Subcellular Relocalization and Positive Selection Play Key Roles in the Retention of Duplicate Genes of *Populus* Class III Peroxidase Family

Lin-Ling Ren, Yan-Jing Liu, Hai-Jing Liu, Ting-Ting Qian, Li-Wang Qi, Xiao-Ru Wang, and Qing-Yin Zeng

Gene duplication is the primary source of new genes and novel functions. Over the course of evolution, many duplicate genes lose their function and are eventually removed by deletion. However, some duplicates have persisted and evolved diverse functions. A particular challenge is to understand how this diversity arises and whether positive selection plays a role. In this study, we reconstructed the evolutionary history of the class III peroxidase (PRX) genes from the *Populus trichocarpa* genome. PRXs are plant-specific enzymes that play important roles in cell wall metabolism and in response to biotic and abiotic stresses. We found that two large tandem-arrayed clusters of PRXs evolved from an ancestral cell wall type PRX to vacuole type, followed by tandem duplications and subsequent functional specification. Substitution models identified seven positively selected sites in the vacuole PRXs. These positively selected sites showed significant effects on the biochemical functions of the enzymes. We also found that positive selection acts more frequently on residues adjacent to, rather than directly at, a critical active site of the enzyme, and on flexible regions rather than on rigid structural elements of the protein. Our study provides new insights into the adaptive molecular evolution of plant enzyme families.

INTRODUCTION

Gene duplication is an important mechanism for the evolution of novel gene function. The accumulation of whole-genome sequences provides substantial evidence for the abundance of duplicate genes in plant species (Tuskan et al., 2006). Over the course of evolution, many duplicate genes may lose their functions and be ultimately removed by deletion. However, some duplicates have persisted and evolved diverse functions. A particular challenge is to understand the evolutionary mechanisms responsible for the initial retention and subsequent divergence of newly created duplicate genes. The neofunctionalization model has been proposed as an important process driving the retention of duplicated genes by evolving novel functions (Byrne and Wolfe, 2007; Sakuma et al., 2013). However, this model cannot satisfactorily explain how a duplicate gene can escape the load of deleterious mutations that would probably accumulate before enough beneficial mutations could confer a new function (Byun-McKay and Geeta, 2007). Because deleterious mutations are more probable than advantageous mutations after gene duplication, it has been assumed that most duplicate copies are lost and a limited number of duplicates that gain new functions are maintained by positive selection. Positive selection is now recognized as playing an important role in the initial retention and subsequent divergence of duplicate genes (Shiu et al., 2006; Barkman et al., 2007; Beisswanger and Stephan, 2008; Lan et al., 2013). However, the lack of appropriate functional assays has generally hindered attempts to elucidate the molecular mechanisms and effects of positive selection on the retention and divergence of duplicate genes.

Protein subcellular relocalization has recently been considered as another form of functional divergence of duplicate genes (Byun-McKay and Geeta, 2007; Marques et al., 2008; Qian and Zhang, 2009; Wang et al., 2009) based on the following reasoning: (1) different subcellular structures play different biological functions in the eukaryotic cell; (2) most eukaryotic proteins are localized to specific subcellular structures to perform their functions; and (3) proteins could alter their functions when relocated to a new subcellular compartment due to altered microenvironments. Even the same protein might have a different biological function when localized to a new subcellular structure (Leissring et al., 2004). Thus, changes in protein subcellular localization may lead to the origin of new functions.

Class III peroxidases (PRXs; EC 1.11.1.7) are plant-specific enzymes that can catalyze the reduction of H$_2$O$_2$ by moving electrons to various donor molecules such as phenolic compounds, lignin precursors, or secondary metabolites (Cosio and
Dunand, 2009). Plant PRX proteins are involved in several important physiological and developmental processes, including lignin and suberin formation, the cross-linking of cell wall components, wound healing, the removal of H₂O₂, the oxidation of toxic reductants, and defense against pathogen or insect attack (Gabaldón et al., 2005; Bindschedler et al., 2006; Costo and Dunand, 2009; Daudi et al., 2012). PRXs are encoded by a large gene family. The Arabidopsis thaliana genome contains 73 and 138 PRX genes, respectively (Tognolli et al., 2002; Passardi et al., 2004). Based on the absence or presence of the C-terminal signal peptide, plant PRXs are divided into cell wall and vacuole types (Matsui et al., 2011). The wall-type PRXs are involved in cell wall metabolism including lignin polymerization, whereas the vacuole-type PRXs are involved in defense responses against abiotic and biotic stresses (Ostergaard et al., 2000; Li et al., 2003; Gabaldón et al., 2005; Bindschedler et al., 2006; Costa et al., 2008). This functional divergence between cell wall and vacuole PRXs makes this gene family a model system for exploring the contribution of protein subcellular relocalization to the expansion of the gene family.

Populus trichocarpa is a model woody perennial plant with a recently sequenced genome (Tuskan et al., 2006). Unlike herbaceous plants such as Arabidopsis and rice, Populus shows tree-specific traits such as wood formation and perennial growth in the field. Some of these traits might be related to the functional evolution of the PRX gene family. Analysis of the Populus genome suggests a recent occurrence of a whole-genome duplication event in the Salicaceae 60 to 65 million years ago that affected ~92% of the genome (Tuskan et al., 2006). In addition, Populus and Arabidopsis share more ancient large-scale duplication events (Tuskan et al., 2006). The complex history of genome duplications and chromosomal rearrangements in Populus provides an opportunity to study the molecular mechanisms of duplicate gene retention over the course of genome evolution. In this study, we conducted genome-wide annotation of PRX genes in the Populus genome and identified the genetic events responsible for their expansion and organization. Using ancestral protein reconstruction, protein subcellular localization, statistical tests for positive selection, site-directed mutagenesis, and enzyme functional assays of ancestral, current, and mutant proteins, this study revealed that protein subcellular relocalization and positive selection likely have played important roles in the retention of duplicate genes.

RESULTS

Identification of the PRX Gene Family in the Populus Genome

Ninety-three full-length genes encoding PRX proteins were identified in the Populus genome (Supplemental Data Set 1). Among these, 10 (PRX5, 8, 18, 32, 38, 47, 60, 66, 69, and 76) were considered putative pseudogenes based on the presence of a frame shift disrupting the coding region or a stop codon occurring prematurely, resulting in a truncated protein. After removing these stop codons or revising the frame shifts by deleting one or two nucleotides, the 10 full-length sequences were included in the phylogenetic and gene expression analyses. In addition to full-length PRX genes, 43 partial PRX fragments were identified (Supplemental Table 1). The length of these fragments ranged from 24 to 279 amino acids. The conserved domain analysis on these fragments showed that they contained a partial version of the peroxidase domain characteristic of PRX; thus, they were considered to be pseudogenes. We were unable to reliably analyze the phylogenetic relationships of these fragments because of their small size.

