

DNA Variation in the Basic Chitinase Locus (*ChiB*) Region of the Wild Plant *Arabidopsis thaliana*

Akira Kawabe and Naohiko T. Miyashita

Laboratory of Plant Genetics, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto, 606-8502 Japan

Manuscript received November 3, 1998

Accepted for publication July 27, 1999

ABSTRACT

Nucleotide variation in a 2.2-kbp region of basic chitinase (*ChiB*) locus in 17 ecotypes of *Arabidopsis thaliana* was compared with previously investigated regions to investigate genetic mechanisms acting on DNA polymorphism. In the *ChiB* region, dimorphic DNA variation was detected, as in the *Adh* and *ChiA* regions. Nucleotide diversity (π) of the entire region was 0.0091, which was similar to those of the two other regions. About half of polymorphic sites (37/87) in the *ChiB* region were observed in only two ecotypes. Tajima's *D* was negative but not significantly, while Fu and Li's *D** was positive. Neither McDonald-Kreitman nor Hudson, Kreitman, Aguadé tests showed a significant result, indicating that these loci were under similar evolutionary mechanisms before and after speciation. Linkage disequilibria were observed within the three regions because of dimorphic polymorphisms. Interlocus linkage disequilibrium was not detected between the *Adh* and the two chitinase regions, but was observed between the *ChiA* and *ChiB* regions. This could be due to epistatic interaction between the two chitinase loci, which are located on different chromosomes.

IN *Arabidopsis thaliana*, DNA polymorphism was first reported in the alcohol dehydrogenase (*Adh*) region (Hanfstingl *et al.* 1994; Innan *et al.* 1996; Miyashita *et al.* 1996). In the *Adh* region, dimorphic DNA variation was observed and considered the result of balancing selection (Hanfstingl *et al.* 1994) and/or fusion of two diverged subpopulations (Innan *et al.* 1996). Similarly, dimorphism was observed in the acidic chitinase (*ChiA*) region (Kawabe *et al.* 1997) and *Cauliflower (Cal)* region (Purugganan and Suddith 1998), suggesting that the presence of the dimorphic variations could be a characteristic of the nuclear genome of *A. thaliana*. However, patterns of polymorphism (*i.e.*, the proportion of singleton and replacement polymorphisms) and results of neutrality tests varied among the regions.

To clarify the mechanisms responsible for the different patterns of polymorphism in the regions, one approach is to compare loci that have similar function. The chitinase system in plants provides a good opportunity, since the system consists of multiple genes of similar function but distinct nucleotide (or amino acid) sequences. Chitinases are enzymes that catalyze hydrolysis of chitin and play an important role in plant defense systems against fungal pathogens. Plant chitinases are classified into four classes with respect to their structures. Class I chitinase has four motifs: signal peptide, cysteine-rich (chitin binding) domain, hinge region, and catalytic domain (Shinshi *et al.* 1990; Collinge *et*

al. 1993; Hamel *et al.* 1997). Class II and IV chitinases are similar to class I chitinase. Class II chitinase lacks the N-terminal cysteine-rich domain, while class IV chitinase is distinguished from class I chitinase by four deletions (Collinge *et al.* 1993; Hamel *et al.* 1997). Class III chitinase is similar to bacterial chitinases and has lysozyme activities (Mettraux *et al.* 1989; Henrissat 1990; Jekel *et al.* 1991; Watanabe *et al.* 1992; Beintema and Terwisscha van Scheltinga 1996). Of the four classes of chitinases, only class III chitinase has no sequence similarities and may have an independent evolutionary origin (Meins *et al.* 1992; Collinge *et al.* 1993; Hamel *et al.* 1997). As with other pathogenesis-related (PR) proteins, chitinases are classified as acidic or basic according to their protein constitutions and expression patterns. The acidic and basic chitinases are expressed in the central vacuole and intracellular space, respectively (Mauch and Staehelin 1989; Mettraux *et al.* 1989). The induction mechanisms are also different between acidic and basic PR proteins. The acidic PR proteins are induced by salicylic acid, while basic PR proteins are induced by ethylene (Ohashi and Ohshiro 1992).

In *A. thaliana*, two chitinases, acidic (*ChiA*) and basic (*ChiB*) chitinases, were cloned and their DNA sequences were reported (Samac *et al.* 1990). The *ChiA* locus has three exons, coding a protein 302 amino acids (aa) long, while the *ChiB* has two exons, whose product is 336 aa long. It is not possible to align nucleotide or amino acid sequences of the two chitinases. *ChiA* and *ChiB* are located on chromosomes 5 and 3, respectively (Bell and Ecker 1994; Sato *et al.* 1998). *ChiA* encodes

Corresponding author: Naohiko Miyashita, Laboratory of Plant Genetics, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan. E-mail: arabis@kais.kyoto-u.ac.jp

a class III chitinase and *ChiB* a class I chitinase. These two chitinases were classified as acidic and basic chitinases, although both chitinases are basic proteins (Beintema and Terwisscha van Scheltinga 1996). This classification is based on their different expression patterns (Samac *et al.* 1990; Samac and Shah 1991, 1994). *ChiB* is expressed in the root tissue constitutively and induced by ethylene (Samac *et al.* 1990; Samac and Shah 1994), as are other basic PR proteins (Ohashi and Ohshiro 1992). *ChiA* is expressed in root, leaf vascular, hydathode, guard cell, and anther tissues (Samac and Shah 1991) and is induced by salicylic acid (Samac and Shah 1991), typical for acidic PR proteins (Ohashi and Ohshiro 1992).

In this study, we report intraspecific nucleotide variation in a 2.2-kbp fragment of the *ChiB* region in *A. thaliana* using the same ecotypes (ecological races) analyzed for the *ChiA* region. The main purpose of this study is to investigate whether or not polymorphic patterns in the *ChiA* region are related to protein function. If the pattern of polymorphism in a locus is related to its function, similar patterns should be found in loci of similar function.

