DOI: 10.1534/genetics.103.023028

Population Genetic Evidence for Rapid Changes in Intraspecific Diversity and Allelic Cycling of a Specialist Defense Gene in Zea

Peter Tiffin,*,[1] Robert Hacker[†] and Brandon S. Gaut[†]

**Department of Plant Biology, University of Minnesota, Saint Paul, Minnesota 55108 and [†]Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697*

Manuscript received October 8, 2003
Accepted for publication May 15, 2004

ABSTRACT

Two patterns of plant defense gene evolution are emerging from molecular population genetic surveys. One is that specialist defenses experience stronger selection than generalist defenses. The second is that specialist defenses are more likely to be subject to balancing selection, *i.e.*, evolve in a manner consistent with balanced-polymorphism or trench-warfare models of host-parasite coevolution. Because most of the data of specialist defenses come from *Arabidopsis thaliana*, we examined the genetic diversity and evolutionary history of three defense genes in two outcrossing species, the autotetraploid *Zea perennis* and its most closely related extant relative the diploid *Z. diploperennis*. Intraspecific diversity at two generalist defenses, the protease inhibitors *wip1* and *mpi*, were consistent with a neutral model. Like previously studied genes in these taxa, *wip1* and *mpi* harbored similar levels of diversity in *Z. diploperennis* and *Z. perennis*. In contrast, the specialist defense *hm2* showed strong although distinctly different departures from a neutral model in the two species. *Z. diploperennis* appears to have experienced a strong and recent selective sweep. Using a rejection-sampling coalescent method, we estimate the strength of selection on *Z. diploperennis hm2* to be ~3.0%, which is approximately equal to the strength of selection on *tb1* during maize domestication. *Z. perennis hm2* harbors three highly diverged alleles, two of which are found at high frequency. The distinctly different patterns of diversity may be due to differences in the phase of host-parasite coevolutionary cycles, although higher *hm2* diversity in *Z. perennis* may also reflect reduced efficacy of selection in the autotetraploid relative to its diploid relative.

MOLECULAR population genetic surveys of plant defense genes are providing novel insight into the evolutionary history of plant defenses and, by extension, plant-enemy interactions (reviewed in DeMeaux and Mitchell-Olds 2003). On the basis of the limited number of genes studied to date, defenses active against one or perhaps few enemies, *i.e.*, specialist defenses, appear on average to experience stronger selection than defenses active or potentially active against a broad array of enemies, *i.e.*, generalist defenses. Four specialist defenses, including three NBS-LRR genes from *Arabidopsis thaliana* (*Rpm1*, *Rps2*, *Rps5*, and *Rpp13*; Caicedo *et al.* 1999; Stahl *et al.* 1999; Tian *et al.* 2002; Mauricio *et al.* 2003; Rose *et al.* 2004) and detoxifying enzyme *hm2* in *Zea mays* ssp. *parviglumis* (Zhang *et al.* 2002) have patterns of intraspecific diversity inconsistent with expectations under a neutral model. Only one specialist defense, *hm1* in *Z. mays* ssp. *parviglumis*, has a pattern of intraspecific diversity consistent with expectations under neutrality. In contrast, the majority of putative generalist defenses surveyed have patterns of intraspecific diversity consistent with a neutral model [chitinases *A. thaliana chiB* (Kawabe and Miyashita 1999); *Z. mays* ssp. *parviglumis chiA*, *chiB*, and *chiI*; and *Z. diploperennis chiB* and *chiI* (Tiffin 2004) and proteinase inhibitors *A. thaliana Atti1*, *A. lyrata Atti1* and *Atti2* (Clauss and Mitchell-Olds 2003), and *Zea wip1* (Tiffin and Gaut 2001a)] while three show evidence of recent positive selection [*A. thaliana Atti2* (Clauss and Mitchell-Olds 2003) and possibly *chiA* (Kawabe *et al.* 1997; Kawabe and Miyashita 2002) and *Z. diploperennis chiA* (Tiffin 2004)].

A second pattern that appears to be emerging from molecular population genetic surveys of plant defense genes is that selection is more likely to maintain diversity at specialist than at generalist defense genes. Four of the five specialist defense genes, *A. thaliana Rpm1*, *Rps2*, *Rps5*, and *Rpp13* show evidence for selection having maintained diversity in populations—*i.e.*, the genes have been under balancing selection (Caicedo *et al.* 1999; Stahl *et al.* 1999; Tian *et al.* 2002; Mauricio *et al.* 2003; Rose *et al.* 2004). In contrast, only a single generalist defense, *A. thaliana mam2* at the *GSL-ELONG* locus, shows evidence of balancing selection (Kroymann *et al.* 2003). Evidence for balancing selection in specialist but not generalist defenses suggests that specialist defenses are

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AY320258–320280, AY52550–AY52559, and AY549598–AY549638.

[1]*Corresponding author:* Department of Plant Biology, University of Minnesota, 1445 Gortner Ave., St. Paul, MN 55108.
E-mail: ptiffin@umn.edu

more likely to evolve in a manner consistent with balanced-polymorphism or trench-warfare models of defense (reviewed in Bergelson *et al.* 2001). These models predict that functionally distinct defense alleles will be maintained in a population by some form of frequency-dependent selection.

The emerging pattern of balancing selection maintaining diversity at specialist but not generalist defenses must be viewed with caution given that the genes showing evidence for balancing selection come from only one species, inbreeding *A. thaliana*. The identification of balancing selection, as opposed to selective sweeps, in *A. thaliana* may be favored for two reasons. First, polymorphisms in *A. thaliana* are biased toward low frequency (*e.g.*, Purugganan and Suddith 1999; Wright *et al.* 2003), perhaps due to selfing and population subdivision (Sharbel *et al.* 2000). Low-frequency variants are also apparent after recovery from selective sweeps (Tajima 1989), and hence it may be difficult to distinguish demographic effects from recent selective sweeps in this system (Clauss and Mitchell-Olds 2003). Second, low effective recombination rates in selfers ensures that sites under balancing selection are in linkage disequilibrium with sites in close physical proximity (Nordborg *et al.* 2002), leading to a pronounced peak of variation around the site under balancing selection (Kreitman and Hudson 1991; Tian *et al.* 2002) that may be relatively easy to detect.

The emerging pattern of stronger balancing selection at specialist defenses also may be biased by the selection of genes that have been surveyed. The four specialist defenses surveyed in *A. thaliana* (*Rpm1*, *Rps2*, *Rps5*, and *Rpp13*) were all identified because they harbor resistant and susceptible genotypes in contemporary populations (Kunkel 1996; Bittner-Eddy *et al.* 1999). Loci that experience recent selective sweeps are much less likely to exhibit phenotypically detectible polymorphisms. The *GSL-ELONG* locus, the one generalist defense to harbor a pattern of balancing selection, was also investigated because it was identified as a candidate gene for a QTL associated with intraspecific variation in glucosinolate production (de Quiros *et al.* 2000). Because QTL will be detectable only if parental lines have functionally distinct alleles, molecular population genetic analyses of QTL candidates will be biased toward finding evidence of genes under balancing selection.

