Dimorphic DNA variation in the anionic peroxidase gene AtPrx53 of Arabidopsis thaliana

Evgenia V. Kupriyanova, Tatiana A. Ezhova, and Sergei V. Shestakov

1Department of Genetics, Faculty of Biology, Moscow State University, Vorob’evy gory 1/12, Moscow, 119992 Russia
2Vavilov Institute of General Genetics, Russian Academy of Sciences, ul. Gubkina 3, Moscow, 119991 Russia

(Received 21 March 2007, accepted 3 September 2007)

The DNA polymorphism in the AtPrx53 gene which encodes anionic peroxidase was analyzed in 20 Arabidopsis thaliana accessions. There are two divergent sequence types (Col and Dj-like haplotypes) in the AtPrx53 gene that differ by 2 indel and 16 non-singleton nucleotide polymorphisms including 5 nucleotide polymorphic sites responsible for 4 deduced amino acid replacements. Two of the amino acid substitutions (Phe/Ser at position 180 and Asp/Asn at position 270) could be responsible for the difference in electrophoretic mobility of AtPrx53 allozymes. One of them (Phe/Ser at position 180) lies within the hypervariable region, indicating that this amino acid polymorphism is subjected to balancing selection. The revealed difference between deduced allozymes is related to the dimorphism in mobility of three major anionic peroxidase isoforms which according to previously established data encoded by AtPrx53 gene. The haplotype Col which included 12 accessions from three different continents is characterized by faster mobility of three isoforms in comparison with the Dj haplotype represented by eight accessions. There is a significant association between the haplotype and several developmental traits: leaf number, flowering time, main stem height etc. Lines of the Dj haplotype have shorter duration of vegetative stages and flower earlier than most of Col haplotype accessions. The reasons of this association are discussed.

Key words: Arabidopsis ecotypes, anionic peroxidase, AtPrx53 dimorphism, nucleotide diversity

INTRODUCTION

Plant peroxidases mediating many vital reactions from germination to senescence are encoded by tens of paralogous genes in plant genomes (Gaspar et al., 1982; Penel et al., 1992; Welinder, 1992). The sequenced genomes of Arabidopsis proved to contain 73 peroxidase genes, distributed among all chromosomes (Tognolli et al., 2002; Welinder et al., 2002); however, the data on the functions of particular peroxidase genes and on the level of their variation remain scarce. Previously we have analyzed nucleotide sequences of five peroxidase genes AtPrx52-AtPrx56 located on the upper arm of chromosome 5 for six Arabidopsis thaliana accessions (ecotypes Col, Dj-M, Bla-M, En-M, lines Ler and K-156). The level of intraspecific nucleotide variation significantly differed between these genes (up to 20 times): the greatest diversity was observed in the tandemly duplicated AtPrx53 and AtPrx54 genes (0.00915 and 0.00703, respectively), while the least variation was observed in AtPrx56 gene (0.00048). AtPrx53 and AtPrx54 genes demonstrated allelic dimorphism; this was expressed for AtPrx53 also on the deduced protein level (Kupriyanova et al., 2006). The AtPrx53 gene codes for highly active anionic peroxidase present in the plant roots and aerial parts and involved in lignification (Ostergaard et al., 1996; Ostergaard et al., 2000). As shown previously, AtPrx53 gene encodes a protein represented by three isoforms (Lebedeva et al., 2003; Lebedeva et al., 2004). In the present study, DNA polymorphism in AtPrx53 gene and anionic peroxidase spectrum in additional 15 ecotypes were analyzed. Dimorphic polymorphism in AtPrx53 related to electrophoretic mobility of three anionic isoforms encoded by this gene was demonstrated.
MATERIALS AND METHODS

Plant material and growth conditions  Arabidopsis thaliana ecotypes were obtained from the seed collection of the Genetics Department of Moscow State University (Yanushkevich, 1985) and the Arabidopsis Biological Resource Center (ABRC, Ohio State University, Columbus) (Table 1).