MATERIALS AND METHODS

Plant materials: Seeds of 16 ecotypes of *A. thaliana* were obtained from Professor Nobuharu Goto, Sendai Arabidopsis Seed Stock Center, Miyagi University of Education, Sendai, Japan and were grown in soil pots in an incubator under 16-hr light and 8-hr dark conditions. These ecotypes were sampled worldwide. Sixteen of the 17 ecotypes were analyzed for *Adh* variation (Innan *et al.* 1996), and the same 17 ecotypes were analyzed for the *ChiA* region (Kawabe *et al.* 1997). To estimate divergence, a related species, *Arabidopsis gemmifera* (sampled at Ashibi, Kyoto prefecture, Japan), was used.

Sequencing: Total DNA was extracted from mature plants by a CTAB method of Weising *et al.* (1991) and used as template for PCR amplification. Primers for PCR amplification are 5' CGT CTA TTT TTA TTT TCT CCA 3' (BCH-1) and 5' TTG GTT TGA TGT TGG TTT TGT 3' (BCH-102), which amplifies a fragment about 2.5 kbp long encompassing the entire coding region of the *ChiB* locus. Two primers, BCH-1 and 5' AAG TTT AGG CTC TGA TTT ATG 3' (BCH-101), were used for *A. gemmifera* to amplify a fragment about 2.5 kbp long containing the entire coding region of the *ChiB* locus. The PCR products were cloned into plasmid pUC 18 as described in Terauchi *et al.* (1997). Sequencing reactions (ALFexpress Auto Read sequence kit, Pharmacia, Piscataway, NJ) followed the dideoxy chain termination method (Sanger *et al.* 1977) using cloned plasmid as the template. Sequencing primers were designed at ~350- to 400-bp intervals. Sequence for each ecotype was determined in both strands by a Pharmacia ALFred sequencer. Singleton variants were confirmed by sequencing more than two clones to eliminate PCR artifacts. Newly obtained nucleotide sequences were deposited in the DDBJ database under nos. AB023448–AB023464.

Sequence analyses: Seventeen sequences including a published sequence of the ecotype Columbia (Col-0; Samac *et al.* 1990) were studied. The analyzed region is located between nucleotide positions 394 and 2640 of the Columbia ecotype (Samac *et al.* 1990). Nucleotide diversity (π , Nei and Li 1979; Tajima and Nei 1984), $4N_e\mu$ (θ , Watterson 1975), the num-

ber of recombination events (R_m , Hudson and Kaplan 1985), and $4N_e c$ (C , Hudson 1987) were estimated by using DnaSP program version 2.5 (Rozas and Rozas 1997). The phylogenetic tree was constructed by the neighbor-joining (NJ) method (Saitou and Nei 1987) based on the genetic distances for the entire region calculated by the Jukes and Cantor (1969) method (PHYLIP v. 3.572; Felsenstein 1996). Nucleotide sequences of the *Adh* and *ChiA* regions were obtained from Miyashita *et al.* (1996, 1998), Innan *et al.* (1996), and Kawabe *et al.* (1997). Between all non-singleton variants in the three regions, linkage disequilibrium was analyzed by the DnaSP program (Rozas and Rozas 1997) and tested by Fisher's exact test.

RESULTS

DNA polymorphism in the *ChiB* region of *A. thaliana*:

There were 109 variants (83 nucleotide substitutions and 26 length variants) in the entire *ChiB* region (Figure 1). Of these, 22 nucleotide substitutions and 5 length variants were singletons. Length variation was observed only in noncoding regions. Five repeat number variants [poly(A) at site 500–506, poly(T) at 1335–1338, poly(A) at 1391–1396, poly(T) at 2598–2603, and poly(AT) at 1306–1333] were excluded from the following analyses.

In the coding region, 21 nucleotide polymorphisms were observed, including four replacement substitutions. The proportion of replacement nucleotide sites (19.1%) was lower than those in the *ChiA* (43.2%) and *Adh* (31.6%) regions. The high proportion of replacement sites in *ChiA* could be specific to the locus. Only one replacement site variant, a non-singleton, was observed in *ChiB*; it caused a nonconservative amino acid change, between negatively charged Asp and positively charged Lys (Miyata *et al.* 1979). However, the site was located at amino acid position 17, which is included in the N-terminal region of 30 amino acids that is cut off from mature protein (Samac *et al.* 1990). Therefore, this nonconservative polymorphic site is not related to chitinase function. The other three amino acid changes, all singletons, were conservative.

Nucleotide diversity (π) of the *ChiB* region was estimated (Table 1). Nucleotide diversity of the entire region was 0.0091. This value was almost the same as those of the *ChiA* (0.0104), *Adh* (0.0080), and *Cal* (0.0070) regions. Table 1 also summarizes the results of Tajima's (1989) and Fu and Li's (1993) tests. The D test statistics were negative in most regions, while D^* was positive in some regions. None of the tests gave a significant result, except for Fu and Li's test on the 3' flanking region. The *ChiB* region had an excess of doublet sites (Figure 2), which contrasted with the other three loci. This excess of doublet polymorphisms caused contrary results for Tajima's and Fu and Li's tests. The significantly positive D^* value under Fu and Li's test for the 3' flanking region reflected that only 1 of 21 polymorphic sites in the region was a singleton polymorphism. In the *ChiA* region, the two test statistics were significantly negative, because of excess singletons (Kawabe *et al.* 1997). These

