DOI: 10.1534/genetics.104.026856

Comparative Evolutionary Histories of Chitinase Genes in the Genus Zea and Family Poaceae

Peter Tiffin[1]

Department of Plant Biology, University of Minnesota, St. Paul, Minnesota 55108

Manuscript received January 27, 2004
Accepted for publication March 23, 2004

ABSTRACT

Patterns of DNA sequence diversity vary widely among genes encoding proteins that protect plants against pathogens and herbivores. Comparative studies may help determine whether these differences are due to the strength of selection acting on different types of defense, in different evolutionary lineages, or both. I analyzed sequence diversity at three chitinases, a well-studied component of defense, in two species of Zea and several Poaceae taxa. Although the Zea species are closely related and these genes code for proteins with similar biochemical function, patterns of diversity vary widely within and among species. Intraspecific diversity at *chiB*, *chiI*, and *Z. mays* ssp. *parviglumis chiA* are consistent with a neutral-equilibrium model whereas *chiA* had no segregating sites within *Z. diploperennis*—consistent with a recent and strong selective sweep. Codons identified as having diverged among Poaceae taxa in response to positive selection were significantly overrepresented among targets of selection in Arabis, suggesting common responses to selection in distantly related plant taxa. Divergence of the recent duplicates *chiA* and *chiB* is consistent with positive selection but relaxed constraint cannot be rejected. Weak evidence for adaptive divergence of these duplicated downstream components of defense contrasts with strong evidence for adaptive divergence of genes involved in pathogen recognition.

MOST studies aimed at understanding the evolution of plant defenses against herbivores and pathogens have examined patterns of selection acting in contemporary populations (reviewed in Rausher 2001; de Meaux and Mitchell-Olds 2003) or used comparative methods to examine broad macroevolutionary patterns (Berenbaum 1983; Farrell *et al.* 1991). Recently, molecular population genetic analyses have been used to examine the patterns of sequence diversity at plant defense genes, thereby providing new insights into the evolutionary history of defense. Although the number of studies using molecular analyses is still relatively small, they provide evidence that at least some defense genes have evolved in response to strong positive selection (Bishop *et al.* 2000; Stotz *et al.* 2000; Tiffin *et al.* 2004), that diversity at other genes is maintained through some form of balancing selection (Caicedo *et al.* 1999; Stahl *et al.* 1999; Tian *et al.* 2002; Mauricio *et al.* 2003), and that other putative defense genes have patterns of diversity that do not deviate significantly from neutral expectations (Kawabe and Miyashita 1999; Clauss and Mitchell-Olds 2003). The results from these investigations clearly show that the evolutionary histories of defense genes are variable. They do not, however, tell us if this variation is the result of variation in the strength or patterns of selection acting on different genes, in different phylogenetic lineages, or a combination of the two. Comparative studies, which examine the evolutionary history of multiple defense genes within a species or orthologous genes in multiple evolutionary lineages, offer an opportunity to differentiate between these possibilities. Surprisingly, few comparative analyses of defense gene evolution have been conducted.

Plant chitinases are a well-studied class of defense, having been the subject of both functional and molecular-evolutionary analyses, and are therefore good candidates for comparative investigation. Chitinases are hydrolytic enzymes that catalyze the degradation of chitin, a major component of fungal cell walls. These enzymes can inhibit the growth of fungal hyphae *in vitro* (Schlumbaum *et al.* 1986; Huynh *et al.* 1992), some chitinases are induced following pathogen infection (Wu *et al.* 1994), and the overexpression of at least some chitinases in transgenic plants causes significant reductions in pathogen damage (Broglie *et al.* 1991). Taken together, these observations support the notion that a primary function of plant chitinases is in defending plants against attack by fungal pathogens, although there is also evidence that chitinases may function as lysozymes degrading bacterial cell walls and may play a role in

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AY532721–AY532788.

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

developmental processes (Passarinho and de Vries 2002).

Analysis of DNA sequence data has revealed that some chitinases have evolved in response to positive selection. In particular, divergence of a class I (glycosyl-hydrolase family 19) chitinase gene among Arabis species is characterized by high d_N:d_S; *i.e.*, there are more substitutions per replacement than per synonymous site (Bishop *et al.* 2000). Interestingly, the codons identified as targets of positive selection were overrepresented in the active-site cleft of the enzyme, a pattern consistent with this gene having coevolved in response to variation in pathogen cell walls or a pathogen-produced chitinase inhibitor (Bishop *et al.* 2000). Although the pattern of interspecific divergence among Arabis species is indicative of positive selection, patterns of diversity at *chiB*, the *Arabidopsis thaliana* gene orthologous to the gene studied in Arabis, are consistent with a neutral model of evolution (Kawabe and Miyashita 1999). The apparent discrepancy between intraspecific and interspecific analyses may indicate that selection acting on chitinases differs among closely related taxa; positive selection acting within Arabis lineages has resulted in the high d_N:d_S, even if selection has not acted within *A. thaliana*. Alternatively, positive selection may have acted on the *A. thaliana* chitinase but selection was not recent and the effect of selection on allelic diversity may be masked by the accumulation of neutral mutations.

The primary objective of this research was to examine the diversity and evolutionary histories of chitinase genes in two species of Zea and among genera in the family Poaceae and to compare these patterns of diversity to those found at the class I chitinase in Arabis. In particular, I ask whether (i) three genes that code for proteins with very similar biochemical function experience similar evolutionary histories within Zea species; (ii) there is evidence for positive selection having driven the divergence of Poaceae chitinases; and (iii) codons detected as diverging in response to positive selection in Arabis are also identified as targets of positive selection in Poaceae. Finding similar evolutionary histories in Zea and Arabis, model monocot and dicot systems, may indicate that evolution in response to positive selection is a general phenomenon of plant family 19 glycosyl-hydrolase chitinases.

The three Zea genes investigated (*chiI*, *chiA*, and *chiB*) are all members of the family 19 glycosyl hydrolases (Huynh *et al.* 1992; Wu *et al.* 1994), and *chiI* is classified as an acidic member of the class I chitinases (Wu *et al.* 1994). *chiA* and *chiB* were originally classified as basic members of the class I chitinases (Huynh *et al.* 1992). (*chiA* and *chiB* were the gene names given by Huynh *et al.* 1992; *chiB* does not indicate orthology with the Arabis/Arabidopis *chiB* gene.) However, several short deletions in the catalytic domain differentiate *chiA* and *chiB* sequences from class I chitinases, making them more similar to members of the class IV chitinases (Neuhaus 1999). With the exception of these short deletions, class I and IV chitinases are structurally similar. The mature proteins coded for by all three genes consist of a cysteine-rich chitin-binding domain that may serve to anchor the enzyme to the pathogen cell walls, a catalytic domain, and a hinge region that connects these two (Neuhaus 1999). Excluding the hinge region, which appears to evolve under low selective constraint, the amino acid sequences of *chiI* and Arabis *chiB* are ~64% identical. Zea *chiA* and *chiB* have amino acid sequence that are ~89% identical to one another. The amino acid sequences of *chiA* and *chiB* are 34–37% identical to Zea *chiI* and Arabis *chiB* sequences.