Expression of Populus PRX Genes under Normal Growth Conditions and Abiotic Stress

We investigated the expression patterns of Populus PRX genes under normal growth conditions and in response to stress treatments (H₂O₂, salicylic acid, salt, and drought applications). We examined the expression of all 93 Populus PRX genes in five tissues, including roots, shoots, leaves, buds, and phloem (Figure 1B). Among the 93 PRX genes, 41 were expressed under all growth conditions and 44 were selectively expressed either in a specific tissue and/or in response to a specific treatment, whereas eight (PRX5, 8, 14, 32, 60, 68, 69, and 76) were not detected in any tissue nor in response to any treatment. To validate whether the lack of expression was due to PCR failure, we used the same sets of primers to amplify these eight genes from genomic DNA. We found that all primers produced the expected results. We further designed additional eight pairs of primers, each specific to one of the eight PRX genes (Supplemental Table 2), and performed amplification detection on both mRNA and genomic DNA. Again, we did not detect expression of these PRX genes in any of the tissues examined, but we did observe positive amplifications from genomic DNA. Thus, we consider that these PRX genes are likely expressed at subdetectable levels, or only induced in response to treatments and/or in tissues not examined in our study, or pseudogenes.

Among the five types of tissue examined, PRX genes showed preferential expression in roots (Supplemental Figure 1). For example, 81 of the 93 PRX genes were expressed in roots under normal growth conditions compared with 63, 52, 61, and 68 genes in shoots, leaves, buds, and phloem, respectively. Under both normal growth and the four stress treatment conditions, 77 genes were expressed in roots, while 52, 46, 56, and 45 genes were expressed in shoots, leaves, buds, and phloem, respectively. This preferential expression of PRX genes in roots was also observed in Arabidopsis and rice (Hiraga et al., 2000; Tognolli et al., 2002) and is thought to be associated with their role in defense responses against soil pathogens (Passardi et al., 2005; Hatsugai et al., 2009). In addition, we found 17, 10, 8, and 5 Populus PRX genes that were not expressed in leaves, buds, phloem, and shoots under normal growth conditions were expressed in response to one or several stress treatments, which indicates that they may play a role in the response to abiotic stresses.

Chromosomal Distribution of the PRX Gene Family

To understand the duplication and expansion history of the Populus PRX family, we examined the distribution of the 93 PRX genes on chromosomes. We found that although each of the 19 Populus chromosomes harbored PRX genes, the distribution was obviously heterogeneous (Supplemental Figure 2). Five clusters (clusters I to V) with high densities of PRX genes (more
Figure 1. Phylogenetic Relationships and Expression Patterns of *Populus* PRXs.
than two copies) were observed on chromosomes 1, 3, 7, 13, and 16, in addition to seven pairs (PRX13/14, PRX26/27, PRX36/37, PRX53/54, PRX62/63, PRX78/79, and PRX86/87) arranged in tandem on chromosomes 1, 4, 5, 8, 10, 15, and 17. A total of 37 PRX genes were arranged in tandem repeats, indicating that tandem duplications contributed significantly to the expansion of this gene family. The remaining PRX genes were dispersed on all 19 chromosomes. The poplar genome underwent a recent whole-genome duplication; 60 to 65 million years ago (Tuskan et al., 2006). Paralogous segments created by this genome duplication were identified in a previous genome annotation (Tuskan et al., 2006). Based on this information, we identified 14 duplicate gene pairs, including PRX9/18, PRX11/88, PRX15/33, PRX16/38, PRX28/65, PRX29/92, PRX30/57, PRX31/48, PRX32/49, PRX34/52, PRX39/91, PRX42/89, PRX55/60, and PRX56/59, which were each located in a pair of paralogous blocks and can thus be considered as direct results of this whole-genome duplication event (Supplemental Figure 2).

Figure 1. (continued).

The maximum-likelihood procedure using the WAG+I+G model with 100 bootstrap replicates was used for phylogeny reconstruction. Numbers at each node in the phylogenetic tree are bootstrap values, and only bootstrap values >50% are shown. Putative pseudogenes are indicated with asterisks. In (B), the gray box indicates positive detection of gene expression in the corresponding tissue under normal growth conditions (NC) and following H$_2$O$_2$ (HO), salicylic acid (SA), salt (ST), and drought (DR) stress treatments.
Clusters I and II were the largest of the five tandem duplication clusters, containing seven and five full-length copies, respectively, and three PRX fragments (Figure 2A; Supplemental Figure 2). Among these 12 copies, PRX5 and PRX8 were considered putative pseudogenes based on the presence of a frame shift disrupting the coding region and their lack of expression. The other 10 PRX genes in these two clusters were expressed under all growth conditions or in a specific tissue (Figure 1B), suggesting that they were functional. In the following sections, we focused on the retention mechanisms of the PRX genes in these two clusters.

Duplication History of PRX Genes in Clusters I and II

Clusters I and II were each located in a pair of paralogous blocks (Figure 2A); we thus considered them as created by the recent whole-genome duplication event. On the phylogenetic tree of all 93 Populus PRX genes, the 12 PRX genes in these two clusters were grouped together (Figure 1A), indicating that they descended from a single common ancestor. Based on the relationships among these copies (Figure 2B), we reconstructed the most parsimonious history of gene duplication, loss, and rearrangement as illustrated in Figure 2C. It seems likely that five rounds of tandem duplications created six ancestral genes (i, k, j, g, e, and c, Figure 2C) before the recent whole-genome duplication, after which this set of genes evolved independently into the organization in clusters I and II, by complex tandem duplication, gene loss, and rearrangement events. If this scenario is valid, we should expect the levels of divergence between paralogs to comport with their hypothetical divergence history. We calculated divergence at synonymous sites \( K_{st} \) for all pairwise comparisons of the 10 functional PRX genes in clusters I and II (Supplemental Table 3). We found the lowest average \( K_{st} \) (0.0315) between copies created most recently after the recent whole-genome duplication, 0.4359 for those due to the recent whole-genome duplication, 0.4359 for those due to the recent round of tandem duplications created six ancestral genes (i, k, j, g, e, and c, Figure 2C) before the recent whole-genome duplication, after which this set of genes evolved independently into the organization in clusters I and II, by complex tandem duplication, gene loss, and rearrangement events. If this scenario is valid, we should expect the levels of divergence between paralogs to comport with their hypothetical divergence history. We calculated divergence at synonymous sites \( K_{st} \) for all pairwise comparisons of the 10 functional PRX genes in clusters I and II (Supplemental Table 3). We found the lowest average \( K_{st} \) (0.0315) between copies created most recently after the recent whole-genome duplication, 0.4359 for those due to the recent whole-genome duplication, and 0.4648 to 1.5264 for duplications more ancient, with the highest average \( K_{st} \) (1.5264) between copies derived from the first and second round of ancestral duplication. These results support the hypothetical history of the PRX genes in these two clusters.

