locus over a model without selection on that locus). Similar results were obtained using a different analytical method, namely the 'fdist' method of Beaumont and Nichols (2), as implemented in LOSITAN (3).

Fig. S4. Microsatellite allele frequencies for apple (open circles) and hawthorn (filled triangles) fly populations plotted against latitude for chromosome 1. Also given are r^2 values (explained variation) for linear regressions between allele frequencies for a locus among hawthorn and apple sites and latitude, as well as whether the locus displayed significant host-related differentiation as determined by the Monte-Carlo bootstrapping ANOVA analysis (designated by ANOVA Host or Host/Latitude Interaction), a F-test for heterogeneity in slopes between the linear regressions for the host races (designated by Slope), and Fisher exact tests for allele frequencies between hawthorn and apple populations at individual sympatric sites (designated by asterisks above hawthorn population triangle symbols). Alleles provide the total number of variants segregating for a locus and the size (in base pairs) of the different alleles for the pooled allele class plotted in the figure. * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$, ***** = $P \leq 0.00001$, ****** = $P \leq 0.000001$.

Fig. S5. Microsatellite allele frequencies for apple (open circles) and hawthorn (filled triangles) fly populations plotted against latitude for chromosome 2. Also given are r^2 values (explained variation) for linear regressions between allele frequencies for a locus among hawthorn and apple sites and latitude, as well as whether the locus displayed significant host-related differentiation as determined by the Monte-Carlo bootstrapping ANOVA analysis (designated by ANOVA Host or Host/Latitude Interaction), a F-test for heterogeneity in slopes between the linear regressions for the host races (designated by Slope), and Fisher exact tests for allele frequencies between hawthorn and apple populations at individual sympatric sites (designated by asterisks above hawthorn population triangle symbols). Alleles provide the total number of variants segregating for a locus and the size (in base pairs) of the different alleles for the pooled allele class plotted in the figure. * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$, ***** = $P \leq 0.00001$, ****** = $P \leq 0.000001$.

Fig. S6. Microsatellite allele frequencies for apple (open circles) and hawthorn (filled triangles) fly populations plotted against latitude for chromosome 3. Also given are r^2 values (explained

variation) for linear regressions between allele frequencies for a locus among hawthorn and apple sites and latitude, as well as whether the locus displayed significant host-related differentiation as determined by the Monte-Carlo bootstrapping ANOVA analysis (designated by ANOVA Host or Host/Latitude Interaction), F-test for heterogeneity in slopes between the linear regressions for the host races (designated by Slope), and Fisher exact tests for allele frequencies between hawthorn and apple populations at individual sympatric sites (designated by asterisks above hawthorn population triangle symbols). Alleles provide the total number of variants segregating for a locus and the size (in base pairs) of the different alleles for the pooled allele class plotted in the figure. * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$, ***** = $P \leq 0.00001$, ****** = $P \leq 0.000001$.

Fig. S7. Microsatellite allele frequencies for apple (open circles) and hawthorn (filled triangles) fly populations plotted against latitude for chromosome 4. Also given are r^2 values (explained variation) for linear regressions between allele frequencies for a locus among hawthorn and apple sites and latitude, as well as whether the locus displayed significant host-related differentiation as determined by the Monte-Carlo bootstrapping ANOVA analysis (designated by ANOVA Host or Host/Latitude Interaction), F-test for heterogeneity in slopes between the linear regressions for the host races (designated by Slope), and Fisher exact tests for allele frequencies between hawthorn and apple populations at individual sympatric sites (designated by asterisks above hawthorn population triangle symbols). Alleles provide the total number of variants segregating for a locus and the size (in base pairs) of the different alleles for the pooled allele class plotted in the figure. * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$, ***** = $P \leq 0.00001$, ****** = $P \leq 0.000001$.

Fig. S8. Microsatellite allele frequencies for apple (open circles) and hawthorn (filled triangles) fly populations plotted against latitude for chromosome 5. Also given are r^2 values (explained variation) for linear regressions between allele frequencies for a locus among hawthorn and apple sites and latitude, as well as whether the locus displayed significant host-related differentiation as determined by the Monte-Carlo bootstrapping ANOVA analysis (designated by ANOVA Host or Host/Latitude Interaction), F-test for heterogeneity in slopes between the linear regressions for the host races (designated by Slope), and Fisher exact tests for allele frequencies between hawthorn and apple populations at individual sympatric sites (designated by asterisks above hawthorn population triangle symbols). Alleles provide the total number of variants segregating for a locus and the size (in base pairs) of the different alleles for the pooled allele class plotted in the figure. * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$, ***** = $P \leq 0.00001$, ****** = $P \leq 0.000001$.