Because *chiA* and *chiB* appear to be the products of a relatively recent duplication, these genes offer an opportunity to test for evidence of positive selection having driven the divergence of recently duplicated defense genes. Analyses of intragenome divergence of plant nucleotide binding site-leucine rich repeat (NBS-LRR) gene family members (R-genes) in several taxa have revealed that many duplicated R-genes have diverged at replacement sites faster than at synonymous sites (*i.e.*, d_N:$d_S > 1$), providing strong evidence that positive selection has driven the divergence of duplicated genes involved in pathogen recognition (Meyers *et al.* 1998; Bergelson *et al.* 2001; Mondragon-Palomino *et al.* 2002). Evidence for positive selection, as well as the large number of NBS-LRR genes present in plant genomes, has been interpreted as evidence for a selective advantage associated with the ability to recognize a wide diversity of parasites (Hulbert *et al.* 2001). However, evidence for positive selection having driven the divergence of downstream components of plant defense (often referred to as pathogenesis-related or PR genes or those components that directly affect parasite growth and fitness) is lacking. Plant chitinases are among these downstream components.

MATERIALS AND METHODS

Sampling DNA sequences: Chitinase genes were amplified and sequenced from 13 accessions of *Zea mays* ssp. *parviglumis* (hereafter referred to as ssp. *parviglumis*), 8 accessions of *Z. diploperennis*, and 1 accession of *Tripsacum dactyloides*. Both ssp. *parviglumis* and *Z. diploperrennis* were sampled from throughout their geographic ranges (appendix). PCR conditions for all genes were 35 cycles of 1 min at 94°, 1 min at 60°, and 2 min at 72°, 1 m betaine [*N,N,N*-trimethylglycine; Sigma (St. Louis)] was added to each reaction. *chiA* and *chiB* primers (*chiA* forward *ctgcagtgttgctatctgttc*, reverse *attagtgagaattcacacatcc*; *chiB* F: *gctcaaatactgatctcactg*, R: *caacatgcatatcacagctgc*) amplified the entire coding region (~840 bases) and ~100 and 160 bases of intron and 3′ flanking sequence, respectively. *chiI* primers (F: *aagatcacaatgatgagagcc*, R: *tgattgctggatctgcgtgtag*) amplified a region that includes the entire coding region (955 bases, the *atg* start site being in the forward primer) and ~125 bases of 3′ flanking sequence. Primers were designed from *Z. mays* ssp. *mays* sequences available in GenBank (*chiI* L00973, *chiA* M84164, and *chiB* M84165).

Z. diploperennis and ssp. *parviglumis* are both outcrossing diploid species and therefore PCR products were cloned into TA vectors (Promega, Madison, WI) before sequencing. Be-

cause cloned products may contain bases that result from misincorporation by *Taq* polymerase, all DNAs that produced alleles with singletons were used as templates in one or more subsequent PCRs. The products of those reactions were cloned and sequenced. Singletons present in the products from more than one PCR reaction were assumed to be true variants; those not confirmed after sequencing the products of multiple PCR reactions were assumed to have resulted from polymerase error and excluded from analysis.

In addition to the Zea and Tripsacum sequences obtained via PCR, *chiA*, *chiB*, and *chiI*-like sequences from other Poaceae genera were obtained by searching GenBank. In addition, one Sorghum sequence was obtained from TIGR. Each of the three chitinases were used as query sequences in a BLASTN search and sequences were included in further analyses if they met the following criteria: had expected values $<1^{-10}$ for one of the Zea sequences, covered at least 80% of the coding region, and had <95% identity to a sequence from the same species that was already included in the analyses (to exclude allelic variants). Sixteen sequences met these criteria when *chiI* was used to query GenBank (*Triticum aestivum* AB029936, X76041, and AB029935; *Secale cereale* AF280437 and AB051578; *Sorghum bicolor* TIGR TC2489; *Poa pratensis* AF000966, AF000965, and AF000964; *Oryza sativa* AK061280, X56063, L40337, X56787, UO2286, and Z29962; *Hordeum vulgare* L34211). Seven sequences met these criteria when *chiA* or *chiB* were used as queries (*Sorghum arundinaceum* AF402938, *S. bicolor* AY047608, *S. halpense* AF402939, *Saccharum officianarum* AF02937, *O. sativa* AB096140, *T. aestivum* AF112966, *Z. mays* ssp. *mays* AY105600). Genealogies revealed that these sequences fell into two distinct groups (not shown), hereafter referred to as the *chiI*- and *chiA-chiB*-like genes. The amino acid sequences of the mature proteins were >68% identical within the *chiI* group and >60% identical within the *chiA-chiB* group. Amino acid sequences belonging to different groups were <41% identical. Poaceae sequences within each of the two groups are more similar to one another than to any non-Poaceae sequences in GenBank, suggesting that sequences within each group may have diverged after the establishment of the grass family; however, it is possible that these sequences are products of duplication events predating the Poaceae.

Analyses: Intraspecific diversity, θ (Watterson 1975), calculated for each of the three chitinase genes on silent (synonymous and intron) sites, were compared to likelihood estimates of genome-wide diversity estimated using the method described in Wright *et al.* (2003). The multi-locus estimate for *Z. diploperennis* was made using data from six loci, *mpi*, *wip1*, *adh1*, *c1*, *glb1*, and *waxy*; the estimate for ssp. *parviglumis* was made using data from nine loci, *adh1*, *c1*, *glb1*, *hm1*, *hm2*, *mpi*, *tb1*, *waxy*, and *wip1*. The *Z. diploperennis* data are from Tiffin and Gaut (2001a,b) and Tiffin *et al.* (2004), ssp. *parviglumis* data are from Tiffin and Gaut (2001b), Zhang *et al.* (2002), or my unpublished results (*mpi*). Patterns of intraspecific diversity at the loci used for the genome-wide diversity estimates reveal no strong evidence for departures from a neutral equilibrium model in either species. Similar estimates of genome-wide diversity for ssp. *parviglumis* were obtained using all nine loci or only those six loci that were used for the estimate of *Z. diploperennis* (not shown).

Several statistical tests, including Tajima's *D* (Tajima 1989), McDonald-Kreitman (MK; McDonald and Kreitman 1991), and Hudson-Kreitman-Aguadé (HKA; Hudson *et al.* 1987) were used to determine if patterns of diversity within Zea are consistent with expectations under a neutral equilibrium model. HKA tests were conducted using sequences from *T. dactyloides* or the population sample from ssp. *parviglumis* (for *Z. diploperennis* analysis) or *Z. diploperennis* (for ssp. *parviglumis* analysis) to estimate interspecific divergence. These tests were conducted pairwise using each of seven (for *T. dactyloides*) or nine reference loci (*adh1*, *c1*, *glb1*, *waxy*, *hm2*, *wip1*, *mpi*, and the two chitinase loci that were not the subject of the analysis).