Plants of different ecotypes were grown in a 2:1 soil-sand mixture under greenhouse conditions. Several quantitative developmental traits were studied: rosette and cauline leaf number (RLN and CLN, respectively), flowering time (FT - the number of days from the date of sowing to the inflorescence emergence when first flower buds are visible), length of main stem at 40 days after sowing (L40), final main stem length (Lmax), rosette shoot number (RSN), senescence time (ST – days from date of sowing to the rosette senescence). From 10 to 15 plants per line were grown in randomized flats in two replicates for each accession. The mean for the traits examined was calculated for combined samples. Seedlings grown in aseptic culture were used to isolate DNA (20–30 seedlings per DNA sample).

Polyacrylamide gel electrophoresis  Proteins were extracted from the leaves (separately from the main ribs and the rest of the leaf blades), stems, flower buds, flowers, and siliques of adult plants. Weighed tissue samples were stirred in 0.4 M Tris-HCl buffer (pH 6.8) with 10% sucrose. The tissue-buffer ratio was 1:1. The homogenate was centrifuged at 13000 g for 15 min. Vertical polyacrylamide gel electrophoresis (PAGE) was carried out as described by Davis (1964) with modifications. The isoforms were numbered according to Lebedeva et al. (2003).

Generation of PCR fragments for sequencing  DNA isolation was performed according to standard protocol (Dallaporta et al., 1983). The AtPrx53 gene was amplified using primers described earlier (Kupriyanova et al., 2006). The AtPrx53 sequence of ecotypes was determined by direct sequencing of the PCR product.

Nucleotide sequence analysis  Translation in silico was performed according to the Translate program (http://cn.expasy.org/tools). Multiple alignments of nucleotide and amino acid sequences were done in program ClustalW (Higgins et al., 1996), alignments were displayed by using GeneDoc 2.6.002. Peroxidase gene sequences from the Columbia ecotype were available at http://www.arabidopsis.org. Sequence of a full-length AtPrx53 gene copy from the line Ler was available at

<table>
<thead>
<tr>
<th>Table 1. Plant materials used in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ecotype</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Dj-M (Dijon)*</td>
</tr>
<tr>
<td>En-M (Enkheim)*</td>
</tr>
<tr>
<td>Bla-M (Blanes)*</td>
</tr>
<tr>
<td>Kl-M (Koln)</td>
</tr>
<tr>
<td>Li-M (Limburg)</td>
</tr>
<tr>
<td>Pet-M (Petersgoph)</td>
</tr>
<tr>
<td>Est-M (Estland)</td>
</tr>
<tr>
<td>Col-M (Columbia)*</td>
</tr>
<tr>
<td>Ler (Landsberg erecta)*</td>
</tr>
<tr>
<td>Ws (Wassilewskija)</td>
</tr>
<tr>
<td>Sav-0 (Slavice)</td>
</tr>
<tr>
<td>Kas-1 (Kashmir)</td>
</tr>
<tr>
<td>Cvi-0 (Cape Verde Islands)</td>
</tr>
<tr>
<td>Co-1 (Coimbra)</td>
</tr>
<tr>
<td>Hi-0 (Hilversum)</td>
</tr>
<tr>
<td>Ita-0 (Ibel Tazekka)</td>
</tr>
<tr>
<td>Gel-1 (Geleen)</td>
</tr>
<tr>
<td>Wei-0 (Weiningen)</td>
</tr>
<tr>
<td>An-1 (Antwerpen)</td>
</tr>
<tr>
<td>Eri-1 (Eriengsbona)</td>
</tr>
</tbody>
</table>

*– analyzed previously (Kupriyanova et al., 2006).
Polymorphism in the AtPrx53 gene of Arabidopsis thaliana

http://www.arabidopsis.org/Cereon/index.jsp. Level of intraspecific nucleotide diversity was estimated as mean pairwise differences (\(\pi\); Nei, 1987). Deviation of the observed nucleotide variation from the neutral evolution was assessed using D (Tajima, 1989) and D* (Fu and Li, 1993) tests. Linkage disequilibrium between polymorphic sites (their nonrandom distribution in genes) was evaluated using Z_{ns} test (Kelly, 1997). The minimum number of recombination events (\(R_m\)) was calculated according to Hudson and Kaplan (1985). All calculations were performed using the DnaSP program version 4.0 (Rozas et al., 2003). One-way analyses of variance (ANOVA) were performed using a standard statistical software package (Statistica version 6.0; StatSoft, Tulsa, OK, USA).