Subcellular Localization of the PRX Proteins in Clusters I and II

Among the 12 PRX genes in clusters I and II, two (PRX5 and 8) were considered putative pseudogenes based on the presence of frame shifts disrupting the coding region and their lack of expression. Thus, we only analyzed subcellular localization of the remaining 10 PRX genes. We generated a C-terminal green fluorescent protein (GFP) fusion for each protein and visualized their subcellular location by confocal microscopy after transient expression of the fusions in Nicotiana benthamiana. Confocal microscopy analysis showed that eight fusion proteins (PRX2, 3, 4, 7, 20, 21, 22, and 23-GFP) resulted in the accumulation of fluorescent signal in the vacuole, as illustrated by PRX4-GFP (Figure 3A), indicating that these eight fusion proteins localize to the vacuole. The fluorescent signal of the other two fusion proteins (PRX6 and 24-GFP) were only detected in the peripheral regions of the N. benthamiana epidermal cells, suggesting plasma membrane and/or cell wall localization. To distinguish between the plasma membrane and/or cell wall localizations, the epidermal cells were plasmolized by treatment with 0.3 g/mL sucrose. After plasmolysis, the fluorescent signal was retained in the cell wall, as illustrated by PRX6-GFP (Figure 3A), indicating that these two PRX-GFP fusion proteins localize to the cell wall. We thus divided these 10 PRX proteins into two types: cell wall PRXs (PRX6 and 24) and vacuole PRXs (PRX2, 3, 4, 7, 20, 21, 22, and 23).

Divergence in Selective Pressure between Clusters I and II PRX Genes

Phylogenetic relationships among PRX genes showed that the eight vacuole PRXs (PRX2, 3, 4, 7, 20, 21, 22, and 23) in clusters I and II grouped together with 100% support and the two cell wall-type PRX6 and PRX24 diverged earlier (Figure 1A). To determine whether there was a significant change in the selective pressure between the two groups of PRXs (clades A and B in Figure 4), we performed maximum likelihood codon models analysis using phylogenetic analysis by maximum likelihood (PAML) package version 4.4 (Yang, 2007). Two assumptions were tested: a one-ratio model that assumed the same \( \omega \) (= \( d_\text{s}/d_\text{var} \)) ratio for both groups of PRXs and a two-ratio model in which
the two PRX types were assigned a different \( \omega \) ratio. The log-likelihood values under the one-ratio and two-ratio models were \( \ln L = -5617.378566 \) and \( -5612.132522 \), respectively (Table 1). The likelihood ratio test showed that the two-ratio model rejected the null model (one-ratio model), indicating that the selective pressure differed significantly between the two PRX groups (\( P < 0.001 \)). Under the two-ratio model, the \( \omega \) values for eight vacuole PRXs and two cell wall PRXs were 0.31267 and 0.16888, respectively, indicating that the vacuole PRXs were under more relaxed selection constraints than the cell wall PRXs.

To test whether some branches in the phylogenetic tree of the 10 PRXs were under positive selection, we first applied a free-ratio model test using PAML (Yang, 2007). The free-ratio model allows an independent assignment of \( \omega \) to each branch. This analysis showed three branches (PRX3, 4, and 7; clade C) had \( \omega > 1 \) (bold branches in Figure 4A). We then applied a dynamic maximum likelihood-based procedure to globally search for optimal models to estimate substitution rates (Zhang et al., 2011). This procedure found three optimal branch-specific models for the 10 PRXs (Figures 4B to 4D; Supplemental Table 4). These models support the pattern revealed by the free-ratio model and consistently identified the same three branches in clade C as \( \omega > 1 \).

To further assess the sites of selection in the sequences identified by the branch models (i.e., clade C in Figure 4), we performed site model tests using PAML (Table 1). This analysis identified the M8 model as the most probable, under which 13 amino acid sites were likely under positive selection (Table 1), including seven with posterior probabilities >0.95 (alignment positions 57, 92, 161, 167, 180, 200, and 247 in Figure 5). We also applied four codon substitution models (M0, M3, M3 + S1, and M3 + S2) in Fitchmodel software (Guindon et al., 2004) to investigate the variation in selection among amino acid sites for the 10 PRXs in Figure 4. The likelihood ratio tests between nested models (M0 versus M3; M3 versus M3+S1; M3+S1 versus M3+S2) suggested that the M3 + S2 model was significantly better than the other models (Table 2). The switching rate between \( \omega_2 \) (moderate purifying selection) and \( \omega_3 \) (positive selection) (\( R_{23} = 7.384 \)) was significantly higher than that between \( \omega_1 \) (strong purifying selection) and \( \omega_2 \) (\( R_{12} = 0.545 \)) and that between \( \omega_1 \) and \( \omega_3 \) (\( R_{13} = 0.002 \)) (Table 2). Under the M3 + S2 model, seven amino acid sites were inferred to be under positive selection (\( \omega_3 = 4.149 \)) in the four PRXs in clade C (Figure 6). These sites corresponded to the seven positive selection sites (with posterior probabilities >0.95) identified by the site model in PAML (Table 1).

**Structural Distribution of the Putative Positively Selected Sites**

To understand the evolutionary changes of the putative positively selected sites, we reconstructed the most recent common ancestral protein (MRCAP) of the four PRXs (PRX2, 3, 4, and 7) using CODEML in the PAML package with an overall accuracy of 0.99171. Based on the x-ray structure of the Armoracia rusticana PRX protein (PDB accession number 7ATJ), we modeled the three-dimensional structures of the five PRX proteins (PRX2, 3, 4, 7, and MRCAP). The simulated structures of the five PRX proteins were superimposed to evaluate the goodness of fit to the overall topologies. The five PRX proteins shared the same conformation of structural elements for the \( \alpha \)-helices and \( \beta \)-sheets (Supplemental Figure 3).

PRX proteins are heme-containing enzymes that can catalyze the oxidative coupling of phenolic compounds using H\(_2\)O\(_2\) as the oxidizing agent (Henriksen et al., 1999). Based on sequence and structural comparisons, seven residues (corresponding to Arg-55, Ser-59, Arg-62, His-66, His-194, Arg-198, and Arg-200 of the MRCAP; red arrows in Figure 5) were predicted as active.
sites that could bind the substrate heme or ferulic acid through hydrogen bonds. Six residues (corresponding to Gly-93, Pro-247, Ala-92, Ser-161, Leu-167, and Gly-180 of the MRCAP; purple arrows in Figure 5) could interact with substrates within 4.0 Å through hydrophobic interactions. These six residues were also considered active sites. Among the seven putative positively selected sites, five (alignment positions 57, 92, 161, 167, and 180, corresponding to Gly-57, Ala-92, Ser-161, Leu-167, and Gly-180 of the MRCAP; Figures 5 and 7) were adjacent to the active sites, one (alignment positions 200, corresponding to Arg-200 of the MRCAP; Figures 5 and 7) was predicted to be an active site for binding heme, and another (alignment position 247, corresponding to Pro-247 of the MRCAP; Figures 5 and 7) was distant from active site regions.

Substrate Activity of the PRX Proteins in Clusters I and II

To examine the biochemical functions of the 10 PRX proteins in clusters I and II, we expressed and purified the corresponding proteins. Plant PRXs can catalyze the reduction of H$_2$O$_2$ by moving electrons to various donor molecules. In this study, six substrates (listed in Table 3) were selected to determine PRX protein activity.

