Xu Hui (Orcid ID: 0000-0002-5212-9012)
Liu Yan-Jing (Orcid ID: 0000-0002-1986-543X)

Re-submitted to *The Plant Journal*
Previous vesrion: TPJ-01469-2021

# Divergence of Active Site Motifs among Different Classes of *Populus* Glutaredoxins Results in Substrate Switches

Hui Xu[1,2], Zhen Li[3,4], Peng-Fei Jiang[1,2], Li Zhao[5], Chang Qu[1], Yves Van de Peer[3,4,6], Yan-Jing Liu[1*] and Qing-Yin Zeng[1,2*]

[1]State Key Laboratory of Tree Genetics and Breeding, Chinese Academy of Forestry, Beijing 100091, China
[2]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, 100093, China
[3]Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium
[4]VIB Center for Plant Systems Biology, 9052 Ghent, Belgium
[5]Department of Ecology and Environmental Science, Umeå University, SE-90187, Umeå, Sweden
[6]Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria 0028, South Africa

**Running title**: Active site motif evolution of *Populus* GRX family

***Key words:*** Active site motif, Gene family, Functional divergence, Glutaredoxin, Enzymatic specificity，*Populus trichocarpa*

***Corresponding author:** Yan-Jing Liu
State Key Laboratory of Tree Genetics and Breeding, Chinese Academy of Forestry
Beijing 100091, China
E-mail: yan.jing.liu@163.com
Phone: +86-13811118971

***Corresponding author:** Qing-Yin Zeng
State Key Laboratory of Tree Genetics and Breeding, Chinese Academy of Forestry
Beijing 100091, China
E-mail: qingyin.zeng@caf.ac.cn
Phone: +86-13699127045

**Summary**

Enzymes are essential components of all biological systems. The key characteristics of proteins functioning as enzymes are their substrate specificities and catalytic efficiencies. In plants, most genes encoding enzymes are members of large gene families. Within such families, the contribution of active site motif to the functional divergence of duplicate genes has not been well elucidated. In this study, we identified 41 glutaredoxin (GRX) genes in *Populus trichocarpa* genome. GRXs are ubiquitous enzymes in plants that play important roles in developmental and stress-tolerance processes. In poplar, GRX genes were divided into four classes based on clear differences in gene structure and expression pattern, subcellular localization, enzymatic activity, and substrate specificity of the encoded proteins. Using site-directed mutagenesis, this study revealed that the divergence of the active site motif among different classes of GRX proteins resulted in substrate switches and thus provided new insights into the molecular evolution of these important plant enzymes.

## Introduction

Proteins functioning as enzymes catalyze biochemical reactions by lowering the activation energy of reactions. A key characteristic of the catalytic activity of an enzyme is its substrate specificity. Enzymes bind to specific substrates and initiate the catalytic process via a so-called active site motif comprising a substrate-binding site and a catalytic site. In plants, genes encoding enzymes involved in responses to biotic or abiotic stimuli tend to belong to large gene families, such as the late embryogenesis abundant protein and glutathione S-transferase families (Lan *et al.*, 2013, Liu *et al.*, 2013). In general, enzymes belonging to the same family share an active site motif that is usually relatively conserved in its amino-acid sequence. Variations in the active site motif among different members of a family may alter substrate specificity and catalytic activity of the enzyme and result in functional divergence. Members of the same enzyme family have been shown to have distinct catalytic activities toward specific substrates (Dixon *et al.*, 2009, Lan *et al.*, 2009). However, the lack of appropriate assays integrating both evolutionary and functional insights has generally hindered attempts to elucidate the influence of variations in the active site motif on substrate specificity and enzyme activity.

Glutaredoxins (GRXs; EC 1.20.4.1) are small ubiquitous oxidoreductases essential for the response to oxidative stress. GRXs can mediate the reversible reduction of disulfide bridges or glutathionylated proteins in the presence of glutathione (GSH) via a dithiol or monothiol mechanism (Lillig *et al.*, 2008). Plant GRXs are encoded by a large gene family. For example, *Arabidopsis thaliana* and *Oryza sativa* genomes contain 33 and 29 GRX genes, respectively (Couturier *et al.*, 2009a). Plant GRXs have several types of active site motifs such as CPYC, CSY[C/S], CGFS, and CCx[C/S]. GRXs with CPYC and CSY[C/S] active site motifs have been shown to be involved in the response to oxidative stress (Fernandes and Holmgren, 2004, Cheng *et al.*, 2006), whereas GRXs with the CGFS motif are involved in the [Fe-S] cluster assembly (Rouhier *et al.*, 2010). GRXs with CCx[C/S] motifs regulate the development of flower organs (e.g., anthers and petals) and are important for defense against pathogens (Ndamukong *et al.*, 2007, Wang *et al.*, 2009). Because enzyme functions differ according to the type of the active site motif, the plant GRX family is an ideal model system for investigating the relationship between variations in the active site motif and functional dynamics.

Here, we consider *Populus trichocarpa* as our plant model, as it is one of the most important perennial tree models with a well-studied genome (Tuskan *et al.*, 2006). As a perennial genus, *Populus* has a long generation time and is widely distributed globally. During their lifespan, *Populus* trees have to tolerate severe stressors such as temperature fluctuations, drought, and pathogen attacks. To deal with most, if not all of these stressors, GRX genes play a crucial role. In this study, we identified members of the GRX family within the *Populus* genome and reconstructed the evolutionary history of this family. We then conducted a comprehensive analysis of the gene sequences, structures, and expression patterns, subcellular localizations, and enzymatic properties of wild-type and mutant GRX proteins. Our results revealed that the variations in active site motif play key roles in the divergence of enzymatic activities of *Populus* GRX subfamilies.

## Results

### Large GRX Gene Family in *Populus* Genome

We identified 41 full-length genes containing the glutaredoxin domain in *P. trichocarpa* genome (Supplementary Table S1). Among these 41 GRX genes, *PtGRXA2* was considered a pseudogene because it contained frame shifts that disrupted the coding region. After deleting several nucleotides to correct the frame shift, this gene was used in the subsequent phylogenetic analyses.

Based on the phylogenetic tree, the 41 *Populus* GRXs were divided into four distinct clades (gray, blue, green, and purple clades in Fig. 1, referred to the alpha, beta, gamma, and delta GRX classes, respectively). Except for the beta clade, the other three clades had a bootstrap support of $\geqslant$ 74%. To verify the classification of *Populus* GRXs, we used 371 GRX genes identified from 13 other land-plant species that represent the four major lineages of land plants (bryophytes, lycophytes, gymnosperms, and angiosperms, Supplementary Table S2), together with the 41 *Populus* GRX genes, to reconstruct the evolutionary history of the GRX gene family (Fig. 2A). All GRXs were grouped into four distinct clades that correspond to the alpha (α), beta (β), gamma (γ) and delta (δ) GRX classes in *P. trichocarpa* with $\geqslant$74% bootstrap support. The nomenclature for plant GRXs is not consist in different studies, which may cause misinterpretations. Thus, we propose that each plant GRX gene should be named based on its abbreviated species name, a gene class identifier, and a number within that gene class. A, B, G and D are used as gene class identifiers corresponding to the alpha (α), beta (β), gamma (γ) and delta (δ) GRX classes, respectively. For example, *PtGRXA1* indicates that this GRX gene is isolated from species *Pt* (*Populus trichocarpa*), while *A1* indicates that this gene is the first member of the alpha class.

*Populus trichocarpa* genome contained 27 alpha, six beta, five gamma, and three delta class GRX members. Two bryophytes, *Marchantia polymorpha* and *Physcomitrella patens*, did not contain any delta class GRXs, which were present in all vascular plants (Fig. 2B). *Selaginella moellendorffii*, *M. polymorpha*, and *P. patens* contained three, two, and two alpha GRXs, respectively. *Pinus taeda* and *Amborella trichopoda* each had eight alpha GRXs. Most angiosperms had more than 17 alpha GRXs, except for *Brachypodium distachyon*, which presented nine alpha GRXs. Thus, the number of alpha GRXs has expanded rapidly in angiosperms.

### Structural Features of *Populus* GRX Genes

Within each GRX class, except for gamma, the gene structure is generally conserved (Fig. 3C). The delta GRX genes contain five introns, and the intron positions are conserved. The alpha GRX genes generally do not contain introns, except for *PtGRXA27*, which has one intron. Of the six beta GRX genes, five have three introns, and one (*PtGRXB4*) has four introns. The structures of the five gamma GRX genes are highly variable. *PtGRXG1*, *2* and *5* contain two introns, *PtGRXG3* five introns, and *PtGRXG4* one intron.

Each *Populus* GRX protein has only one GRX domain, except for PtGRXG1, PtGRXG2, PtGRXD1, PtGRXD2 and PtGRXD3 (Fig. 3D). PtGRXG1 and PtGRX2 contain three GRX domains and one thioredoxin domain. PtGRXD1, PtGRXD2 and PtGRXD3 have one GRX domain and two other protein domains (DEP and DUF547 domains). The active site motifs of the four

classes of GRXs are located in the N-terminus of the GRX domain. *Populus* alpha, beta, gamma and delta GRXs possess the CCx[C/S], Cxx[C/S], CGFS and C[Q/R]DC types of active site motifs, respectively. The class-specific active site motif type further supports the class designations among the 41 *Populus* GRXs.

**Genomic Distribution of *Populus* GRX Gene Family**

We investigated the distribution of GRX genes among *Populus* chromosomes to elucidate the duplication mechanism and expansion history of the GRX family (Supplementary Fig. S1). All GRX genes are located on 15 *Populus* chromosomes. Chromosomes 5, 9, 13 and 19 do not harbor any GRX genes. Two GRX gene clusters (clusters I and II) are found on chromosomes 2 and 14. Clusters I and II comprise 6 and 7 alpha GRX genes, respectively. The *Populus* genome underwent a whole-genome duplication (WGD) event 60–65 million years ago (Tuskan *et al.*, 2006). Clusters I and II were resulted from this recent WGD event. Based on the phylogenetic tree and the genomic positions of these 13 alpha GRX genes, we reconstructed the expansion history of these two clusters. The most parsimonious scenario for gene duplication is presented in Fig. 4. It is likely that four ancestral genes created by two rounds of tandem duplications existed in *Populus* genome before the recent WGD event. After the WGD event, one ancestral gene in cluster I and two ancestral genes in cluster II underwent tandem duplication (Fig. 4C). Based on the evolutionary history of clusters I and II, *PtGRXA19/25* and *20/26* are identified as duplicate pairs created by the recent WGD event (Fig. 4C). In addition, we also found that seven duplicate pairs (*PtGRXA1/2*, *5*/*6*, *7*/*8*, *9*/*11*, *PtGRXB1*/*2*, *PtGRXG1*/*2*, and *PtGRXD1*/*2*) are each located in a pair of paralogous blocks created by the recent WGD event (Supplementary Fig. S1). These seven duplicate pairs were also considered as direct results of the recent WGD event.

**Expression Patterns of *Populus* GRX Genes**

The expression of the 41 *Populus* GRX genes was examined in five tissues (internodes, nodes, roots, young leaves, and mature leaves) using the PopGenIE database (Fig. 3B). No expression data for *PtGRXA3*, *4*, *22*, or *PtGRXD1* could be found in the database. Expression data were available for the other 37 GRX genes. Thus, this study only analyzed the expression patterns of these 37 GRX genes with available expression data.

*Populus* beta GRX genes showed significant differences in gene expression patterns with alpha class GRXs ($P < 0.03$, MRPP nonparametric test). The expression patterns of *Populus* gamma GRX genes were significantly different from those of alpha and delta class GRXs, respectively ($P < 0.05$, MRPP nonparametric test). We also found that gene expression varied within the classes. In the alpha class, the expression levels of *PtGRXA1*, *2*, *12*, *13* and *25* were higher than those of the other alpha genes ($P$ = 1e-4, MRPP nonparametric test). In the beta class, the expression level of *PtGRXB2* was lower than those of the other five beta class genes. GRX genes also exhibited greater expression in some specific tissues. For example, *PtGRXA12* and *13* were expressed at higher levels in the roots and mature leaves than in the internodes, nodes, and young leaves.

**Subcellular Localizations of *Populus* GRX Proteins**

Proteins must localize to the appropriate subcellular compartments to ensure their proper function. Of the 41 *Populus* GRX genes, one (*PtGRXA2*) was a pseudogene, and four (*PtGRXA3*, *8*, *17* and *27*) were not successfully cloned in this study. These five GRX genes were therefore omitted from the analysis of protein subcellular localizations. The remaining 36 *Populus* GRXs (22 alpha, 6 beta, 5 gamma and 3 delta GRXs) were selected to investigate their subcellular localization by transiently expressing C-terminal GFP fusion GRXs and organelle markers in epidermal cells of *Nicotiana benthamiana* and visualizing their expression using confocal microscopy.

All the alpha GRX proteins were located in the cytosol and nucleus as illustrated by PtGRXA1–GFP in Fig. 5 (Supplementary Fig. S2). The six beta GRX proteins were located in different subcellular compartments. PtGRXB1, 2, and 3 were located in the cytosol and nucleus. PtGRXB5 and 6 were found in the endoplasmic reticulum as illustrated by PtGRXB5-GFP in Fig. 5 (Supplementary Fig. S2). The fluorescence of PtGRXB4 was coincided with the autofluorescence of the chlorophyll, thus, PtGRXB4 was considered as a chloroplast-located protein (Fig. 5). Of the five gamma GRXs, PtGRXG1 and 2 were located in the cytosol and nucleus. The fluorescent signal of PtGRXG3 was merged with the mCherry signal of the mitochondria marker. Thus, PtGRXG3 was considered as a mitochondria-located protein (Fig. 5). PtGRXG4 and G5 were located in the chloroplasts (Supplementary Fig. S2). All three delta GRXs were detected as having weak fluorescent signal overlapping with the mCherry signal of the endoplasmic reticulum marker, indicating that these three proteins were localized to the endoplasmic reticulum (Supplementary Fig. S2).