Fig. S9. Genetic responses for loci on chromosomes 1-5 in the selection experiment. Shown are allele frequency differences between flies eclosing in the 7 day diapause minus 35 day non-diapause treatment. Sign of the response indicates if flies in the diapause treatment had higher frequencies of alleles more common to the apple race at the Grant, MI site (+) or hawthorn race (-). Of the 26 loci displaying significant allele frequency responses in the selection experiment, 21 shifted in a direction favoring apple race alleles at the Grant, MI site in surviving diapausing flies. Bars indicate sets of loci showing significant linkage disequilibrium among themselves in the study, as determined by the composite method of Weir (7). * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$ for significance level of allele frequency difference between diapause vs. non-diapause rearing treatments, as determined by Fisher exact tests.

Fig. S10. Spearman rank correlation (r) between mean F_{ST} values for sympatric hawthorn and apple fly populations for microsatellite and allozyme markers (a metric for the degree of genetic differentiation in nature) vs. the magnitude of the genetic response of loci in the selection experiment, as measured by marginal fitness value estimates for the favored allele class at a locus (i.e., the allele class having higher frequency in the 7 day diapause treatment).

Fig. S1

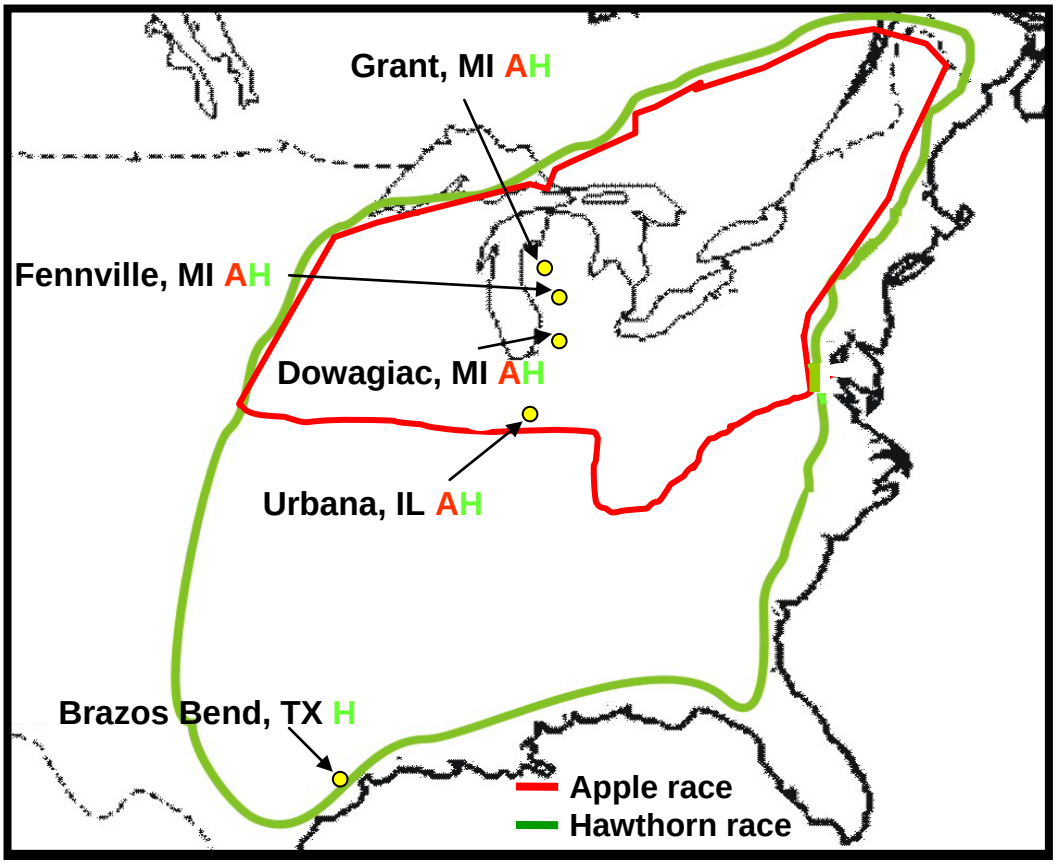
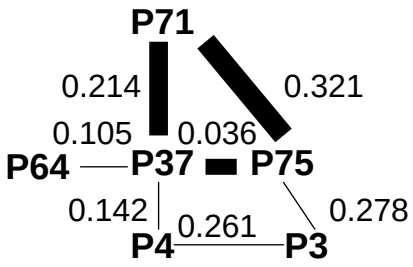

Grant, MI AH
Fennville, MI AH
Dowagiac, MI AH
Urbana, IL AH
Brazos Bend, TX H
Apple race
Hawthorn race

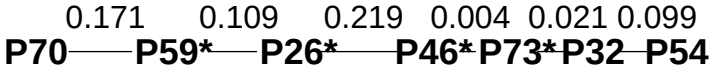
Fig. S2

A) Chromosome 1 (7 crosses)