The only specialist defenses studied from a taxon other than *A. thaliana* are the Zea *hm1* and *hm2* genes. *Hm1* and *hm2* code for nitrate reductases that inhibit the HC toxin produced by the fungal pathogen *Cochliobolus carbonum* (syn. *Helminthosporium maydis*), thereby protecting plants from infection (Johal and Briggs 1992; Multani *et al.* 1998). Work in contemporary populations of *Z. mays* ssp. *mays* (maize) indicates that *hm1* is primarily responsible for defense against *C. carbonum*, although *hm2* also confers partial resistance (Nelson and Ullstrup 1964). Molecular population genetic analyses of diversity in maize and the teosinte *Z. mays* ssp. *parviglumis* indicate that *hm2* is the likely target of an ongoing selective sweep in this species (Zhang *et al.* 2002). Because this apparent sweep is ongoing, it is unclear if the positively selected allele will go to fixation, as predicted by classic arms-race models, or whether the apparent selective advantage associated with this allele will weaken as the allele increases in frequency, as predicted by dynamic-polymorphism/trench-warfare models.

Surveying *hm2* diversity in taxa closely related to *Z. mays* ssp. *parviglumis* may provide further insight into the evolutionary dynamics of this specialist defense. If *hm2* is experiencing similar evolutionary dynamics in closely related species, then sampling closely related species may provide a snapshot of this dynamic at different points in a coevolutionary cycle or selective sweep. Of course, *hm2* in closely related taxa may be experiencing different coevolutionary dynamics, but this too would be informative, providing insight into the variation in selection experienced by defense genes in closely related taxa. Two taxa, *Z. diploperennis* and *Z. perennis*, are particularly promising for a comparative population genetic survey of *hm2*. These species are morphologically similar, perennial species with restricted geographic distributions in the state of Jalisco in southwestern Mexico (Doebley 1990; Sanchez *et al.* 1998). However, these taxa are clearly distinct because *Z. diploperennis* is diploid and *Z. perennis* is a polyploid, which morphological and molecular data indicate originated from a *Z. diploperennis*-like progenitor (Iltis *et al.* 1979; Doebley *et al.* 1987; Buckler and Holtsford 1996; Tiffin and Gaut 2001b). Because these species have similar geographic distributions and life histories they may be likely to be exposed to similar pathogen pressures. Like other Zea species, all of which are native to Mexico and Central America, *Z. diploperennis* and *Z. perennis* are both wind-pollinated outcrossing taxa. Comparing *hm2* diversity in *Z. perennis* and *Z. diploperennis* is also facilitated by these species having similar patterns of DNA diversity at apparently neutral genes (Tiffin and Gaut 2001b).

The primary objective of this research is to examine the evolutionary history of *hm2* in two closely related species *Z. diploperennis* and *Z. perennis*, to gain a greater understanding of the long-term evolution of plant defense. We were particularly interested in determining if *hm2* would provide support for the apparent pattern that specialist defenses are more likely to experience stronger selection than generalist defenses and whether the pattern of selection acting on specialist defenses is more likely to be balancing than positive. Examining intraspecific diversity in closely related taxa also provides an opportunity to examine whether defense genes have similar evolutionary dynamics in closely related taxa.

Because we are interested in comparing the selective histories of specialist and generalist defenses we also investigated intraspecific diversity at two putative generalist defenses, the protease inhibitors *mpi* (maize prote-

ase inhibitor) and *wip1* (wound induced protein). Evidence for plant protease inhibitors having a primary role in defense against herbivores and pathogens (Ryan 1990) includes induction following pathogen infection and/or herbivore damage (Rohrmeier and Lehle 1993; Cordero *et al.* 1994; Tamayo *et al.* 2000), reduced growth and reproduction of some herbivores and pathogens when fed a diet or grown on medium that contains protease inhibitors (Tamayo *et al.* 2000), and increased resistance in plants expressing transgenic protease inhibitors (Hilder *et al.* 1987; Johnson *et al.* 1989). Because proteases are digestive enzymes present in most if not all herbivore guts, are secreted by many parasites, and are encoded by many viruses, protease inhibitors are potentially active against a broad array of enemies. No defense genes have been previously investigated in *Z. perennis* and only chitinase genes had been previously investigated in *Z. diploperennis* (Tiffin 2004).

MATERIALS AND METHODS

Sampling DNA sequences: We PCR amplified ~630 bp of *mpi*, 660 bp of *wip1*, and 1450 bp of *hm2* from eight accessions of *Z. diploperennis*, six accessions of *Z. perennis*, and one accession of *Tripsacum dactyloides* (see appendix). PCR conditions for *mpi* were 35 cycles of 1 min at 94°, 1 min at 50°, and 2 min at 72°; conditions for *hm2* were similar except the annealing temperature was 55° and 1 M betaine (*N,N,N*-trimethylglycine, Sigma, St. Louis) was added to each reaction. *Wip1* alleles, including the entire 280-bp coding region, a 90-bp intron, and 290 bp of flanking sequence, were amplified using primers and conditions described previously (Tiffin and Gaut 2001a). *Mpi* primers (F, ctgcagtgttgctatctgttc; R, attagtgagaattcacacatcc) amplified the entire coding region (220 bp in Zea) and ~280 and 60 bp of upstream and downstream flanking regions, respectively. *hm2* primers (F, tagcagtgaagtgcaggtg; R, attatgaga catggctggag) amplified a region extending from 12 bp 3′ of the *atg* start site to ~250 bp 3′ of the predicted stop codon. This region extends ~100 bp 5′ and 300–425 bp 3′ (depending on indels) beyond the region investigated by Zhang *et al.* (2002). *Oryza sativa* and *Sorghum bicolor* sequences used in relative rate tests were obtained from GenBank [*mpi*: AC079022, *O. sativa* (genomic) and BE917718, *S. bicolor* (EST); *wip1*: AP002526, *O. sativa* and AW680689, *S. bicolor*]. Sequences new to this study have been submitted to GenBank (*mpi*, AY549598–AY549627; *wip1*, AY52550–AY52559 and AY549628–AY549638; and *hm2*, AY320258–320280; appendix).

Analyses of evolutionary history are sensitive to the frequency of segregating sites, especially unique single-base-pair variants (singletons). To ensure that all singletons in our data set represented true variants and did not result from misincorporation of a nucleic acid during *Taq* amplification, all DNAs that yielded alleles with singletons were used as templates in one or more subsequent PCRs (a minimum of two for tetraploids) and the products of those reactions were cloned and sequenced. For the tetraploids, at least six clones from each reaction were sequenced. If the tetraploid contained four distinct alleles, sampling 12 isolates results in >95% probability of resampling the allele that initially contained the singleton, assuming no amplification bias (Tiffin and Gaut 2001b). Singletons not confirmed by this approach were assumed to have resulted from polymerase error and were excluded from the analyses.