**Expression and Purification of *Populus* GRX Proteins**

The catalytic characteristics of the *Populus* GRX proteins may be related to the response to oxidative stress as antioxidants. Among the 41 *Populus* GRX genes, except for the pseudogene *PtGRXA2* and the four genes (*PtGRXA3*, *8*, *17* and *27*) that could not be successfully cloned, 36 *GRX*s (22 alpha, 6 beta, 5 gamma and 3 delta *GRX*s,) were subcloned into *Escherichia coli* for protein expression.

Twenty two alpha GRXs with N-terminal 6-histidine (6×His) tag were subcloned into *E. coli* for protein expression. Among these 22 GRXs, five (PtGRXA6, 7, 9, 11, and 18) were not expressed in *E. coli*, 12 (PtGRXA4, 5, 10, 16, 19, 20–26) were expressed as inclusion bodies, and five (PtGRXA1, 12–15) could be expressed as soluble proteins. After purification by Nickel–Sepharose High Performance column, these five soluble GRX proteins were not stable and easily precipitated in the enzyme assay buffer. We then also used GST Fusion and Purification System to express and purify the GRX proteins. Based on *Populus* GRX phylogenetic tree, we selected eight alpha GRXs (PtGRXA1, 2, 5, 11, 12, 16, 20, and 21) to construct GST-tagged protein expression vectors. Unfortunately, seven of the eight tested GRXs were expressed as inclusion bodies or even unexpressed, and only PtGRXA12 was soluble. We performed purification and activity analysis on GST-tagged PtGRXA12. The PtGRXA12 did not show any activity toward four substrates. In this case, to obtain soluble and stable GRX proteins, we used MBP (maltose binding

protein) Fusion and Purification System to express 22 alpha GRXs. All the 22 alpha GRXs could be expressed as soluble proteins in *E. coli* when fused with N-terminal MBP tag (Supplementary Fig. S3). These 22 recombinant alpha GRXs were purified using the Amylose Resin column. Purified recombinant alpha GRXs were stable in the enzyme assay buffer. However, after deleting the MBP tag by protease factor Xa, the alpha GRX proteins were unstable and easily precipitated in the enzyme assay buffer. Thus, we have to use purified MBP-tagged recombinant proteins to test substrate specificity and enzymatic activity of 22 alpha GRXs.

All *Populus* beta, gamma and delta GRXs with N-terminal 6×His tag were expressed as soluble proteins in *E. coli*, except for PtGRXB2 and PtGRXD1, which could not be expressed in *E. coli* (Supplementary Fig. S3). After purification by Nickel-Sepharose High Performance column, these 5 beta, 5 gamma and 2 delta GRX proteins were stable in the enzyme assay buffer. Thus, the substrate specificity and enzymatic activity of *Populus* beta, gamma and delta GRXs were tested using purified His-tagged recombinant proteins.

Previous studies showed that some GRX proteins, such as *Arabidopsis* GRXC1, could form homodimers by ligating iron–sulfur clusters, leading to the GRXs inactivation (Riondet *et al.*, 2012). When this kind of GRXs was expressed in *E. coli*, part of the recombinant GRXs existed as active monomeric proteins, and the other part as inactive iron–sulfur homodimers or/and tetramers (Lillig *et al.*, 2005, Riondet *et al.*, 2012, Abdalla *et al.*, 2018). In this study, to obtain monomeric GRXs, all GRX proteins purified by affinity chromatography column (Amylose Resin column or Nickel–Sepharose High Performance column) were further separated by size-exclusion chromatography (Supplementary Fig. S3). All isolated monomeric GRX proteins were used to determine their enzymatic activities.

**Substrate Activity of *Populus* GRX Proteins**

To investigate substrate specificity and enzymatic activity, we performed *in vitro* catalytic actions on the purified *Populus* GRXs toward four substrates commonly used to detect the redox characteristics of GRX: bis(2-hydroxyethyl) disulfide (HED), L-cystine, dehydroascorbate (DHA), and cumene hydroperoxide (Cum–OOH).

Among the 22 alpha GRX proteins examined, 14 showed enzymatic activities only toward substrate Cum–OOH, eight did not show any activity toward any of the substrates. This indicates that *Populus* alpha GRXs may have only peroxidase activity. Although all these 14 alpha GRXs reacted with Cum–OOH, their activities varied 6.24-fold at the highest.

All five beta GRXs examined had enzymatic activities toward substrates HED, L-cystine and DHA. Except for PtGRXB1, four beta GRXs (PtGRXB3–6) also showed enzymatic activities toward Cum–OOH. This indicates the beta GRXs not only have peroxidase activities, but also have thioltransferase activities and DHA reductase activities.

All gamma and delta GRXs (PtGRXG1–5, and PtGRXD2 and 3) did not show any enzymatic activity toward four substrates.

This study identified nine duplicate gene pairs formed by a recent WGD event. Among these nine duplicate pairs, enzyme specificities of five duplicate pairs (PtGRXA5/6, 9/11, 19/25 and

20/26 and PtGRXG1/2) could be examined. PtGRXA9/11 and PtGRXG1/2 duplicate pairs did not show any activity toward the four substrates examined. For PtGRXA5/6, 19/25, and 20/26 duplicate pairs, although both enzymes of each pair only showed enzymatic activities toward Cum–OOH, their activities were significantly different ($P < 0.05$, Mann–Whitney U test; Fig. 6A).

In this study, GST-tagged or His-tagged GRX proteins were mainly expressed as inclusion bodies or even unexpressed, while MBP-tagged GRXs could be expressed as soluble proteins in *E. Coli*. To compare the difference in substrate specificity and enzyme activity of GRX fused with different tags, we expressed and purified PtGRXB4 proteins tagged with His, GST, and MBP tags, respectively (Supplementary Table S3). Compared to MBP-tagged PtGRXB4, His-tagged or GST-tagged PtGRXB4 showed much higher enzymatic activities toward the substrates HED, DHA, and L-cystine. PtGRXB4 proteins with His tag showed very weak enzymatic activity toward Cum–OOH, whereas PtGRXB4 protein with GST or MBP tag had no enzymatic activity toward Cum–OOH.

**Mutagenesis Analysis**

This study revealed that GRXs belonging to different classes have substrate preferences. To investigate whether substrate preference is related to the active site motif, this study exchanged active site motifs among different classes of GRX proteins for biochemical assays. PtGRXB4 showed activity to all four tested substrates. The crystal structure of the GRX domain in GRXS12 protein (PDB: 3FZ9) from *Populus tremula* × *tremuloides* was analyzed (Couturier *et al.*, 2009b). The GRX domains of PtGRXB4 and GRXS12 shared identical protein sequences, indicating that we could clearly understand the three-dimensional structure of the GRX domain of PtGRXB4. Thus, we selected PtGRXB4 for the mutagenesis analysis, which assists the understanding of changes in biochemical functions at the structural level. After deleting the signal sequence, the active site motif (CSYS) of PtGRXB4 was located at positions 28–31. The $^{28}$CSYS$^{31}$ of PtGRXB4 was mutated to the active site motifs present in the alpha, beta, gamma, and delta class GRXs (Fig. 7, Supplementary Table S4).

*Populus* alpha GRXs have four types of active site motifs: CCMC, CCMS, CCLC, and CYMS. Among the 27 *Populus* alpha GRXs, 14 have CCMC motifs, 10 have CCMS motifs, two have CCLC motifs, and one (PtGRXA27) has a CYMS motif. Because CCMC and CCMS motifs are most common in alpha GRXs, we selected CCMC and CCMS as the target motifs for mutagenesis analysis. Compared with wild-type PtGRXB4, $^{28}$CCMC$^{31}$ and $^{28}$CCMS$^{31}$ mutants showed higher enzymatic activity toward Cum–OOH ($P < 0.05$, Mann–Whitney U test) and lower enzymatic activities toward substrates HED and DHA ($P < 0.05$, Mann–Whitney U test). Consistent with the decrease of substrate activities, the substrate affinities ($1/K_m$) and catalytic efficiencies ($k_{cat}/K_m$) of $^{28}$CCMC$^{31}$ or $^{28}$CCMS$^{31}$ mutants for GSH, HED and DHA were lower than those of wild-type PtGRXB4 ($P < 0.05$, Mann–Whitney U test; Table 1).

*Populus* beta GRXs have four types of active site motifs: CPYC, CGYC, CPFC, and CSYS. PtGRXB5 and PtGRXB6, both of which possess the active site motif CPYC, showed much higher activities than those of PtGRXB4 toward the four test substrates ($P < 0.05$, Mann–Whitney U test;

Fig. 6A). When the $^{28}$CSYS$^{31}$ active site motif of PtGRXB4 was replaced with CPYC, compared with wild-type PtGRXB4, the $^{28}$CPYC$^{31}$ mutant protein exhibited higher activities toward all the four tested substrates; it also exhibited higher substrate affinities and catalytic efficiencies for GSH, HED, L-cystine and DHA ($P < 0.05$, Mann–Whitney U test; Fig. 7, Table 1).

PtGRXB1, of which CGYC was the active site motif, showed much lower activities toward HED, L-cystine, and DHA than did other beta GRXs and had and no activity toward Cum–OOH ($P < 0.05$, Mann–Whitney U test; Fig. 6A). When the $^{28}$CSYS$^{31}$ motif of PtGRXB4 was replaced with CGYC, the $^{28}$CGYC$^{31}$ mutant showed lower enzymatic activities, substrate affinities, and catalytic efficiencies toward substrates HED, L-cystine, and DHA compared with the wild-type proteins ($P < 0.05$, Mann–Whitney U test; Fig. 7, Table 1). Further, the $^{28}$CGYC$^{31}$ mutant did not show any activity to Cum–OOH.

PtGRXB3 had the active site motif of CPFC, and it showed much higher activities toward four test substrates than those of PtGRXB4 ($P < 0.05$, Mann–Whitney U test; Fig. 6A). Compared with the wild-type PtGRXB4, the $^{28}$CPFC$^{31}$ mutant protein exhibited higher enzymatic activity with HED ($P < 0.05$, Mann–Whitney U test; Fig. 7). Correspondingly, the substrate affinities and catalytic efficiencies of the $^{28}$CPFC$^{31}$ mutant for HED were higher than those of PtGRXB4 ($P < 0.05$, Mann–Whitney U test; Table1).

*Populus* gamma GRXs possess the CGFS motif. When the $^{28}$CSYS$^{31}$ active site motif of PtGRXB4 was replaced with CGFS, compared to wild-type PtGRXB4, the mutant protein showed much higher enzymatic activities toward substrates HED, L-cystine, and Cum–OOH ($P < 0.05$, Mann–Whitney U test), and lower activity toward substrate DHA ($P < 0.05$, Mann–Whitney U test). Consistent with the changes in substrate activities, the substrate affinities and catalytic efficiencies of the $^{28}$CGFS$^{31}$ mutant for GSH and L-cystine, and the catalytic efficiencies for HED were increased when compared with those of the wild type ($P < 0.05$, Mann–Whitney U test); conversely, the substrate affinity and catalytic efficiency of the $^{28}$CGFS$^{31}$ mutant for DHA was significantly decreased ($P < 0.05$, Mann–Whitney U test; Fig. 7, Table 1).

*Populus* delta GRXs have two types of active site motifs: CQDC and CRDC. *Populus* delta GRXs showed no activity to any of the four substrates. When the $^{28}$CSYS$^{31}$ active site motif of PtGRXB4 was replaced with CQDC or CRDC, both $^{28}$CQDC$^{31}$ and $^{28}$CRDC$^{31}$ mutants did not show any activity to Cum–OOH and had much lower enzymatic activities, substrate affinities, and catalytic efficiencies toward substrates HED, L-cystine, and DHA compared with wild-type PtGRXB4 ($P < 0.05$, Mann–Whitney U test, Fig. 7, Table 1).

We also mutated the active site motif of one alpha GRX (PtGRXA13) to the active site motifs presented in other classes of GRXs. Among *Populus* alpha GRXs, the proteins with the CCMC active site motif were the most (14 of 27 alpha GRXs). PtGRXA13 had the highest catalytic activity toward Cum–OOH in the proteins with the CCMC active site motif (Fig. 6A). Thus, we selected PtGRXA13 for the mutagenesis analysis (Table 2). The $^{21}$CCMC$^{24}$ of PtGRXA13 was mutated to CSYS, present in beta GRXs, CGFS, present in gamma GRXs; and CRDC, present in delta GRXs. The $^{21}$CGFS$^{24}$ mutant and $^{21}$CRDC$^{24}$ mutant showed no catalytic activity toward any substrate. The $^{21}$CSYS$^{24}$ mutant showed weak activity toward L-cystine but no activity toward Cum–OOH.

**Structural Variation of the GSH–binding Pocket Among PtGRXB4 and Its Mutants**

In this study, we investigated structural differences between PtGRXB4 and its mutants, which contribute to differences in enzymatic characteristics. Based on the crystal structure of PtGRXB4 (PDB:3FZ9), we modeled the structures of the $^{28}$CCMC$^{31}$, $^{28}$CPYC$^{31}$, $^{28}$CGFS$^{31}$, and $^{28}$CRDC$^{31}$ mutants of PtGRXB4. An L-shaped pocket binding GSH was observed in the GRX domain in both wild-type and mutated PtGRXB4 (Fig. 8A). The first ($C^{28}$) and third amino acids ($M^{30}$, $Y^{30}$, $F^{30}$ or $D^{30}$) in the active site motifs were located at one arm of the L-shaped pocket. Owing to the different structures and sizes of the third amino acid in different active site motifs, the edge shape and size of the L-shaped GSH–binding pocket of the wild-type and mutated proteins were slightly different (Fig. 8A, 8B).