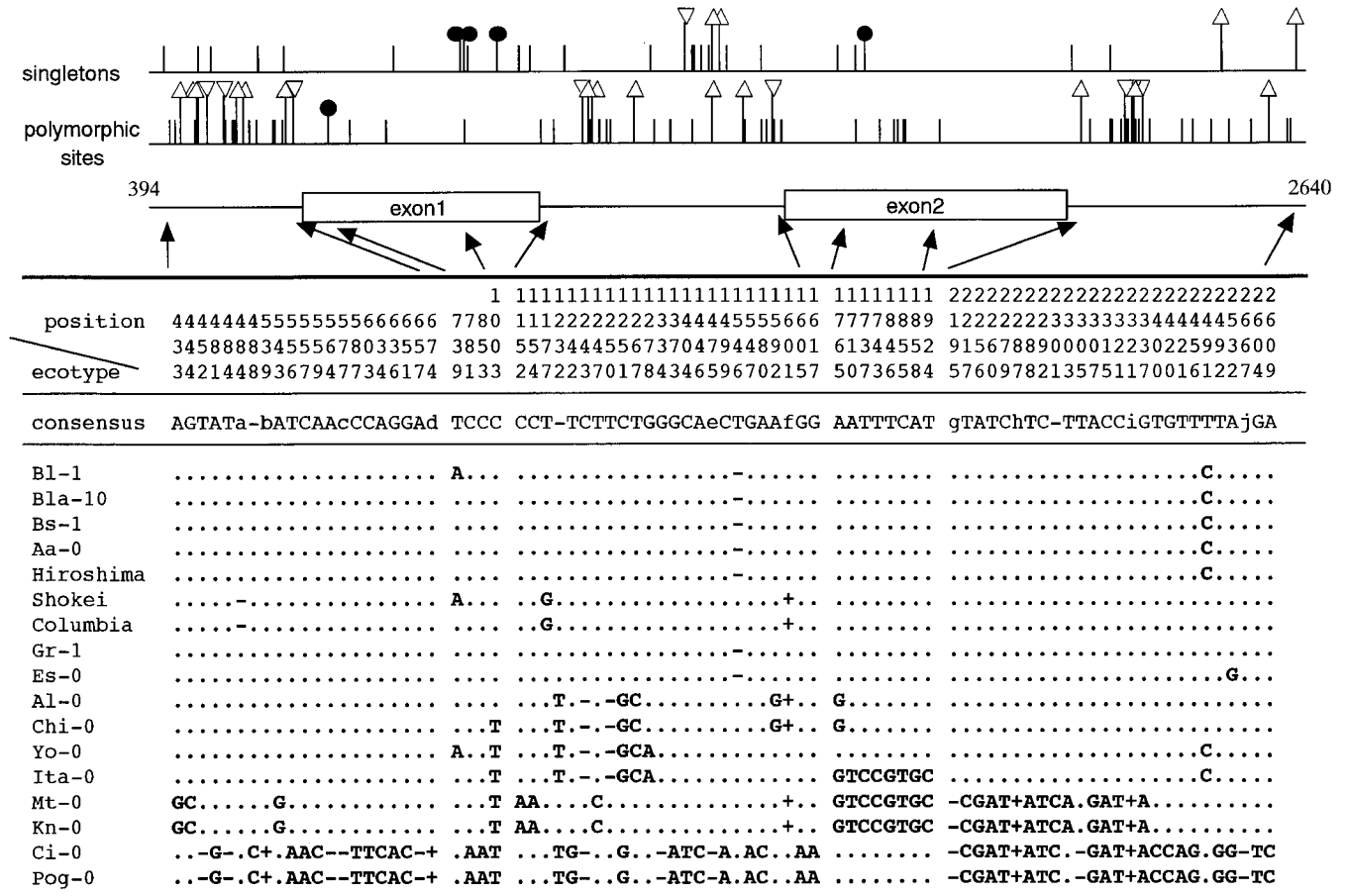


Figure 1.—Summary of DNA variations detected in the *ChiB* region of *A. thaliana*. Vertical bars indicate nucleotide variations, of which replacement changes are shown with solid circles. Open triangles mean indels. Non-singleton changes are summarized at the bottom, where dot indicates nucleotide identical to consensus, + indicates presence of indel, and minus indicates absence. Sequences of indels, indicated by lowercase letters, are a, TTA; b, AA; c, ACTGT; d, CTGGATACTAC; e, TCTTGATTAATCTAAAC GCAAT; f, TTAA; g, AATATAATAC; h, AT; i, ATATT; and j, GATTC.

chitinase loci also differed in the proportion of singleton nucleotide sites.

Intragenic recombination in the *ChiB* region of *A. thaliana*: As in the *Adh*, *ChiA*, and *Cal* regions, non-singleton polymorphisms found throughout the *ChiB*

region were dimorphic (Figure 1), confirming previous results that dimorphism was a characteristic of DNA variation in the *A. thaliana* nuclear genome. It was certain that intragenic recombination had occurred in the *ChiB* region, especially in exon 2. Two ecotypes, Ci-0

TABLE 1

Summary of polymorphism in the *ChiB* region of *A. thaliana*

Region	Size	Polymorphic sites	Indels	θ	π	D	D^*
All regions	2168	82	25	0.0112	0.0091	-0.810 NS	0.319 NS
5' flanking	287	16	9	0.0165	0.0109	-1.323 NS	0.312 NS
Coding region	1008	21	0	0.0062	0.0049	-0.795 NS	-0.573 NS
Synonymous	240.29	17	0	0.0209	0.0179	-0.558 NS	-0.189 NS
Replacement	767.71	4	0	0.0015	0.0009	-1.353 NS	-1.569 NS
Exon 1	460	10	0	0.0064	0.0043	-1.222 NS	-1.285 NS
Intron	427	24	8	0.0166	0.0117	-1.187 NS	0.115 NS
Exon 2	548	11	0	0.0059	0.0055	-0.289 NS	0.190 NS
3' flanking	446	21	8	0.0129	0.0146	0.203 NS	1.337 *

D and D^* are test measures of Tajima's and Fu and Li's tests. *, significant at 5%; NS, not significant.

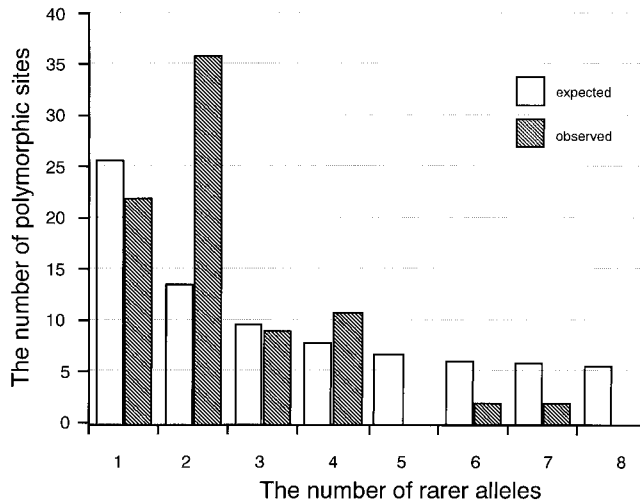


Figure 2.—Frequency spectrum of polymorphic nucleotide sites in the *ChiB* region. The expected value was obtained according to Tajima (1989).