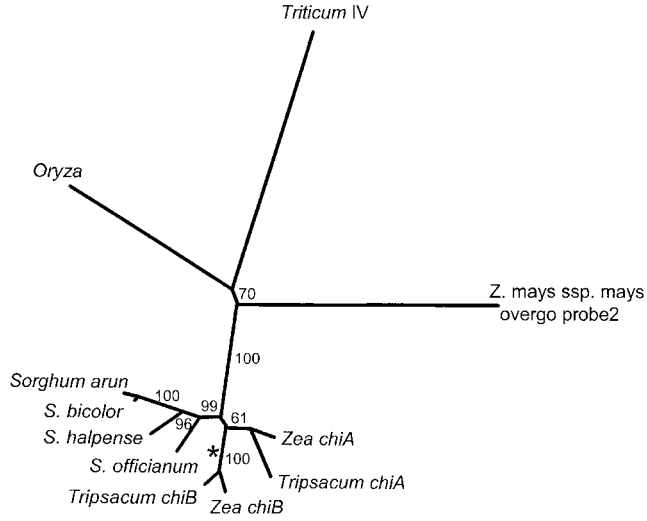

Figure 1.—Neighbor-joining genealogy showing relationships among *chiA-chiB*-like sequences from Poaceae. Bootstrap values are shown next to the branches. The *chiA* and *chiB* sequences from Zea and Tripsacum are from this study; others were obtained from GenBank (see materials and methods for accession numbers). The asterisk indicates the branch with the highest estimated d_N:d_S.

McDonald-Kreitman tests on *chiA* and *chiB* were conducted using two different outgroups. First, to test for differences in divergence between taxa, MK tests were conducted using the orthologous sequence from *T. dactyloides*. Second, to test for nonneutral divergence within the Zea genome, MK tests were conducted using the chitinase paralog as an outgroup. For *chiI*, for which there is no known recent duplicate within Zea and for which repeated efforts failed to produce a Tripsacum sequence, MK tests were conducted using a sequence from *S. bicolor* (TIGR TC2489) as an outgroup. The significance of MK tests was determined using a G-test with a William's correction for small sample size. All estimates of diversity and tests of nonneutral evolution were conducted using DnaSP version 3.53 (Rozas and Rozas 1999). Prior to analyses sequences were visually aligned using BioEdit (Hall 1999).

Relative rate tests using the methods of Fitch (1976) and Tajima (1993) as implemented in MEGA2.1 (Kumar *et al.* 2001) were used to examine heterogeneity in the rates of *chiA* and *chiB* divergence since duplication. These analyses were conducted using a sequence from *O. sativa* (AB096140) as an outgroup. The Zea *chiA* (AY532768) and *chiB* (AY532722) sequences came from the same ssp. *parviglumis* individual; using other Zea sequences produced similar results. *O. sativa* was used because the *chiA* and *chiB* sequences from Zea are more similar to one another than either is to the Oryza sequence, indicating that the duplication that resulted in *chiA* and *chiB* occurred after the divergence of the Zea and Oryza lineages. In contrast, there is only weak support for this duplication event having occurred after the divergence of the Zea and Sorghum lineages, and genealogies with likelihood scores not significantly different from the maximum-likelihood genealogy (Figure 1) included either the *chiA* or the *chiB* sequences with the Sorghum sequences, rather than with each other.

Heterogeneity in d_N:d_S among branches in the genealogies of *chiI*-like sequences and *chiA-chiB*-like sequences was tested by comparing the goodness of fit of a model that allowed separate d_N:d_S for each branch of the genealogy with a model that fit a single d_N:d_S for the entire genealogy (model 1 *vs.*

model 0 in CODEML in PAML; Yang 1998). Likelihood models, as implemented in PAML, were also used to test for evidence of positive selection having acted on specific codons (Yang *et al.* 2000). Specifically, a model that allowed for three classes of d_N:d_S (NSsites = 3) was compared to a model that assumes a single d_N:d_S value for all codons (NSsites = 0). The goodness of fit of these models were compared using a likelihood-ratio test. The genealogies used for PAML analyses were constructed using a heuristic search under the criterion of maximum parsimony in PAUP* (Swofford 1998). To assure that results from PAML analyses were not dependent on the hypothesized genealogical relationships, I analyzed the data using the five most parsimonious trees as well as all trees with likelihood scores not significantly less than the likelihood of the maximum-likelihood tree. Results from these analyses were qualitatively similar to those obtained with the most parsimonious tree, differing primarily in the number of amino acids identified as having low posterior probabilities of having been targets of positive selection (data not shown). A comparison of models allowing for approximately continuous distributions of ω, PAML M7 *vs.* M8 conducted with the maximum-likelihood trees, produced results qualitatively similar to the categorical models and only the results from the discrete class models are reported. Reported results from PAML analyses were obtained using sequences from the same ssp. *parviglumis* individual (*chiI* accession AY532743, *chiA* AY532768, and *chiB* AY532722) to represent Zea; using other Zea sequences produced similar results.

After putative targets of positive selection were identified, 2 × 2 contingency tests were used to determine if codons identified as possible targets of selection were overrepresented among (1) codons identified as targets of positive selection in Arabis or (2) the active-site cleft of the molecule, a pattern detected in Arabis (Bishop *et al.* 2000). For these analyses the active-site cleft, defined as regions of the molecule that come within 0.6 nm of a bound substrate, was based on the regions of the molecule used by Bishop *et al.* (2000). This region was based on a *H. vulgare* crystal structure (Hart *et al.* 1995) and is referred to as the chitinase binding site by Brameld and Goddard (1998). Amino acids identified as possible targets of positive selection were mapped onto the three-dimensional structure of barley class II chitinase (Protein Data Bank code 1CNS-A) with bound hexa-*N*-acetyl-D-glucosamine (Brameld and Goddard 1998) using VMD software (Humphrey *et al.* 1996).

Interspecific analyses were conducted on the chitin-binding domain and catalytic regions of the gene only. As such, two short regions of coding sequence were removed: the signal sequence, which is not part of the mature protein, and the "hinge" region that connects the chitin-binding and catalytic domains (Huynh *et al.* 1992). These regions were excluded because it was not possible to align these regions from different genera with any confidence, consistent with the expectations that they evolve under low selective constraint.

RESULTS

Intraspecific diversity: With one exception, the three Zea chitinase genes, *chiA*, *chiB*, and *chiI*, exhibit levels of diversity in both ssp. *parviglumis* and *Z. diploperennis* that are typical of other nuclear genes that have been sampled from these species (Figure 2). Therefore, levels of diversity at these five loci (exempting *chiA* in *Z. diploperennis*) reveal no evidence for nonneutral evolution. Similarly, tests of nonneutral evolution, including Tajima's *D* (Table 1) and HKA (all $P > 0.4$), reveal no evidence for positive selection or long-lived polymorphisms.