To further understand the catalyzed mechanism of GRXs to the four substrates tested in this study, this study used molecular docking to predict the structures of PtGRXB4 in complex with the substrates HED, DHA, L-cystine, and Cum–OOH, respectively (Supplementary Fig. S4). All four substrates can bind to the active site pocket of the GRX protein. And the four substrates bind to GRX near the first cysteine ($C^{28}$) of active site motif.

**Discussion**

**Classification and Rapid Expansion of the GRX Gene Family in Angiosperms**

GRXs are glutathione dependent oxidoreductases essential for responding to oxidative stress (Fernandes and Holmgren, 2004). They belong to large gene families, presenting 29 and 33 members in *Oryza* and *Arabidopsis*, respectively (Couturier *et al.*, 2009a). In a previous study, researchers identified 38 GRXs from *Populus trichocarpa* genome (Couturier *et al.*, 2009a). in this study, we identified 41 full-length GRX genes from the latest version of *P. trichocarpa* genome. As a large gene family, clear naming and classification of members are important for understanding the functional characteristics of the family. Although there is already a reference nomenclature for GRXs (Couturier *et al.*, 2009a), it should be noted that the current nomenclature for plant GRXs is confusing. The GRXs belonging to the same subclass had two different naming formats. For example: *At4g28730* in *A. thaliana* and *Os08g45140* in *O. sativa* belonged to the same subclass, but they were named as *AtGrxC5* and *OsGrxS12*, respectively (Couturier *et al.*, 2009a). In addition, the same genes had different names in different studies. For example, a rice GRX gene (*Os04g42930*) was named *OSGRXC2.2* or *OsGRX14* (Garg *et al.*, 2010, Morita *et al.*, 2015). The principle of Greek alphabet designations was widely used for other plant large gene families, such as GST, cyclin, and expansin gene families (Renaudin *et al.*, 1996, Dixon *et al.*, 2002, Li *et al.*, 2002). In this context, we propose a new nomenclature system for plant GRX genes (Please see "Large GRX Gene Family in *Populus* Genome" in the results section). Based on the phylogenetic relationship of land plant GRXs, plant GRXs have been divided into four classes: alpha (α), beta (β), gamma (γ) and delta (δ) GRX classes (Fig. 2A). The class-specific active site motifs, features of gene structure, and the class-specific enzymatic activities further supported the class designations among plant GRXs.

The GRX gene family in angiosperms expanded more rapidly than in bryophytes, lycophytes, and gymnosperms. The expansion of GRX family in angiosperms is mainly due to the expansion of alpha GRXs (Fig. 2B). The rapid expansion of alpha GRX genes in angiosperms can likely be explained by functional requirements. Indeed, plant alpha GRX genes have diverse functions. For example, *At5g14070* and *At3g02000*, two *Arabidopsis* alpha GRX genes, are important for anther development (Xing and Zachgo, 2008, Yang *et al.*, 2008); four *Arabidopsis* GRX genes (*At4g15700*, *At4g15680*, *At4g15690*, and *At4g15660*) control primary root growth (Patterson *et al.*, 2016), whereas *At1g03850* (GRXS13) plays a key role in protecting the plant against photoxidative stress (Laporte *et al.*, 2012). Compared with bryophytes and lycophytes, angiosperms have more complex organ systems and structures, such as flowers. New organs might require more alpha GRXs to maintain their biological functions. The expansion of alpha GRX likely provided new raw genetic material necessary for the evolution of new functions for these new organs.

**Functional Divergence of Duplicate Gene Pairs Created by Whole-genome Duplication**

Clear divergence in expression patterns and enzymatic specificity was observed among duplicate pairs formed by a WGD event (Fig. 3B and 6A). Several studies showed that the duplicate pairs created by a WGD event tend to maintain similar expression patterns, whereas duplicate pairs created by tandem duplications tend to diverge rapidly (Ganko *et al.*, 2007, Yang *et al.*, 2013). However, in this study, *Populus GRX* duplicate pairs created by the WGD event exhibited rapid divergence in gene expression patterns. In a study on *Arabidopsis*, > 50% of the duplicate gene pairs formed by the most recent polyploidy event showed divergent expression profiles (Blanc and Wolfe, 2004). Casneuf *et al* (2006) reported a strong bias in the divergence of gene expression toward gene function. Genes associated with stress responses diverge quicker after duplication, than do genes associated with macromolecular metabolism (Casneuf *et al.*, 2006). GRXs are largely associated with cellular response to oxidative stress. The characteristics of GRX genes related to ess response may be the reason why *Populus* GRX duplicate pairs created by the WGD event had rapid divergence in gene expression patterns.

This study examined enzymatic specificities of *Populus* GRX duplicate pairs created by the *Populus* specific WGD event. Although three duplicate pair (PtGRXA5/6, 19/25 and 20/26) had similar substrate spectrum, they showed differences in enzymatic activities (Fig. 6A). This result indicates that the divergences of biochemical functions had occurred in WGD-derived *Populus* GRX duplicate pairs. From this perspective, the divergences of enzymatic activities of coding proteins might have contributed to the retention of WGD-derived GRX genes in *Populus* genome.

**Variation in Subcellular Localizations of *Populus* GRXs**

Most proteins produced by eukaryotes are synthesized in the cytosol, and many need to be further sorted to other subcellular compartments. Proteins could alter their functions when relocated to a new subcellular compartment (Leissring *et al.*, 2004). Even the same protein might have a different biological function when localized to different subcellular structure (Leissring *et al.*, 2004). Among the 41 *Populus* GRXs, previous studies only showed the subcellular localizations of five

GRX proteins through experiments (Rouhier *et al.*, 2007, Bandyopadhyay *et al.*, 2008, Couturier *et al.*, 2009b), whereas the subcellular localizations of other proteins were only determined based on in silico prediction. In this study, we systematically showed the subcellular localizations of *Populus* GRXs through experiments. All *Populus* alpha GRX proteins were located in the cytosol and nucleus, whereas all *Populus* delta GRXs were localized to the endoplasmic reticulum (Supplementary Fig. S2 and Fig. 6C). This result indicates that the biological functions of delta GRXs are different from those of alpha GRXs. Among the six *Populus* beta GRX proteins, three different subcellular localizations were observed. Similar to beta GRX, the six *Populus* gamma GRX also showed three different subcellular localizations (Supplementary Fig. S2 and Fig. 6C). Different subcellular localizations among different members of the same subfamily indicated functional differentiation. Many studies showed that diversity of protein subcellular localizations within gene families is common (Liu *et al.*, 2013, Ren *et al.*, 2014). The subcellular relocalization of duplicate proteins might facilitate functional diversification or might lead to the origin of new functions (Byun-McKay and Geeta, 2007).

**Divergence of Active Site Motifs of *Populus* GRXs Resulted in Substrate Switches**

We found that different classes of *Populus* GRXs had different preferred substrates. For example, the enzymatic activities of *Populus* beta GRX proteins toward the substrates HED, L-cystine, and DHA were significantly higher than their activities toward the substrate Cum–OOH, whereas *Populus* alpha GRXs had only enzymatic activities toward Cum–OOH. The results of the mutagenesis analysis showed that when the active site motif of a beta GRX (PtGRXB4) was exchanged between different classes, the substrates preference was correspondingly exchanged. When the active motif of an alpha GRX (PtGRXA13) was mutated to that of a beta GRX, the mutants showed substrate preference similar to that of beta GRXs. These results indicate that divergence of active site motifs of *Populus* GRXs resulted in substrate switches.

GRXs can act as thioltransferase using GSH as the electron donor and disulfide substrate (RSSR) as electron acceptor (Fig. 9). HED has been used as a classical RSSR substrate, and L-cystine has also been used as a substrate in lots of studies (Holmgren and Aslund, 1995, Zaffagnini *et al.*, 2008, Manta *et al.*, 2019, Wang *et al.*, 2021). Both HED and L-cystine have the intramolecular disulfide bonds and GRX can catalyze thioldisulfide exchanges between GSH and disulfide substrates (HED and L-cystine). The newest mechanism showed HED is firstly reduced to a 2-mercaptoethanol (2-ME: EtOH-SH) by GRX; at the same time, a mixed disulfide between GRX (E) and the remaining part of HED (-S-EtOH) is formed as a reaction intermediate (E·S-S-EtOH). The disulfide then reacts with GSH to generate GS-SEtOH and an active free GRX enzyme (thiolate form). Sequentially, the glutathionylated GRX (E·S-S-G) and the second 2-ME are generated. At last, the glutathionylated GRX is reduced by another GSH molecule and produced GSSG (Begas *et al.*, 2015, Begas *et al.*, 2017).

GRX is also known to have GSH-dependent DHA reductase activity and is involved in ascorbate–glutathione pathway (Washburn and Wells, 1999). The proposed catalytic mechanisms of the DHA reductase activity of GRX starts with the formation of a thiohemiketal intermediate between GRX and DHA (Fig. 9). A molecule of GSH then displaces the ascorbic acid from GRX,

and a glutathionylated GRX is formed. At last, a second molecule of GSH attacks the glutathionylated GRX, forming GSSG and an active free GRX (Washburn and Wells, 1999). The ascorbate–glutathione pathway is recognized to play important roles in $H_2O_2$ metabolism, signaling, development and environmental responses (Foyer and Noctor, 2011). Therefore, this catalytic reaction is also a typical test for assessing the enzymatic function of GRX (Rouhier and Jacquot, 2003, Couturier *et al.*, 2009b, Couturier *et al.*, 2011, Kim *et al.*, 2020). In addition, GRX is reported to have GSH peroxidase activity, which can directly reduce hydroperoxides (ROOH) to corresponding alcohol (ROH) (Fig. 9). Cum–OOH is commonly used as a classical substrate to detect the peroxidase activities (Collinson *et al.*, 2002, Collinson and Grant, 2003).

GRXs are multifunctional oxidoreductases that can reduce a variety of substrates. GRXs catalyze reactions via dithiol or monothiol mechanisms, which both rely on the inherent affinity between GRXs and GSH moiety (Lillig *et al.*, 2008). The four compounds (HED, L-cystine, DHA, and Cum–OOH) selected in this study represent four kinds of substrates that can be chemically catalyzed by GRXs. The substrate preference switching indicates the role transition of GRXs among thioltransferase, DHA reductase, and peroxidase. Subsequently, it might indicate that different members of the GRX family regulate plant responses to oxidative stress differently. The switches of preferred substrates might be due to the difference in affinities ($1/K_m$) and catalytic efficiencies ($k_{cat}/K_m$) of the GRXs with different active site motifs for different substrates.

According to the catalytic mechanism of the GRX proteins, the first cysteine of the active site motif addressed in this study is the site where GRX forms disulfide bonds with GSH (Couturier *et al.*, 2009b, Couturier *et al.*, 2011, Begas *et al.*, 2017). It was the thiol of the first cysteine residue of the GRX active site motif that attacked the substrates. The first cysteine residue of the active site motif was on the surface of the protein, and the p*K*a of its thiol was significantly lower than that of the free cysteine (Dillet *et al.*, 1998, Lillig *et al.*, 2008). The value of this cysteine thiol p*K*a could affect both nucleophilicity and leaving group ability of this cysteine residue, thereby affecting the catalytic activity of GRXs (Foloppe *et al.*, 2001, Foloppe and Nilsson, 2004, Foloppe *et al.*, 2012). The first cysteine in the active site motif of GRX proteins with different active site motifs possessed different p*K*a values. The catalytic ability of GRXs was attributed to the difference of the p*K*a between its first cysteine thiol moiety in the active site motif and the product thiol (Jao *et al.*, 2006). Divergence of active site motifs of *Populus* GRXs might lead to change of the p*K*a of the first cysteine thiol in the active site motif, thereby resulting in substrate switches.

The catalytic mechanisms of GRX on the above four substrates indicates that the formation of glutathionated GRX played key roles in all the reactions. This means that the binding ability of GRX to GSH would affect the catalytic efficiency of GRX to substrates. The structure of the substrate–binding pocket could affect the binding of the enzyme to the substrates. As the active site motif of GRXs located at one arm of the L-shaped GSH–binding pocket and different amino acids have different structures and sizes, the change of active site motifs could directly influence the structure of the substrate–binding pocket (Fig. 8A, 8B). The change in the structure of the L-shaped pocket would directly lead to changes in enzymatic activity, and even changes in substrate specificity. These changes could also explain that the active site motif divergence of *Populus* GRXs

resulted in substrate switches.

## Experimental Procedures

### Identification of GRX Genes from *P. trichocarpa* Genome

*Populus trichocarpa* genome assembly version 3.0 (https://phytozome.jgi.doe.gov) was searched with 62 full-length GRX protein sequences of *O. sativa* and *A. thaliana* (Supplementary Table S2) using the TBLASTN program with default algorithm parameters. *P. trichocarpa* GRX candidates were then searched against the National Center for Biotechnology Information (NCBI) Conserved Domains Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and Pfam database (https://pfam.xfam.org) to confirm the presence of glutaredoxin domains in their protein structures. Candidates containing the glutaredoxin domain were identified as GRXs (Garg *et al.*, 2010). In this study, 41 GRX genes were identified from *P. trichocarpa* genome. The GRX genes were amplified from *P. trichocarpa* cDNA, cloned into the pEASY-T3 vector (TransGen, Beijing, China), and sequenced in both directions (Supplementary Table S1). The confirmed sequences of *P. trichocarpa* GRX genes were mapped to the genome of *P. trichocarpa* to verify their intron/exon structures. For genes that PCR did not detect (5 of 41 in this study), their intron/exon structures were gained from the annotation in *P. trichocarpa* genome assembly version 3.0.