and Pog-0, which have 39 distinct variants, were diverged from other ecotypes. But these two ecotypes are not obvious parental types of any evidently recombinant ecotypes. Complicated partitioning in the *ChiB* region and the absence of an informative outgroup sequence made it impossible to use the methods previously used for detecting the number and times of recombination events (Innan *et al.* 1996; Kawabe *et al.* 1997). Instead, we estimated the minimum number of recombination events (R_m , Hudson and Kaplan 1985) and $4Nc$ (C, Hudson 1987) to describe intragenic recombination in the *ChiB* region, together with the *Adh* and *ChiA* regions (Table 2). The values per site for *ChiB* were intermediate between those of the two regions. These two chitinase loci again differed in the rate of recombination. The estimated C values for each of the three regions were always smaller than estimates of θ , indicating that recombination has occurred less frequently than nucleotide mutation. Although estimates of $4N\mu$ (θ) were relatively constant over the three regions, those of C varied widely. This result meant that recombination rate in *A. thaliana* was not constant, but strongly depended on region.

Genealogical relationship of ecotypes based on nucleotide variations in the *ChiB* region:

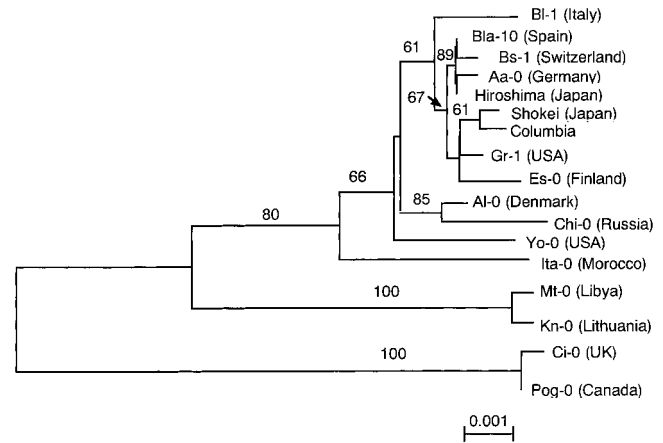


Figure 3.—Neighbor-joining tree of ecotypes based on the nucleotide variation in the *ChiB* region. Bootstrap probabilities >50% are shown above branches.

otide variations in the *ChiB* region: An estimate of genealogical relationships among ecotypes was obtained by the NJ method, based on nucleotide variation in the entire *ChiB* region (Figure 3). Because there have been some recombination events, this NJ tree may not be greatly informative. However, there were clearly three distinct clusters. The first cluster included 13 ecotypes, the second cluster included Mt-0 and Kn-0, and the third cluster consisted of Ci-0 and Pog-0. Nucleotide diversities within clusters II and III were low (0.0005 and 0.0009, respectively), while that of cluster I was 0.0075, which was similar to the overall estimate. This result suggested that sequences in cluster I were relatively older than those in clusters II and III. Average nucleotide distances between clusters were 0.0147, 0.0221, and 0.0223 for I and II, I and III, and II and III, respectively. For the *Adh* and *ChiA* regions, average nucleotide distances between the two most divergent sequence types were 0.0127 and 0.0148, respectively. Thus, the *ChiB* region contained more divergent sequences than the other two regions.

As in the *Adh* and *ChiA* regions, no association between phylogenetic clustering and sample locations was observed (Figure 3). In particular, the two ecotypes in each of clusters II and III, which had low nucleotide

TABLE 2
Summary of estimated recombination at the three regions

Region	Size length	Per gene			Per site		
		θ	C	R_m	θ	C	C/θ
<i>Adh</i>	2380	19.12 ± 3.08^a	4.2 ± 1.90	6	0.0080 ± 0.0013	0.0017 ± 0.0008	0.2197
<i>ChiA</i>	1751	19.02 ± 1.95	0.001 ± 0.47	2	0.0109 ± 0.0011	$0.0000^b \pm 0.0003$	0.0001
<i>ChiB</i>	2168	19.62 ± 1.99	1.3 ± 1.02	5	0.0090 ± 0.0009	0.0006 ± 0.0005	0.0663

^a All the standard deviations were obtained by bootstrap of 100 replications.

^b 5.6×10^{-7} .

TABLE 3

Result of McDonald-Kreitman test for the *ChiB* locus

Comparisons	Replacement	Synonymous
Polymorphic in <i>A. thaliana</i>	4	17
Fixed between <i>A. thaliana</i> and <i>A. gemmifera</i>	18	27

$$G = 2.91; P = 0.088.$$

diversity, came from different continents. These results support the hypothesis that the *A. thaliana* population has spread over the world recently (King *et al.* 1993; Price *et al.* 1994; Innan *et al.* 1996, 1997).

Interspecific variation between *A. thaliana* and *A. gemmifera* in the *ChiB* coding region: A McDonald-Kreitman test (McDonald and Kreitman 1991) was conducted to examine the ratio of polymorphic replacement and silent sites in *A. thaliana* relative to divergence between *A. thaliana* and *A. gemmifera* (Table 3). Because only one sequence was determined in *A. gemmifera*, the divergence could include both fixed sites between *A. thaliana* and *A. gemmifera* and some polymorphic sites within *A. gemmifera*. None of the comparisons gave significant results, although the ratio of replacement to silent changes for the interspecific comparison was higher than that for the intraspecific comparison. This result indicated that ratio of replacement to synonymous substitutions was constant in intra- and interspecific variations.

To examine the difference in levels of nucleotide variation between polymorphism and divergence among the *Adh*, *ChiA*, and *ChiB* regions, a Hudson, Kreitman, Aguadé test (Hudson *et al.* 1987) was conducted, comparing the number of polymorphic sites in coding regions and using *A. gemmifera* as a reference species (Table 4). None of the comparisons were significant, indicating that the level of polymorphism was consistent with divergence among the three regions.