RESULTS

Nucleotide variation at the AtPrx53 gene of A. thaliana We examined nucleotide variation in the AtPrx53 gene in the samples of 20 accessions (19 ecotypes and Ler line) of A. thaliana. The sequenced region was 1.29 kb, including the entire transcriptional unit. A total of 29 nucleotide polymorphisms (20 informative sites, five singleton variable sites, four indels) were detected in the AtPrx53 gene (Fig. 1). Length variation was observed only in noncoding region of the 2nd intron.

Half of the informative nucleotide polymorphisms (13 of 20) were clustered in exon 3, introns 2 and 3 (in a 279-bp window). Of these, five occurred within a 42-bp region of exon 3. Twenty nucleotide substitutions and four indels were found more than once; all these polymorphic sites have two alternative nucleotides. Based on these polymorphisms, the 20 investigated accession sequences were classified into two haplotypes in AtPrx53 gene – Dj-or Col-like. All studied alleles were either identical to the Col or Dj sequences or differed by no more than four SNP or indel (singleton and nonsingleton) from either Col or Dj. In addition to singleton mutations specific to individual ecotype (Eri-l, Wei-o, Di-M, Sav-0, Kas-1), there was another form of nonsingleton deviation from the strict Col or Dj haplotype. Ecotypes Pet-M, Eri-l, Cvi-o, Ita-0, Kas-1, An-1, Bla-M, Kl-M that belong to Col haplotype have Dj-specific nucleotides in 2–3 polymorphic sites (429, 576, 595). Ecotypes Gel-l, En-M (Dj haplotype) have one Col–specific nucleotide or/and 1-bp deletion (Fig. 1). The group of samples An-1, Bla-M, Kl-M that has similar nucleotide substitutions (mutations) at position 935 where Col and Dj sequences are identical. These base-pair changes between haplotypes are all silent and may conceivably represent parallel mutations, the products of conversion replacement between Col and Dj haplotypes or residual “ancestral” haplotypes. No association between sequence type and geographic origin of ecotype was observed.

![Fig. 1. Polymorphic variations in the gene AtPrx53 of A. thaliana.](http://www.arabidopsis.org/Cereon/index.jsp)
in the \textit{AtPrx53} region (Table 1 and Fig. 1). The species-
wide nucleotide diversity, \(\pi\), for \textit{AtPrx53} was found to be
0.00768; the level of DNA variation in the coding region
was lower than in noncoding regions (Table 2).

The two haplotypes slightly differ in the level of nucle-
otide variation; i.e., the Col haplotype has a higher vari-
ation than the Dj haplotype; the difference is connected
with higher level of DNA variation in the coding region
of the Col haplotype (Table 2). This observation suggests
that the Col haplotype was relatively older than the Dj
haplotype. The presence of distinct haplotypes is consist-
ten with positive (although not significantly different
from 0) values of Tajima’s \(D\) and Fu and Li’s \(D^*\) statistics
estimated using 20 \textit{AtPrx53} sequences (Table 2). In con-
trast, average levels of Tajima’s \(D\) and Fu and Li’s \(D^*\) sta-
tistics for each of haplotypes are negative and indicate an
excess of low-frequency polymorphisms (Table 2).

The proportion of pairwise comparisons (excluding sin-
gleton mutations) with a significant association between
variants (\(P < 0.05\)), i.e., sites in linkage disequilibrium, is
137 of significant comparisons from total 190 (72\%) using
the \(\chi^2\) test. The level of intragenic linkage disequilib-
rium was also estimated by using the \(Z_{og}\) statistic (Kelly,
1997). The levels of \(Z_{og}\) (0.725) was higher than expected under neutrality (\(P < 0.025\)). The observed pat-
tern of nucleotide variation at \textit{AtPrx53} is compatible with
both balanced polymorphism and the possibility of ance-
stral subdivision. In contrast to previously noted evi-
dence of moderate reciprocal recombination in \textit{Adh}
(Innan et al., 1996) and other \textit{Arabidopsis} genes
(Kuittinen et al., 2000), the estimate of \(R_m\) (minimum
number of recombination events) for \textit{AtPrx53} is zero.