### Identification of GRX Genes in Land Plants

The proteome databases of *Glycine max*, *Eucalyptus grandis*, *Solanum lycopersicum, Arabidopsis thaliana*, *Zea mays*, *Setaria italica*, *Brachypodium distachyon*, *Oryza sativa*, *Amborella trichopoda*, *Selaginella moellendorffii*, *Marchantia polymorpha*, and *Physcomitrella patens* in Phytozome (https://phytozome.jgi.doe.gov) were searched with the 41 full-length *P. trichocarpa* GRX protein sequences using the BLASTP program with default algorithm parameters. In addition, the proteome database of gymnosperm *Pinus taeda* in congenie (http://congenie.org/blast) were also searched using the same method. A total of 371 GRXs were identified from these thirteen land plant species (Supplementary Table S2). Each GRX protein contained at least one glutaredoxin domain.

### Phylogenetic Analysis

Full-length *P. trichocarpa* GRX protein sequences were aligned using the MUSCLE software (Edgar, 2004), and the alignment was adjusted manually using the BioEdit software (Hall, 1999) (Supplementary Data set 1). The Jones, Taylor, and Thornton (JTT) amino acid substitution model was selected as the optimal substitution model by the model generator version 0.85 (Keane *et al.*, 2006). A phylogenetic tree was constructed using the maximum-likelihood method with the PhyML software. One-hundred bootstrap replicates were performed to obtain the confidence support. Eight TRX proteins were used as outgroups for phylogenetic analysis. The transcript names of these eight TRX proteins in Phytozome were Potri.001G159000, Potri.001G416500, Potri.005G193400, Potri.006G123100, Potri.007G074000, Potri.013G132200, Potri.014G029200, and Potri.016G138800.

The phylogenetic tree of 412 GRX protein sequences from the 14 land plants was constructed

with the RaxML software (Stamatakis, 2006) (Supplementary Data set 2). The amino acid substitution model was JTT, and one-hundred bootstrap replicates were performed. Four *E. coli* TRXs (NCBI node: ASF79911.1, pdb|5HR0|, pdb|2TIR|A, pdb|5IKN|K) were used as outgroups.

**Expression of *Populus* GRX Genes Under Normal Conditions**

To investigate the expression patterns of *P. trichocarpa* GRX genes under normal conditions, the absolute expression values of each gene in the root, nodes, internodes, young leaves, and mature leaves of *P. trichocarpa* were retrieved from the PopGenIE database (http://popgenie.org/eximage). Each absolute value was divided by the minimum value in the expression data set, and then was log10 transformed for normalization. After normalizing the data, a heatmap of relative gene expression was created using the Heml software (version 1.0.3.7), with the expression levels represented by different colors.

Multiple Response Permutation Procedure (MRPP) is a nonparametric test that is applicable to data sets with unequal sample sizes and violations of the normality assumption. The test was performed with the Vegan package in R. The sampling frequency was 10,000 times.

**Subcellular Localizations of *Populus* GRX Proteins**

The signal sequence of each *Populus* GRX was predicted by SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/), Predisi (http://www.predisi.de/), Phobius (http://phobius.sbc.su.se/cgi-bin/predict.pl), TargetP 2.0 Server (http://www.cbs.dtu.dk/services/TargetP/), Signal-BLAST (http://sigpep.services.came.sbg.ac.at/signalblast.html) and ProP 1.0 Server (http://www.cbs.dtu.dk/services/ProP/) with default algorithm parameters. Predicted signal sequences are shown in italic and underlined in Supplementary Table S1. Full-length GRX genes were subcloned into a modified pCAMBIA1302 vector (Supplementary Fig. S5). This process added a C-terminal GFP tag to each GRX. The primers used to construct the vectors are listed in Supplementary Table S5. The reconstructed vectors were verified by sequencing and transformed into *Agrobacterium tumefaciens* EHA105. The subcellular localization of each GRX–GFP protein was preliminarily analyzed by infiltrating *Agrobacterium* cultures into epidermal cells of tobacco (*Nicotiana benthamiana*) leaves (Sparkes *et al.*, 2006). To further verify the subcellular localizations of GRXs, the co-expression assays of GRX–GFP proteins and organelle marker proteins were performed. Three vectors expressing specific organelle markers, respectively were selected: CD3-959 (expressing the endoplasmic reticulum marker), CD3-991 (expressing the mitochondria marker) and p1302-H2A-mcherry (expressing the nuclear marker). *Agrobacterium* strain expressing fusion GRX–GFP protein and the *Agrobacterium* strain expressing the specific organelle marker protein were mixed to a final $OD_{600}$ of 0.6 (Sparkes *et al.*, 2006, Yao *et al.*, 2015). The mixed culture was then infiltrated into epidermal cells of tobacco leaves. The leaf tissues were examined under a confocal laser microscope (Zeiss LSM880, Carl Zeiss GmbH, Jena, Germany) every 24 h from days 2–4 after cell infiltration. GFP fusion fluorescence and mCherry fluorescence were excited with a laser at 488 nm and 584nm, respectively. Chlorophyll autofluorescence was

imaged using a 543-nm laser.

**Expression and Purification of *Populus* GRX Proteins**

The signal sequence of each *Populus* GRX was predicted by the six databases mentioned in the subsection "Subcellular Localizations of *Populus* GRX Proteins" of Experimental Procedures. However, there were some proteins whose signal sequences could not be predicted using the aforementioned websites. GRX domains of these proteins were predicted by Conserved Domains Database in NCBI. This study considered the N-terminal sequence 12 amino acids away from the GRX domain as the signal sequence of these proteins.

After deleting the signal sequences, the *Populus* GRX genes were subcloned into the modified ΔpET-30a expression system (Yang *et al.*, 2009) to obtain an N-terminal 6×His tag. The primers used to construct the expression vectors are listed in Supplementary Table S6. Colonies with the appropriate inserts were identified by sequencing. Recombinant plasmids were inserted into *E. coli* BL21 (DE3). *Escherichia coli* BL21 (DE3) cells with the correct recombinant plasmids were cultured in Luria–Bertani liquid medium at 37°C overnight. The cells were then diluted 1:100 with fresh Luria–Bertani liquid medium and cultured until an optical density ($OD_{600}$) of 0.6 was reached. Isopropyl-b-D-thiogalactopyranoside was added to the cultures at a final concentration of 0.1 mM to induce the synthesis of recombinant fusion proteins. After induction (12 h at 20°C), cells were centrifuged at 8,000 × g at 4°C for 5 min and harvested. The harvested cells were resuspended in binding buffer (20 mM Tris–HCl, pH 7.9, 500 mM NaCl, 20 mM imizole). After cold sonication and centrifugation, the supernatant was transferred to a Nickel–Sepharose High Performance column (GE Healthcare Bio–Ssciences) and the GRXs bond to the column were eluted by the elution buffer (20 mM Tris–HCl, pH 7.9, 500 mM NaCl, 500 mM imizole).

Regarding *Populus* alpha GRXs with 6×His tag, some were not expressed in *E. coli*, whereas others were expressed as inclusion bodies or were unstable in the enzyme assay buffer. We then tested the expression of GST–tagged alpha GRX proteins in *E. coli*. We selected eight alpha GRXs (PtGRXA1, 2, 5, 11, 12, 16, 20 and 21) to construct GST-tagged protein expression vectors. The eight selected alpha genes were subcloned into the pGEX-4T-1 expression system to obtain an N-terminal GST tag. Expression of the recombinant proteins was performed as described above. The harvested cells were lysised in 50 mM Tris–HCl, pH 7.9, 150 mM NaCl by cold sonication. The purification was performed with the Glutathione Sepharose$^{TM}$ 4 Fast Flow column (GE Healthcare Bio–Sciences) and the GRXs bound to the column were eluted by the elution buffer (50 mM Tris–HCl, pH 7.9, 150 mM NaCl,15mM GSH). Seven of these eight GST–tagged alpha GRXs were expressed as inclusion bodies or even unexpressed, and only PtGRXA12 was soluble. At last, we used a modified pMAL-c5X expression system with an N-terminal MBP tag to express *Populus* alpha GRXs (Supplementary Fig. S6). The primers used to construct expression vectors of the MBP fused alpha GRXs are listed in Supplementary Table S6. The expression methods of MBP-tag fused protein and His–tag fused protein were the same; the purification methods of MBP-tag fused protein and His–tag fused protein were different. The purification method of MBP-tag fused proteins was

following: The harvested cells containing the MBP-tag fused protein were resuspended in binding buffer (20 mM Tris–HCl, pH 7.9, 200 mM NaCl, 1 mM EDTA) and disrupted by cold sonication. The homogenate was centrifuged again at 10,000 × g at 4°C for 10 min. The supernatant was transferred to an Amylose Resin column (New England Biolabs, Ipswich, MA, USA) that had been pre-equilibrated with binding buffer. Fusion proteins bound to the column were eluted with the elution buffer (20 mM Tris–HCl, pH 7.9，200 mM NaCl, 1 mM EDTA, 10 mM maltose).

The total cell lysate of IPTG-induced recombinant *E. coli* BL21, the supernatant, the resultant particulate material, and the purified recombinant GRX proteins from affinity chromatography column were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE; Supplementary Fig. S3).

**Isolation of Monomeric GRXs**

Some of the *Populus* GRXs purified from the *E.coli* system could form not only monomers, but also homodimer or/and tetramer by ligating iron–sulfur clusters. Thus, each GRX protein purified from the affinity chromatography column was then subjected to size-exclusion chromatography by a Superdex 200pg 16/600 GL column (GE Healthcare, Sweden, code: 28989335) or a Superdex 75pg 16/600 GL column (GE Healthcare, Sweden, code: 28989333) to obtain the monomeric GRX. The size-exclusion chromatography column was equipped in an ÄKTA FPLC system (GE Healthcare Europe GmbH Succursale France, Orsay, France). Based on the calibration curve of molecular weight for Superdex 200 pg column or 75pg column, the peak position of the monomeric GRX proteins were determined in the chromatography profiles. The separated monomeric GRX proteins were eluted with 20 mM Tris–HCl, pH 7.9, 200 mM NaCl. The obtained monomeric GRX were analyzed by SDS–PAGE (Supplementary Fig. S3).

**Site-directed Mutagenesis**

The PtGRXB4 and PtGRXA13 protein were used for mutagenesis analysis. Site-directed mutagenesis was performed using methods described in our previous study (Zeng and Wang, 2005). All mutant primers are listed in Supplementary Table S7. The mutantant *GRXs* of *PtGRXB4*were subcloned into the ΔpET-30a expression system. The expression, purification, and isolation of the monomeric form of the proteins were performed as those of PtGRXB4. The samples of each step of the expression and purification were analyzed by SDS–PAGE (Supplementary Fig. S3). The mutantant *GRXs* of *PtGRXA13* were subcloned into the modified pMAL-c5X expression system. The expression, purification, and isolation of the monomeric form of the proteins were performed as those of PtGRXA13.

**Enzyme Assays and Kinetic Constants of *Populus* GRX Enzyme**

The enzymatic activities of wild-type and mutant GRXs were assessed using HED, DHA, L-cystine, and Cum–OOH as substrates. The activities toward HED and DHA were detected as described by Zaffagnini et al. (2008). Activities toward L-cystine were detected as described by Ahn et al (1992). Activities toward Cum–OOH were detected using a modified method based on that

described by Collinson et al (2002). A mixture of 1 mM GSH, 0.2 mM NADPH, and 6ug/mL yeast glutathione reductase was prepared in 50 mM $K_3PO_4$, pH 7.0. Equal concentration and volume of GRX and BSA proteins (dissolved in 20 mM Tris-HCl, pH 7.9, 200 mM NaCl) were then added to the sample cuvette and reference cuvette, respectively. At last, 1.38 mM Cum–OOH was added to the mixture to a final volume of 3 mL. The decrease of the mixture in absorbance at 340 nm in 3 minutes was followed using Evolution 300 (Thermo Scientific, USA). Each reaction was repeated at least 25 times to measure the activities of GRX proteins toward Cum–OOH. Protein concentration was determined by measuring its absorption at 280 nm. All the assays were performed at 25 °C.

The kinetic constants of PtGRXB4 and its mutants were determined under various concentrations of GSH and HED, DHA, or L-cystine. The affinity for GSH ($1/K_m$) was measured with a GSH concentration range from 0.5 to 8.0 mM at a fixed HED concentration of 0.7mM. The affinities for HED, DHA, or L-cystine were determined with a fixed GSH concentration of 1.0 mM and different concentrations of HED (0.175 to 2.8mM), DHA (0.25 to 4mM), and L-cystine (0.078 to 2.5mM), respectively. The kinetic parameters were analyzed using the nonlinear regression using the HYPER32 software (http://hyper32.software. informer.com/). The Mann–Whitney U test was one of the most widely used nonparametric tests to test the difference in the sum of ranks between two independent samples that violate the normality assumption. It was performed with default algorithm parameters of the SPSS software (https://www.ibm.com/analytics/spss-statistics-software).

**Homology Modeling**

Owing to the identical protein sequences of GRXS12 from *Populus tremula* × *tremuloides* and PtGRXB4, the crystal structure of the GRX domain in GRXS12 proteins (PDB: 3FZ9) was used as a template for modeling the structures of the mutant proteins of PtGRXB4. Optimal structures were tomatically modeled by SWISS–MODEL with default parameters (https://swissmodel.expasy.org/interactive). The comparison, analysis, and display of modeled structures were processed with the PyMol software (DeLano Scientifific LLC, San Carlos, CA, USA).