Linkage disequilibrium within and between loci: Because most of the ecotypes studied for the *ChiB* region were also used for analyses of the *Adh* and *ChiA* regions (Innan *et al.* 1996; Kawabe *et al.* 1997), it was possible to analyze intra- and interlocus linkage disequilibrium between polymorphic DNA variants in the three regions (Figure 4). Since all three regions showed dimorphism, significant linkage disequilibria were observed within each region. The proportions of linkage disequilibria that were significant at least at the 5% level were 368 of 1035 tests (36%), 173 of 351 (49%), and 1651 of 3403 (49%) for *Adh*, *ChiA*, and *ChiB*, respectively. Sign tests (Lewontin 1995) were applied to test the overall degree of linkage disequilibrium (Table 5). In each region, an excess of coupling (positive) linkage disequilibrium was detected. Between the *Adh* and either of the two chitinase regions, almost no pairs of sites showed linkage disequilibrium significant at the 5% level [12 of 5060 tests (0.2%)]. On the other hand, between the two chitinase regions, 428 of 2241 tests (19%) were significant at the 5% level. The significant interlocus linkage disequilibria were caused by two divergent ecotypes Ci-0 and Pog-0, which had unique polymorphisms in the two regions. In the *ChiB* region, 27 nucleotide substitutions and 12 indels were uniquely found in Ci-0 and Pog-0. Almost all of these variants were located in noncoding regions, with the exception of two synonymous substitutions in exon 1. It should be mentioned that because of the enormous number of comparisons, none of the tests were significant when the Bonferroni correction was applied.

DISCUSSION

Patterns of DNA variation among the *Adh*, *ChiA*, and *ChiB* regions: One of the main purposes of this study was to distinguish whether or not the polymorphic patterns in the *ChiA* region, *i.e.*, excess of rare alleles and higher replacement polymorphism, were related to the

TABLE 4

Summary of the results of Hudson, Kreitman, Aguadé test for the *ChiB* region

Region	<i>r</i> ^a	No. of sites		No. of poly sites		<i>T</i>	θ_1	θ_2	χ^2	<i>P</i>	
		Within	Between	Within	Between						
<i>ChiB</i>	Total	17	1008	1008	22	52.47					
	Synonymous	17	240.29	240.30	17	34					
	Replacement	17	767.71	767.70	4	18.47					
vs. <i>Adh</i>	Total	17	1137	1137	19	44.21	6.97	0.007	0.005	0.00	0.97
	Synonymous	17	252.67	253	13	30.06	6.21	0.020	0.016	0.04	0.83
	Replacement	17	884.33	884	6	14.14	10.03	0.002	0.002	0.56	0.45
vs. <i>ChiA</i>	Total	17	906	906	44	48.82	4.19	0.009	0.012	1.38	0.24
	Synonymous	17	203.98	205.16	25	25.94	3.81	0.026	0.030	0.88	0.35
	Replacement	17	702.02	700.84	19	22.88	5.08	0.003	0.006	2.46	0.12

^a The number of *A. thaliana* ecotypes.

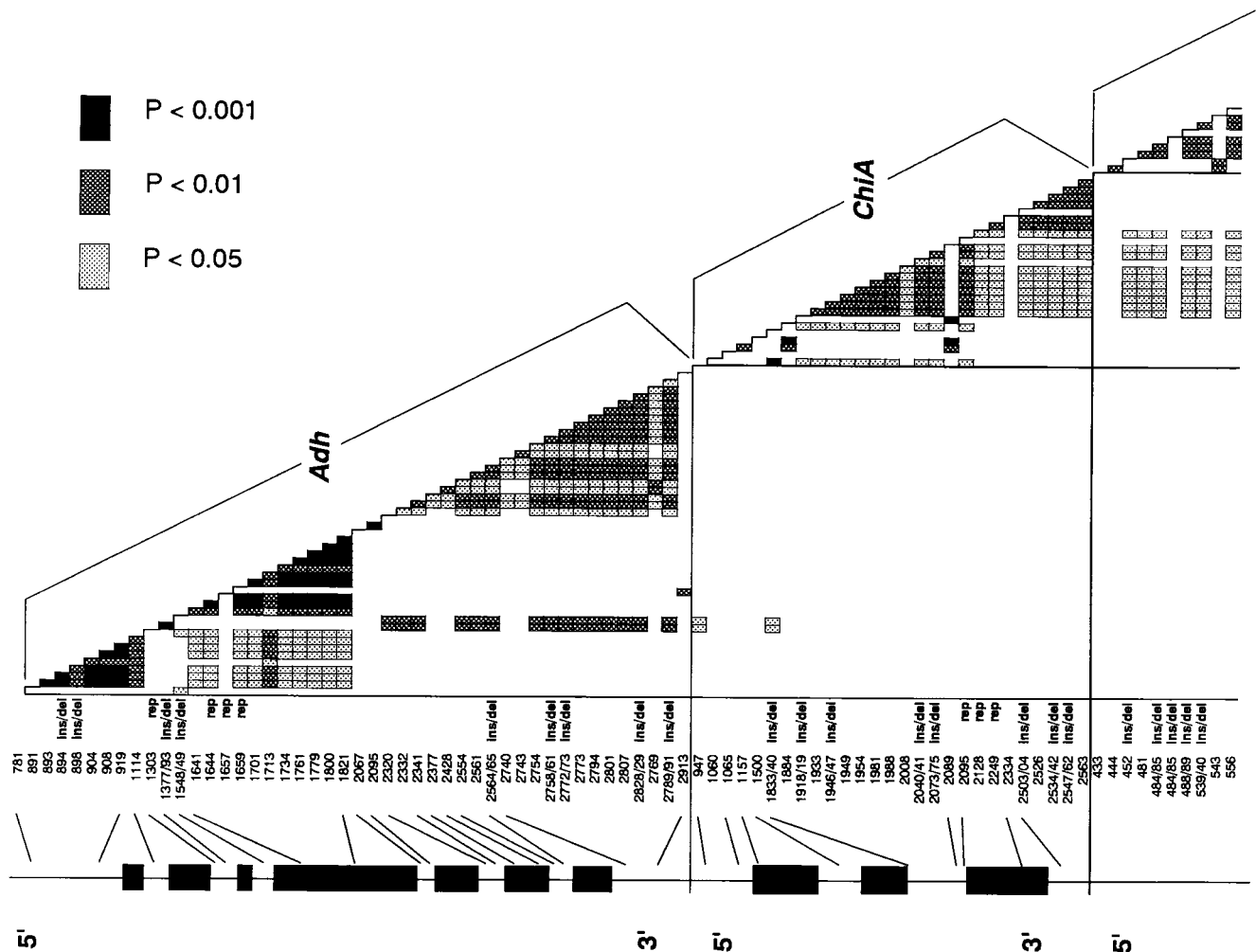


Figure 4.—Intra- and interlocus linkage disequilibrium between polymorphic DNA variations in the *Adh*, *ChiA*, and *ChiB* regions. Polymorphic variations in each region are arranged from 5' to 3'. Statistical significance based on Fisher's exact test is shown.