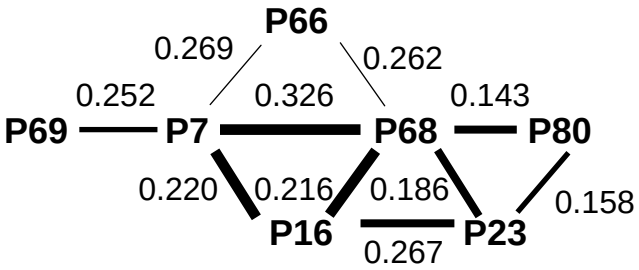
1. (P71-P37-P75)

B) Chromosome 2 (7 crosses)



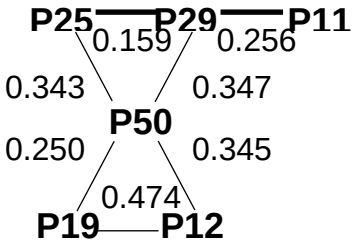
1. (P46-P73) 2. (P70-P59-P26) 3. (P46-P73-P32)
4. (P70-P59-P26-P46-P73-P32)

C) Chromosome 3 (7 crosses)

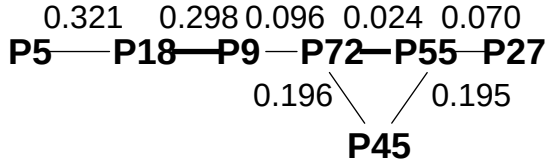


1. (P23-P68-P80) 2. (69-7-66-23-80)

E) Chromosome 4 (7 crosses)



F) Chromosome 5 (7 crosses)



1. (P72-P45-P55-P27) 2. (P9-P55) 3. (P72-P55)

Fig. S3

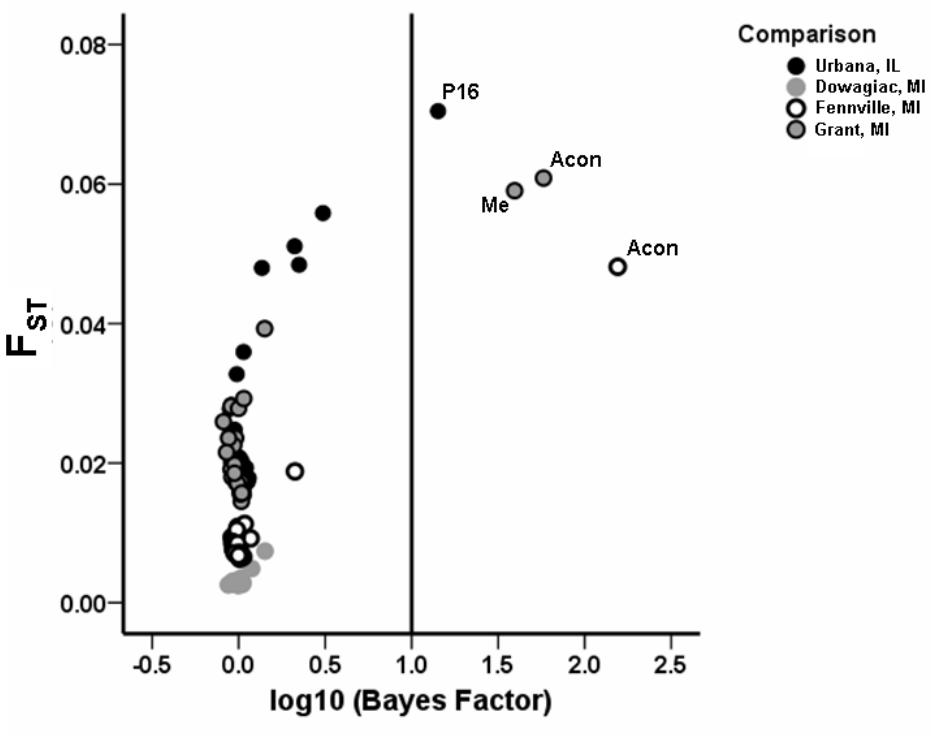

Comparison
Urbana, IL
Dowagiac, MI
Fennville, MI
Grant, MI
P16
Acon
Me
Acon
F_{ST}
0.08
0.06
0.04
0.02
0.00
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
log10 (Bayes Factor)