**Molecular Docking**

According to the catalytic mechanism of GRXs to HED, L-cystine, DHA, and Cum-OOH, the potential binding site of the four substrates on GRXs was the first cysteine of the active site motif. The molecular docking was performed using Autodock 4.2.6 software package (Sanner, 1999). The crystal structure of PtGRXB4 protein (PDB:3FZ9) was selected as the receptor for molecular docking. The structures of HED, L-cystine, DHA, and Cum–OOH were selected as ligands. The molecular structures of HED, L-cystine, DHA, and Cum–OOH were obtained from Pubchem (https://pubchem.ncbi.nlm.nih.gov/) and were optimized by MOPAC program (Stewart, 1990). The structures of the receptor and ligands were then processed using Autodock Tools 1.5.6 to obtain a pdbqt file (Morris *et al.*, 2009). The coordinates of the center point of the boxes in which the protein

and each ligand were docked were -4.748, 3.478, 2.129. The number of grid points in each direction of X×Y×Z was set to four specifications: 50×50×50, 40×40×40, 30×30×30 and 24 ×24×24. Each of the ligands were docked into the receptor with four specifications of the docking box, respectively. The grid spacing is 0.375 Å, and the rest of the parameters were default values. The number of molecular docking was set to 100 in each program. Finally, based on the estimated free energy of binding and the positions of binding, the optimal docking results are shown in Supplementary Fig. S4.

**Data Availability Statement**

Sequencing data of *P. trichocarpa GRX*s can be found in GenBank databases under the accession numbers listed in Supplementary Table S1. Sequencing data of other 13 land plant *GRX*s can be found in Phytozome or congenie databases under the accession numbers listed in Supplementary Table S2. All the other relevant data can be found within the paper and its supporting information.

**Acknowledgement**

This study was supported by the National Science Foundation of China (31822011, 31425006) and the Chinese Academy of Forestry (CAFYBB2018ZX001). Zhen Li is funded by a postdoctoral fellowship from the Special Research Fund of Ghent University (BOFPDO2018001701).

**Author contributions**

Q.-Y.Z. and Y.-J.L. designed research; H.X. P.-F.J. L.Z. and C.Q. performed the experiments; H.X. and Y.-J.L. analyzed data; Q.-Y.Z., Y.-J.L. and H.X. wrote the article. Q.-Y.Z. and Y.-J.L. agree to serve as the author responsible for contact and ensure communication. All authors read, revised and approved the final manuscript.

**Conflict of interests**

All the authors declare that there are no competing interests.

**Supporting Information**

**Supplementary Figure S1**. Genomic localization of full-length *Populus* GRX genes.

**Supplementary Figure S2.** Subcellular localizations of *Populus* GRX proteins.

**Supplementary Figure S3.** SDS-PAGEs of *Populus* GRXs and mutants of PtGRXB4.

**Supplementary Figure S4.** Predicted structure of PtGRXB4 in complex with its substrate by molecular docking.

**Supplementary Figure S5**. Modified pCAMBIA1302 vector.

**Supplementary Figure S6**. Modified pMAL-c5x vector.

**Supplementary Table S1**. Full-length GRX genes identified from *Populus trichocarpa* genome.

**Supplementary Table S2**. Full-length GRXs identified from 13 land plant species.

**Supplementary Table S3.** Enzymatic activities of PtGRXB4 with different tags.

**Supplementary Table S4.** Enzymatic activities of PtGRXB4 and its mutants showed in Figure 7A.

**Supplementary Table S5.** Primers used to construct the vectors for analysis of the subcellular localization of *Populus* GRXs.

**Supplementary Table S6.** Primers used to construct protein expression vectors for *Populus* GRXs.

**Supplementary Table S7.** Primers used to amplify *Populus* mutated GRX genes and construct mutated protein expression vectors.

**Supplementary Data Set 1.** Sequence alignment used to construct the phylogenetic tree in Figure 1.

**Supplementary Data Set 2.** Sequence alignment used to construct the phylogenetic tree in Figure 2A.



**References**

**Abdalla, M., Eltayb, W.A., El-Arabey, A.A., Mo, R., Dafaalla, T.I.M., Hamouda, H.I., Bhat, E.A., Awadasseid, A. and Ali, H.A.A.** (2018) Structure analysis of yeast glutaredoxin Grx6 protein produced in *Escherichia coli*. *Genes and Environ*, **40**, 15.

**Ahn, B.Y. and Moss, B.** (1992) Glutaredoxin homolog encoded by vaccinia virus is a virion-associated enzyme with thioltransferase and dehydroascorbate reductase activities. *Proc. Natl Acad. Sci. USA* **89**, 7060-7064.

**Bandyopadhyay, S., Gama, F., Molina-Navarro, M.M., Gualberto, J.M., Claxton, R., Naik, S.G., Huynh, B.H., Herrero, E., Jacquot, J.P., Johnson, M.K. and Rouhier, N.** (2008) Chloroplast monothiol glutaredoxins as scaffold proteins for the assembly and delivery of [2Fe–2S] clusters. *EMBO J* **27**, 1122–1133.

**Begas, P., Liedgens, L., Moseler, A., Meyer, A.J. and Deponte, M.** (2017) Glutaredoxin catalysis requires two distinct glutathione interaction sites. *Nat Commun*, **8**, 14835.

**Begas, P., Staudacher, V. and Deponte, M.** (2015) Systematic re-evaluation of the bis(2-hydroxyethyl) disulfide (HEDS) assay reveals an alternative mechanism and activity of glutaredoxins. *Chem Sci*, **6**, 3788-3796.

**Blanc, G. and Wolfe, K.H.** (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell*, **16**, 1679-1691.

**Byun-McKay, S.A. and Geeta, R.** (2007) Protein subcellular relocalization: a new perspective on the origin of novel genes. *Trends Ecol Evol*, **22**, 338-344.

**Casneuf, T., Bodt, S.D., Raes, J., Maere, S. and Peer, Y.V.d.** (2006) Nonrandom divergence of gene expression following gene and genome duplications in the flowering plant *Arabidopsis thaliana*. *Genome Bio*, **7**, R13.

**Cheng, N.H., Liu, J.Z., Brock, A., Nelson, R.S. and Hirschi, K.D.** (2006) AtGRXcp, an *Arabidopsis* chloroplastic glutaredoxin, is critical for protection against protein oxidative

damage. *J Biol Chem.*, **281**, 26280-262888.

**Collinson, E.J. and Grant, C.M.** (2003) Role of yeast glutaredoxins as glutathione S-transferases. *J Biol Chem*, **278**, 22492-22497.

**Collinson, E.J., Wheeler, G.L., Garrido, E.O., Avery, A.M., Avery, S.V. and Grant, C.M.** (2002) The yeast glutaredoxins are active as glutathione peroxidases. *J Biol Chem*, **277**, 16712-16717.

**Couturier, J., Jacquot, J.P. and Rouhier, N.** (2009a) Evolution and diversity of glutaredoxins in photosynthetic organisms. *Cell Mol Life Sci*, **66**, 2539–2557

**Couturier, J., Koh, C.S., Zaffagnini, M., Winger, A.M., Gualberto, J.M., Corbier, C., Decottignies, P., Jacquot, J.P., Lemaire, S.D., Didierjean, C. and Rouhier, N.** (2009b) Structure-function relationship of the chloroplastic glutaredoxin S12 with an atypical WCSYS active site. *J Biol Chem*, **284**, 9299-9310.

**Couturier, J., Ströher, E., Albetel, A.N., Roret, T., Muthuramalingam, M., Tarrago, L., Seidel, T., Tsan, P., Jacquot, J.P., Johnson, M.K., Dietz, K.J., Didierjean, C. and Rouhier, N.** (2011) Arabidopsis chloroplastic glutaredoxin C5 as a model to explore molecular determinants for iron-sulfur cluster binding into glutaredoxins. *J Biol Chem*, **286**, 27515-27527.

**Dillet, V., Dyson, H.J. and Bashford, D.** (1998) Calculations of electrostatic interactions and p$K_a$s in the active site of *Escherichia coli* thioredoxin. *Biochemistry* **37**, 10298-10306.

**Dixon, D.P., Hawkins, T., Hussey, P.J. and Edwards, R.** (2009) Enzyme activities and subcellular localization of members of the *Arabidopsis* glutathione transferase superfamily. *J Exp Bot*, **60**, 1207-1218.

**Dixon, D.P., Lapthorn, A. and Edwards, R.** (2002) Plant glutathione transferases. *Genome Biol*, **3**, 1-10.

**Edgar, R.C.** (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792-1797.

**Fernandes, A.P. and Holmgren, A.** (2004) Glutaredoxins: glutathione-dependent redox enzymes with functions far beyond a simple thioredoxin backup system. *Antioxid Redox Signal.*, **6**, 63-74.

**Foloppe, N. and Nilsson, L.** (2004) The glutaredoxin -C-P-Y-C- motif: influence of peripheral residues. *Structure*, **12**, 289-300.

**Foloppe, N., Sagemark, J., Nordstrand1, K., Berndt, K.D. and Nilsson, L.** (2001) Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from *Escherichia coli*: comparison with functionally related proteins. *J Mol Biol* **310**, 449-470.

**Foloppe, N., Vlamis-Gardikas, A. and Nilsson, L.** (2012) The -Cys-X1-X2-Cys- motif of reduced glutaredoxins adopts a consensus structure that explains the low p$K_a$ of its catalytic cysteine. *Biochemistry*, **51**, 8189-8207.

**Foyer, C.H. and Noctor, G.** (2011) Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol*, **155**, 2-18.

**Ganko, E.W., Meyers, B.C. and Vision, T.J.** (2007) Divergence in expression between duplicated genes in *Arabidopsis*. *Mol Biol Evol*, **24**, 2298-2309.

**Garg, R., Jhanwar, S., Tyagi, A.K. and Jain, M.** (2010) Genome-wide survey and expression analysis suggest diverse roles of glutaredoxin gene family members during development and response to various stimuli in rice. *DNA Res*, **17**, 353-367.

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

**Holmgren, A. and Aslund, F.** (1995) Glutaredoxin. *Methods Enzymol*, **252**, 283-292.

**Jao, S.C., Ospina, S.M.E., Berdis, A.J., Starke, D.W., Post, C.B. and Mieyal, J.J.** (2006) Computational and mutational analysis of human glutaredoxin (thioltransferase). *Biochemistry* **45**, 4785-4796.

**Keane, T.M., Creevey, C.J., Pentony, M.M., Naughton, T.J. and McLnerney, J.O.** (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol Biol*, **6**, 29.

**Kim, G., Omeka, W.K.M., Liyanage, D.S. and Lee, J.** (2020) Molecular characterization, redox regulation, and immune responses of monothiol and dithiol glutaredoxins from disk abalone (*Haliotis discus discus*). *Fish Shellfish Immunol*, **107**, 385-394.

**Lan, T., Gao, J. and Zeng, Q.-Y.** (2013) Genome-wide analysis of the LEA (late embryogenesis abundant) protein gene family in *Populus trichocarpa*. *Tree Genet Genomes*, **9**, 253-264.

**Lan, T., Yang, Z.L., Yang, X., Liu, Y.J., Wang, X.R. and Zeng, Q.Y.** (2009) Extensive functional diversification of the *Populus* glutathione S-transferase supergene family. *Plant Cell*, **21**, 3749-3766.

**Laporte, D., Olate, E., Salinas, P., Salazar, M., Jordana, X. and Holuigue, L.** (2012) Glutaredoxin GRXS13 plays a key role in protection against photooxidative stress in *Arabidopsis*. *J Exp Bot*, **63**, 503-515.

**Leissring, M.A., Farris, W., Wu, X., Christodoulou, D.C., Haigis, M.C., Guarente, L. and Selkoe, D.J.** (2004) Alternative translation initiation generates a novel isoform of insulin-degrading enzyme targeted to mitochondria. *Biochem J* **383**, 439-446.

**Li, Y., Darley, C.P., Ongaro, V.n., Fleming, A., Schipper, O., Baldauf, S.L. and McQueen-Mason, S.J.** (2002) Plant expansins are a complex multigene family with an ancient evolutionary origin. *Plant Physiol*, **128**, 854-864.

**Lillig, C.H., Berndt, C. and Holmgren, A.** (2008) Glutaredoxin systems. *Biochim Biophys Acta*, **1780**, 1304-1317.

**Lillig, C.H., Berndt, C., Vergnolle, O., Lonn, M.E., Hudemann, C., Bill, E. and Holmgren, A.** (2005) Characterization of human glutaredoxin 2 as iron–sulfur protein: A possible role as redox sensor. *Proc Natl Acad Sci USA,* **102**, 8168–8173.

**Liu, Y.J., Han, X.M., Ren, L.L., Yang, H.L. and Zeng, Q.Y.** (2013) Functional divergence of the glutathione S-transferase supergene family in *Physcomitrella patens* reveals complex patterns of large gene family evolution in land plants. *Plant physiol*, **161**, 773-786.

**Manta, B., Moller, M.N., Bonilla, M., Deambrosi, M., Grunberg, K., Bellanda, M., Comini, M.A. and Ferrer-Sueta, G.** (2019) Kinetic studies reveal a key role of a redox-active glutaredoxin in the evolution of the thiol-redox metabolism of trypanosomatid parasites. *J Biol Chem*, **294**, 3235-3248.

**Morita, S., Yamashita, Y., Fujiki, M., Todaka, R., Nishikawa, Y., Hosoki, A., Yabe, C., Nakamura, J., Kawamura, K., Suwastika, I.N., Sato, M.H., Masumura, T., Ogihara, Y., Tanaka, K. and Satoh, S.** (2015) Expression of a rice glutaredoxin in aleurone layers of developing and mature seeds: subcellular localization and possible functions in antioxidant defense. *Planta*, **242**, 1195-1206.

**Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. and Olson, A.J.** (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, **30**, 2785-2791.

**Ndamukong, I., Abdallat, A.A., Thurow, C., Fode, B., Zander, M., Weigel, R. and Gatz, C.** (2007) SA-inducible Arabidopsis glutaredoxin interacts with TGA factors and suppresses

JA-responsive *PDF1.2* transcription. *Plant J*, **50**, 128-139.