function of chitinase protein. None of the *ChiA*-specific polymorphic patterns were observed in the *ChiB* region, except for asymmetric partitioning of dimorphism. In both regions, dimorphism was maintained at low frequencies. These results suggested that the patterns of polymorphism in the *ChiA* region were not related to chitinase function, but might be specific to the locus. Therefore, another approach is necessary to determine the causes of higher replacement polymorphism or high proportion of rarer alleles in the *ChiA* region.

In *Drosophila melanogaster*, recombination rate is related to level of intraspecific DNA variations (Aguadé *et al.* 1991; Begun and Aquadro 1992) as a result of hitchhiking and/or background selection (Maynard Smith and Haigh 1974; Charlesworth *et al.* 1993). *A. thaliana* is a self-pollinated plant and its outcrossing rate was estimated to be only $\sim 0.3\%$ (Abbott and Gomes 1989). In such highly selfing species, the whole genome may act as a linked locus and the effects of hitchhiking or background selection may be similar for

all loci (Charlesworth *et al.* 1993). Estimated recombination rates were higher in the *Adh* and *ChiB* regions than in the *ChiA* region (Table 2). The reason why the recombination rate in the *ChiA* region was reduced could not be clarified in this study. However, the rates of synonymous substitution were also related to estimated recombination rates (C) of the three regions in *A. thaliana* (d.f. = 1, $r = 0.99996$, $P < 0.01$), as shown in *D. melanogaster* (Moriyama and Powell 1996). This relationship indicates that recombination may play an important role in determining levels of intraspecific variation in *A. thaliana*.

In the three regions, C/θ , which reduced to c/μ , ranged from 0.0001 to 0.2197 and the average was 0.0823. In *D. melanogaster*, this ratio was estimated as 1.6 in the *Adh* region (Hudson 1987). If the mutation rate has been similar between *D. melanogaster* and *A. thaliana*, the lower value of C/θ in *A. thaliana* suggests less recombination in this plant species and is consistent with the difference in breeding system of these organisms.

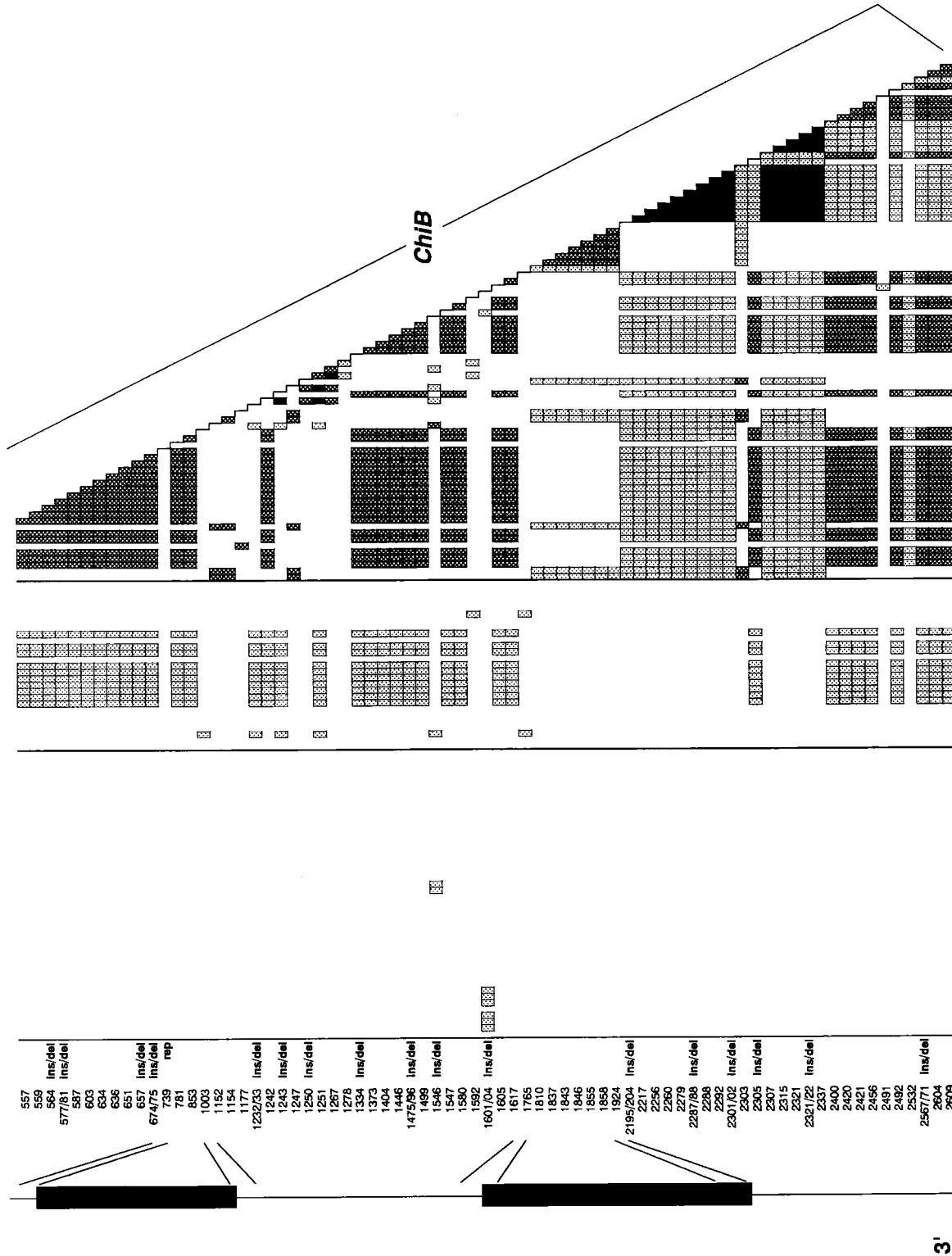


Figure 4.—(Continued)