All 10 purified PRX proteins showed enzymatic activities toward these six substrates with much higher enzymatic activities toward coniferyl alcohol than toward the other five substrates (paired-samples t test, P < 0.003), except for PRX21 toward sinapyl alcohol and PRX23 toward pyrogallol. We conducted a multiple response permutation procedure (MRPP) test to evaluate the significance of differences in the enzyme activity profiles between the cell wall and vacuole PRXs. MRPP is a nonparametric test that is flexible over unequal sample sizes and the violation of normality assumptions. The MRPP test showed no significant differences in the enzyme activity profiles between the two PRX types (P > 0.55). However, the Kendall’s W nonparametric test for related samples indicated a significant difference in enzyme activity among the eight vacuole PRXs (P = 0.00). Thus, diversification in enzyme activity toward different substrates has apparently evolved among vacuole PRXs.

Kinetics of the PRX Proteins

The apparent kinetic constants of the 10 PRX proteins were determined using coniferyl alcohol, ferulic acid, and pyrogallol as substrates (Table 4). The 10 PRX proteins showed much higher affinities (higher $1/K_m$), turnover number (higher $k_{cat}$), and catalytic efficiency (higher $k_{cat}/K_m$) toward coniferyl alcohol and ferulic acid (paired-samples t test, P < 0.01) than toward pyrogallol. The MRPP test showed no significant difference in the enzyme kinetics between the cell wall and vacuole PRXs (P > 0.30). However, the apparent kinetic constants of the eight vacuole PRX proteins showed significant variation (Kendall’s W test, P = 0.00) toward the substrates (Table 4). For example, the catalytic efficiency toward coniferyl alcohol varied from 600.83 to 64,632.39 mM$^{-1}$ s$^{-1}$.

Functional Significance of Positive Selection

Branch model tests indicated that four vacuole PRXs (PRX2, 3, 4, and 7) in cluster I were under positive selection pressure (Figures 4 and 6). These four PRXs were derived after the recent whole-genome duplication in Populus. The enzymatic activities of these four vacuole PRXs varied 8-fold for sinapyl alcohol and 41-fold for ascorbic acid (Table 3). To understand the evolution of this enzymatic activity, we expressed and purified the MRCAP and the four vacuole PRXs. Wilcoxon nonparametric tests for two related samples indicated significant differences in enzyme activity profiles between the MRCAP and each of the four PRXs (P < 0.046). PRX2 and 3 showed much higher enzymatic activity toward all six substrates than the MRCAP (Table 3), except PRX2 toward ascorbic acid. The apparent kinetic constant values for coniferyl alcohol, ferulic acid, and pyrogallol were similar among MRCAP and PRX2 and 3 (Wilcoxon nonparametric tests,
P > 0.21). PRX4 and 7 showed much lower enzymatic activity toward all six substrates than MRCAP (Table 3). Significant differences in the apparent kinetic constants toward the three substrates listed in Table 4 were observed between the MRCAP and each of PRX4 and 7 (Wilcoxon nonparametric tests, P < 0.01). Compared with the MRCAP, except for affinity of PRX4 toward substrate coniferyl alcohol, PRX4 and 7 showed much lower affinities (lower 1/K_m), turnover number (lower k_cat), and catalytic efficiency (lower k_cat/K_m) toward substrates coniferyl alcohol, ferulic acid, and pyrogallol (Table 4).

**Figure 5.** Alignment of PRX2, 3, 4, and 7 and Their MRCAP. 
α-Helices and β-strands are represented as blue cylinders and yellow arrows, respectively. Conserved residues in these five PRXs are marked in gray. Positive selection sites (with posterior probabilities >0.95) calculated using the site models in the PAML package are marked in green. The predicted active sites for binding substrates are indicated with red and purple arrows. The numbers in parentheses correspond to positive selection sites (Figure 6) predicted by Fitmodel software.

**Table 2.** Likelihood Ratio Test of the Models in the Fitmodel Program for *Populus* PRXs

<table>
<thead>
<tr>
<th></th>
<th>M0 (No Heterogeneity)</th>
<th>M3 (Heterogeneity across Sites)</th>
<th>M3+S1 (Shifting across Branches)</th>
<th>M3+S2 (Unequal Switching Rates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln L</td>
<td>-6136.029</td>
<td>-5950.128</td>
<td>-5875.428</td>
<td>-5863.143</td>
</tr>
<tr>
<td>ν1, ν2, ν3</td>
<td>0.340</td>
<td>0.000 0.314 1.417</td>
<td>0.000 0.305 1.713</td>
<td>0.001 0.184 4.149</td>
</tr>
<tr>
<td>̂Ω1, ̂Ω2, ̂Ω3</td>
<td>1.000</td>
<td>0.302 0.474 0.224</td>
<td>0.496 0.295 0.208</td>
<td>0.416 0.471 0.113</td>
</tr>
<tr>
<td>R_{12} R_{13} R_{23}</td>
<td>1.605 1.605 1.605</td>
<td>1.605 1.605 1.605</td>
<td>0.545 0.002 7.384</td>
<td></td>
</tr>
<tr>
<td>χ^2 (models; P)</td>
<td>371.802 (M0 vs. M3; P &lt; 0.001)</td>
<td>149.400 (M3 vs. M3+S1; P &lt; 0.001)</td>
<td>24.570 (M3+S1 vs. M3+S2; P &lt; 0.001)</td>
<td></td>
</tr>
</tbody>
</table>
positively selected sites (alignment positions 92, 161, 200, and 247 in Figure 5) in the MRCAP were mutated to the corresponding amino acid residues present in PRX2, 3, 4, and 7, each of the MRCAP mutants (A92G, A92N, S161F, R200Q, R200K, P247V, and P247L) showed a diverse response to the tested substrates (Figure 8). For example, the P247V mutant showed decreased enzymatic activity toward ascorbic acid and coniferyl alcohol but increased enzymatic activity toward pyrogallol, guaiacol, ferulic acid, and sinapyl alcohol (Table 3).

DISCUSSION

Gene duplication is thought to play an important role in providing raw material for plant evolution (Ohno, 1970). However, in the event of a redundancy, duplicate genes may lose their functions and ultimately be removed by deletion over the course of evolution. In this study, we found 43 PRX fragments from the Populus genome, indicating that massive gene losses have occurred in this gene family. Nonetheless, 93 full-length PRXs have been retained in the Populus genome, bringing us to a central issue in the evolution of duplicate genes: determining which evolutionary mechanisms play a central role in the retention and functional diversification of duplicate genes. The evolutionary mechanisms of the retention of duplicate genes might vary among different gene families. Previous studies have shown that divergence in gene expression or protein function contributes to the retention of duplicate genes (Ganko et al., 2007; Lan et al., 2009; Yang et al., 2013). Protein subcellular relocation has recently been considered another mechanism for gene retention because proteins could alter their functions when relocated to a new subcellular structure (Byun-McKay and Geeta, 2007; Marques et al., 2008; Qian and Zhang, 2009). Plant PRXs are involved in a broad range of physiological processes throughout the plant life cycle (Passardi et al., 2005). These proteins are localized to the cell wall or vacuole. The cell wall PRXs are involved in cell wall metabolism, including lignin polymerization, whereas the vacuole PRXs are involved in defense responses to abiotic and biotic stresses and the oxidation of toxic reductants and other secondary metabolites (Ostergaard et al., 2000; Li et al., 2003; Gabaldón et al., 2005; Bindschedler et al., 2006; Costa et al., 2008). This divergence in subcellular localization might be an important first step in the evolution of new PRX functions following gene duplication. This study examined the functional evolution of the Populus PRX gene family and the possible drivers for it.