**Patterson, K., Walters, L.A., Cooper, A.M., Olvera, J.G., Rosas, M.A., Rasmusson, A.G. and Escobar, M.A.** (2016) Nitrate-regulated glutaredoxins control *Arabidopsis* primary root growth. *Plant Physiol*, **170**, 989-999.

**Ren, L.L., Liu, Y.J., Liu, H.J., Qian, T.T., Qi, L.W., Wang, X.R. and Zeng, Q.Y.** (2014) Subcellular relocalization and positive selection play key roles in the retention of duplicate genes of *Populus* class III peroxidase family. *Plant Cell*, **26**, 2404-2419.

**Renaudin, J.P., Doonan, J.H., Freeman, D., Hashimoto, J., Hirt, H., Inze, D., Jacobs, T., Kouchi, H., Rouze, P. and Sauter, M.** (1996) Plant cyclins: a unified nomenclature for plant A-, B- and D-type cyclins based on sequence organization. *Plant Mol Biol*, **32**, 1003-1018.

**Riondet, C., Desouris, J.P., Montoya, J.G., Chartier, Y., Meyer, Y. and Reichheld, J.-P.** (2012) A dicotyledon-specific glutaredoxin GRXC1 family with dimer-dependent redox regulation is functionally redundant with GRXC2. *Plant Cell Environ*, **35**, 360-373.

**Rouhier, N., Couturier, J., Johnson, M.K. and Jacquot, J.P.** (2010) Glutaredoxins: roles in iron homeostasis. *Trends Biochem Sci*, **35**, 43-52.

**Rouhier, N. and Jacquot, J.P.** (2003) Molecular and catalytic properties of a peroxiredoxin-glutaredoxin hybrid from *Neisseria meningitidis*. *FEBS Lett*, **554**, 149-153.

**Rouhier, N., Unno, H., Bandyopadhyay, S., Masip, L., Kim, S.K., Hirasawa, M., Gualberto, J.M., Lattard, V., Kusunok, M., Knaff, D.B., Georgiou, G., Hase, T., Johnson, M.K. and Jacquot, J.P.** (2007) Functional, structural, and spectroscopic characterization of a glutathione-ligated [2Fe–2S] cluster in poplar glutaredoxin C1. *Proc Natl Acad Sci USA,* **104**, 7379-7384.

**Sanner, M.F.** (1999) Python: a programming language for software integration and development. *J Mol Graph Model*, **17**, 57-61.

**Sparkes, I.A., Runions, J., Kearns, A. and Hawes, C.** (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat Protoc*, **1**, 2019-2025.

**Stamatakis, A.** (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688-2690.

**Stewart, J.J.** (1990) MOPAC: a semiempirical molecular orbital program. *J Computer-Aided Mol Des*, **4**, 1-103.

**Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I. and Hellsten, U.** (2006) The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, **313**, 1596-1604.

**Wang, Y., Hou, Y. and Wang, Q.** (2021) Cloning, expression, characterization, and antioxidant protection of glutaredoxin3 from Psychrophilic Bacterium *Psychrobacter* sp. ANT206. *Front Microbio*, **12**, 633362.

**Wang, Z., Xing, S., Birkenbihl, R.P. and Zachgo, S.** (2009) Conserved functions of *Arabidopsis* and rice CC-type glutaredoxins in flower development and pathogen response. *Mol Plant*, **2**, 323-335.

**Washburn, M.P. and Wells, W.W.** (1999) The catalytic mechanism of the glutathione-dependent dehydroascorbate reductase activity of thioltransferase (glutaredoxin). *Biochemistry*, **38**, 268-274.

**Xing, S. and Zachgo, S.** (2008) *ROXY1* and *ROXY2*, two Arabidopsis glutaredoxin genes, are required for anther development. *Plant J*, **53**, 790-801.

**Yang, X., Kalluri, U.C., Jawdy, S., Gunter, L.E., Yin, T., Tschaplinski, T.J., Weston, D.J., Ranjan, P. and Tuskan, G.A.** (2008) The F-box gene family is expanded in herbaceous annual plants relative to woody perennial plants. *Plant Physiol*, **148**, 1189-1200.

**Yang, X., Sun, W., Liu, J.P., Liu, Y.-J. and Zeng, Q.Y.** (2009) Biochemical and physiological characterization of a tau class glutathione transferase from rice (*Oryza sativa*). *Plant Physiol and Biochem*, **47**, 1061-1068.

**Yang, Z.L., Liu, H.J., Wang, X.R. and Zeng, Q.Y.** (2013) Molecular evolution and expression divergence of the *Populus* polygalacturonase supergene family shed light on the evolution of increasingly complex organs in plants. *New Phytol*, **197**, 1353-1365.

**Yao, X., Li, J., Liu, J. and Liu, K.** (2015) An *Arabidopsis* mitochondria-localized RRL protein mediates abscisic acid signal transduction through mitochondrial retrograde regulation involving ABI4. *J Exp Bot*, **66**, 6431-6445.

**Zaffagnini, M., Michelet, L., Massot, V., Trost, P. and Lemaire, S.D.** (2008) Biochemical characterization of glutaredoxins from *Chlamydomonas reinhardtii* reveals the unique properties of a chloroplastic CGFS-type glutaredoxin. *J Biol Chem*, **283**, 8868-8876.

**Zeng, Q.Y. and Wang, X.R.** (2005) Catalytic properties of glutathione-binding residues in a tau class glutathione transferase (PtGSTU1) from *Pinus tabulaeformis*. *FEBS Lett*, **579**, 2657-2662.

**Table 1. Kinetic constants of PtGRXB4 and its mutants for HED, L-cystine, and DHA conjugation reactions.** Means ± SD obtained from at least three independent replicates.

| GRX proteins | GSH | | | HED | | | L-cystine | | | DHA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $1/K_m$ ($mM^{-1}$) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}\ s^{-1}$) | $1/K_m$ ($mM^{-1}$) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}\ s^{-1}$) | $1/K_m$ ($mM^{-1}$) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}\ s^{-1}$) | $1/K_m$ ($mM^{-1}$) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}\ s^{-1}$) |
| Wild type | 0.32±0.04 | 23.90 | 7.46 | 4.55±0.27 | 7.58 | 34.30 | 14.54±0.81 | 4.49 | 65.07 | 2.41±0.37 | 7.68 | 18.23 |
| $^{28}$CCMC$^{31}$ mutantt | 0.20±0.01 | 12.91 | 2.56 | 3.77±0.24 | 2.04 | 7.67 | 6.29±0.39 | 1.60 | 10.06 | 0.19±0.02 | 0.72 | 0.14 |
| $^{28}$CCMS$^{31}$ mutant | 0.26±0.02 | 8.31 | 2.17 | 2.59±0.19 | 2.17 | 5.61 | 5.82±0.34 | 8.91 | 51.80 | 0.22±0.02 | 1.19 | 0.26 |
| $^{28}$CPYC$^{31}$ mutant | 0.87±0.09 | 17.74 | 15.36 | 7.94±0.25 | 8.88 | 70.48 | 28.13±3.96 | 4.14 | 115.00 | 7.27±0.47 | 5.76 | 41.77 |
| $^{28}$CGYC$^{31}$ mutant | 0.29±0.01 | 17.90 | 5.17 | 3.42±0.24 | 5.99 | 20.44 | 11.27±0.44 | 1.70 | 19.10 | 1.84±0.13 | 5.56 | 10.20 |
| $^{28}$CPFC$^{31}$ mutant | 0.42±0.02 | 24.52 | 10.17 | 9.50±0.17 | 9.44 | 89.90 | 10.38±0.69 | 3.20 | 32.99 | 4.01±0.43 | 7.14 | 28.44 |
| $^{28}$CGFS$^{31}$ mutant | 0.41±0.04 | 32.86 | 14.12 | 4.70±0.31 | 10.83 | 50.85 | 24.23±1.82 | 9.63 | 234.88 | 0.18±0.01 | 2.45 | 0.45 |
| $^{28}$CQDC$^{31}$ mutant | 0.14±0.01 | 2.03 | 0.28 | 3.25±0.23 | 0.39 | 1.26 | 5.96±0.41 | 0.38 | 2.26 | 0.40±0.05 | 0.15 | 0.06 |
| $^{28}$CRDC$^{31}$ mutant | 0.24±0.01 | 2.09 | 0.50 | 3.11±0.07 | 0.65 | 2.02 | 9.27±0.90 | 0.98 | 8.99 | 0.16±0.01 | 0.31 | 0.05 |

**Table 2. Enzymatic activities of PtGRXA13 and its mutants.** Means ± SD obtained from at least three independent replicates. n.d., no activity detected.

| | Specific activity (μmol/s per μmol) | | | |
|---|---|---|---|---|
| **GRX proteins** | **HED** | **L-cystine** | **DHA** | **Cum–OOH** |
| Wild type | n.d. | n.d. | n.d. | 0.188±0.009 |
| $^{21}$CSYS$^{24}$ mutant | n.d. | 0.020±0.002 | n.d. | n.d. |
| $^{21}$CGFS$^{24}$ mutant | n.d. | n.d. | n.d. | n.d. |
| $^{21}$CRDC$^{24}$ mutant | n.d. | n.d. | n.d. | n.d. |

**Figure Legends**

**Figure 1. Phylogenetic tree of *Populus* GRXs.** Numbers at each node represent the bootstrap values. Alpha (α), beta (β), gamma (γ), and delta (δ) GRX classes are shaded in gray, blue, green, and purple, respectively. Sequence alignments used to reconstruct the tree is available as Supplementary Data Set S1.

**Figure 2. Phylogenetic tree (A) and copy numbers (B) of GRXs of 14 land plant species.** In (A), numbers at the node in the phylogenetic tree represent bootstrap values. The alpha (α), beta (β), gamma (γ) and delta (δ) GRXs are indicated by black, blue, green, and purple lines, respectively. In (B), the copy number of the GRX genes in each species is shown on the box side. The sequence alignments used to reconstruct the tree in (A) is available as Supplementary Data Set S2.

**Figure 3. Phylogenetic relationships (A), expression patterns (B), gene structures (C), and conserved domains (D) of *Populus* GRXs.** In (A), α, β, γ and δ class GRXs are represented by black, blue, green, and purple letters, respectively. Numbers at each node in the phylogenetic tree represent bootstrap values. The pseudogene is indicated by asterisks. In (B), the bar indicates the reference expression level. Gray boxes denote genes that do not have expression information in the PopGenIE database. IN, internodes; ND, nodes; RT, roots; ML, mature leaves; YL, young leaves. In (C), exons and introns are indicated by boxes and lines, respectively. In (D), gray lines indicate the full-length GRX protein sequence. In (C) and (D), GRX, TRX, DEP, and DUF547 domains are highlighted in green, red, purple, and blue, respectively.

**Figure 4. Genomic localization (A), phylogenetic relationships (B), and putative evolutionary histories (C) of *Populus* GRX genes in clusters I and II.** In (A), regions that correspond to homologous genome blocks are shaded in gray and connected with lines. Paralogous clusters I and II are indicated by red dashed lines within gray-shaded trapezoids. In (B) and (C), phylogenetic tree is reconstructed using the JTT model. Numbers on branches indicate bootstrap values. Letters in the black and colored boxes represent ancestral GRX genes. T, putative tandem duplication; W, putative whole-genome duplication.

**Figure 5. Subcellular localizations of *Populus* GRX proteins.** GFP signal (green) of GRX proteins, mCherry signal (red) of nuclear marker (NC marker), endoplasmic reticulum marker (ER marker), and mitochondrial marker (MT marker), and chlorophyll autofluorescence (red) were detected using confocal laser-scanning microscopy. An overlay is shown in yellow. Nucleus, endoplasmic reticulum, chloroplasts, and mitochondria are indicated by the purple, white, yellow, and blue arrows, respectively. Bars = 20 μm.

**Figure 6. Phylogenetic tree and enzymatic activities (A), active site motifs (B), and subcellular localizations (C) of *Populus* GRX proteins.** In (A), α, β, γ, and δ class GRXs are represented by black, blue, green, and purple letters, respectively. Numbers at each node in the phylogenetic tree

represent bootstrap values. The pseudogene is indicated by asterisk. Suffixes attached to each GRX label: C, successfully cloned; A, purified GRX assayed; U, recombinant protein were not expressed; -, analysis not performed. Values shown are mean ± SD, as calculated from at least three replicates. n.d., no activity detected. In (B), the active site motif of each GRX protein is shown. In (C), CY, cytosol; NC, nucleus; ER, endoplasmic reticulum; MT, mitochondria; CP, chloroplast. Dash indicates that analysis is not performed.

**Figure 7. Changes in enzymatic activities induced by mutations at the active site motif of PtGRXB4 protein.** The enzymatic activities are means as calculated from at least three replicates. The enzymatic activity of PtGRXB4 toward each substrate is set as baseline for comparison with the mutants specified on the x axis.

**Figure 8. GRX domain structures (A) and side chains of the amino acids in the active site motifs (B) of PtGRXB4 and its mutants**. In (A), glutathione (GSH) is shown as green balls. The GSH–binding pocket is indicated by a green L-shaped box. In (B), amino acids in the active site motif of both PtGRXB4 and its mutants are shown as sticks; GSH is shown as green sticks.

**Figure 9. Mechanistic models of reactions catalyzed by GRXs.** Alpha and beta GRXs are shown in black and blue ovals, respectively.