Fig. S4

Chromosome 1

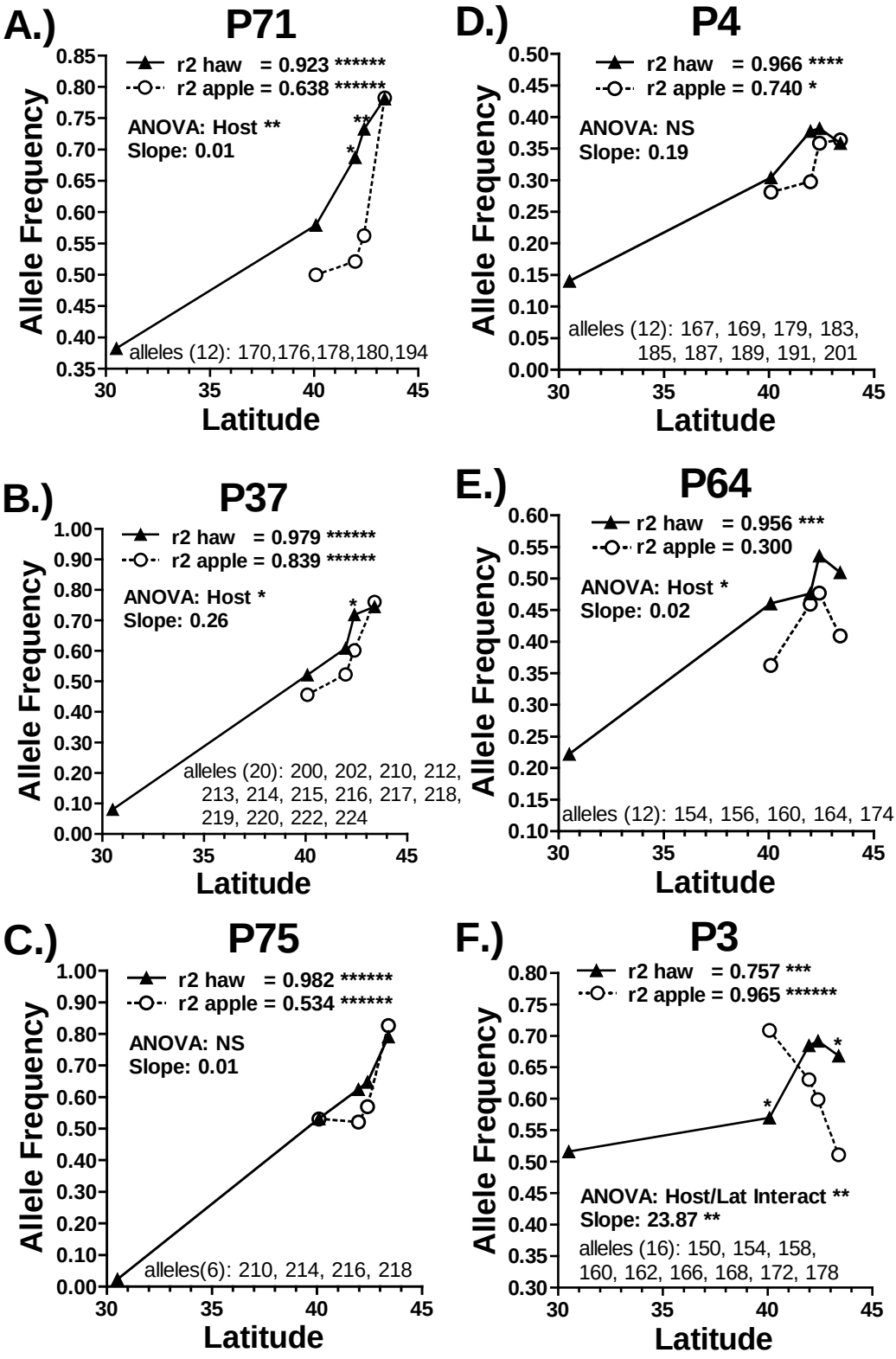


Fig. S5

Chromosome 2

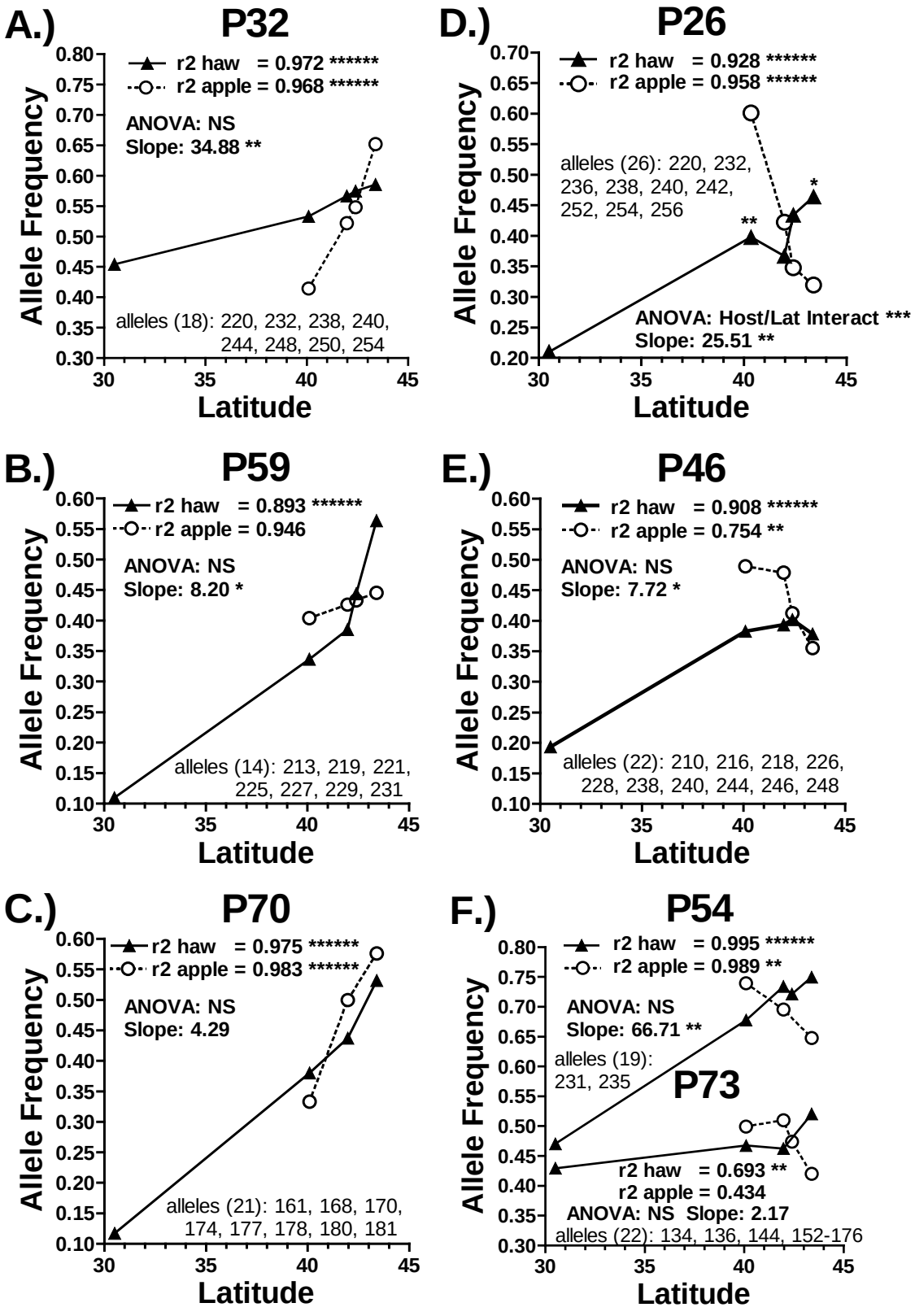


Fig. S6 **Chromosome 3**

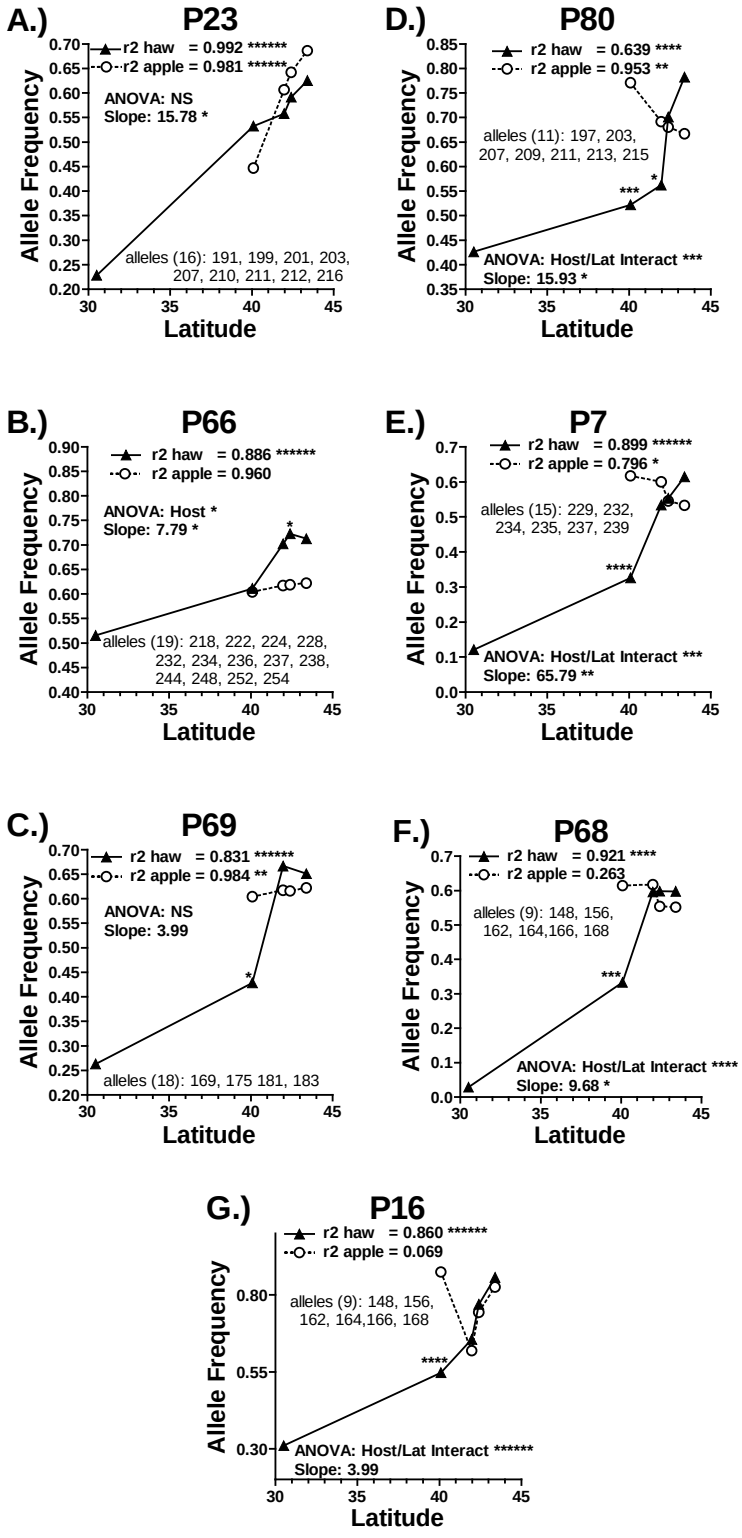


Fig. S7

Chromosome 4