Divergence Aided by C-Terminal Signal Peptide

In the eukaryotic cell, most proteins are synthesized in the cytosol, and many need to be further sorted to the cell membrane or various membrane-bound cellular organelles. In the plant cell, when the final destination is the chloroplast, mitochondrion, or secretory pathway, sorting usually relies on the presence of an N- or C-terminal targeting sequence (signal peptide) (Emanuelsson et al., 2000). In most cases, the signal peptide sequence will be removed after the proteins are directed to their appropriate subcellular organelles. Thus, the signal peptide does not usually participate directly in protein function. The loss
Figure 7. The Predicted Model of Interaction between MRCAP and Its Ligands.

Heme and ferulic acid are denoted by red and purple sticks, respectively. The predicted active sites for binding heme and ferulic acid are indicated with white sticks. The positively selected site residues are indicated by green balls and hydrogen bonds with green dashed lines.

or gain of a signal peptide can easily occur in duplicate genes. If an incomplete duplication event produced only a partial copy with a partial or lost signal peptide, the protein coded by this incomplete duplication gene may immediately alter its subcellular localization. Alternative translation initiation or point mutations occurring in the stop codon can also result in the addition of amino acids in the N- or C-terminal sequence of the proteins (Leissring et al., 2004). Because altering the subcellular target by changing signal peptide sequences is rapid, the resulting functional divergence among duplicate genes through subcellular relocation should also be rapid. In plant PRXs, a highly hydrophobic C-terminal extension acts as a vacuolar sorting signal (Carter et al., 2004; Costa et al., 2008). In a horseradish PRX (HRP C1a), this vacuolar sorting signal consists of a 15-amino acid C-terminal peptide that includes a highly conserved hexapeptide that plays a key role in vacuolar sorting (Matsui et al., 2011). When the Populus PRXs in clusters I and II were aligned with HRP C1a and Arabidopsis vacuole PRXs, the vacuolar sorting signal peptide was not observed in the two cell wall PRXs (PRX6 and 24) (Figure 3B). The eight vacuole PRXS (PRX2, 3, 4, 7, 20, 21, 22, and 23) had the vacuolar sorting signal peptide as well as a highly conserved C-terminal hexapeptide (Figure 3B, black shading). Based on the putative duplication history of the PRXs in clusters I and II, the ancestor (Figure 2C, ancestral copy a) of clusters I and II was likely a cell wall type, and the ancestor (Figure 2C, ancestral copy d) of vacuole PRXs may have evolved by gaining the vacuolar sorting signal peptide.

Functional Divergence of Vacuole PRX Proteins

The Populus vacuole PRX proteins examined in this study showed different enzymatic activities and/or catalytic efficiency toward different oxidative substrates, indicating divergence in their biochemical properties. To understand the process and direction of the enzymatic divergence, we reconstructed the

### Table 3. Specific Activities of the Populus PRXs and Mutants to Six Substrates (Mean ± sO Obtained from at Least Three Independent Determinations)