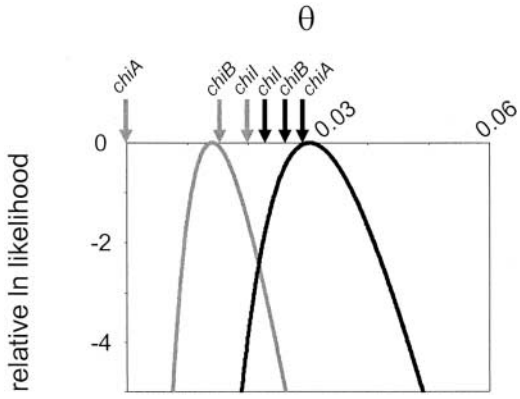

Figure 2.—Maximum-likelihood estimates of genome-wide nucleotide diversity, θ, in *Z. diploperennis* (shaded line and shaded arrows) and ssp. *parviglumis* (solid line and solid arrows). Estimates for *Z. diploperennis* were based on silent sites at six loci; nine loci were used for the ssp. *parviglumis* estimate. Diversity detected at the chitinase genes were not used in making the genome-wide estimates and are designated with arrows at the top.

McDonald-Kreitman tests, however, revealed significant departures from a neutral-equilibrium model for two genes, *chiI* in ssp. *parviglumis* and *chiB* within *Z. diploperennis* (Table 2).

In contrast to the typical levels of diversity found at chitinase loci in ssp. *parviglumis* and *chiB* and *chiI* in *Z. diploperennis*, all of the *chiA* sequences sampled from *Z. diploperennis* were identical. The absence of polymorphic sites results in diversity at *chiA* that is lower than diversity found at any other gene that has been sampled from this or any other nondomesticated Zea species (White and Doebley 1999; Tiffin and Gaut 2001a,b; Whitt *et al.* 2002; Zhang *et al.* 2002; Tiffin *et al.* 2004). Diversity at *chiA* is, however, similar to diversity at *hm2* within *Z. diploperennis*, a gene that appears to have experienced a recent and strong selective sweep (Tiffin *et al.* 2004). HKA tests used to compare diversity and divergence at *Z. diploperennis chiA* to other nuclear genes confirmed that the evolution of this gene is not typical of other nuclear loci (all $P < 0.01$ when using ssp. *Parviglumis*; all $P < 0.001$ when using Tripsacum as the outgroup). An HKA test conducted with ssp. *parviglumis hm2* data was, however, not significant, suggesting that the pattern of *chiA* diversity and divergence is not significantly different from the evolution of this apparent target of strong positive selection. Because of the absence of segregating sites, Tajima's *D* and MK tests were not conducted on the *Z. diploperennis chiA* sample.

Diversification of Poaceae chitinase genes: Previous analyses revealed that adaptive changes that differentiate chitinase genes from different Arabis species are concentrated within the active-site cleft (Bishop *et al.* 2000). To determine if a similar pattern is observed among Poaceae sequences, codon-based likelihood analyses similar to those used for analyses of the Arabis data were used to identify which codons, if any, appear to have

TABLE 1

Number of sequences, segregating sites, haplotypes, estimates of diversity calculated on silent (θ_S) and replacement (θ_R) sites, and values of Tajima's *D*

Gene	Species	No. of sequences	Segregating sites	Haplotypes	θ_S	θ_R	Tajima's *D*
chiA	ssp. *parviglumis*	13	51	13	0.029	0.007	−0.74
	Z. diploperennis	8	0	1	0	0	–
chiB	ssp. *parviglumis*	14	54	14	0.026	0.008	−0.22
	Z. diploperennis	7	32	6	0.020	0.0065	−0.51
chiI	ssp. *parviglumis*	15	27	14	0.022	0.013	−0.63
	Z. diploperennis	9	19	6	0.014	0.003	−0.81

diverged in response to positive selection. For both *chiI* and *chiA-chiB* data sets, models that included a class of codons with $d_N:d_S > 1$ fit the data significantly better than models with only a single codon class (PAML model 3 *vs.* model 0; *chiI*, $-2 \ln L = 402.8$, $P_{df=4} < 0.001$; *chiA-chiB*, $-2 \ln L = 106.3$, $P_{df=4} < 0.001$). The estimate of $d_N:d_S$ for the elevated *chiI* class was, however, relatively low ($d_N:d_S = 1.244$). The low $d_N:d_S$ value suggests that if divergence at these sites reflects positive selection and not relaxed selective constraint, then the strength of selection was not strong. Alternatively, selection may have been strong but the signal of selection is obscured by high divergence at synonymous sites (the *chiI*-like sequences from Poaceae had an average d_S of 0.44). The elevated codon class for *chiA-chiB* had $d_N:d_S = 2.58$ although only 3 codons had posterior probabilities >0.9 of being members of this class. In contrast, 26 of the *chiI* codons had posterior probabilities >0.95 of belonging to the positively selected class.

Taken together, the codon-specific tests provide only weak evidence for positive selection having driven the interspecific divergence of chitinase genes within Poaceae. Nevertheless, these analyses identify codons that are most likely to have evolved in response to positive selection, which can be compared to the 15 sites identified as having evolved in response to positive selection in Arabis. For *chiI*, 6 of the 26 codons with posterior probabilities of belonging in the selected class in Poaceae were also identified as targets of positive selection in Arabis, significantly more than expected by chance (Figures 3 and 4; Fisher's exact test, $P < 0.02$). However, unlike the pattern detected in Arabis, the positively selected codons were not significantly overrepresented in the active-site cleft of the molecule (Fisher's exact test, $P > 0.5$; Figures 3 and 4). For *chiA-chiB*, none of the three codons with posterior probabilities >0.9 of belonging in the positively selected class was identified as a target of positive selection in Arabis or in the *chiI*-like sequences nor did any of these candidates of positive selection in *chiA-chiB* fall into the active-site cleft. Moreover, there were no fixed differences between ssp. *parviglumis* and *Z. diploperennis* at these sites, as may be expected if positive selection at one of these sites were responsible for the absence of diversity at the *Z. diploperennis chiA* locus.

Divergence of recent duplicates: Relative rate tests used to compare the evolutionary rates of the *chiA* and

TABLE 2

Distribution of fixed and polymorphic synonymous (S) and replacement (R) sites and the significance of MK tests

Gene	Species	Outgroup	Fixed synonymous	Fixed replacement	Polymorphic synonymous	Polymorphic replacement	*P*
chiA	ssp. *parviglumis*	Tripsacum	10	19	11	10	0.21
		Paralog	16	36	11	10	0.09
	Z. diploperennis	Tripsacum	10	22	0	0	–
		Paralog	17	38	0	0	–
chiB	ssp. *parviglumis*	Tripsacum	7	9	8	10	0.97
		Paralog	19	29	8	11	0.85
	Z. diploperennis	Tripsacum	7	12	9	3	0.022
		Paralog	17	31	9	3	0.015
chiI	ssp. *parviglumis*	Sorghum	55	23	15	1	0.034
	Z. diploperennis	Sorghum	56	23	3	3	0.33