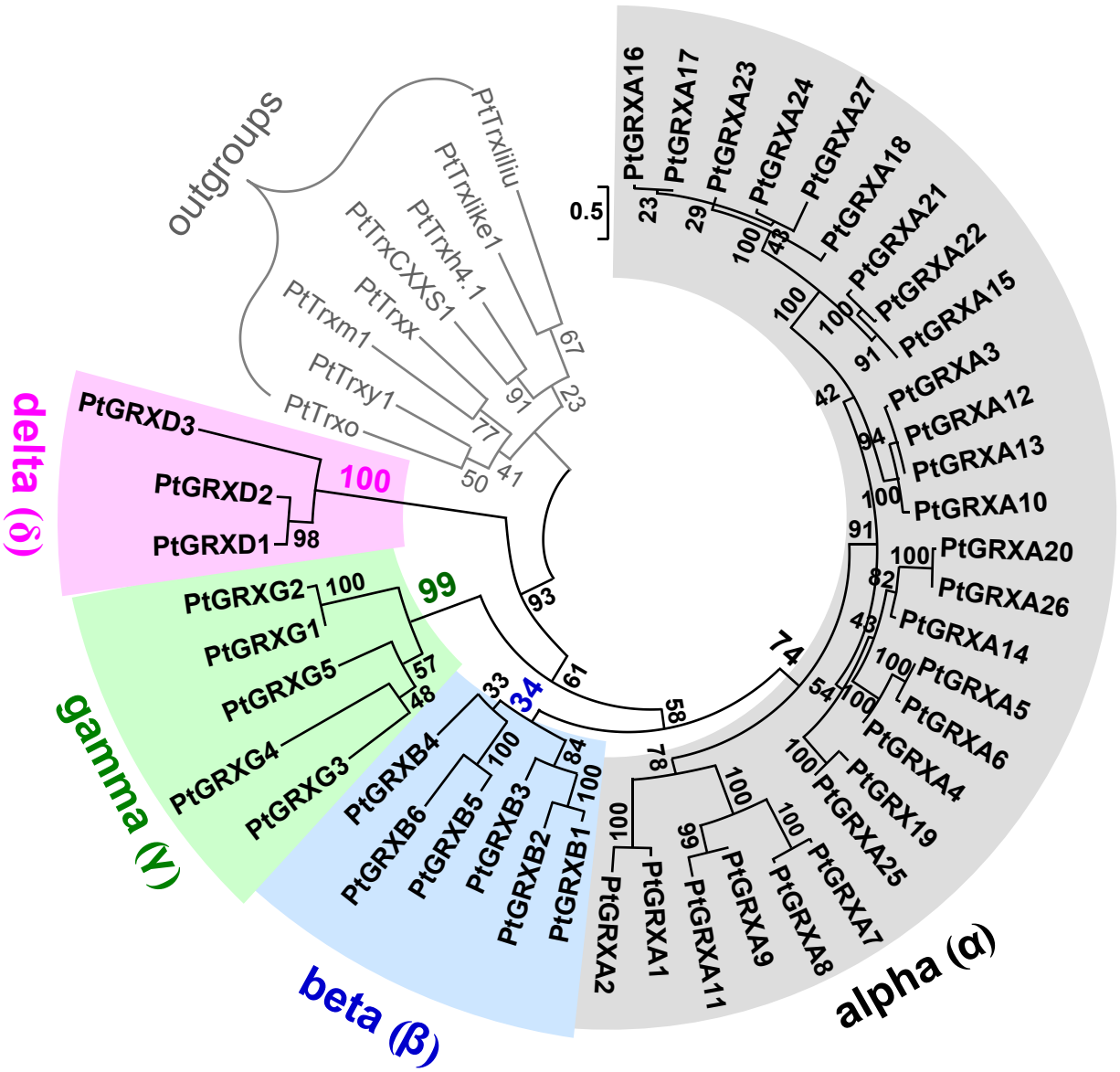



**Figure 1. Phylogenetic tree of *Populus* GRXs**. Numbers at each node represent the bootstrap values. Alpha (α), beta (β), gamma (γ), and delta (δ) GRX classes are shaded in gray, blue, green, and purple, respectively. Sequence alignments used to reconstruct the tree is available as Supplementary Data Set S1.

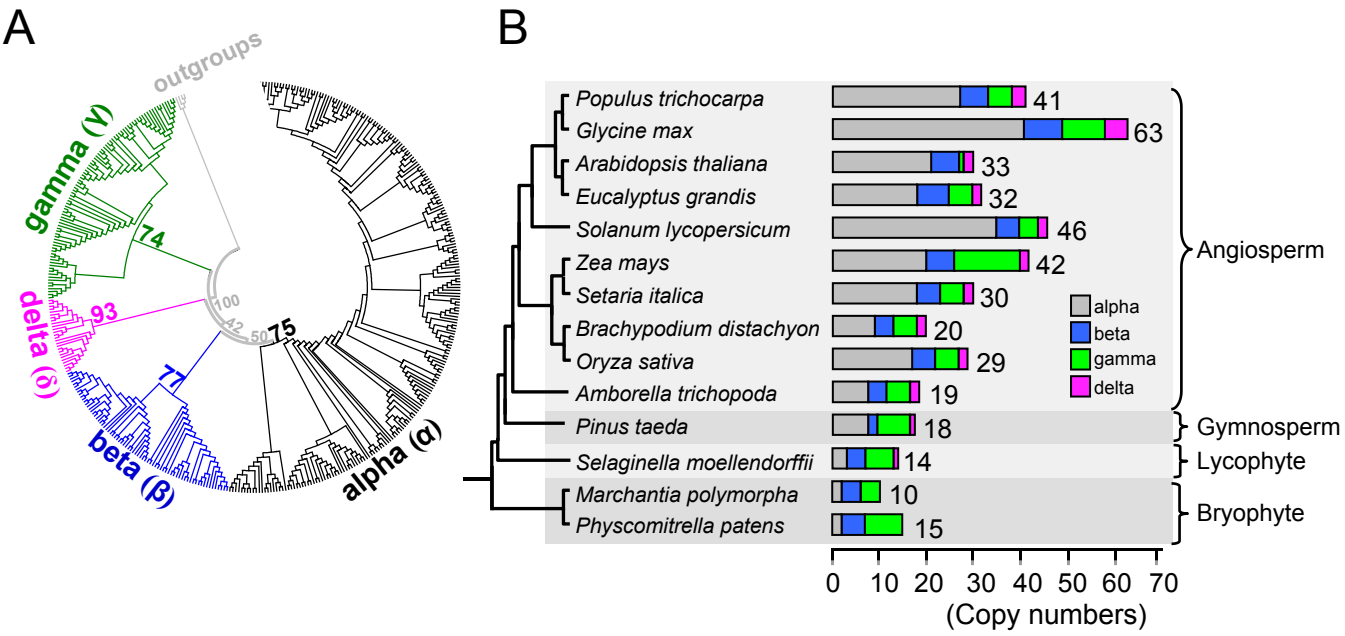



**Figure 2. Phylogenetic tree (A) and copy numbers (B) of GRXs of 14 land plant species.** In (A), **n**umbers at the node in the phylogenetic tree represent bootstrap values. The alpha (α), beta (β), gamma (γ) and delta (δ) GRXs are indicated by black, blue, green, and purple lines, respectively. In (B), the copy number of the GRX genes in each species is shown on the box side. The sequence alignments used to reconstruct the tree in (A) is available as Supplementary Data Set S2.

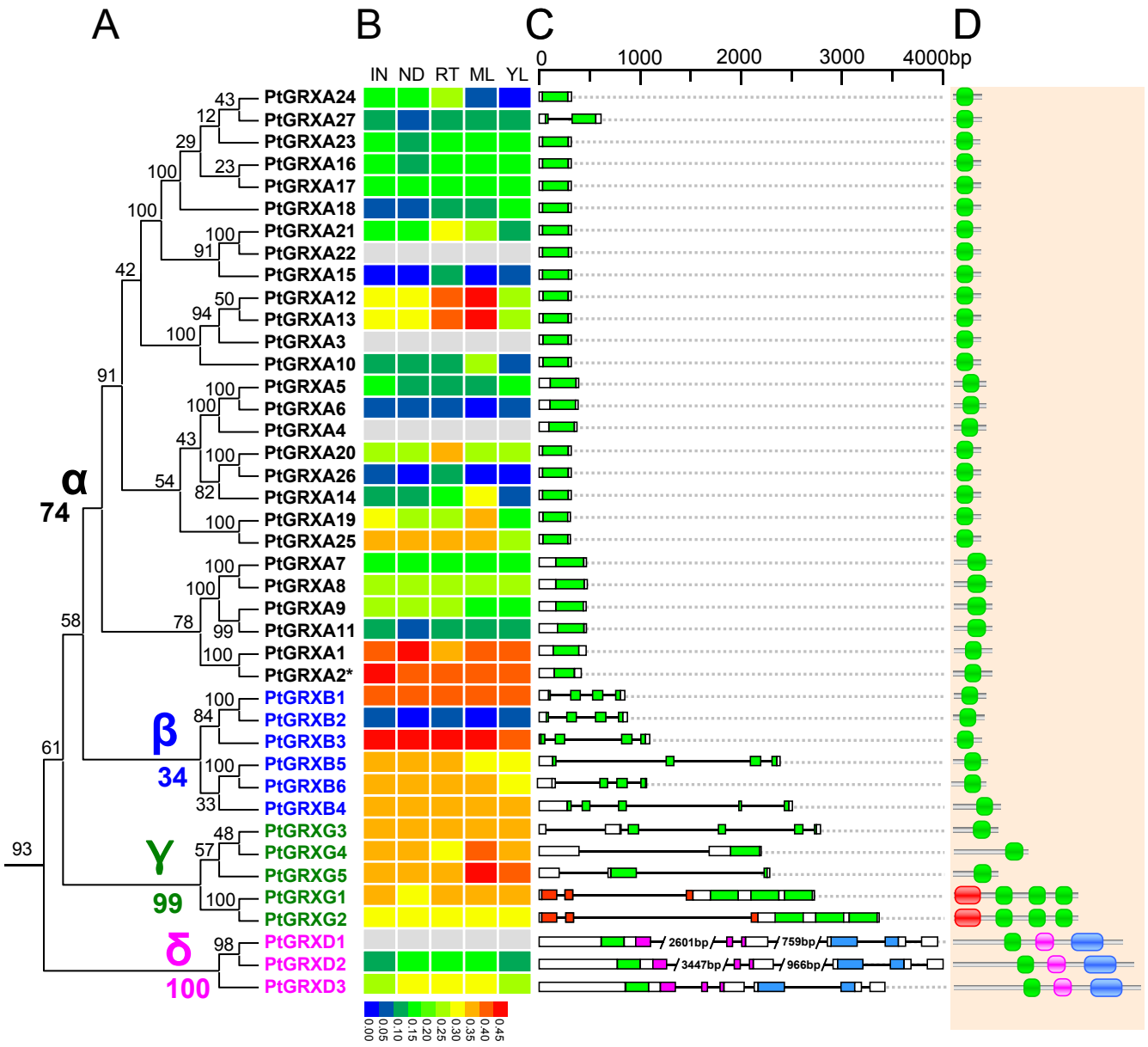



**Figure 3. Phylogenetic relationships (A), expression patterns (B), gene structures (C), and conserved domains (D) of *Populus* GRXs.** In (A), α, β, γ and δ class GRXs are represented by black, blue, green, and purple letters, respectively. Numbers at each node in the phylogenetic tree represent bootstrap values. The pseudogene is indicated by asterisks. In (B), the bar indicates the reference expression level. Gray boxes denote genes that do not have expression information in the PopGenIE database. IN, internodes; ND, nodes; RT, roots; ML, mature leaves; YL, young leaves. In (C), exons and introns are indicated by boxes and lines, respectively. In (D), gray lines indicate the full-length GRX protein sequence. In (C) and (D), GRX, TRX, DEP, and DUF547 domains are highlighted in green, red, purple, and blue, respectively.