Nucleotide diversity at silent sites (synonymous sites
and noncoding positions) for the \textit{AtPrx53} exhibits the
same level of nucleotide diversity (\(\pi = 0.01374,\) Table 2)
as the mean level of \(\pi\) in previously studied genes \textit{Adh},
\textit{ChiA}, \textit{ChiB}, \textit{FAH1}, \textit{F3H} represented by two major haplo-
types (Aguadé, 2001).

To examine change of nucleotide variation along the
\textit{AtPrx53} region, sliding window analysis was conducted
(Fig. 2). The highest variation of silent nucleotide diver-
sextists in exon 3 and intron 3. However, no clear peak
was observed within each sequence types (Fig. 2). Ac-
gording to Strobeck (1983) and Hudson and Kaplan
(1988), one expects the accumulation of differences
between the alleles to peak around the position of the
selected site. The highest peak of nucleotide variation is
clearly associated with the region of exon 3 where replace-
ment polymorphic site 745–746 is located (related to
a amino acid replacement Phe/Ser\(^{180}\), see below).

\textbf{Amino acid replacements at \textit{AtPrx53} and allozyme
variation} The nucleotide dimorphism leads to dimor-
phism in deduced amino acids. In the coding region
of \textit{AtPrx53}, 5 nucleotide polymorphisms were observed
responsible for 4 deduced amino acid replacements, all
non-singletons (Fig. 1). Two amino acid replacements
are fairly conserved: one Leu/Val replacement site variant
in the N-terminal region (signal peptide that is cut off
from mature protein), another is Ile/Val change - in the
middle of mature protein. The Phe/Ser replacement
results in a radical substitution (from aromatic group to
hydroxyl group) in the immediate proximity to amino acid
triplet involved in putative site of glycosylation (Welinder
et al., 2002). The replacement in the C-terminal region
causes a charge change from negative Asp to neutral Asn
at position 270. Since the adjacent glycine at position
271 is conserved and presumably governs the protein

\begin{table}[h]
\centering
\caption{Summary of DNA polymorphism in the \textit{AtPrx53} gene of \textit{A. thaliana}}
\begin{tabular}{|l|c|c|c|}
\hline
 & All the accessions & Col- & Dj-
\hline
 & (n = 20) & haplotype fast & haplotype slow
\hline
\textbf{Entire region} & & &
\hline
S & 25 (20/5) & 5 (3/2) & 4 (1/3)
\hline
\(\pi\) & 0.00768 & 0.00127 & 0.00091
\hline
\(\pi\) silent & 0.01374 & 0.00200 & 0.00223
\hline
Tajima’s \(D\) & 1.56713 NS & -0.04304 NS & -1.02972 NS
\hline
Fu and Li’s \(D^*\) & 0.52975 NS & -0.13525 NS & -0.92081 NS
\hline
\textbf{Intronic region} & & &
\hline
S & 10 (9/1) & 2 (2/0) & 2 (1/1)
\hline
\(\pi\) & 0.01447 & 0.00228 & 0.00326
\hline
\textbf{Coding region} & & &
\hline
All sites & & &
\hline
\(\pi\) & 0.00576 & 0.00098 & 0.00025
\hline
Synonymous & & &
\hline
\(\pi\) & 0.01289 & 0.00168 & 0.00102
\hline
Nonsynonymous & & &
\hline
\(\pi\) & 0.00347 & 0.00087 & 0.00000
\hline
\end{tabular}
\end{table}

NS—not significant, \(\pi\) – nucleotide diversity, S – number of segregating sites (numbers of mutations/
singletons are shown in parentheses), n – number of sequences, indels were ignored.
Polymorphism in the *AtPrx53* gene of *Arabidopsis thaliana*

**Fig. 2.** Sliding-window plot of silent nucleotide diversity for the *AtPrx53* gene. The window size is 50 bp, and step size is 25 bp. The structure of the *AtPrx53* gene is presented below the plot. The arrow below the third exon indicates the nonsynonymous polymorphism at site 745–746.