Sequence analyses: Two measures of genetic diversity, the number of segregating sites (S) and Watterson's θ (Watterson 1975), were calculated separately on silent (synonymous and intron) and replacement sites. To determine if the diversity estimates for the defense genes that were the subject of this investigation were different from diversity at apparently neutral genes, we calculated maximum likelihood multilocus estimates of θ using the method of Wright *et al.* (2003) and data from previously investigated genes (*adh1*, *c1*, *glb1*, and *waxy*; Tiffin and Gaut 2001a). This method assumes no intragenic recombination, free recombination between loci, and a constant mutation rate. We also performed several tests of neutral evolution including Tajima's D (Tajima 1989), Fu's F_s (Fu 1997), McDonald-Kreitman (MK; McDonald and Kreitman 1991), and HKA (Hudson *et al.* 1987). For *wip1* and *mpi*, MK and HKA tests were conducted with a *T. dactyloides* sequence as an outgroup. Despite repeated attempts, we were unable to obtain an *hm2* sequence from *T. dactyloides* and therefore MK and HKA tests on *hm2* samples were conducted using the ssp. *parviglumis* sequences analyzed by Zhang *et al.* (2002) as an outgroup. The significance of MK tests was evaluated using a G-test with a William's correction. Confidence intervals (95%) around estimates of θ ($\hat{\theta}$), the probability of obtaining the number of sampled haplotypes (H), and the significance of F_s were estimated by running 1000 coalescent simulations of the neutral model. Simulations based on either S or $\hat{\theta}$ produced similar results and only results from simulations run with fixed S are shown. To be conservative in determining confidence intervals and testing for departures from a neutral model, all simulations were run with no recombination. Measures of polymorphism, tests of neutral evolution, and coalescent simulations were calculated using DnaSP v.3.53 (Rozas and Rozas 1999). Relative rates tests were conducted using the method of Fitch (1976) and Tajima (1993) as implemented by MEGA2 (Kumar *et al.* 2000).

To estimate the selection coefficient, s, and fixation time, T, measured in N_e generations, we used a rejection-sampling method based on selective sweeps simulated using the coalescent model of Przeworski (2003). This method employs three summary statistics estimated from the data (Tajima's D, S, and H) as well as assumed values of the mean mutation rate μ: the mean recombination rate per base pair c; the distance in base pairs, d, between the sequenced region and the site under selection; and the mean diploid population size N. In brief, this method simulates selective sweeps and then samples from the posterior distribution of T and s to obtain a sample that is consistent with the summary statistics calculated from the data. We assumed $\mu = 6.5 \times 10^{-9}$ mutations/site/year (Gaut *et al.* 1996), $c = 4 \times 10^{-7}$ (Wang *et al.* 1999), and $d = 1$ or 1000 and calculated N from the multilocus estimate of θ ($= 4N\mu$) based on neutral loci (Figure 1). Because these estimates, N, c, and μ, may be imprecise, uncertainty in their values is modeled by sampling from prior distributions of these parameters, as per Przeworski (2003).

RESULTS

Genetic diversity in *mpi* and *wip1*: For *mpi* $\hat{\theta}$'s fall into the 95% credibility interval (CI) of the multilocus likelihood $\hat{\theta}$ calculated using data from four other apparently neutrally evolving genes, *adh1, c1, glb1*, and *waxy* (Figure 1). Moreover, tests of nonneutral evolution that rely on intraspecific diversity—Fu's F_s and Tajima's D—were not significant when applied to *mpi* data (Table 1). In contrast, $\hat{\theta}$ in *Z. diploperennis wip1* is slightly higher than the 95% CI of the multilocus likelihood estimate. None-

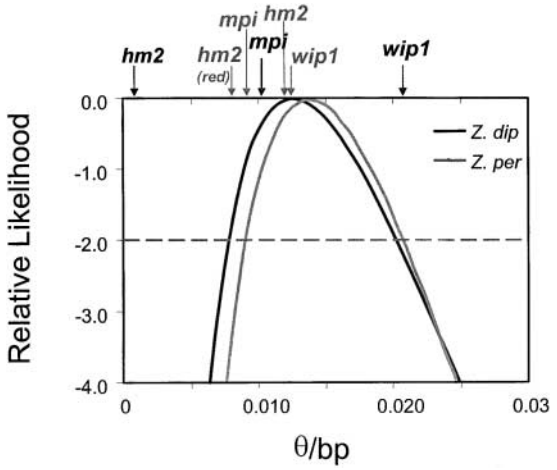

FIGURE 1.—Maximum likelihood estimates of θ, derived from apparently neutrally evolving genes in *Z. diploperennis* and *Z. perennis*. Arrows indicate the estimates of θ at *Hm2*, which was not used in making the likelihood estimates. The horizontal line marks the 95% credibility interval. *hm2* (red) was calculated after removing the *per7b* sequence from the *Z. perennis* data.

theless, Fu's F_s and Tajima's *D* tests were not significant when applied to *wip1* (Table 1), indicating that *wip1* has not experienced a recent selective sweep or been subject to balancing selection.

Results from relative rates and MK tests are also consistent with *mpi* having evolved neutrally (all $P > 0.45$). Similarly, MK tests conducted on *wip1* data are consistent with a neutral-equilibrium model ($P > 0.05$). In contrast, relative rates tests conducted on *wip1* sequences were significant for both species ($P < 0.01$ for all sequences), indicating rapid evolutionary change within the lineage leading to Zea. The significant relative rates tests are consistent with *wip1* having evolved in response to previous episodes of positive selection (TIFFIN and GAUT 2001a), although the evidence of selection is no longer evident in the frequency spectrum of intraspecific polymorphisms. Consistent with earlier analyses of apparently neutrally evolving loci (TIFFIN and GAUT 2001b), $\hat{\theta}$'s for *wip1* and *mpi* do not differ significantly between *Z. diploperennis* and *Z. perennis* (Table 1).

Natural selection on *hm2*: In contrast to the *mpi* and *wip1* data, patterns of *hm2* intraspecific diversity are inconsistent with neutrality. In *Z. diploperennis*, seven of the sampled alleles are identical and four others differ from these seven at only a single site (Figure 2). Moreover, *hm2* diversity in *Z. diploperennis* is an order of magnitude lower than that in any other *Z. diploperennis* gene sampled to date (Figure 1) and lower than diversity at any other gene within any of the teosintes (summarized in ZHANG *et al.* 2002), with the exception of a *Z. diploperennis* chitinase gene (*chiA*) that appears to have been subject to strong positive selection (TIFFIN 2004).

HKA tests on *Z. diploperennis hm2* data, conducted with *adh1*, *c1*, *glb1*, *waxy*, *mpi*, and *wip1* as the second locus and ssp. *parviglumis* as an outgroup, all had *P*-values <0.07, and all but two were significant at $P < 0.03$. In contrast, no HKA test among the other genes was significant (all $P > 0.3$). These results suggest that *hm2* or a closely linked region has evolved in response to selection within *Z. diploperennis*. Other tests of nonneutral evolution [Tajima's *D*, Fu's F_s, (Table 1), and Fay and Wu's *H* (FAY and WU 2000); data not shown] were not significant but this may not be surprising given that these tests rely on the frequency of segregating sites to identify departures from a neutral model. Although we sequenced >1450 bp for each of 11 *hm2* alleles, we detected only three segregating sites, providing little power for rejecting a neutral model.