<table>
<thead>
<tr>
<th>Specific Activity (μmol/min per mg)</th>
<th>Ascorbic acid</th>
<th>Pyrogallol</th>
<th>Guaiacol</th>
<th>Ferulic Acid</th>
<th>Coniferyl Alcohol</th>
<th>Sinapyl Alcohol</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRX2</td>
<td>272.42 ± 8.93</td>
<td>2138.40 ± 44.55</td>
<td>821.06 ± 13.69</td>
<td>1993.73 ± 59.47</td>
<td>4192.93 ± 10.93</td>
<td>206.72 ± 10.63</td>
</tr>
<tr>
<td>PRX3</td>
<td>379.54 ± 15.58</td>
<td>1968.64 ± 32.13</td>
<td>561.22 ± 6.45</td>
<td>2037.17 ± 29.96</td>
<td>3369.13 ± 100.62</td>
<td>181.81 ± 1.70</td>
</tr>
<tr>
<td>PRX4</td>
<td>81.96 ± 4.91</td>
<td>671.70 ± 6.60</td>
<td>219.32 ± 6.52</td>
<td>865.33 ± 15.73</td>
<td>2394.40 ± 3.64</td>
<td>55.63 ± 0.72</td>
</tr>
<tr>
<td>PRX6</td>
<td>29.61 ± 0.76</td>
<td>520.02 ± 12.64</td>
<td>115.89 ± 2.73</td>
<td>704.60 ± 43.05</td>
<td>3379.71 ± 262.40</td>
<td>37.96 ± 2.61</td>
</tr>
<tr>
<td>PRX7</td>
<td>9.37 ± 0.45</td>
<td>58.59 ± 4.59</td>
<td>43.86 ± 1.11</td>
<td>218.30 ± 5.90</td>
<td>250.56 ± 1.82</td>
<td>26.57 ± 1.61</td>
</tr>
<tr>
<td>PRX20</td>
<td>8.83 ± 0.50</td>
<td>561.60 ± 10.80</td>
<td>598.50 ± 0.50</td>
<td>953.44 ± 28.32</td>
<td>4271.29 ± 36.44</td>
<td>31.93 ± 0.94</td>
</tr>
<tr>
<td>PRX21</td>
<td>120.98 ± 4.36</td>
<td>72.90 ± 8.10</td>
<td>171.08 ± 1.79</td>
<td>446.04 ± 2.36</td>
<td>507.49 ± 3.64</td>
<td>531.30 ± 15.46</td>
</tr>
<tr>
<td>PRX22</td>
<td>65.47 ± 1.92</td>
<td>3073.14 ± 53.46</td>
<td>3569.20 ± 55.20</td>
<td>1789.82 ± 22.66</td>
<td>5352.96 ± 104.96</td>
<td>579.60 ± 15.46</td>
</tr>
<tr>
<td>PRX23</td>
<td>61.15 ± 3.75</td>
<td>1248.75 ± 12.15</td>
<td>486.92 ± 2.92</td>
<td>364.62 ± 15.34</td>
<td>1120.67 ± 16.40</td>
<td>18.84 ± 1.93</td>
</tr>
<tr>
<td>PRX24</td>
<td>42.11 ± 1.37</td>
<td>171.25 ± 2.55</td>
<td>129.25 ± 5.90</td>
<td>189.19 ± 8.37</td>
<td>977.32 ± 31.03</td>
<td>41.26 ± 1.57</td>
</tr>
<tr>
<td>MRCAP</td>
<td>310.92 ± 5.06</td>
<td>1034.04 ± 13.79</td>
<td>428.40 ± 1.17</td>
<td>1768.84 ± 63.17</td>
<td>2899.64 ± 92.26</td>
<td>156.23 ± 2.31</td>
</tr>
<tr>
<td>MRCAP(G57A)</td>
<td>381.75 ± 10.47*</td>
<td>1276.02 ± 22.68*</td>
<td>493.75 ± 10.25*</td>
<td>2410.50 ± 34.46*</td>
<td>2988.63 ± 72.16*</td>
<td>182.90 ± 5.80*</td>
</tr>
<tr>
<td>MRCAP(A92G)</td>
<td>151.54 ± 1.46*</td>
<td>1991.11 ± 45.46*</td>
<td>404.29 ± 2.96*</td>
<td>1541.22 ± 38.53*</td>
<td>3547.76 ± 66.94*</td>
<td>93.97 ± 2.30*</td>
</tr>
<tr>
<td>MRCAP(A92N)</td>
<td>263.84 ± 2.72*</td>
<td>1588.37 ± 16.97*</td>
<td>510.14 ± 1.14*</td>
<td>2184.16 ± 37.02*</td>
<td>2909.05 ± 27.33*</td>
<td>137.01 ± 1.26*</td>
</tr>
<tr>
<td>MRCAP(S161F)</td>
<td>317.39 ± 10.59</td>
<td>781.71 ± 1.32*</td>
<td>277.81 ± 4.52*</td>
<td>1204.36 ± 14.46*</td>
<td>1700.51 ± 41.99*</td>
<td>175.33 ± 4.57*</td>
</tr>
<tr>
<td>MRCAP(L167F)</td>
<td>223.39 ± 7.07*</td>
<td>892.76 ± 10.91*</td>
<td>310.25 ± 8.76*</td>
<td>1293.56 ± 17.76*</td>
<td>1364.77 ± 11.37*</td>
<td>77.47 ± 2.87*</td>
</tr>
<tr>
<td>MRCAP(G180S)</td>
<td>221.64 ± 1.49*</td>
<td>541.69 ± 3.54*</td>
<td>197.03 ± 2.13*</td>
<td>1011.26 ± 4.72*</td>
<td>1327.63 ± 55.61*</td>
<td>129.93 ± 5.31*</td>
</tr>
<tr>
<td>MRCAP(R200Q)</td>
<td>80.89 ± 2.66*</td>
<td>1383.32 ± 19.53*</td>
<td>630.00 ± 1.91*</td>
<td>1408.47 ± 22.60*</td>
<td>3040.11 ± 82.87*</td>
<td>322.69 ± 6.17*</td>
</tr>
<tr>
<td>MRCAP(R200K)</td>
<td>82.73 ± 2.40*</td>
<td>1529.65 ± 3.12*</td>
<td>558.92 ± 13.96*</td>
<td>2033.23 ± 36.31*</td>
<td>2880.51 ± 52.56*</td>
<td>103.41 ± 2.17*</td>
</tr>
<tr>
<td>MRCAP(P247V)</td>
<td>262.34 ± 2.70*</td>
<td>2125.64 ± 8.59*</td>
<td>564.72 ± 13.81*</td>
<td>2370.73 ± 21.45*</td>
<td>2586.73 ± 59.64*</td>
<td>202.71 ± 16.83*</td>
</tr>
<tr>
<td>MRCAP(P247L)</td>
<td>246.54 ± 2.44*</td>
<td>1728.00 ± 37.38*</td>
<td>448.97 ± 0.90*</td>
<td>1791.18 ± 48.81*</td>
<td>2410.94 ± 28.03*</td>
<td>133.75 ± 4.95*</td>
</tr>
</tbody>
</table>

Asterisk indicates significant difference (P < 0.05) in enzymatic activity between the MRCAP wild type and its mutants.
MRCAP of the four PRXs (PRX2, 3, 4, and 7) that were generated after the recent whole-genome duplication in Populus and examined the biochemical characteristics of the MRCAP. This paleo-molecular biochemistry approach has unraveled the evolutionary mysteries of some duplicate genes (Zhang and Rosenberg, 2002; Huang et al., 2012; Smith et al., 2013). We found that the enzymatic activities of the MRCAP toward six oxidative substrates were significantly lower than that of the PRX2 and 3 (except for PRX2 toward ascorbic acid) but higher than that of PRX4 and 7 (Table 3). The Escape from Adaptive Conflict (EAC) model might be the best explanation for this divergence pattern. Under the EAC model, after duplication each daughter gene is free to specialize for either the original or the novel function and to improve those functions (Des Marais and Rausher, 2008). If so, we might expect to observe that the different copies of PRXs perform different functions and that the ancestral protein MRCAP performed both of these functions, albeit with reduced efficiency. Increased enzymatic activities toward five oxidative substrates indicated that PRX2 and 3 improved the ancestral protein’s response to these substrates. On the other hand, PRX4 and 7 showed reduced function for these substrates and might be specialized for other substrates not tested in this study. Thus, EAC seems a clear feature of vacuole PRX evolution.

The Gene Balance Hypothesis, also known as the gene dosage hypothesis, has been proposed as a general mechanism to explain long-term retention of duplicate genes (Birchler and Veitia, 2007, 2010). The Gene Balance Hypothesis posits that genes encoding components of multisubunit complexes are more likely to evolve in concert because the dosage change in the quantities of subunits affects the interaction and function of the whole complex (Birchler et al., 2011). Such genes, e.g., ribosomal proteins, protein kinases, and transcription factors that have extensive functional interactions with other biomolecules, are preferentially retained following whole-genome duplication (Birchler et al., 2001; Edger and Pires, 2009; Freeling, 2009). In contrast, genes encode more poorly connected products in a network or complex, such as PRX and those involved in stress responses in plants, would have an elevated probability of retention following tandem duplication (Hanada et al., 2008; Freeling, 2009). The Populus vacuole PRXs expanded by tandem duplication. The increase of PRX protein abundance dictated by increased gene copy number may be beneficial in defense responses especially in changing environments. Selection for increased gene dosage allows duplications to be selectively advantageous at birth, which increases the likelihood of subsequent functional divergence in duplicate genes (Conant and Wolfe, 2008). This dosage selection may also have played a role in the retention of the tandem duplicates for novel vacuole PRX genes.