**Fig. 3.** Anionic peroxidase isoforms from inflorescences in Dj and Col haplotypes. Numbers designated isoforms coded by *AtPrx53*. Faster isoform 8 has minor activity. Figure represents peroxidase isoforms in 9 ecotypes. The isoform electrophoretic mobility in Kas-1, An-1, Eri-1, Ita-0, Co-1 and Cvi-0 was the same as in Col-M ecotype. The isoform mobility in Gel-1, Ler, Sav-0, Hi-0 and Wei-0 was the same as in Dj-M ecotype (data not shown).
structure (Welinder et al., 2002), the charge difference of the polymorphic amino acids can also have an effect on the protein conformation. Thus, \textit{AtPrx53} haplotypes code proteins differing by four fixed amino acids residues. Two of them could be potentially responsible for different properties of allozymes of \textit{AtPrx53}.

In this study anionic isofrom spectrum in 15 ecotypes was analyzed. The special attention was paid to three isoforms encoded by \textit{AtPrx53} and had been designated previously by numbers 8–10 (Lebedeva et al., 2003). According to Northern blot analysis (data not shown), these isoforms are not derived from alternatively spliced transcripts but more probably represent posttranslationally modified protein isoforms. A clear difference in electrophoretic mobility of three isoforms between plants of Col and Dj haplotypes is observed (Fig. 3).

Among them, the fastest isofrom 8 along with intermediate isofrom 9 were expressed in leaf blade, the slow isoform 10 and isofrom 9 were expressed in leaf central vein (data not shown). All three isoforms were active in inflorescence (Fig. 3). Only slight displacement of isofrom 10 was detected in ecotypes Kl-M and Li-M that may be caused by ecotype-specific content of substances capable of modifying peroxidase electrophoretic mobility. Thus, the nucleotide and deduced amino acid dimorphism was associated with dimorphism mobility of three isoforms.

\textit{AtPrx53} haplotypes associations with developmental traits To examine possible functional divergence between haplotypes, we studied several quantitative developmental traits. All ecotypes with Dj haplotype show shorter duration of vegetative stages and flower earlier (develop less rosette and cauline leaves and form inflorescence earlier) than most of ecotypes with the Col haplotype (An-1 is an exception). They also have shorter reproductive stage (develop shorter stem and demonstrate earlier senescence, Table 3).

Col-M ecotypes showed a much greater range of variation for most quantitative traits studied (Table 4). This variation is due to late flowering ecotypes Kas-1 and Ita-0 with highest RLN, SLN and ecotype An-1, which demonstrates features typical for both Dj and Col haplotypes. Among all ecotypes studied, An-1 plants have less RLN and FT, but develop greatest RSN (Table 3) and resemble bushy-like plants of auxine-insensitive mutants. Late flowering ecotypes with greater RLN have greater potentials for additional rosette shoot formation. The fact that An-1 disrupts usual correlation between quantitative characters suggests that this ecotype may have some mutation.

Associations between \textit{AtPrx53} haplotypes and quantitative traits were tested using one-way analyses of variance. There is a significant association between