The pattern of *hm2* polymorphism in *Z. perennis* is distinctly different from the pattern in *Z. diploperennis* but also appears indicative of nonneutral evolution. In contrast to *Z. diploperennis*, in which there were only 3 segregating sites among 11 sampled alleles, the 12 alleles sampled from *Z. perennis* contained 38 segregating sites distributed among 3 haplotypes (excluding two apparent recombinants *per3b* and *per3c*, Figure 2). The proba-

TABLE 1

Number of alleles (*N*), haplotypes (*H*), and segregating sites (*S*), measures of genetic diversity calculated on silent and replacement (θ_N) sites, and results from Tajima's *D* tests

Gene	Species	*N*	*H*	*S*	θ_{silent}	θ_N	*D*
mpi	*Z. diploperennis*	13	9	13	10.5 (4.6–17.4)	2.0	−1.19
	Z. perennis	16	8	12	9.1 (4.4–15.3)	0.0	0.33
wip1	*Z. diploperennis*	10	8	27	21.4 (10.7–34.2)	6.1	1.03
	Z. perennis	10	6	16	12.2 (5.4–19.8)	6.4	−0.08
hm2	*Z. diploperennis*	11	4	3	0.41 (0.0–1.7)	1.0	−1.11
	Z. perennis	12	5	38	12.7 (12.97–22.5)	4.5	1.20
	Z. perennis, reduced	11	4	27	8.1 (2.39–17.8)	4.16	2.68***

θ-values are per site ×1000 and 95% confidence intervals around these estimates are in parentheses. The *per*7b sequence has been eliminated from the *Z. perennis* reduced data set. ***$P < 0001$.

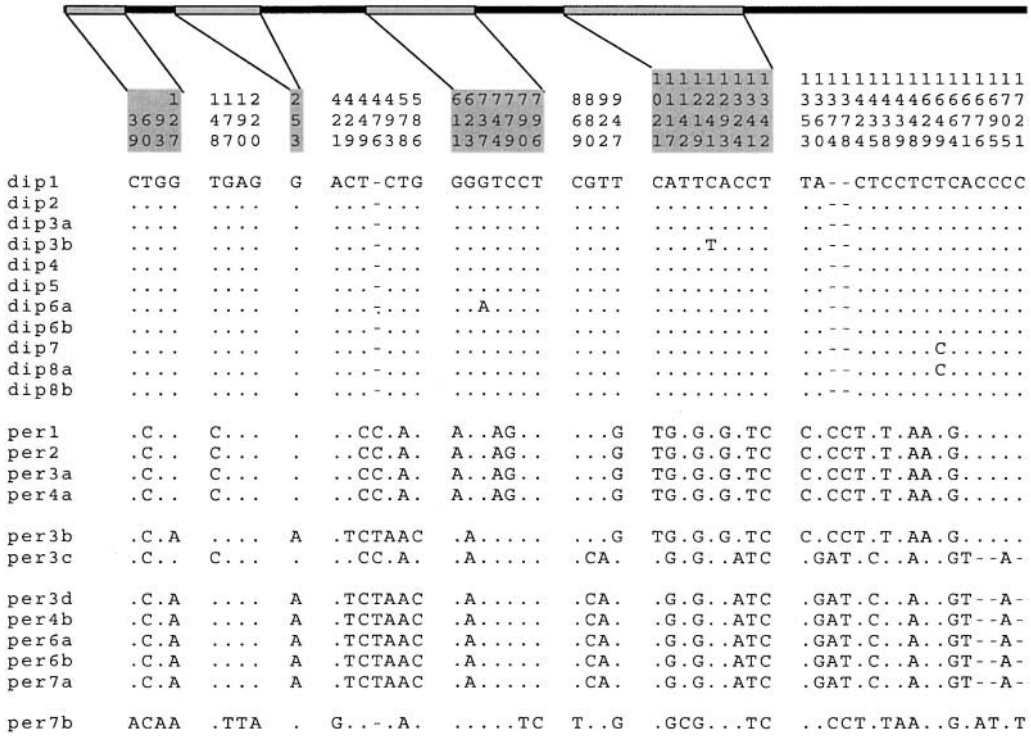

FIGURE 2.—The *hm2* exon-intron structure and sequences from *Z. diploperennis* and *Z. perennis*. The shaded regions represent exons, the position of polymorphic sites is shown at the top, and dashes indicate indels.

bility of ≤3 haplotypes in a sample with 38 segregating sites is extremely low under the standard neutral model ($P < 0.002$), and the result holds when the two recombinant haplotypes are included ($P < 0.02$ of ≤5 haplotypes). In contrast, haplotype numbers at six other genes sampled from the same collection of DNAs were consistent with neutral expectations (all $P > 0.25$). Interestingly, the coding regions of the two main haplotypes differ from one another primarily at replacement sites, suggesting that the haplotypes may be functionally diverged; although, d_N:d_S is not significantly >1 (8 replacement *vs.* 1 synonymous difference, d_N:d_S = 2.8, $P = 0.27$).

Although this pattern of diversity is suggestive of nonneutral history, Tajima's *D* was not significant when calculated on the entire *Z. perennis hm2* data set (Table 1). However, one of the three haplotypes is represented by only a single sequence (*per7b*, Figure 2), resulting in a high proportion of singletons, which could strongly affect tests of neutrality. Eliminating the third haplotype from the data eliminated all singletons and resulted in a substantially lower $\hat{\theta}$ and significantly positive values of *D* (Table 1) and F_s ($F_s = 8.1$, $P < 0.001$). F_s was also significant when all sequences were included ($F_s = 6.97$, $P < 0.01$) although HKA tests were not significant when conducted on either full or reduced *Z. perennis* data sets.

Two aspects of the *hm2* data indicate this gene has diverged between taxa in an atypical manner. First, 95% confidence intervals around $\hat{\theta}$ in *Z. diploperennis* and *Z. perennis* do not overlap (Table 1), unlike other genes from these taxa (Figure 1; Table 1). Second, *hm2* appears to be responsible for significant heterogeneity in the distribution of fixed differences and shared polymorphisms among loci ($G = 36.2$; $P < 0.001$, Table 2). To determine which locus was responsible for the significant result, we jackknifed contingency tests, removing one locus at a time. Six of these tests were significant at $P < 0.05$ (all remained significant after a Bonferroni correction), but the hypothesis of homogeneity was not rejected when *hm2* was removed ($P > 0.2$).

Simulations of selection on *hm2*: Statistical tests indicate that *hm2* has evolved nonneutrally in Zea taxa and the pattern of diversity at *hm2* in *Z. diploperennis* appears to be consistent with a recent selective sweep. To estimate the strength of selection that acted on *hm2* during this apparent sweep, as well as the time when the sweep occurred, we examined the posterior distributions of these parameters (*T* and *s*, the selection coefficient) that are consistent with the pattern of diversity

TABLE 2

Number of sites fixed and polymorphisms shared between *Z. diploperennis* and *Z. perennis*