Interchromosomal linkage disequilibrium between the two chitinase regions: Although statistical significance was not confirmed for linkage disequilibrium, it is clear that the two ecotypes have a divergent sequence type in the two chitinase loci. To explain this pattern in the chitinase loci, which lie on different chromo-

somes, two hypotheses could be considered. First, the pattern may represent a consequence of random assortment of the chromosomes, where the two loci are located, after formation of a heterozygote between two divergent sequence types. Although this plant species is highly selfing, previous studies of nucleotide variation

TABLE 5

The sign test applied to *Adh*, *ChiA*, and *ChiB* regions of *A. thaliana*

Region	Test based on sign of <i>D</i> ^a			Test based on negative <i>D</i> ^a			
	No. of (k, m) classes ^a	Summed <i>G</i>	<i>P</i>	No. of negative <i>D</i>		<i>G</i> (d.f. = 1)	<i>P</i>
				Obs.	Exp.		
<i>Adh</i>	13	21.56	<i>P</i> < 0.10	5	21.13	14.80	<i>P</i> < 0.005
<i>ChiA</i>	9	18.11	<i>P</i> < 0.05	9	15.39	3.09	<i>P</i> < 0.10
<i>ChiB</i>	11	33.56	<i>P</i> < 0.005	20	47.68	19.71	<i>P</i> < 0.005

^a See Lewontin (1995).

in the *Adh* and *ChiA* loci (Innan *et al.* 1996; Kawabe *et al.* 1997) showed that recombination events had occurred. Because recombination events can be detected only in the presence of heterozygous individuals, some outcrossing must have occurred in the evolutionary history of *A. thaliana*. Once this combination is formed in the two chitinase loci, it will be maintained by selfing at the species level. The second explanation for the pattern in the chitinase loci could be epistatic interaction between the genes. Since *ChiA* and *ChiB* are functionally related, this hypothesis is attractive. At this moment, we do not have any supporting evidence. To distinguish the possibilities, it may be necessary to study DNA variation in surrounding regions of the chitinase loci and compare chitinase activity between ecotypes.

We are grateful to N. Goto, Sendai Arabidopsis Stock Center, Miyagi University of Education, for *A. thaliana* seeds and to F. Tajima, M. Uyenoyama, and E. Stahl for comments and improving the manuscript. This article is contribution number 556 from the Laboratory of Plant Genetics, Graduate School of Agriculture, Kyoto University.

LITERATURE CITED

- Abbott, R. J., and M. F. Gomes, 1989 Population genetic structure and outcrossing rate of *Arabidopsis thaliana* (L.) Heynh. *Heredity* **62**: 411–418.
- Aguadé, M., N. T. Miyashita and C. H. Langley, 1991 Reduced variation in the *yellow-achaete-acute* region in natural populations of *Drosophila melanogaster*. *Genetics* **122**: 607–616.
- Begun, D. J., and C. E. Aquadro, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**: 519–520.
- Beintema, J. J., and A. C. Terwisscha van Scheltinga, 1996 Plant lysozyme, pp. 75–86 in *Lysozyme: Model Enzymes in Biochemistry and Biology*, edited by P. Jolles. Birkhauser Verlag, Basel.
- Bell, C. J., and J. R. Ecker, 1994 Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* **19**: 137–144.
- Charlesworth, B., M. T. Morgan and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–1303.
- Collinge, D. B., K. M. Kragh, J. D. Mikkelsen, K. K. Nielsen, U. Rasmussen *et al.*, 1993 Plant chitinases. *Plant J.* **3**: 31–40.
- Felsenstein, J., 1996 PHYLIP (Phylogeny inference package), ver. 3.572. University of Washington Press, Seattle.
- Fu, Y. X., and W. H. Li, 1993 Statistical tests of neutrality of mutation. *Genetics* **133**: 693–709.
- Hamel, F., R. Bolvin, C. Tremblay and G. Bellemere, 1997 Structural and evolutionary relationships among chitinases of flowering plants. *J. Mol. Evol.* **44**: 614–624.
- Hanfstingl, U., A. Berry, E. A. Kellog, J. T. Costa III, W. Rudiger *et al.*, 1994 Haplotype divergence coupled with lack of diversity at the *Arabidopsis thaliana* alcohol dehydrogenase locus: role for both balancing and directional selection? *Genetics* **138**: 811–828.
- Henrissat, B., 1990 Weak sequence homologies among chitinases detected by clustering analysis. *Protein Sequences Data Anal.* **3**: 523–526.
- Hudson, R. R., 1987 Estimating the recombination parameter of finite population model without selection. *Genet. Res.* **50**: 245–250.
- Hudson, R. R., and N. L. Kaplan, 1985 Statistical properties of the number of recombination events in the history of a sample of DNA sequence. *Genetics* **111**: 147–164.
- Hudson, R. R., M. Kreitman and M. Aguadé, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153–159.
- Innan, H., F. Tajima, R. Terauchi and N. T. Miyashita, 1996 Intra-genic recombination in the *Adh* locus of the wild plant *Arabidopsis thaliana*. *Genetics* **143**: 1761–1770.
- Innan, H., R. Terauchi and N. T. Miyashita, 1997 Microsatellite polymorphism in natural population of the wild plant *Arabidopsis thaliana*. *Genetics* **146**: 1441–1452.
- Jekel, P. A., J. B. H. Hartmann and J. J. Beintema, 1991 The primary structure of hevamine, an enzyme with lysozyme/chitinase activity from *Heves brasiliensis* latex. *Eur. J. Biochem.* **200**: 123–138.
- Jukes, T. H., and C. R. Cantor, 1969 Evolution of protein molecules, pp. 21–132 in *Mammalian Protein Metabolism*, edited by H. Munro. Academic Press, New York.
- Kawabe, A., H. Innan, R. Terauchi and N. T. Miyashita, 1997 Nucleotide polymorphism in the acidic chitinase locus (*ChiA*) region of the wild plant *Arabidopsis thaliana*. *Mol. Biol. Evol.* **14**: 1303–1315.
- King, G., J. Nienhuis and C. Hussey, 1993 Genetic similarity among ecotypes of *Arabidopsis thaliana* estimated by analysis of restriction fragment length polymorphisms. *Theor. Appl. Genet.* **86**: 1028–1032.
- Lewontin, R. C., 1995 The detection of linkage disequilibrium in molecular sequence data. *Genetics* **140**: 377–388.
- Mauch, F., and L. A. Staehelin, 1989 Functional implications of the subcellular localization of ethylene-induced chitinase and β -1,3-glucanase in bean leaves. *Plant Cell* **1**: 447–457.
- Maynard Smith, J., and J. Haigh, 1974 The hitchhiking effect of a favorable gene. *Genet. Res.* **23**: 23–55.
- McDonald, J. H., and M. Kreitman, 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654.
- Meins, F., Jr., J. M. Neuhaus, C. Sperisen and J. Ryals, 1992 The primary structure of plant pathogenesis-related glucanohydrolases and their genes, pp. 245–282 in *Genes Involved in Plant Defense*, edited by T. Boller and F. Meins, Jr. Springer Verlag, New York.
- Mettraux, J. P., W. Burkhardt, M. Moyer, S. Dincher, W. Middeledt *et al.*, 1989 Isolation of a complementary DNA encoding a chitinase with structural homology to a bifunctional lysozyme/chitinase. *Proc. Natl. Acad. Sci. USA* **86**: 896–900.
- Miyashita, N. T., H. Innan and R. Terauchi, 1996 Intra- and interspecific variation in the alcohol dehydrogenase locus region of wild plants *Arabis gemmifera* and *Arabidopsis thaliana*. *Mol. Biol. Evol.* **13**: 433–436.