For *chiA* and *chiB*, sequences from *T. dactyloides* and the paralog within the species (*i.e.*, ssp. *parviglumis chiA* is the paralog for ssp. *parviglumis chiB*) were used as the outgroup.

```
                                                    XXX XXXXXXXXX
Arabis  RQAGGALCPNGLCCSEFGWC GDTEPYCKQPGCQSQCTPGG TPPGPTPTGDLSIIISSSQF DDMLMHRNDAACPARGFYTY NAFITAAKSFPNFGTTGDTT
chiI    S.....L...C....Q.... .S.SD..GS-......SGSC GS--------VAS..PE.L. NQ..L....A..P.N..... AG..A..NA..G.....APD

                                 *                                *
Arabis  TRKKEIAAFFGQTSHETTGG WASAPDGPFSWGYCFKQEVN PNS--DYCEPSATWPCAPG KRYYGRGPMQLSWNYNYGLC GRAIGVDLLNNPDLVANDAV
chiI    VQ.R.L...LA..... ..T.....YA......E.QG GA.GPD..E.SAQ....A. .K......I.I.Y.....PA .Q...AGI.A......T.PT

Arabis  IAFKAAIWFWMTAQPPKPSC HAVIAGQWQPSDADSAAGRL PGYGVITNIINGGLECGRGQ DGRVADRIGFYQRYCNIF GVNPGGNLDCYNQRPFGNGLLDAAI*
chiI    VS.ET.V.....P.S..... .D.MT...T..AA.T..... .....V..........H.A .S.........K...DLL ..SY.D....A..T..-NG*
```

FIGURE 3.—Aligned Arabis *chiB* and *Z. mays* ssp. *parviglumis chiI* amino acid sequences of mature chitinase proteins. Amino acids identified as targets of positive selection are shaded; the active-site residues are underlined and in boldface type; the two catalytic residues are designated by an asterisk above the amino acid. Indels are indicated with a "–"; amino acids identical in the *chiI* and *Arabis* sequences are indicated by a period. The hinge region is designated by *XXX*, the region of the molecule before the hinge. The Zea *chiI* sequence is GenBank AY532743; the Arabis sequence is GenBank AAF69783. The positively selected amino acids in Arabis were identified by BISHOP *et al.* (2000).

chiB sequences from ssp. *parviglumis* revealed that *chiB* has evolved significantly faster than *chiA* when either all sites or only replacement sites were analyzed (Figure 5; $P = 0.022$, $P = 0.008$, respectively). Likelihood analyses also detected significant heterogeneity in the d_N:d_S ratios along branches of the genealogy describing the relationships among *chiA-chiB*-like genes (PAML model 1 fit the data significantly better than M0; $-2 \ln L = 51.01$; $P_{df=19} < 0.001$). This analysis identified the branch that separates the node connecting the *chiA-chiB* divergence to the divergence of *chiB* in the Zea and Tripsacum lineages (designated by an asterisk in Figure 1) as having higher d_N:d_S than other branches in the genealogy (21.9 replacement, 5.3 synonymous substitutions, d_N:$d_S = 1.47$). However, neither likelihood nor 2×2 contingency tests reject the hypothesis that this branch had d_N:$d_S = 1$ and thus the elevated rate is consistent with relaxed constraint. Direct estimates of d_N:d_S between Zea *chiA-chiB* genes ranged from 0.51 to 0.80 (32.5–40.5 replacement and 21.5–13.5 synonymous changes, depending on the sequences compared).

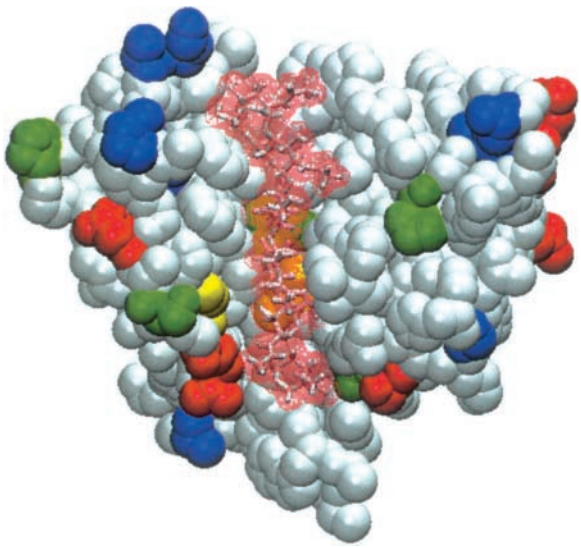

FIGURE 4.—Chitinase structure with bound hexa-*N*-acetyl-D-glucosamine ligand. Putative positively selected amino acids identified in Poaceae only are blue, those identified in Arabis only are green, and those identified in Poaceae and Arabis are red; the catalytic residues are yellow.

DISCUSSION

Diversity and evolutionary histories: Examination of three chitinase genes in two Zea species revealed that the pattern or strength of selection acting on defense may be highly variable over time between closely related species and among proteins with similar biochemical functions. Two of these genes, *chiB* and *chiI*, harbor levels of diversity in both *Z. diploperennis* and ssp. *parviglumis* that are typical of other nuclear genes. However, there is some evidence that both of these genes have evolved nonneutrally; MK tests reveal significant depar-

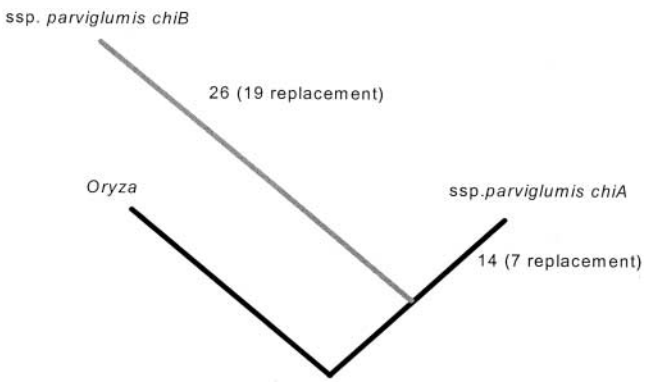

FIGURE 5.—Number of total and replacement changes that differentiate *chiA* and *chiB*. The placement of the changes was inferred using a sequence from Oryza as an outgroup. Relative rate tests conducted on all changes ($P = 0.022$) or replacement changes ($P = 0.019$) indicate significant heterogeneity in evolutionary rates following duplication. Similar results were obtained with all *Z. diploperennis* and *Z. mays* ssp. *parviglumis* sequences.

tures from a neutral-equilibrium model for *chiB* in *Z. diploperennis* and for *chiI* in ssp. *parviglumis*.

The results from MK tests should be viewed with some caution for two reasons. First, both of these genes exhibit high codon bias (ENC = 31–34, GC_{3rd} = 93–96%). When the synonymous class used in a MK test is subject to purifying selection, as expected if codon bias is maintained by selection, then a significant MK test is evidence that replacement sites experience less negative selection coefficients than synonymous sites—but it does not indicate whether replacement sites experience negative, positive, or no selection (Akashi 1995). The second reason to be cautious is that significant MK tests may result from a combination of slightly deleterious replacement mutations and recent increases in population sizes (McDonald and Kreitman 1991; Eyre-Walker 2002). Although neither of these possibilities can be excluded, previously analyzed Zea genes also have high codon bias and these genes do not routinely result in significant MK tests (Hilton and Gaut 1998; White and Doebley 1999; Tenaillon *et al.* 2001; Tiffin and Gaut 2001a,b; Zhang *et al.* 2002), as may be expected if the forces are responsible for the significant MK tests detected for *chiI* and *chiB* are not gene specific.