	Fixed	Shared
adh1	0	23
c1	0	5
glb1	2	7
waxy	0	10
mpi	0	7
wip1	0	12
hm2	5	0

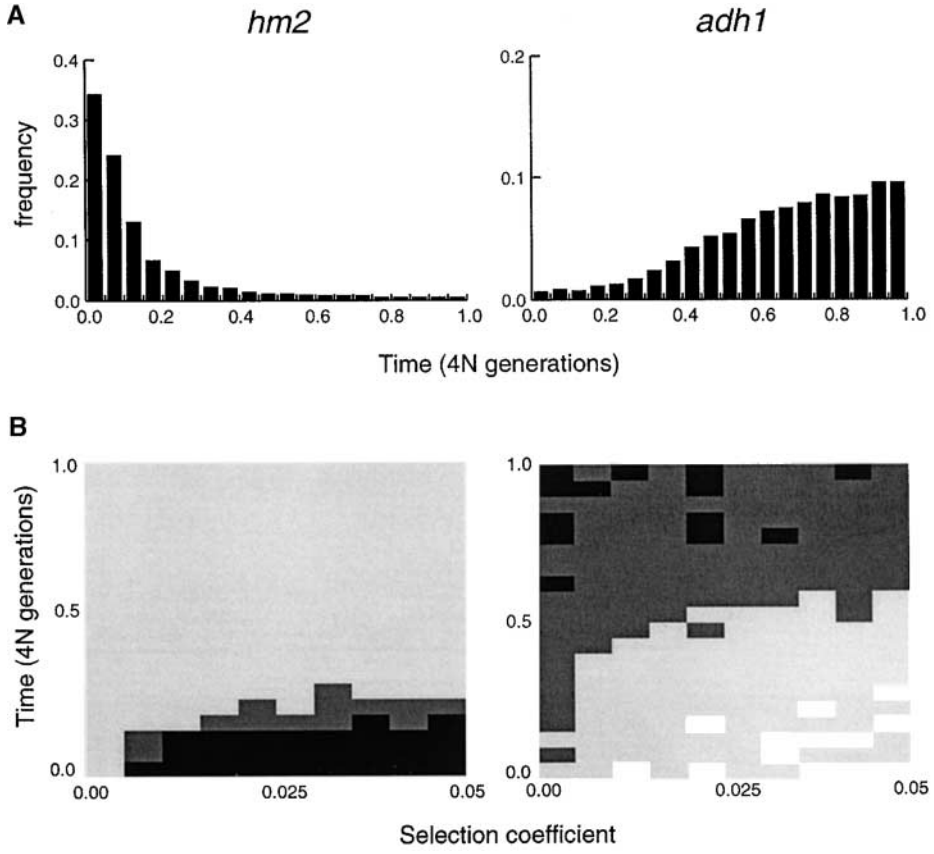

FIGURE 3.—Results of selective sweep simulations and rejection sampling. (A) Sample from the posterior distribution of T, the time of the fixation event, based on *hm2* and *adh1* data (based on 10,000 and 6781 simulations, respectively). (B) The joint posterior distribution of T and the selection coefficient s of the favored allele. Shading represents the frequency of simulated data sets; black > 1.0% data sets, 1.0% > dark gray > 0.5%, 0.5% > light gray > 0.0%, white = 0.0%.

we found in our sample. Assuming that the distance between the sequenced region and the site under selection equaled one base, a sample from the posterior distribution of T clearly favors a recent selective sweep in *hm2* (Figure 3), with the mode of the distribution $T < 0.05$ and 77.9% of the support on $T < 0.2$ a value that, following PRZEWORSKI (2003), we used as an arbitrary cutoff value consistent with a strong selective sweep. In contrast, when this method was used on an apparently neutral gene, *adh1*, only 3.4% of the support was on $T < 0.2$, and the mode of the distribution was $T > 0.9$—consistent with no sweep or a very old sweep that is no longer evident in a sample of alleles from present day populations. In addition to providing clear evidence for a strong selective sweep in *Z. diploperennis hm2*, these analyses provide a basis for estimating the strength of selection ($\hat{s}$) that acted on *hm2*. From the joint posterior distribution of s and T, mean s is 0.032 for *hm2* $T < 0.2$, suggesting the selection coefficient during the *hm2* sweep was ~3%. Similar results were obtained when we assumed that $d = 1000$ (*hm2*, 65.6% support for $T < 0.2$, $\hat{s} = 0.035$; *adh*, 9.0% support for $T < 0.2$).

We also applied the rejection-sampling method to *Z. perennis hm2* data. Ten days of computer time produced no posterior data that were consistent with summaries of *Z. perennis hm2* data. (In contrast, hundreds of data sets were simulated for the *Z. diploperennis* data in 1–2 days.) These results suggest that the *Z. perennis hm2* data do not fit a selective sweep model for any values of s and T. Moreover, because a sample taken from a neutral locus should, on average, have values of $T = 1$ and $s = 0$, the inability of the algorithm to produce any posterior data consistent with the pattern of diversity we found in *Z. perennis* suggests that these results are also inconsistent with the neutral model.

DISCUSSION

In this study we investigated the evolutionary history of three plant defense genes, one a specialist defense active against the fungal pathogen *C. carbonum* and two that are potentially active against a wide array of plant enemies. The two generalist defenses (*wip1* and *mpi*) show no evidence of having evolved in response to recent selection, and estimates of diversity at these genes fall within the range of diversity found at other presumably neutrally evolving loci in *Z. diploperennis* and *Z. perennis* (Figure 1). These data also corroborate our earlier finding that these two species harbor similar levels of genetic diversity (θ) at loci with patterns of intraspecific diversity that are consistent with a neutral

equilibrium model (Tiffin and Gaut 2001b). It should be noted, however, that the expectation of θ is $4N\mu$ for diploids, where N is population size and μ is mutation rate, but $8N\mu$ for tetraploids. Thus, although θ is similar between the two taxa, these estimates suggest that the long-term effective population size of the tetraploid is roughly half that of the most closely related extant diploid. Nonetheless, there is little indication, in either the level of sequence diversity or the values of Tajima's *D*, that tetraploid formation included a severe bottleneck.

In contrast to the generalist defenses, the specialist defense *hm2* appears to have evolved in response to powerful and recent selection in both *Z. diploperennis* and *Z. perennis*. This difference is consistent with an emerging pattern of specialist plant defenses showing stronger evidence for selection than generalist defenses. The patterns of diversity and evidence for selection are, however, distinctly different in the two species. In *Z. diploperennis*, *hm2* appears to have experienced a recent selective sweep as evident by extremely low diversity (Figure 1; significant HKA tests and coalescent simulations, see Figure 3). The selection that acted on *hm2* in *Z. diploperennis* appears to have been strong with an estimated selection coefficient, *s*, of ~3%, assuming conservatively that the selected site is very near the sequenced region. To corroborate the estimate of *s* obtained using Przeworski's (2003) rejection-sampling method, we also estimated *s* using the equation $\hat{s} = 100dc$ (Kaplan *et al.* 1989). Assuming $c = 4 \times 10^{-7}$ (Wang *et al.* 1999), $\hat{s}$ ranges from 0.029 to 0.058, depending on whether the selective site is assumed to be in the middle ($d = 725$) or the end ($d = 1450$) of the sequenced region. Thus, lower-range estimates of $\hat{s}$ from both methods are ~0.03. Moreover, both methods produce estimates of $\hat{s}$ that are similar to estimates of selection on *tb1* (Wang *et al.* 1999), a gene responsible for major differences in plant architecture that differentiate maize from the teosinte *Z. mays* ssp. *parviglumis* (Doebley *et al.* 1995, 1997; Clark *et al.* 2004), that were made using the equation of Kaplan *et al.* (1989). Although estimates for both genes are based on many assumptions that are difficult to verify, this comparison makes the important point that selection on disease resistance genes in the wild can be on the order of the strength of selection on artificially selected genes like *tb1*.