---

**Table 3. Developmental characters of different plant accessions**

<table>
<thead>
<tr>
<th>Ecotypes</th>
<th>RLN</th>
<th>CLN</th>
<th>RSN</th>
<th>L40</th>
<th>Lmax</th>
<th>FT</th>
<th>ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col-M</td>
<td>9.7 ± 0.2</td>
<td>3.6 ± 0.2</td>
<td>1.3 ± 0.4</td>
<td>13.5 ± 0.8</td>
<td>45.5 ± 1.5</td>
<td>38.0 ± 0.8</td>
<td>56.3 ± 0.4</td>
</tr>
<tr>
<td>Li-M</td>
<td>12.1 ± 0.3</td>
<td>3.7 ± 0.1</td>
<td>1.9 ± 0.5</td>
<td>18.1 ± 1.2</td>
<td>43.4 ± 1.4</td>
<td>39.1 ± 0.1</td>
<td>57.1 ± 0.2</td>
</tr>
<tr>
<td>Est-M</td>
<td>11.3 ± 0.5</td>
<td>3.8 ± 0.3</td>
<td>1.0 ± 0.3</td>
<td>20.3 ± 1.2</td>
<td>43.3 ± 1.2</td>
<td>38.8 ± 0.5</td>
<td>57.4 ± 0.1</td>
</tr>
<tr>
<td>Kas-1</td>
<td>33.1 ± 0.5</td>
<td>11.5 ± 0.4</td>
<td>3.5 ± 0.4</td>
<td>0.0</td>
<td>36.9 ± 3.5</td>
<td>63.6 ± 1.2</td>
<td>78.3 ± 0.2</td>
</tr>
<tr>
<td>Bla-M</td>
<td>13.9 ± 0.2</td>
<td>4.6 ± 0.2</td>
<td>1.7 ± 0.4</td>
<td>14.2 ± 0.4</td>
<td>40.9 ± 0.6</td>
<td>40.0 ± 0.1</td>
<td>58.3 ± 0.3</td>
</tr>
<tr>
<td>Kl-M</td>
<td>11.5 ± 0.5</td>
<td>2.8 ± 0.2</td>
<td>1.5 ± 0.4</td>
<td>12.7 ± 1.0</td>
<td>40.8 ± 2.4</td>
<td>42.7 ± 0.2</td>
<td>60.7 ± 0.2</td>
</tr>
<tr>
<td>An-1</td>
<td>5.8 ± 0.1</td>
<td>3.0 ± 0.1</td>
<td>3.6 ± 0.2</td>
<td>13.1 ± 0.3</td>
<td>33.3 ± 0.4</td>
<td>29.6 ± 0.4</td>
<td>40.7 ± 0.1</td>
</tr>
<tr>
<td>Pet-M</td>
<td>13.7 ± 0.5</td>
<td>6.0 ± 0.2</td>
<td>0.0</td>
<td>21.4 ± 1.8</td>
<td>44.8 ± 1.7</td>
<td>43.5 ± 0.2</td>
<td>60.4 ± 0.1</td>
</tr>
<tr>
<td>Eri-1</td>
<td>15.7 ± 0.3</td>
<td>6.9 ± 0.4</td>
<td>1.3 ± 0.2</td>
<td>7.3 ± 1.2</td>
<td>36.8 ± 1.2</td>
<td>43.4 ± 0.2</td>
<td>62.1 ± 0.1</td>
</tr>
<tr>
<td>Co-1</td>
<td>16.9 ± 0.3</td>
<td>6.7 ± 0.6</td>
<td>1.2 ± 0.3</td>
<td>6.0 ± 1.8</td>
<td>36.7 ± 1.2</td>
<td>46.5 ± 0.3</td>
<td>59.2 ± 0.7</td>
</tr>
<tr>
<td>Cvi-o</td>
<td>11.1 ± 0.3</td>
<td>3.9 ± 0.1</td>
<td>0.1 ± 0.05</td>
<td>7.1 ± 1.2</td>
<td>25.6 ± 0.7</td>
<td>40.8 ± 0.1</td>
<td>58.4 ± 0.4</td>
</tr>
<tr>
<td>Gel-1</td>
<td>9.8 ± 0.1</td>
<td>3.5 ± 0.1</td>
<td>1.6 ± 0.3</td>
<td>2.4 ± 0.6</td>
<td>37.6 ± 1.3</td>
<td>32.5 ± 0.1</td>
<td>52.6 ± 0.1</td>
</tr>
<tr>
<td>En-M</td>
<td>7.5 ± 0.2</td>
<td>3.2 ± 0.1</td>
<td>0.1 ± 0.08</td>
<td>12.8 ± 1.3</td>
<td>36.4 ± 1.7</td>
<td>33.2 ± 0.1</td>
<td>45.9 ± 0.1</td>
</tr>
<tr>
<td>Ws</td>
<td>8.3 ± 0.4</td>
<td>3.5 ± 0.2</td>
<td>0.7 ± 0.2</td>
<td>23.1 ± 1.5</td>
<td>34.2 ± 1.3</td>
<td>29.6 ± 0.2</td>
<td>39.6 ± 0.2</td>
</tr>
<tr>
<td>Sav-0</td>
<td>7.5 ± 0.1</td>
<td>3.2 ± 0.1</td>
<td>2.1 ± 0.2</td>
<td>13.7 ± 0.8</td>
<td>34.9 ± 0.8</td>
<td>31.4 ± 0.2</td>
<td>52.0 ± 0.1</td>
</tr>
<tr>
<td>Dj-M</td>
<td>7.6 ± 0.2</td>
<td>3.6 ± 0.2</td>
<td>0.5 ± 0.1</td>
<td>25.4 ± 1.1</td>
<td>45.3 ± 1.5</td>
<td>30.8 ± 0.3</td>
<td>46.1 ± 0.2</td>
</tr>
<tr>
<td>Hi-0</td>
<td>9.4 ± 0.3</td>
<td>4.3 ± 0.3</td>
<td>1.1 ± 0.4</td>
<td>20.0 ± 2.7</td>
<td>28.6 ± 2.4</td>
<td>33.6 ± 0.1</td>
<td>51.4 ± 0.4</td>
</tr>
<tr>
<td>Wei-o</td>
<td>8.5 ± 0.3</td>
<td>4.0 ± 0.4</td>
<td>2.3 ± 0.4</td>
<td>31.3 ± 3.6</td>
<td>37.3 ± 2.1</td>
<td>31.6 ± 0.1</td>
<td>50.2 ± 0.2</td>
</tr>
<tr>
<td>Ita-0</td>
<td>25.5 ± 2.1</td>
<td>NA</td>
<td>NA</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td>76.5 ± 1.2</td>
</tr>
</tbody>
</table>