If the significant MK tests are due to selection, then the apparent discrepancy between MK and tests that rely on patterns of intraspecific diversity may be indicative of episodic selection. For example, selective events may have driven the divergence among species but the most recent event has been masked by the accumulation of neutral mutations. Alternatively, the apparent discrepancy may be due to a pattern of selection that is not easily detected by tests that rely solely on intraspecific diversity (de Meaux and Mitchell-Olds 2003).

In contrast to the lack of evidence for recent selection on either *chiB* or *chiI*, an absence of intraspecific diversity indicates that *chiA* or a very tightly linked locus has experienced a recent and strong selective sweep in *Z. diploperennis*. There was, however, no evidence for nonneutral evolution of this gene in ssp. *parviglumis*. Several other putative defense genes also harbor distinctly different patterns of diversity in closely related taxa. For example, positive selection appears to have driven divergence of a basic chitinase among Arabis species (Bishop *et al.* 2000) but the apparent ortholog harbors neutral levels of diversity in *A. thaliana* (Kawabe and Miyashita 1999). Similarly, *hm2*, which encodes a nitrate reductase that protects plants from infection by the fungal pathogen *Cochliobolus carbonum*, appears to have experienced a recent and strong selective sweep in *Z. diploperennis*, harbors three classes of highly diverged alleles in the closely related species *Z. perennis* (Tiffin *et al.* 2004), and may be the subject of an ongoing sweep in ssp. *parviglumis* (Zhang *et al.* 2002).

The variation in the selective histories of defense genes that is being revealed by molecular population genetic analyses is consistent with our knowledge of the evolutionary ecology of host-parasite interactions. Temporal variation in selection may result from the evolution of mechanisms that allow parasites to circumvent host defense, a basic assumption of coevolutionary models, or parasite population outbreaks and host shifts, both of which occur irregularly (Thompson 1994). Interspecific variation may also result from host shifts or differences in parasite loads, which can be highly variable even between closely related species. Finally, variation in the strength of selection acting on specific defense genes may result from variation in the efficacy of defense against different parasite species or genotypes.

Although patterns of diversity suggest that natural selection acting on Zea chitinases is highly variable, this conclusion must be made with one caveat regarding sampling: The collection of ssp. *parviglumis* DNAs sampled for this study came from throughout the species range. Estimating diversity from samples taken from a subdivided population may skew the frequency distribution of segregating sites within a sample and bias tests of nonneutral evolution (Slatkin 1987; Hudson 1990). Moreover, if there is strong population structure and/or local adaptation of defense traits, as predicted by the geographic mosaic theory of coevolution (Thompson 1999), then this sampling scheme may obscure selective events that occurred within geographically separated subpopulations. *Z. mays* ssp. *parviglumis* is distributed in three geographically separated areas of Mexico (Sanchez and Ordaz 1987; Doebley 1990) but whether population structuring strongly affects the distribution of DNA sequence polymorphisms in this species has not been thoroughly investigated. Nevertheless, tests based on sequence polymorphism detected little evidence for nonneutral evolution in ssp. *parviglumis* and thus the sampling strategy used in this study does not seem to have caused erroneous evidence for selection on these genes. The sampling strategy is also unlikely to have biased the comparison of chitinase genes to genome-wide diversity, given that genome-wide diversity was estimated using data collected from similar species-wide samples. In contrast to the geographically widespread distribution of ssp. *parviglumis*, *Z. diploperennis* is confined to a relatively small geographic area in Jalisco, Mexico (Sanchez and Ordaz 1987) and population structure is unlikely to have a strong effect on patterns of sequence diversity in this species.

Comparison with Arabis: Striking results from Bishop *et al.*'s (2000) analyses of Arabis class I chitinases were that these genes diverged rapidly and that most positively selected amino acids, as revealed by analysis of interspecific divergence, were located in the active-site cleft of the molecule. This pattern was interpreted as evidence for coevolution of plant chitinases with components of pathogen cell walls or, more likely, chitinase inhibitors. In contrast, neither *chiA-chiB* nor *chiI* sites identified as possible targets of positive selection in Poaceae were more likely to fall within than outside of

the active-site cleft. However, the *chiI* codons identified as targets of selection were significantly overrepresented among the codons previously identified as targets of selection in Arabis. However, whether these amino acids evolve as rapidly in Poaceae as in Arabis is unclear; the Arabis data analyzed by Bishop *et al.* (2000) were sampled from recently diverged species (mean K_s <0.03) whereas the Poaceae data analyzed here were sampled from more distantly related taxa (mean K_s ~0.44). Determining whether Poaceae chitinases diverged as rapidly as they diverged in Arabis will require more extensive sampling of chitinase sequences from closely related Poaceae species.

Although positively selected amino acids in Poaceae are not significantly overrepresented in the active-site cleft, three of the six codons identified as being positively selected in both taxonomic groups were located within the active-site cleft. Two of these three were near the Glu89 (Glu138 in Figure 3) reactive site in a region of the molecule predicted to comprise a flexible loop important for bringing the substrate into close proximity to the Glu89 active residue (Brameld and Goddard 1998). Three others fell outside of it: two within the catalytic region and one within the CBD, a region of the molecule that appears to be important for anchoring chitinase to the chitin substrate (Taira *et al.* 2002). Functional analyses have confirmed that amino acid changes at one of these putatively positively selected residues (Asp146 in Figure 3) can cause up to 50% reduction in chitinase activity (Ohnuma *et al.* 2002). The effects that substitutions at the other two residues have on chitinase activity are not available; however, they appear to be promising targets for functional analyses.

Regardless of the function of specific residues, finding a set of positively selected codons in both the dicot genus Arabis and monocot family Poaceae suggests that some of the adaptive responses of chitinase enzymes may be similar across a wide range of taxa. The evolutionary response common to both Poaceae and Arabis presumably reflects responses to selection imposed by pathogens utilizing similar mechanisms to circumvent plant chitinases. Nevertheless, the majority of putative positively selected sites were not common to both taxa, possibly reflecting adaptive response to selective pressures that are lineage specific. Lineage-specific adaptive responses in defense proteins were also observed by Stotz *et al.* (2000) who found no overlap among positively selected codons in polygalacturonase inhibitor proteins from legume and nonlegume dicots (polygalacturonases are pathogen enzymes that degrade plant cell walls).