In contrast to *Z. diploperennis hm2*, which appears to have experienced a recent and strong sweep, diversity at *hm2* in *Z. perennis* is characterized by the presence of three alleles, each of which differs from alleles in the other classes by at least 27 sites (~2%). The distinct alleles segregating at *hm2* as well as positive values of Tajima's *D* and Fu's F_s and incompatibility with both selective sweep and neutral simulations (see RESULTS) could indicate a long-lived polymorphism in *Z. perennis*. The diversity at *Z. perennis hm2* is not, however, consistent with expectations of a balanced polymorphism in which allele frequencies have been stable. Simulation studies (Innan and Tajima 1999) show that when two haplotypes are maintained at stable frequencies through balancing selection the sum of the within-haplotype $\hat{\theta}$'s is expected to be equal to θ estimated using all data. We find, however, that the sum of the within-class $\hat{\theta} = 0$, well below the 95% confidence intervals around $\hat{\theta}$ calculated from the *Z. perennis* data. The lack of intrahaplotype polymorphism suggests that the two main allelic classes have not been maintained as a long-term stable polymorphism but rather that both alleles have recently increased in frequency due to positive selection. Positive selection could also explain the comparatively high interspecific divergence of *hm2*, relative to other loci (Table 2). It may be possible that the *Z. perennis* data also reflect population structuring but this seems unlikely given that *Z. perennis* is wind pollinated with a geographically limited range (Sanchez *et al.* 1998). Moreover, the six other loci we have examined from this same collection of *Z. perennis* DNAs reveal no obvious patterns that indicate population structure.

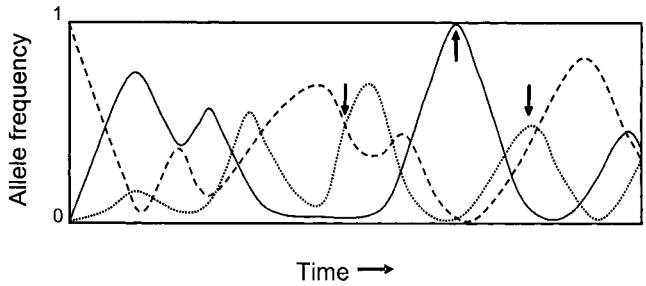

Figure 4.—Hypothetical dynamics of three defense alleles coevolving with parasite avirulence alleles (after Seger 1990). *Hm2* may exhibit very low diversity in *Z. diploperrenis* because it has been the object of a recent selective sweep (up arrow). Two common alleles may be present in *Z. perennis* because each has recently increased in frequency (down arrows).

Because *Z. diploperennis* and *Z. perennis* are closely related with similar life histories and geographic distributions it seems likely that the distinct patterns of diversity found at *hm2* in these species reflect a common underlying host-parasite coevolutionary processes—with the two species being at different phases of coevolutionary cycles. These patterns appear inconsistent, however, with two-allele models that predict regular cyclic fluctuations in gene frequencies (Jayakar 1970; Seger 1988, 1990; Stahl *et al.* 1999). These data do, however, appear consistent with three-allele models that exhibit highly irregular fluctuations in gene frequencies (Seger 1988) and may more appropriately describe allelic variation at many defense genes in natural populations. In *Z. diploperennis*, *hm2* may have low diversity because it is at the peak of a cycle whereas an apparently long-lived polymorphism may be detected in *Z. perennis* because *hm2* is experiencing simultaneous increase of two alleles (Figure 4). The apparent ongoing sweep detected at *hm2*

in *Z. mays* ssp. *parviglumis* may also be consistent with these more complicated models of host-parasite coevolution.

This explanation for the distinctly different patterns of diversity found in *Z. diploperennis* and *Z. perennis* is clearly speculative. Moreover, many patterns of diversity are consistent with these three-allele coevolutionary models, depending on the strength of selection, the amplitude of cycles, and the phase of a cycle from which alleles were sampled. As such, it is unclear what data would provide definitive evidence against these models. We note that other dynamic-polymorphism/trench-warfare models suffer similar drawbacks. Nevertheless, if these three-allele models correctly describe the evolutionary dynamics of defense genes we might expect that other defense genes will also harbor distinctly different patterns of diversity in closely related species or isolated populations, just as we have documented here. Sequence polymorphism at the recently sampled *A. thaliana RPP13* locus is also inconsistent with simple two-allele models of balancing selection, but does appear consistent with multiallele frequency-dependent selection (Rose *et al.* 2004).

One caveat in interpreting our data is that the higher diversity in *Z. perennis* may be related to beneficial alleles spreading more slowly through autopolyploid than through diploid species. Theoretical models show that the number of alleles that segregate at meiosis affects the rate at which selectively favored alleles spread through a population (Hill 1971; Otto and Whitton 2000). *Z. perennis* is an autotetraploid and thus segregates four rather than two alleles per locus. Therefore, beneficial alleles will be slower to fix in the autotetraploid *Z. perennis* and, at loci that experience strong positive selection, this species may harbor greater diversity than the diploid *Z. diploperennis* (assuming other things are equal, such as population size). The *hm2* data showing higher diversity in *Z. perennis* than in *Z. diploperennis* are consistent with the prediction that autotetraploids "mask" beneficial alleles from selection. Nevertheless, under a simple masking model there is no reason to expect that the allele sweeping through the population would be highly diverged from other alleles found within the species—the selectively advantageous allele might differ from other alleles only at recent mutations that altered the fitness effects. Moreover, the alleles that are masked from selection would be expected to be under relaxed selective constraint and therefore accumulate mutations. Neither of these expectations is met in *Z. perennis* in which each of the three haplotypes differ from one another at ~2% of sites, there are no polymorphic sites within either haplotype that is represented by more than a single sequence, and the relatively high d_N:d_S between the two main haplotypes suggests that they may be diverged functionally. Given the inconsistencies between our data and the expectations of tetravalent inheritance slowing the spread of positively selected alleles, we think selective pressure imposed by the parasite against which *hm2* is active is at least partially responsible for the patterns of diversity we see at this locus. Unfortunately direct tests of this hypothesis may be difficult, given that the parasite genotypes potentially responsible for past selection on *hm2* may no longer be common or even present in contemporary populations.

We know of only two other studies that have examined defense gene diversity in closely related taxa; Clauss and Mitchell-Olds (2003) examined intraspecific diversity of duplicated trypsin inhibitor genes in the closely related *A. thaliana* and *A. lyrata* ssp. *petraea* and Tiffin (2004) examined diversity of three chitinase genes in *Z. diploperennis* and *Z. mays* ssp. *parviglumis*. Both of these studies revealed evidence for interspecific differences in evolutionary histories; in particular, both detected evidence for positive selection in one taxon, but neutral patterns of diversity in the second taxon. Unlike *Z. diploperennis* and *Z. perennis*, the taxa investigated in Clauss and Mitchell-Olds (2003) and Tiffin (2004) differ in life history traits and geographic ranges and it is therefore unclear whether interspecific differences in diversity found in those taxa are due to selective forces acting on the defense genes or due to demographic forces with genome-wide effects. Here we show that defense genes in closely related taxa with similar life histories and geographic ranges may have substantially different evolutionary histories and levels of sequence diversity. Moreover, the *hm2* data suggest that simple two-allele models are unlikely to adequately capture the evolution of all defense genes in natural populations.