RLN – rosette leaf number; CLN – cauline leaf number, correspondingly; FT – flowering time; L40 – length of main stem at 40 days after sowing; Lmax – final main stem length; RSN – rosette shoot number; ST – senescence time. Values are means ± s.e.m. Ita-0 does not produce an inflorescence without vernalization, therefore CLN, RSN, Lmax and FT were not scored. NA means not analysed.
haplotype and all developmental traits studied (Table 4).

**DISCUSSION**

The level of the *AtPrx53* nucleotide variation is higher (0.00768 for all sites and 0.01374 for silent sites) than in nearby peroxidase genes (Kupriyanova et al., 2006). This result agrees with highly differentiated *AtPrx53* haplotypes, each of which has a low level of nucleotide variation. Members of each sequence type came from at least two different continents or localities that were far apart. These results support the hypothesis that the present *A. thaliana* population has spread over the world recently, probably after fusion of divergent populations (King et al., 1993; Price et al., 1994; Innan et al., 1996, 1997).

Two distinct haplotypes were present at intermediate frequencies, resulting in positive Tajima’s *D* and Fu and Li’s *D* *s* values. The directions of the Tajima’s *D* and Fu and Li’s *D* *s* tests statistics, as well as the results of the linkage disequilibrium analysis show deviations from the neutral-equilibrium model.

The revealed pattern of polymorphism at the *AtPrx53* peroxidase and the strong haplotype structure in the absence of detectable recombination across *AtPrx53* (*R* *m* = 0) may be interpreted as evidence of balancing selection on this locus or other nearby sites.

Both the Col and Dj haplotypes may have originated from locally adapted ancestral populations and may be currently maintained by local selection on alternative alleles of *AtPrx53* or nearby locus despite the widespread dispersal of this species. It is unclear whether *AtPrx53* locus was the actual target of selection or some closely linked genes may have been the selective target. There are several genes determined flowering time surrounding *AtPrx53* gene; among them *TPL1* (approximately, 500 kb above *AtPrx53*) and *FLC* (400 kb above *AtPrx53*) are linked more tightly to *AtPrx53*. To study the association of *AtPrx53* haplotypes with polymorphism of these genes, further investigation is needed. The association of *AtPrx53* haplotype and developmental characters could be due also to population structure or genomewide divergence between haplogroups (Aranzana et al., 2005).