Evolution of duplicates *chiA* and *chiB*: Because plants, and other hosts, are attacked by a wide variety of parasites, each of which is likely to express a unique combination of ligands and counterdefense mechanisms, it seems reasonable to expect that selection will favor the functional divergence of duplicated defense genes. The strong evidence for plant R-gene family members having diverged in response to positive selection supports the idea that variation in defense mechanisms is selectively advantageous (Ellis *et al.* 2000; Bergelson *et al.* 2001). Similarly, balancing selection appears to be important in maintaining diversity at MHC in vertebrates (Edwards and Hedrick 1998; Hughes and Yeager 1998). Selection also appears to have driven the divergence of the reactive centers of duplicated serine protease inhibitor genes in mammals (Hill and Hastie 1987).

The evidence for the duplicated chitinases *chiA* and *chiB* having diverged in response to positive selection is, however, equivocal. On the one hand, *chiB* has evolved significantly faster than *chiA*, particularly at replacement sites, the likelihood estimate of d_N:d_S along the branch that separates the *chiB* sequence sample from Zea and Tripsacum from the *chiA-chiB* ancestor is >1, and four of the five codons identified as targets of selection within Poaceae also differ between *chiA* and *chiB* from Zea and Tripsacum. The high d_N:d_S after duplication is consistent with a burst of positive selection following duplication, followed by purifying selection once a new function has evolved (Hughes 1994). On the other hand, the estimate of d_N:d_S was not significantly >1 and the identification of positively selected codons is not independent of the pairwise differences between *chiA* and *chiB*. Therefore, the possibility that *chiB* has evolved in response to relaxed selective constraint rather than positive selection cannot be rejected.

Functional analyses of *chiA* and *chiB* activity offer little help in trying to differentiate the action of relaxed and positive selection. A limited study involving only five fungal pathogens revealed growth of three pathogens to be more effectively inhibited by the *chiA* than by the *chiB* protein (Huynh *et al.* 1992). The lower efficacy of *chiB* may suggest relaxed selective constraint; however, this conclusion would be extremely tentative, given that few pathogens were investigated. If positive selection has driven the evolution of *chiB*, then this enzyme may have evolved greater activity against some fungal pathogens, perhaps involving tradeoffs with the ability to inhibit the growth of fungi used in Huynh *et al.* (1992).

If selection has driven the divergence of *chiA* and *chiB*, it appears the strength of selection is weaker than that which has driven the divergence of some duplicated R-genes. Selection acting on duplicated chitinases may be weaker because of the wide diversity of defenses that are induced following pathogen attack. If several different defenses interact additively to inhibit fungal growth, then the selective pressure acting on any specific mechanism may be fairly weak. In contrast, most R-genes are thought to be involved in gene-for-gene interactions (Stahl *et al.* 1999; Ellis *et al.* 2000; Bergelson *et al.* 2001) in which only one plant R-gene effectively recognizes any particular pathogen.

I thank John Bishop, Joy Bergelson, and three anonymous reviewers for comments that improved an earlier version of this manuscript, John Bishop and Tony Dean for discussions about the evolution and

function of chitinase enzymes, Stephen Wright for providing the computer program used to calculate likelihood estimates of diversity, Tony Dean for assistance in mapping positively selected sites onto the protein structure, Brandon Gaut for providing seeds used for this study, and John Doebley for clarifying accession information. This research was supported by National Science Foundation grant DEB-0235027.

LITERATURE CITED

Akashi, H., 1995 Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in Drosophila DNA. Genetics **139:** 1067–1076.

Berenbaum, M. R., 1983 Coumarins and caterpillars: a case for coevolution. Evolution **37:** 163–179.

Bergelson, J., M. Kreitman, E. A. Stahl and D. C. Tian, 2001 Evolutionary dynamics of plant R-genes. Science **292:** 2281–2285.

Bishop, J. G., A. M. Dean and T. Mitchell-Olds, 2000 Rapid evolution in plant chitinases: molecular targets of selection in plant-pathogen coevolution. Proc. Natl. Acad. Sci. USA **97:** 5322–5327.

Brameld, K. A., and W. A. I. Goddard, 1998 The role of enzyme distortion in the single displacement mechanisms of family 19 chitinases. Proc. Natl. Acad. Sci. USA **95:** 4276–4281.

Broglie, K., I. Chet, M. Holliday, R. Cressman, P. Biddle *et al.*, 1991 Transgenic plants with enhanced resistance to the fungal pathogen *Rhizoctonia solani*. Science **254:** 1194.

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.

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

de Meaux, J., and T. Mitchell-Olds, 2003 Evolution of plant resistance at the molecular level: ecological context of species interactions. Heredity **91:** 345–352.

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

Edwards, S. V., and P. W. Hedrick, 1998 Evolution and ecology of MHC molecules: from genomics to sexual selection. Trends Ecol. Evol. **13:** 305–311.

Ellis, J., P. Dodds and T. Pryor, 2000 Structure, function and evolution of plant disease resistance genes. Curr. Opin. Plant Biol. **3:** 278–284.

Eyre-Walker, A., 2002 Changing effective population size and the McDonald-Kreitman test. Genetics **162:** 2017–2024.

Farrell, B. D., D. E. Dussourd and C. Mittler, 1991 Escalation of plant defense: Do latex and resin canals spur plant diversification? Am. Nat. **138:** 881–900.

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

Hall, T. A., 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41:** 95–98.

Hart, P. J., H. D. Pfluger, A. F. Monzingo, T. Hollis and J. D. Robertus, 1995 The refined crystal structure of an endochitinase from *Hordeum vulgare* L. seeds at 1.8 A resolution. J. Mol. Biol. **248:** 402–413.

Hill, R. E., and N. D. Hastie, 1987 Accelerated evolution in reactive centre regions of serine protease inhibitors. Nature **326:** 96–99.

Hilton, H., and B. S. Gaut, 1998 Speciation and domestication in maize and its wild relative: evidence from the globulin-1 gene. Genetics **150:** 863–872.

Hudson, R. R., 1990 Gene genealogies and the coalescent process. Oxf. Surv. Evol. Biol. **7:** 1–44.

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

Hughes, A. L., 1994 The evolution of functionally novel proteins after gene duplication. Proc. R. Soc. Lond. Ser. B **256:** 119–124.

Hughes, A. L., and M. Yeager, 1998 Natural selection at major histocompatibility complex loci of vertebrates. Annu. Rev. Genet. **32:** 415–435.

Hulbert, S. H., C. A. Webb, S. M. Smith and Q. Sun, 2001 Resistance gene complexes: evolution and utilization. Annu. Rev. Phytopathol. **39:** 285–312.

Humphrey, W., A. Dalke and K. Schulten, 1996 VMD: visual molecular dynamics. J. Mol. Graph. **14:** 33.

Huynh, Q. K., C. M. Hironaka, E. B. Levine, C. E. Smith, J. R. Borgmeyer *et al.*, 1992 Antifungal proteins from plants: purification, molecular cloning, and antifungal properties of chitinases from maize seed. J. Biol. Chem. **267:** 6635–6640.

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.

Kumar, S., K. Tamura, I. B. Jakobsen and M. Nei, 2001 MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17:** 1244–1245.

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.