Regardless of the evolutionary mechanisms responsible for the higher *hm2* diversity in *Z. perennis*, if greater diversity at functionally important resistance genes is a general phenomenon in polyploids, then this may, in part, explain ecological observations that tetraploids are often more resistant to pathogens and herbivores than are their diploid relatives (Levin 1983; Nuismer and Thompson 2001).

We thank Molly Przeworski for her generous help with estimating *s* and *T*; Eli Stahl, Deborah Charlesworth, and an anonymous reviewer for comments that improved the manuscript; and John Doebley for providing seed that made this project possible. This research was supported by a U.S. Department of Agriculture Award 99-35301-8076 to P.T. and National Science Foundation grants DEB 9815855 to B.S.G. and DEB 0235027 to P.T.

LITERATURE CITED

Bergelson, J., G. Dwyer and J. J. Emerson, 2001 Models and data on plant-enemy coevolution. Annu. Rev. Genet. **35:** 469–499.

Bittner-Eddy, P., C. Can, N. Gunn, M. Pinel, M. Tor *et al.*, 1999 Genetic and physical mapping of the RPP13 locus in Arabidopsis responsible for specific recognition of several *Peronopora parasitca* (downy mildew) isolates. Mol. Plant-Microbe Interact. **12:** 792–802.

Buckler, E. S., and T. P. Holtsford, 1996 *Zea* systematics: ribosomal ITS evidence. Mol. Biol. Evol. **13:** 612–622.

Caicedo, A. L., B. A. Schaal and B. A. Kunkel, 1999 Diversity and molecular evolution of the *RPS2* resistance gene in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **96:** 302–306.

Clark, R. M., E. Linton, J. Messing and J. F. Doebley, 2004 Pattern of diversity in the genomic region near the maize domestication gene *tb1*. Proc. Natl. Acad. Sci. USA **101:** 700–707.

Clauss, M. J., and T. Mitchell-Olds, 2003 Population genetics of tandem trypsin inhibitor genes in *Arabidopsis* species with contrasting ecology and life histories. Mol. Ecol. **12:** 1287–1299.

Cordero, M. J., D. Raventos and B. San Segundo, 1994 Expression of a maize proteinase inhibitor gene is induced in response to wounding and fungal infection: systemic wound-response of a monocot gene. Plant J. **6:** 141–150.

de Meaux, J., and T. Mitchell-Olds, 2003 Evolution of plant resistance at the molecular level: ecological context of species interactions. Heredity **94:** 343–352.

de Quiros, H. C., R. Magrath, D. McCallum, J. Kroymann, D. Scnabelrauch *et al.*, 2000 Alpha-keto acid elongation and glucosinolate biosynthesis in Arabidopsis thaliana. Theor. Appl. Genet. **101:** 429–437.

Doebley, J., 1990 Molecular evidence and the evolution of maize. Econ. Bot. **44:** 6–27.

Doebley, J., A. Stec and C. Gustas, 1995 *Teosinte branched 1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics **141:** 333–346.

Doebley, J., A. Stec and L. Hubbard, 1997 The evolution of apical dominance in maize. Nature **386:** 485–488.

Doebley, J. F., W. Renfroe and A. Blanton, 1987 Restriction site variation in the *Zea* chloroplast genome. Genetics **117:** 139–147.

Fay, J. C., and C.-I Wu, 2000 Hitchhiking under positive Darwinian selection. Genetics **155:** 1405–1413.

Fitch, W. M., 1976 Molecular evolutionary clocks, pp. 160–178 in *Molecular Evolution*, edited by F. J. Ayala. Sinauer Associates, Sunderland, MA.

Fu, Y. X., 1997 Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics **147:** 915–952.

Gaut, B. S., B. R. Morton, B. M. McCaig and M. T. Clegg, 1996 Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. Proc. Natl. Acad. Sci. USA **93:** 10274–10279.

Hilder, V. A., A. M. R. Gatehouse, S. E. Sheerman, R. F. Barker and D. Boulter, 1987 A novel mechanism of insect resistance engineered into tobacco. Nature **330:** 160–163.

Hill, R. R., 1971 Selection in autotetraploids. Theor. Appl. Genet. **41:** 181–186.

Hudson, R. R., M. Kreitman and M. Aguade, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

Iltis, H. H., J. F. Doebley, R. Guzman and B. Pazy, 1979 *Zea diploperennis* (Gramineae): a new teosinte from Mexico. Science **203:** 186–188.

Innan, H., and F. Tajima, 1999 The effect of selection on the amounts of nucleotide variation within and between allelic classes. Genet. Res. **73:** 15–28.

Jayakar, S. D., 1970 A mathematical model for interaction of gene frequencies in a parasite and its host. Theor. Pop. Biol. **1:** 140–164.

Johal, G. S., and S. P. Briggs, 1992 Reductase activity encoded by the *HM1* disease resistance gene in maize. Science **258:** 985–987.

Johnson, R., J. Narvaez, G. An and C. A. Ryan, 1989 Expression of proteinase inhibitors I and II in transgenic tobacco plants: effects on natural defense against *Manduca sexta* larvae. Proc. Natl. Acad. Sci. USA **86:** 9871–9875.

Kaplan, N. L., R. R. Hudson and C. H. Langley, 1989 The "Hitchiking Effect" revisited. Genetics **123:** 887–899.

Kawabe, A., and N. T. Miyashita, 1999 DNA variation in the basic chitinase locus (*ChiB*) region of the wild plant *Arabidopsis thaliana*. Genetics **153:** 1445–1453.

Kawabe, A., and N. T. Miyashita, 2002 DNA variation in the acidic chitinase locus (*ChiA*) region in *Arabis gemmifera* and its related species. Genes Genet. Syst. **77:** 167–175.

Kawabe, A., H. Innan, R. Terauchi and N. T. Miyashita, 1997 Nucleotide polymorphism in the acidic chitinase focus (*ChiA*) region of the wild plant *Arabidopsis thaliana*. Mol. Biol. Evol. **14:** 1303–1315.

Kreitman, M., and R. R. Hudson, 1991 Inferring the evolutionary histories of *Adh* and the *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics **127:** 565–582.

Kroymann, J., S. Donnerhacke, D. Schnabelrauch and T. Mitchell-Olds, 2003 Evolutionary dynamics of an Arabidopsis insect resistance quantitative trait locus. Proc. Natl. Acad. Sci. USA **100:** 14587–14592.

Kunkel, B. N., 1996 A useful weed put to work: genetic analysis of disease resistance in *Arabidopsis thaliana*. Trends Genet. **12:** 63–69.

Kumar, S., K. Tamura, I. Jakobsen and M. Nei, 2000 MEGA: molecular evolutionary genetics analyses. Pennsylvania State University/Arizona State University, University Park, PA/Tempe, AZ.

Levin, D. A., 1983 Polyploidy and novelty in flowering plants. Am. Nat. **122:** 1–25.