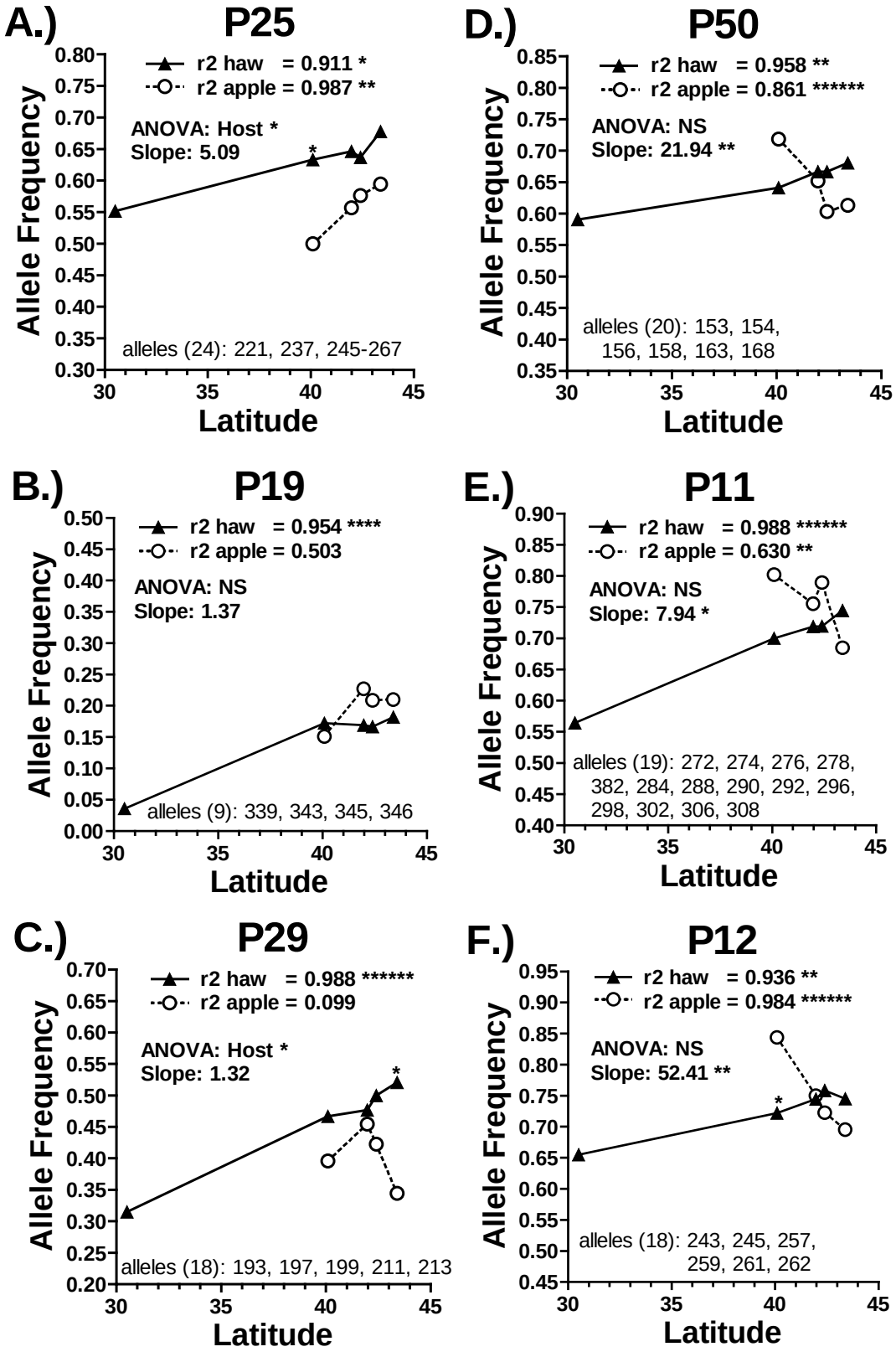


Fig. S8

Chromosome 5

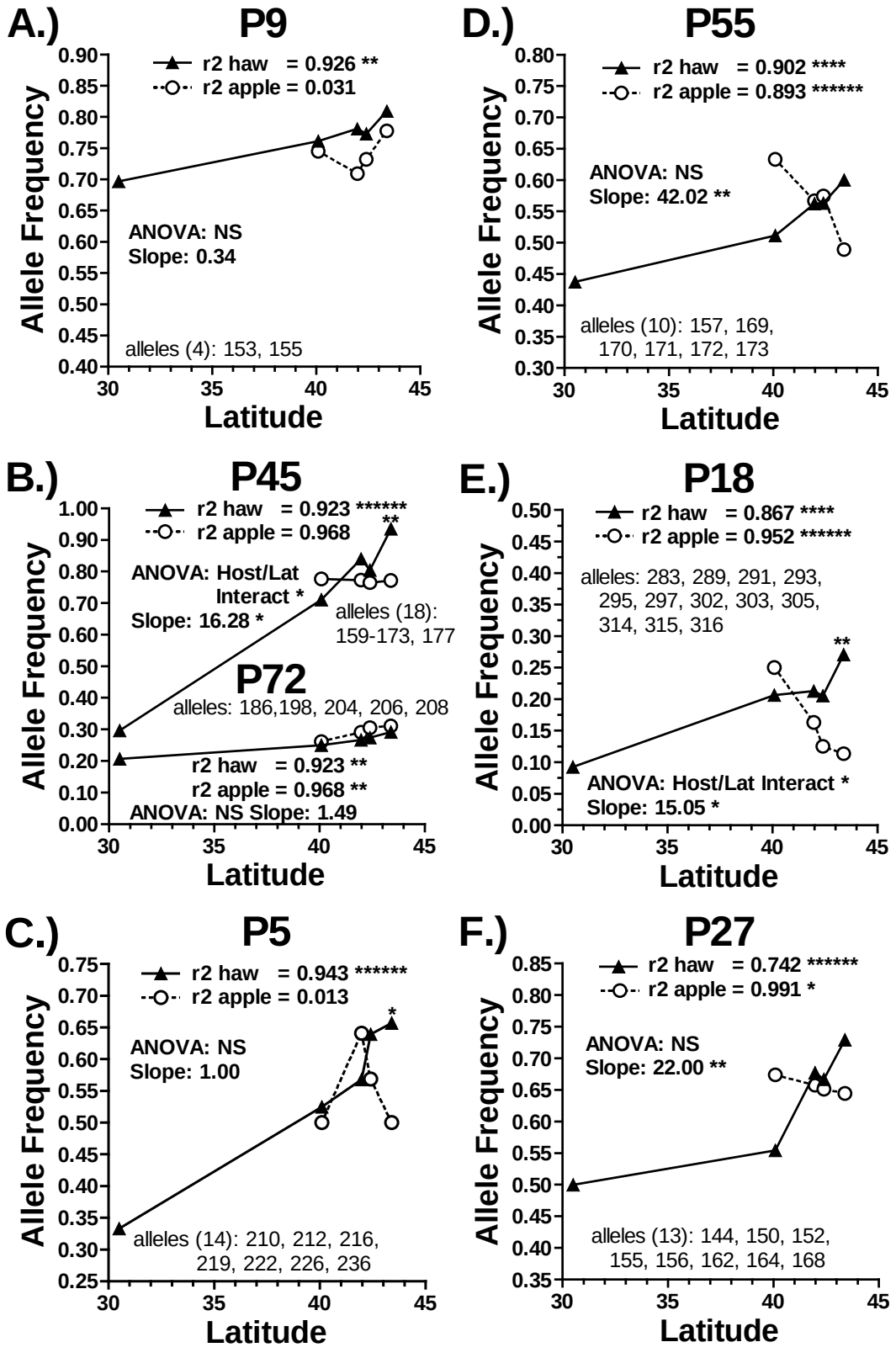


Fig. S9

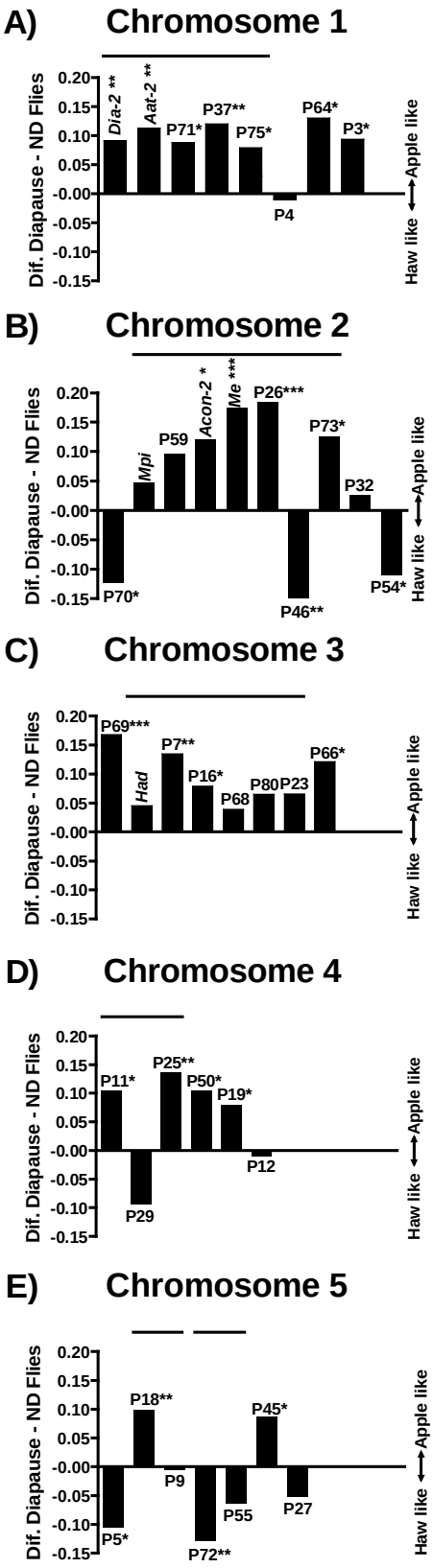
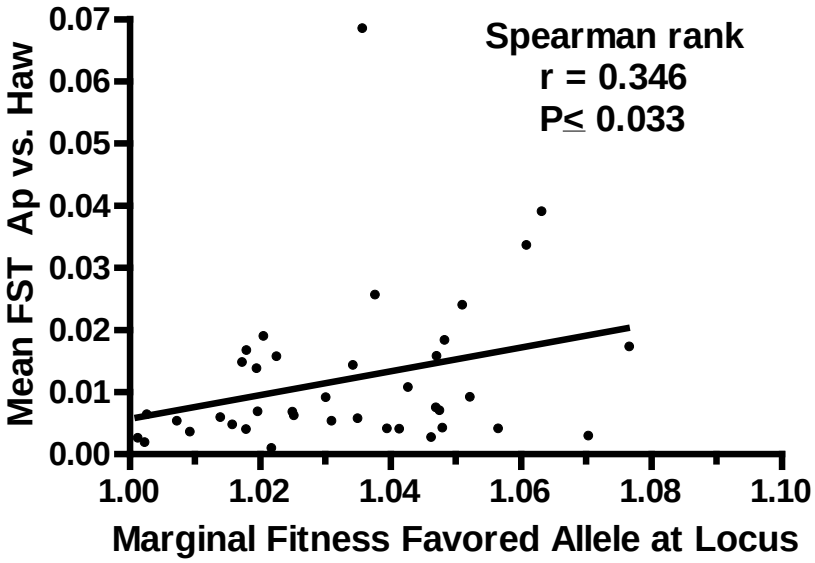

A)
Chromosome 1
Dif. Diapause - ND Flies
0.20
0.15
0.10
0.05
-0.00
-0.05
-0.10
-0.15
Dia-2 **
Aat-2 **
P71*
P37**
P75*
P4
P64*
P3*
Apple like
Haw like
B)
Chromosome 2
Dif. Diapause - ND Flies
Mpi
P59
Acon-2 *
Me ***
P26***
P73*
P32
P70*
P46**
P54*
Apple like
Haw like
C)
Chromosome 3
Dif. Diapause - ND Flies
P69***
Had
P7**
P16*
P68
P80
P23
P66*
Apple like
Haw like
D)
Chromosome 4
Dif. Diapause - ND Flies
P11*
P29
P25**
P50*
P19*
P12
Apple like
Haw like
E)
Chromosome 5
Dif. Diapause - ND Flies
P5*
P18**
P9
P72**
P55
P45*
P27
Apple like
Haw like

Fig. S10


Spearman rank
r = 0.346
P≤ 0.033
Mean FST Ap vs. Haw
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
1.00
1.02
1.04
1.06
1.08
1.10
Marginal Fitness Favored Allele at Locus