An analysis of the biochemical functions of positive selection sites provided further evidence of adaptive changes in the Populus vacuole PRXs. We used site-directed mutagenesis to investigate the functional profiles of seven positively selected sites in PRX2, 3, 4, and 7 relative to MRCAP. For example, one positive selection site (alignment position 57 in Figure 5) is Gly in PRX4, 7, and MRCAP, whereas it is Ala in PRX2 and 3. When this Gly site of MRCAP was replaced by Ala, the mutant protein G57A showed increased enzymatic activities toward five oxidative substrates (Table 3), indicating that the substitution of Ala for Gly at this amino acid site for PRX2 and 3 might confer an adaptive advantage. Conventional theory emphasizes changes in function produced by amino acid replacements at active sites (Chen et al., 1995). In this study, only one site (alignment position 200 in Figure 5) among the seven positive selection sites in PRX2, 3, 4, and 7 was an active site for binding heme, and five

<table>
<thead>
<tr>
<th></th>
<th>Coniferyl Alcohol</th>
<th>Ferulic Acid</th>
<th>Pyrogallol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/Km (mM⁻¹)</td>
<td>kcat (s⁻¹)</td>
<td>kcat/Km (mM⁻¹ s⁻¹)</td>
<td>1/Km (mM⁻¹)</td>
</tr>
<tr>
<td>PRX2</td>
<td>3.70</td>
<td>11481.57</td>
<td>42481.81</td>
</tr>
<tr>
<td>PRX3</td>
<td>1.75</td>
<td>15997.90</td>
<td>72821.33</td>
</tr>
<tr>
<td>PRX4</td>
<td>2.94</td>
<td>5969.78</td>
<td>17257.15</td>
</tr>
<tr>
<td>PRX6</td>
<td>3.85</td>
<td>9225.69</td>
<td>35518.91</td>
</tr>
<tr>
<td>PRX7</td>
<td>0.95</td>
<td>632.45</td>
<td>600.83</td>
</tr>
<tr>
<td>PRX20</td>
<td>1.89</td>
<td>12129.29</td>
<td>22924.36</td>
</tr>
<tr>
<td>PRX21</td>
<td>5.00</td>
<td>878.46</td>
<td>4392.30</td>
</tr>
<tr>
<td>PRX22</td>
<td>5.56</td>
<td>11624.53</td>
<td>64632.39</td>
</tr>
<tr>
<td>PRX23</td>
<td>4.17</td>
<td>2713.45</td>
<td>11315.09</td>
</tr>
<tr>
<td>PRX24</td>
<td>3.23</td>
<td>4199.54</td>
<td>13564.51</td>
</tr>
<tr>
<td>MRCAP</td>
<td>2.22</td>
<td>10298.32</td>
<td>22862.27</td>
</tr>
</tbody>
</table>

Figure 8. Changes in Enzymatic Activity Induced by Mutations at the Seven Selected Sites Relative to the MRCAP Protein.

The six substrates examined are each denoted by a different color. The enzymatic activity of the MRCAP toward each substrate is set as the baseline for comparison with the 10 mutants specified on the x axis.
plants are attacked by pathogens (Henriksen et al., 1999). Ferulic acid can enhance the antioxidant system protects plants against oxidative damage PRX activity. Ascorbic acid is a major metabolite in plants. It is important (Ferreres et al., 2011). In this study, we selected four environmental stresses. Thus, they have evolved various plastic their environment and are constantly exposed to a multitude of challenges they are likely to encounter.

The analysis of complete genome sequences has revealed an unexpected prevalence of gene duplications, particularly in plants. The molecular mechanisms of retention and subsequent functional divergence of duplicate genes are still poorly understood. In this study, after conducting genome-wide gene family annotation and expression analyses, we focused on two tandem-arrayed PRX gene clusters in the Populus genome. Using molecular evolutionary analyses, subcellular localization experiments, reconstruction of the ancestral protein, site-directed mutagenesis, and enzyme functional assays of ancestral and current proteins, we elucidated the evolutionary process that has contributed to the divergence of the PRX gene family in Populus and possibly other plant species. Our findings suggest that protein subcellular relocalization and positive selection play key roles in the retention and subsequent functional divergence of duplicate genes.

**METHODS**

**Identification of PRX Genes in the Populus Genome**

To identify *Populus* PRX genes, the *Populus trichocarpa* genome database version 2.2 (http://www.phytozome.net/) was searched with 73 *Arabidopsis thaliana* PRX protein sequences (Supplemental Data Set 1) using the TBLASTN program with default algorithm parameters. This search found 93 full-length PRX candidates, of which 10 (PRX5, 8, 18, 32, 38, 47, 60, 66, 69, and 76) had frame shifts disrupting the coding region or a stop codon occurring prematurely, resulting in a truncated protein. After removing these stop codons or revising the frame shifts by deleting one or two nucleotides, these 10 full-length sequences were included in the following analyses. These 93 PRX candidates were further analyzed by a search for conserved domains in the National Center for Biotechnology Information database to confirm the presence of peroxidase domains in their protein structures. All 93 *Populus* PRXs contained peroxidase domains, confirming that these genes belong to the peroxidase family. Also, we performed a joint phylogenetic analysis of all 73 *Arabidopsis* PRXs and 93 *Populus* PRXs (Supplemental Figure 5). The phylogenetic tree grouped PRXs from both species with no clear taxonomic division, further confirming that the 93 *Populus* PRXs belong to the PRX gene family. Then, the *Populus* PRX genes were amplified from mRNA from *P. trichocarpa*, cloned into the pEASY-T3 vector (TransGen), and sequenced in both directions to verify the gene sequences.

**Phylogenetic and Molecular Evolution Analyses**

The full-length PRX protein sequences were aligned using the MUSCLE version 3.8.31 (Edgar, 2004) and further adjusted manually using BioEdit version 7.0.5.3 (Hall, 1999) before all phylogenetic analysis. The optimal substitution model of amino acid substitution was selected using the program ModelGenerator version 0.84 (Keane et al., 2006). Phylogenetic
trees were constructed using the maximum likelihood method with PhyML software version 2.4.4 (Guindon and Gascuel, 2003). Two PRX proteins from *Ostreococcus lucimarinus* and *Chlamydomonas reinhardtii* were used as outgroups. One hundred bootstrap replicates were performed in each analysis to obtain the confidence support level.

To evaluate variation in selective pressures between the vacuole and cell wall PRXs in clusters I and II (Figure 4), maximum likelihood codon models in PAML version 4.4 (Yang, 2007) were used to estimate the \( \omega = (d_{S}/d_{S} + d_{I}) \) ratio under two assumptions: a one-ratio model that assumes the same \( \omega \) ratio for two types (clades) of PRXs and a two ratio model in which the two clades were assigned to different \( \omega \) ratios. To test whether some branches in the phylogenetic tree of these PRXs were under positive selection, a free-ratio model in the PAML package was used. Furthermore, the maximum-likelihood procedure that globally searches for optimal substitution models for a given phylogeny (Zhang et al., 2011) was employed as an additional model search strategy. To detect variation in site-specific selection among branches, we performed the Fitmodel test (Guindon et al., 2004) on the phylogenetic tree of the vacuole and cell wall PRXs (Figure 6). We also performed the site model analyses in the PAML package to detect positive selection sites in *PRX2, 3, 4*, and 7. Six site models in PAML were explored: the one-ratio model (M0), the discrete model (M3), the nearly neutral model (M1a), the positive-selection model (M2a), the beta model (M7), and the beta and \( \omega \) model (M8). To verify which of the models best fit the data, likelihood ratio tests were performed by comparing twice the difference in log likelihood values between pairs of the models using a \( \chi^2 \) distribution (Yang et al., 2000).