Mauricio, R., E. A. Stahl, T. Korves, D. C. Tian, M. Kreitman *et al.*, 2003 Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana*. Genetics **163:** 735–746.

McDonald, J. H., and M. Kreitman, 1991 Adaptive protein evolution at the *Adh1* locus in *Drosophila*. Nature **351:** 652–654.

Multani, D. S., R. B. Meeley, A. H. Paterson, J. Gray, S. P. Briggs *et al.*, 1998 Plant-pathogen micro evolution: molecular basis for the recent origin of a disease in maize. Proc. Natl. Acad. Sci. USA **95:** 1686–1691.

Nelson, O. E., and A. J. Ullstrup, 1964 Resistance to leaf spot in maize. J. Hered. **55:** 195–199.

Nordborg, M., J. O. Borevitz, J. Bergelson, C. C. Berry, J. Chory *et al.*, 2002 The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat. Genet. **30:** 190–193.

Nuismer, S. L., and J. N. Thompson, 2001 Plant polyploidy and non-uniform effects on insect herbivores. Proc. R. Soc. Lond. Ser. B Biol. Sci. **268:** 1937–1940.

Otto, S. P., and J. Whitton, 2000 Polyploid incidence and evolution. Annu. Rev. Genet. **34:** 401–437.

Przeworski, M., 2003 Estimating the time since the fixation of a beneficial allele. Genetics **164:** 1667–1676.

Purugganan, M. D., and J. I. Suddith, 1999 Molecular population genetics of floral homeotic loci: departures from the equilibrium-neutral model at the *APETALA3* and *PISTILLATA* gene of *Arabidopsis thaliana*. Genetics **151:** 839–848.

Rohrmeier, T., and L. Lehle, 1993 *WIP1*, a wound-inducible gene from maize with homology to Bowman-Birk proteinase inhibitors. Plant Mol. Biol. **22:** 783–792.

Rose, L. E., P. D. Bittner-Eddy, C. H. Langley, E. B. Holub, R. W. Michemore *et al.*, 2004 Maintenance of extreme amino acid diversity at the disease resistance gene *RPP13* in *Arabidopsis thaliana*. Genetics **166:** 1517–1527.

Rozas, J., and R. Rozas, 1999 DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15:** 174–175.

Ryan, C. A., 1990 Protease inhibitors in plants: genes for improving defenses against insects and pathogens. Annu. Rev. Phytopathol. **28:** 425–449.

Sanchez, G. J. J., T. A. K. Yamakake, M. A. Sanmiguel, J. M. H. Casillas, A. L. Rodriguez *et al.*, 1998 *Distribucion y characterizacion del teocintle*. Instituto Nacional de Investigaciones Forestales, Agricolas, y Pecuarias, Mexico City.

Seger, J., 1988 Dynamics of some simple host-parasite models with more than two genotypes in each species. Philos. Trans. R. Soc. Lond. B **319:** 541–555.

Seger, J., 1990 Evolution of exploiter-victim relationships, pp. 3–25 in *Natural Enemies: The Population Biology of Predators, Parasites and Diseases*, edited by M. J. Crawley. Blackwell Scientific, London.

Sharbel, T. F., B. Haubold and T. Mitchell-Olds, 2000 Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. Mol. Ecol. **9:** 2109–2118.

Stahl, E. A., G. Dwyer, R. Mauricio, M. Kreitman and J. Bergelson, 1999 Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. Nature **400:** 667–671.

Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585–595.

Tajima, F., 1993 Simple methods for testing the molecular evolutionary clock hypothesis. Genetics **135:** 599–607.

Tamayo, M. C., M. Rufat, J. M. Bravo and B. San Segundo, 2000 Accumulation of a maize proteinase inhibitor in response to

wounding and insect feeding, and characterization of its activity toward digestive proteinases of *Spodoptera littoralis* larvae. Planta **211:** 62–71.

Tian, D., H. Araki, E. A. Stahl, J. Bergelson and M. Kreitman, 2002 Signature of balancing selection in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **99:** 11525–11530.

Tiffin, P., 2004 Comparative evolutionary histories of chitinase genes in the genus Zea and family Poaceae. Genetics **167:** 1331–1340.

Tiffin, P., and B. S. Gaut, 2001a Molecular evolution of the wound-induced serine protease inhibitor *wip1* in *Zea* and related genera. Mol. Biol. Evol. **18:** 2092–2101.

Tiffin, P., and B. S. Gaut, 2001b Sequence diversity in the autotetraploid *Zea perennis* and the closely related diploid *Z. diploperennis*: insights from four nuclear loci. Genetics **158:** 401–412.

Wang, R. L., A. Stec, J. Hey, L. Lukens and J. Doebley, 1999 The limits of selection during maize domestication. Nature **398:** 236–239.

Watterson, G. A., 1975 On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7:** 188–193.

Wright, S. I., B. Lauga and D. Charlesworth, 2003 Subdivision and haplotye structure in natural populations of *Arabidopsis lyrata*. Mol. Ecol. **12:** 1247–1263.

Zhang, L., A. S. Peek, D. Dunamas and B. S. Gaut, 2002 Population genetics of duplicated disease-defense genes, *hm1* and *hm2*, in maize (*Zea mays* ssp. *mays* L.) and its wild ancestor (*Zea mays* ssp. *parviglumis*). Genetics **162:** 851–860.

Communicating editor: D. Charlesworth

APPENDIX

Source and GenBank accession numbers for *Z. diploperennis* and *Z. perennis* sequences that are new to this study

	GenBank		
Accession	*mpi*	*wip1*	*hm2*
Z. diploperennis			
M005	*AY549598*, *AY549599*	*AY52550*	*AY320270* (1)
9476	*AY549600*	*AY52551*	*AY320271* (2)
10003	*AY549601*	*AY52552*, *AY52553*	*AY320272*, *AY300273* (3a,b)
Ames 2317	*AY549602*	*AY52554*	*AY320274* (4)
PI 441932	*AY549603*, *AY549604*	*AY52555*, *AY52556*	*AY320275* (5)
PI 462368	*AY549605*, *AY549606*	*AY52557*	*AY320276*, *AY300277* (6a,b)
Ames 21884	*AY549607*, *AY549608*	*AY52558*	*AY320278* (7)
PI 441931	*AY549609*, *AY549610*	*AY52559*	*AY320279*, *AY300280* (8a,b)
Z. perennis			
Ames 21869	*AY549611*, *AY549612*, *AY549613*	—	*AY320258* (1)
Ames 21870	*AY549614*, *AY549615*, *AY549616*	*AY549629*, *AY549630* *AY549631*	*AY320259* (2)
Ames 21873	*AY549617*, *AY549618*, *AY549619*	*AY549632*, *AY549633*	*AY320260*, *AY320261* (3a,b) *AY320262*, *AY320263* (3c,d)
Ames 21874	*AY549620*, *AY549621* *AY549622*, *AY549623*	—	*AY320264*, *AY320265* (4a,b)
Jal-88	*AY549624*	*AY549634*, *AY549635*, *AY549636*	*AY320268*, *AY320269* (6a,b)
Mo10	*AY549625*, *AY549626*	*AY549637*, *AY549638*	*AY32026*, *AY320267* (7a,b)

hm2 sample numbers used in Figure 2 are in parentheses following the GenBank number.