- Miyashita, N. T., A. Kawabe, H. Innan and R. Terauchi, 1998 Intra- and interspecific DNA variation and codon bias of alcohol dehydrogenase (*Adh*) locus in *Arabis* and *Arabidopsis* species. *Mol. Biol. Evol.* **15**: 1420–1429.
- Miyata, T., S. Miyazawa and T. Yasunaga, 1979 Two types of amino acid substitutions in protein evolution. *J. Mol. Evol.* **12**: 219–236.
- Moriyama, E. N., and J. R. Powell, 1996 Intraspecific nuclear DNA variation in *Drosophila*. *Mol. Biol. Evol.* **13**: 261–277.
- Nei, M., and W.-S. Li, 1979 Mathematical model for studying genetic variation in terms of restriction endonuclease. *Proc. Natl. Acad. Sci. USA* **76**: 5269–5273.
- Ohashi, Y., and M. Ohshiro, 1992 Stress-induced expression of genes for pathogenesis-related protein in plants. *Plant Cell Physiol.* **33**: 819–826.
- Price, R. A., J. D. Palmer and I. A. Al-Shehbaz, 1994 Systematic relationships of *Arabidopsis*: a molecular and morphological perspective, pp. 7–20 in *Arabidopsis*, edited by E. M. Meyerowitz and C. R. Somerville. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Purugganan, M. D., and J. I. Suddith, 1998 Molecular population genetics of the *Arabidopsis* *CAULIFLOWER* regulatory gene: non-neutral evolution and naturally occurring variation in floral homeotic function. *Proc. Natl. Acad. Sci. USA* **95**: 8130–8134.
- Rozas, J., and R. Rozas, 1997 DnaSP version 2.0: a novel software package for extensive molecular population genetics analysis. *Comp. Appl. Biosci.* **13**: 307–311.
- Samac, D. A., and D. M. Shah, 1991 Developmental and pathogen induced activation of the *Arabidopsis* acidic chitinase protein. *Plant Cell* **3**: 1063–1072.
- Samac, D. A., and D. M. Shah, 1994 Effect of chitinase antisense RNA expression on disease susceptibility of *Arabidopsis* plants. *Plant Mol. Biol.* **25**: 587–596.
- Samac, D. A., C. M. Hirochika, P. E. Yalalaly and D. M. Shah, 1990 Isolation and characterization of the genes encoding basic and acidic chitinase in *Arabidopsis thaliana*. *Plant Physiol.* **93**: 907–914.
- Saitou, N., and M. Nei, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
- Sanger, F., S. Nicklen and A. R. Coulson, 1977 DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463–5467.
- Sato, S., T. Kaneko, H. Kotani, Y. Nakamura, E. Asamizu *et al.*, 1998 Structural analysis of *Arabidopsis thaliana* chromosome 5. IV. Sequence features of the regions of 1,456,315 bp covered by nineteen physically assigned P1 and TAC clones. *DNA Res.* **5**: 41–54.
- Shinshi, H., J. M. Neuhaus, J. Ryals and F. Meins, Jr., 1990 Structure of tobacco endochitinase gene: evidence that different chitinase genes can arise by transposition of sequence encoding cysteine-rich domain. *Plant Mol. Biol.* **14**: 357–368.
- Tajima, F., 1989 Statistical test for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- Tajima, F., and M. Nei, 1984 Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* **1**: 269–285.
- Terauchi, R., T. Terachi and N. T. Miyashita, 1997 DNA polymorphism at the *Pgi* locus of a wild yam, *Dioscorea tokoro*. *Genetics* **147**: 1899–1914.
- Watanabe, T., W. Oyanagi, K. Suzuki, K. Ohnishi and H. Tanaka, 1992 Structure of the gene encoding chitinase D of *Bacillus circulans* WL-12 and possible homology of the enzyme to other prokaryotic chitinases and class III plant chitinases. *J. Bacteriol.* **174**: 408–414.
- Watterson, G. A., 1975 On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* **7**: 256–276.
- Weising, K., D. Kaemmer, F. Weigand, J. T. Epplen and G. Kahl, 1991 Oligonucleotide fingerprinting reveals various probe-dependent levels of informativeness in chickpea (*Cicer arietinum*). *Genome* **35**: 436–442.

Communicating editor: M. K. Uyenoyama