**Expression of PRX Genes under Different Treatments**

To investigate the expression patterns of the *Populus* PRX genes under normal and abiotic stress conditions, cultivars of *P. trichocarpa* were cultivated in potting soil at 25°C under a 14/10-h light/dark cycle in a growth chamber for 2 months before treatment. For *H$_2$O$_2$, salicylic acid, and NaCl* treatments, cuttings were then irrigated and sprayed using 0.5% *H$_2$O$_2* solution for 12 h, 3.5 mM salicylic acid for 24 h, and 150 mM NaCl solution once a day for 1 week. Drought stress was conducted by withholding water for 2 weeks. Each treatment consisted of three replicates. After treatment, total RNAs were isolated from root, shoot, leaf, bud, and phloem tissues of stem using an Aurum Total RNA kit (Bio-Rad cates). After treatment, total RNAs were isolated from root, shoot, leaf, withholding water for 2 weeks. Each treatment consisted of three repli-

**Expression and Purification of Recombinant Populus PRX Proteins**

To investigate the enzymatic functions of the *Populus* PRXs, 10 members in clusters I and II and the MRCP of PRX2, 3, 4, and 7 were selected for protein expression and purification (Table 3). Predictions of possible signal peptides and cleavage sites in the protein sequences were performed using the SignalP 3.1 server before gene cloning. Then, the coding sequences without the predicted N-terminal signal peptides were cloned into the pET-30a expression vector (Novagen) using the primers listed in Supplemental Table 6. The resulting models were evaluated using the profile-3D program.
alcohol was measured as described by Barceló et al. (2000). Protein concentrations were determined by measuring A280 (Layne, 1957).

The apparent $K_m$ values for pyrogallol were determined using pyrogallol concentrations ranging from 2 to 20 mM and a fixed H$_2$O$_2$ concentration of 2 mM. The apparent $K_m$ values for coniferyl alcohol were determined using coniferyl alcohol concentrations ranging from 0.02 to 0.20 mM and a fixed H$_2$O$_2$ concentration of 0.5 mM. The apparent $K_m$ values for ferulic acid were determined using ferulic acid concentrations ranging from 0.016 to 0.160 mM at a fixed H$_2$O$_2$ concentration of 0.5 mM. The kinetic parameters were derived from nonlinear regression analysis using the Hyper32 program available at http://homepage.ntlworld.com/john.easterby/hyper32.html.

**Accession Numbers**

Sequence data from this study can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the accession numbers listed in Supplemental Data Set 1 for *P. trichocarpa* and *Arabidopsis*. Other sequences used in this study included the *Populus Actin* gene (GenBank accession number XM_002316253), the *A. rusticana* PRX gene (GenBank accession number AAA33377), the *O. lucimarinus* PRX gene (GenBank accession number XM_001419012), and the *C. reinhardtii* PRX gene (GenBank accession number XM_001695424).

**Supplemental Data**

The following materials are available in the online version of this article.

**Supplemental Figure 1.** Summary of the *Populus* PRX Genes Expressed in Different Tissues under Normal Growth Conditions (NC) and Following H$_2$O$_2$ (HO), Salicylic Acid (SA), Salt (ST), and Drought (DR) Stress Treatments.

**Supplemental Figure 2.** Genomic Localization of the Full-Length *Populus* PRX Genes.

**Supplemental Figure 3.** Structural Comparison of Five PRX Proteins.

**Supplemental Figure 4.** Sequence Alignment of the *Populus* PRXs in Clusters I and II.

**Supplemental Figure 5.** Phylogenetic Relationships of the *Populus* and *Arabidopsis* PRXs.

**Supplemental Figure 6.** Modified pBI121 Vector.

**Supplemental Table 1.** PRX Fragments Identified from the *Populus trichocarpa* Genome.

**Supplemental Table 2.** PCR Primers Used to Detect the Expression of *Populus* PRX Genes.

**Supplemental Table 3.** $K_m$ Values among Pairwise Comparisons of the PRX Genes in Clusters I and II.

**Supplemental Table 4.** Substitution Rates and In L Values of the Three Optimal Branch-Specific Models Found by the Global Maximum-Likelihood Search Procedure of Zhang et al. (2011).

**Supplemental Table 5.** Primers Used to Construct the *Populus* PRX Subcellular Localization Vectors.

**Supplemental Table 6.** Primers Used to Construct the *Populus* PRX Protein Expression Vectors.

**Supplemental Data Set 1.** Full-length PRX Genes Identified from the *Populus trichocarpa* Genome and *Arabidopsis* PRX Protein Sequences.

**Supplemental Data Set 2.** Text File of the Alignment Corresponding to the Phylogenetic Tree in Figure 1.

**Supplemental Data Set 3.** Text File of the Alignment Corresponding to the Phylogenetic Tree in Figure 2B.

**Supplemental Data Set 4.** Text File of the Alignment Corresponding to the Phylogenetic Tree in Figure 4.

**Supplemental Data Set 5.** Text File of the Alignment Corresponding to the Phylogenetic Tree in Supplemental Figure 5.

**ACKNOWLEDGMENTS**

We thank Chengjun Zhang of University of Chicago for assistance in branch model tests. This study was supported by grants from the Natural Science Foundation of China (NSFC 91231103 and 31330017).

**AUTHOR CONTRIBUTIONS**


Received February 26, 2014; revised April 18, 2014; accepted May 24, 2014; published June 16, 2014.

**REFERENCES**


Byrne, K.P., and Wolfe, K.H. (2007). Consistent patterns of rate asymmetry and gene loss indicate widespread neofunctionalization...


Subcellular Relocalization and Positive Selection Play Key Roles in the Retention of Duplicate Genes of *Populus* Class III Peroxidase Family

Lin-Ling Ren, Yan-Jing Liu, Hai-Jing Liu, Ting-Ting Qian, Li-Wang Qi, Xiao-Ru Wang and Qing-Yin Zeng

*Plant Cell* 2014;26:2404-2419; originally published online June 16, 2014;
DOI 10.1105/tpc.114.124750

This information is current as of January 15, 2018

| Supplemental Data | /content/suppl/2014/06/09/tpc.114.124750.DC1.html |
| References | This article cites 66 articles, 23 of which can be accessed free at: /content/26/6/2404.full.html#ref-list-1 |
| eTOCs | Sign up for eTOCs at: http://www.plantcell.org/cgi/alerts/ctmain |
| CiteTrack Alerts | Sign up for CiteTrack Alerts at: http://www.plantcell.org/cgi/alerts/ctmain |
| Subscription Information | Subscription Information for *The Plant Cell* and *Plant Physiology* is available at: http://www.aspb.org/publications/subscriptions.cfm |