Meyers, B. C., K. A. Shen, P. Rohani, B. S. Gaut and R. W. Michelmore, 1998 Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell **11:** 1833–1846.

Mondragon-Palomino, M. M., R. W. Michelmore, B. C. Meyers and B. S. Gaut, 2002 Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. Genome Res. **12:** 1305–1315.

Neuhaus, J. M., 1999 Plant chitinases (PR-3, PR-4, PR-8, PR-11), pp. 77–105 in *Pathogenesis-Related Proteins in Plants*, edited by S. K. Datta and S. Muthukrishnan. CRC Press, Boca Raton, FL.

Ohnuma, T., M. Yagi, T. Yamagami, T. Taira, Y. Aso *et al.*, 2002 Molecular cloning, functional expression, and mutagenesis of cDNA encoding rye (*Secale cereale*) seed chitinase-c. Biosci. Biotechnol. Biochem. **66:** 277–284.

Passarinho, P. A., and S. C. de Vries, 2002 *Arabidopsis* chitinases: a genomic survey, p. doi/10.1199/tab.0023 in *The Arabidopsis Book*, edited by C. R. Somerville and E. M. Meyerowitz. American Society of Plant Biologists, Rockville, MD (http://www.aspb.org/publications/arabidopsis).

Rausher, M. D., 2001 Co-evolution and plant resistance to natural enemies. Nature **411:** 857–864.

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

Sanchez, J. J., and L. Ordaz, 1987 El teosinte in Mexico: distribucion actual de las poblaciones. International Board for Plant Genetic Resources, Systematic and Ecogeographic Studies on Crop Genepools, No. 2, Rome.

Schlumbaum, A., F. Mauch, U. Vogelli and T. Boller, 1986 Plant chitinases are potent inhibitors of fungal growth. Nature **324:** 365–367.

Slatkin, M., 1987 The average number of sites separating DNA sequences drawn from a subdivided population. Theor. Popul. Biol. **32:** 42–49.

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.

Stotz, H. U., J. G. Bishop, C. W. Bergmann, M. Koch, P. Albersheim *et al.*, 2000 Identification of target amino acids that affect interactions of fungal polygalacturonases and their plant inhibitors. Physiol. Mol. Plant Pathol. **56:** 117–130.

Swofford, D. L., 1998 *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland MA.

Taira, T., T. Ohnuma, T. Yamagami, Y. Aso, M. Ishiguro *et al.*, 2002 Antifungal activity of rye (*Secale cereale*) seed chitinases: the different binding manner of class I and class II chitinases to the fungal cell walls. Biosci. Biotechnol. Biochem. **66:** 970–977.

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.

Tenaillon, M. I., M. C. Sawkins, A. D. Long, R. L. Gaut, J. F. Doebley *et al.*, 2001 Patterns of DNA sequence polymorphism

along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). Proc. Natl. Acad. Sci. USA **98:** 9161–9166.

Thompson, J. N., 1994 *The Coevolutionary Process.* University of Chicago Press, Chicago.

Thompson, J. N., 1999 Specific hypotheses on the geographic mosaic of coevolution. Am. Nat. **153:** S1–S14.

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., 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.

Tiffin, P., R. Hacker and B. S. Gaut, 2004 Population genetic evidence for rapid changes in intraspecific diversity and allelic cycling of a specialist defense gene in Zea. Genetics **168** (in press).

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

White, S. E., and J. F. Doebley, 1999 The molecular evolution of terminal *ear1*, a regulatory gene in the genus Zea. Genetics **153:** 1455–1462.

Whitt, S. R., L. M. Wilson, M. I. Tenaillon, B. S. Gaut and E. S. Buckler, 2002 Genetic diversity and selection in the maize starch pathway. Proc. Natl. Acad. Sci. USA **99:** 12959–12962.

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

Wu, S. C., A. L. Kriz and J. M. Widholm, 1994 Molecular analysis of two cDNA clones encoding acidic class I chitinase in maize. Plant Physiol. **105:** 1097–1105.

Yang, Z. H., 1998 Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. **15:** 568–573.

Yang, Z. H., R. Nielsen, N. Goldman and A. K. Pedersen, 2000 Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics **155:** 431–449.

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: J. Bergelson

APPENDIX

Species names, accessions, geographic origin, collector, and GenBank accession numbers

Accession	Geographic origin	Collector[a]	GenBank accession nos. *chiI*	*chiA*	*chiB*
		Z. mays ssp. *parviglumis*			
PI 384070	Guerrero, MX	HGW	AY532743	AY532768	AY532722
PI 384064	Guerrero, MX	HGW	AY532744	AY532769	AY532723
PI 384062	Guerrero, MX	HGW	AY532745	AY532770	AY532724, AY532725
PI 384061	Guerrero, MX	HGW	AY532746	AY532771	AY532726
PI 331789	Michoacan, MX	CIMMYT	AY532747	AY532772	AY532727, AY532728
PI 331788	Michoacan, MX	CIMMYT	AY532748	AY532773	AY532729
PI 331786	Mexico, MX	CIMMYT	AY532749	AY532774	AY532730
PI 331783	Guerrero, MX	CIMMYT	AY532750	AY532775	—
BFB 967	Jalisco, MX	BFB	AY532751	AY532776	AY532731, AY532732
HHI site 3	Guerrero, MX	HHI and TC	AY532752	AY532777	AY532733
GWB site 4	Guerrero, MX	GWB and TK	AY532753, AY532754	AY532788, AY532789	AY532734
PI 566691	Michoacan, MX	HGW and JS	AY532755, AY532756	AY532780	
PI 566688	Guerrero, MX	HGW and JS	AY532757	—	AY532735
		Z. diploperennis (all from Jalisco, MX)			
HHI 1250		HHI	AY532758	AY532781	AY532736
9476A		CIMMYT	AY532759	AY532782	AY532737
10003		CIMMYT	AY532760	AY532783	—
Ames 2317		—	AY532761	AY532784	AY532738
PI 441932		RG	AY532762, AY532763	AY532785	AY532739
PI 462368		RG	AY532764	AY532786	AY532740
Ames 21884		HHI	AY532765	AY532787	AY532741
PI441931		HHI	AY532766	AY532788	AY532742
		Tripsacum dactyloides			
WW2120	Oklahoma, USA	—	—	AY532767	AY532721

All seeds were provided by B. S. Gaut (accession numbers preceded by PI or Ames are from the U.S. Department of Agriculture germplasm center in Ames, IA; others are from Centro Internacional para Mejoramiento de Maiz y Trigo (CIMMYT), John Doebley, or C. L. Dewalt (*T. dactyloides*). Collector information was obtained from http://www.panzea.org, http://www.ars-grin.gov/npgs/acc/acc_queries.html, or John Doebley.

[a] HGW, H. G. Wilkes; CIMMYT, Centro Internacional para Mejoramiento de Maiz y Trigo; BFB, B. F. Benz; HHI, H. H. Iltis; TC, T. Cochrane; GWB, G. W. Beadle; TK, T. A. Kato; JS, J. J. Sanchez; RG, R. Guzman.