There are several reasons to believe, however, that balancing selection is acting directly on the *AtPrx53* gene. First, the peroxidase encoded by the *AtPrx53* gene is the major form of anionic (pI 3.5) peroxidases in stem tissue (Østergaard et al., 1996) involved in lignification that relates to secondary cell wall formation. According to our data, three isoforms correspond to the gene *AtPrx53*, their total activity in the leaves, stem, and pods amounts to over 45, 60, and 80% of the activity of all anionic peroxidases (Lebedeva et al., 2003). These data indicate functional importance of *AtPrx53* peroxidase for plant development. The nucleotide dimorphism of the *AtPrx53* gene was associated with deduced amino acid dimorphism and difference in electrophoretic mobility of three peroxidase isoforms. Revealed amino acid polymorphism can influence peroxidase reactions or change their substrate specificity.

Second, the pattern of divergence between the Dj and Col haplotypes (Fig. 2) is consistent with prediction of elevated levels of polymorphism in the vicinity of a balanced polymorphism (Strobeck, 1983). The peak disappeared when only each haplotype was analyzed. This result can be taken as evidence of balancing selection (Hudson and Kaplan, 1988; Tian et al., 2002). Only one of two non-conservative amino acid differences between the Col and Dj haplotypes (Phe/Ser*) is* lies within the hypervariable region. This replacement change may be responsible for allozyme functional dimorphism, since it causes a radical substitution. It may well be that this amino acid polymorphism is subject to balancing selection. Similar peaks of polymorphism associated with amino acid replacement and changes in electrophoretic mobility of encoded isoforms were identified in the *Adh* and *PgIC* loci of *A. thaliana* (Miyashita et al., 1998; Kawabe et al., 2000).

Third, the revealed haplotypes are associated with several developmental traits. Studied quantitative traits

---

**Table 4.** Associations between Col and Dj *AtPrx53* haplotypes and quantitative traits (ANOVA F-test results)

<table>
<thead>
<tr>
<th>Haplotype Col</th>
<th>Haplotype Dj</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosette leaf number</td>
<td>15.0</td>
<td>7.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Cauline leaf number</td>
<td>5.1</td>
<td>2.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Rosette shoot number</td>
<td>1.6</td>
<td>1.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Lenght of main stem (40 days)</td>
<td>12.5</td>
<td>6.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Final main stem length</td>
<td>38.9</td>
<td>5.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Flowering time</td>
<td>42.4</td>
<td>8.2</td>
<td>2.4</td>
</tr>
<tr>
<td>Senescence time</td>
<td>60.4</td>
<td>9.6</td>
<td>2.5</td>
</tr>
</tbody>
</table>

The F-test p-value is less than 0.05; the null hypothesis of equal means is rejected.
(leaf number, flowering time, stem length, rosette shoot number, senescence time) are highly integrative characters that depend not only on the environmental conditions and the activity of genetical factors determining emerging organ identity but also on the activity of factors determining cell behavior (rate of cell division and rate and extent of cell elongation). Cell elongation results from the balance between cell wall loosening and stiffening. Peroxidase can favor elongation by generating oxygen radicals or by regulating the local concentration of hydrogen peroxide (Penel, 2000). Opposite process of cell wall rigidification leading to the inhibition of cell growth can be the result of the peroxidase-mediated cross-linking of several compounds. Among them, polysaccharidelinked ferulates (bound to lignin or not), extensins and lignin monomers form a complex network that solidifies the plant wall (Dunand et al., 2002). One can propose that allozymes encoded by two haplotypes differ by their activity in cell wall lignification and cell growth restriction. Further phenotypic analysis of nearly isogenic lines only differing in AtPrx53 is necessary to specify if the dimorphism is related to balancing selection at this locus or not.

The authors gratefully acknowledge support from the Russian Foundation for Basic Research projects, Leading Scientific Schools Program (SS - 4202.2006.4), RAS Programs “Gene Pool Dynamics” and “Origin and Evolution of Biosphere”.

REFERENCES


Polymorphism in the AtPrx53 gene of *Arabidopsis thaliana*