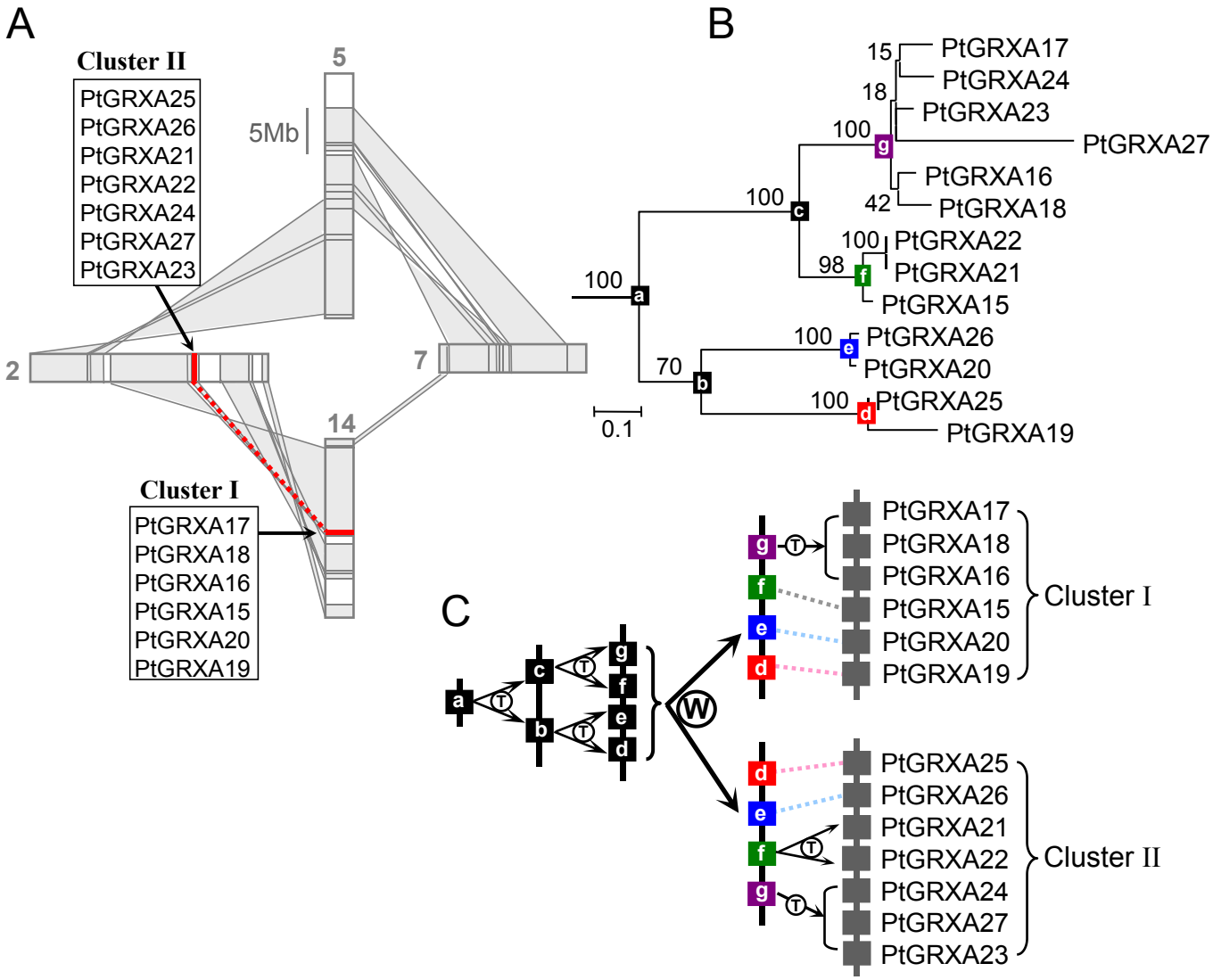



**Figure 4. Genomic localization (A), phylogenetic relationships (B), and putative evolutionary histories (C) of *Populus* GRX genes in clusters I and II.** In (A), regions that correspond to homologous genome blocks are shaded in gray and connected with lines. Paralogous clusters I and II are indicated by red dashed lines within gray-shaded trapezoids. In (B) and (C), phylogenetic tree is reconstructed using the JTT model. Numbers on branches indicate bootstrap values. Letters in the black and colored boxes represent ancestral GRX genes. T, putative tandem duplication; W, putative whole-genome duplication.

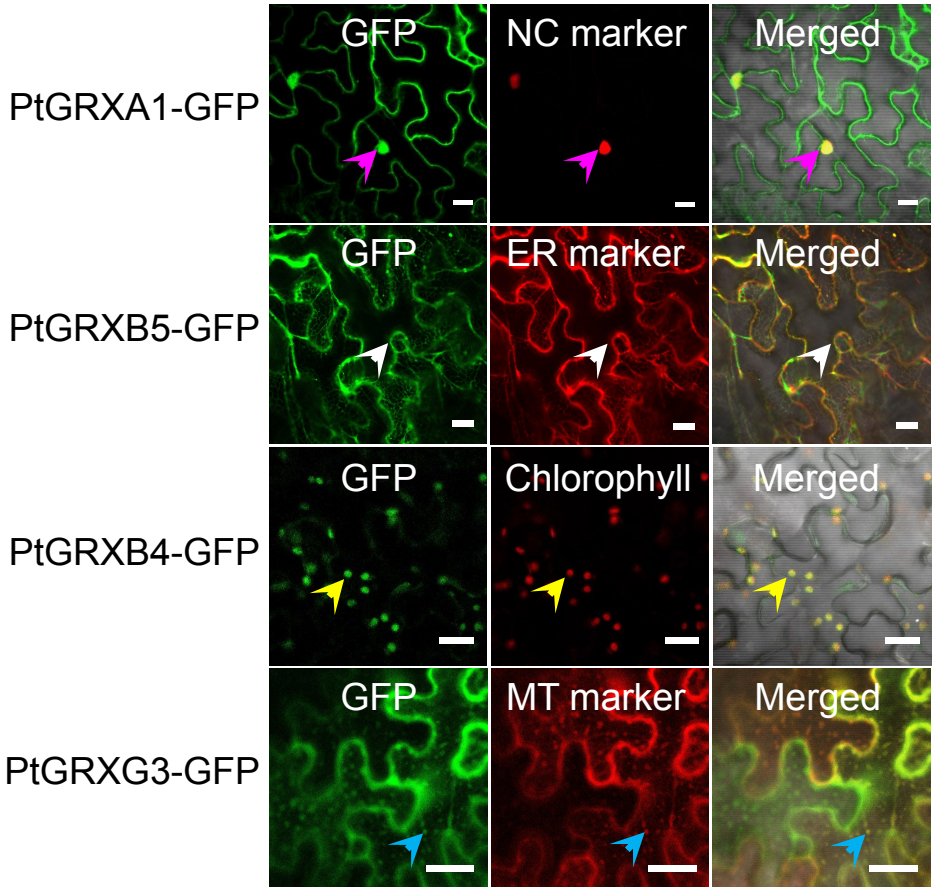
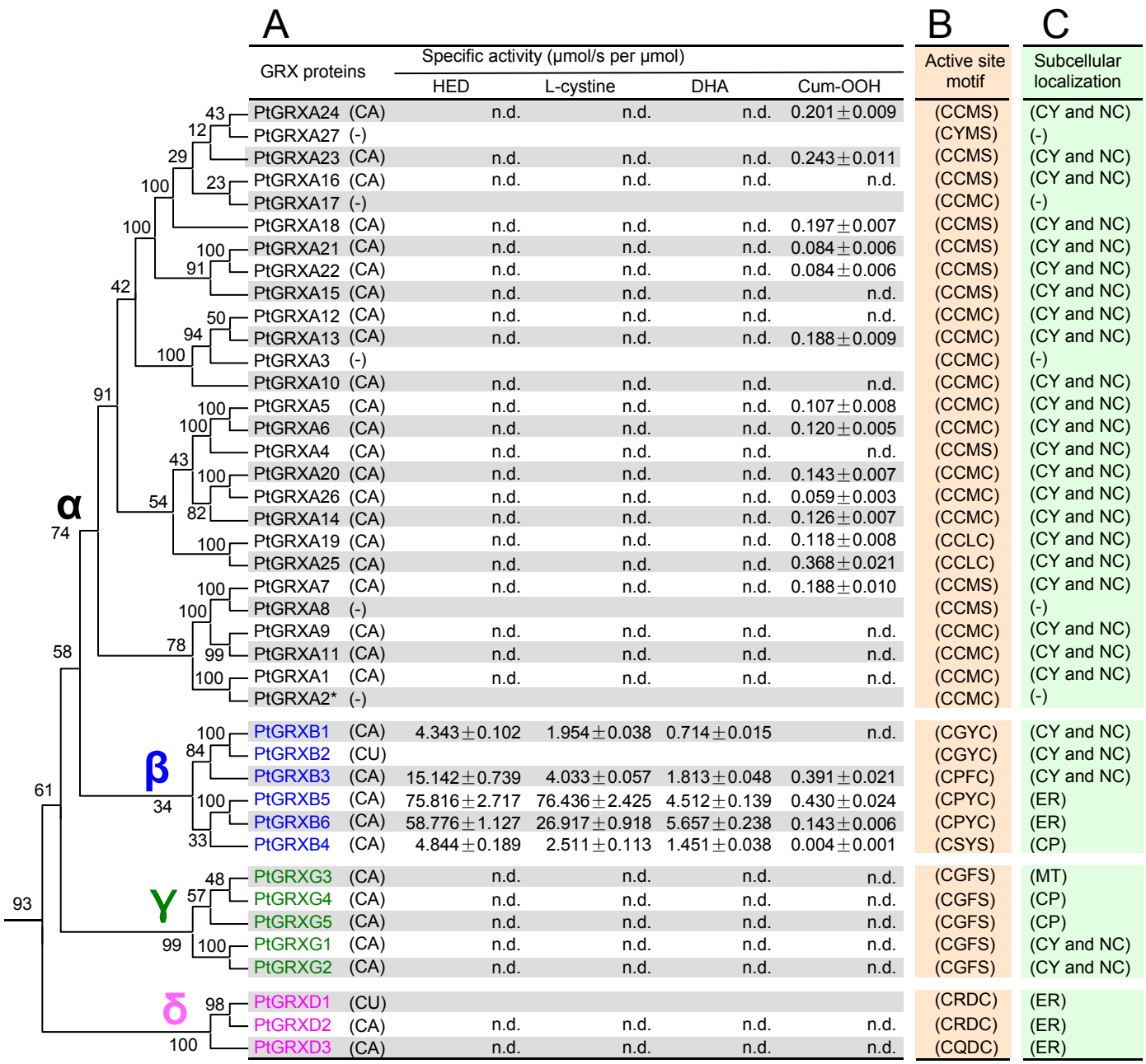



**Figure 5. Subcellular localizations of *Populus* GRX proteins.** GFP signal (green) of the GRX proteins, mCherry signal (red) of nuclear marker (NC marker), endoplasmic reticulum marker (ER marker)， and mitochondrial marker (MT marker), and chlorophyll autofluorescence (red) were detected using confocal laser-scanning microscopy. An overlay is shown in yellow. Nucleus, endoplasmic reticulum, chloroplasts, and mitochondria are indicated by the purple, white, yellow, and blue arrows, respectively. Bars = 20 μm.




| GRX proteins | | Specific activity (μmol/s per μmol) HED | L-cystine | DHA | Cum-OOH | Active site motif | Subcellular localization |
|---|---|---|---|---|---|---|---|
| PtGRXA24 | (CA) | n.d. | n.d. | n.d. | 0.201±0.009 | (CCMS) | (CY and NC) |
| PtGRXA27 | (-) | | | | | (CYMS) | (-) |
| PtGRXA23 | (CA) | n.d. | n.d. | n.d. | 0.243±0.011 | (CCMS) | (CY and NC) |
| PtGRXA16 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMS) | (CY and NC) |
| PtGRXA17 | (-) | | | | | (CCMC) | (-) |
| PtGRXA18 | (CA) | n.d. | n.d. | n.d. | 0.197±0.007 | (CCMS) | (CY and NC) |
| PtGRXA21 | (CA) | n.d. | n.d. | n.d. | 0.084±0.006 | (CCMS) | (CY and NC) |
| PtGRXA22 | (CA) | n.d. | n.d. | n.d. | 0.084±0.006 | (CCMS) | (CY and NC) |
| PtGRXA15 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMS) | (CY and NC) |
| PtGRXA12 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMC) | (CY and NC) |
| PtGRXA13 | (CA) | n.d. | n.d. | n.d. | 0.188±0.009 | (CCMC) | (CY and NC) |
| PtGRXA3 | (-) | | | | | (CCMC) | (-) |
| PtGRXA10 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMC) | (CY and NC) |
| PtGRXA5 | (CA) | n.d. | n.d. | n.d. | 0.107±0.008 | (CCMC) | (CY and NC) |
| PtGRXA6 | (CA) | n.d. | n.d. | n.d. | 0.120±0.005 | (CCMC) | (CY and NC) |
| PtGRXA4 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMS) | (CY and NC) |
| PtGRXA20 | (CA) | n.d. | n.d. | n.d. | 0.143±0.007 | (CCMC) | (CY and NC) |
| PtGRXA26 | (CA) | n.d. | n.d. | n.d. | 0.059±0.003 | (CCMC) | (CY and NC) |
| PtGRXA14 | (CA) | n.d. | n.d. | n.d. | 0.126±0.007 | (CCMC) | (CY and NC) |
| PtGRXA19 | (CA) | n.d. | n.d. | n.d. | 0.118±0.008 | (CCLC) | (CY and NC) |
| PtGRXA25 | (CA) | n.d. | n.d. | n.d. | 0.368±0.021 | (CCLC) | (CY and NC) |
| PtGRXA7 | (CA) | n.d. | n.d. | n.d. | 0.188±0.010 | (CCMS) | (CY and NC) |
| PtGRXA8 | (-) | | | | | (CCMS) | (-) |
| PtGRXA9 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMC) | (CY and NC) |
| PtGRXA11 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMC) | (CY and NC) |
| PtGRXA1 | (CA) | n.d. | n.d. | n.d. | n.d. | (CCMC) | (CY and NC) |
| PtGRXA2* | (-) | | | | | (CCMC) | (-) |
| PtGRXB1 | (CA) | 4.343±0.102 | 1.954±0.038 | 0.714±0.015 | n.d. | (CGYC) | (CY and NC) |
| PtGRXB2 | (CU) | | | | | (CGYC) | (CY and NC) |
| PtGRXB3 | (CA) | 15.142±0.739 | 4.033±0.057 | 1.813±0.048 | 0.391±0.021 | (CPFC) | (CY and NC) |
| PtGRXB5 | (CA) | 75.816±2.717 | 76.436±2.425 | 4.512±0.139 | 0.430±0.024 | (CPYC) | (ER) |
| PtGRXB6 | (CA) | 58.776±1.127 | 26.917±0.918 | 5.657±0.238 | 0.143±0.006 | (CPYC) | (ER) |
| PtGRXB4 | (CA) | 4.844±0.189 | 2.511±0.113 | 1.451±0.038 | 0.004±0.001 | (CSYS) | (CP) |
| PtGRXG3 | (CA) | n.d. | n.d. | n.d. | n.d. | (CGFS) | (MT) |
| PtGRXG4 | (CA) | n.d. | n.d. | n.d. | n.d. | (CGFS) | (CP) |
| PtGRXG5 | (CA) | n.d. | n.d. | n.d. | n.d. | (CGFS) | (CP) |
| PtGRXG1 | (CA) | n.d. | n.d. | n.d. | n.d. | (CGFS) | (CY and NC) |
| PtGRXG2 | (CA) | n.d. | n.d. | n.d. | n.d. | (CGFS) | (CY and NC) |
| PtGRXD1 | (CU) | | | | | (CRDC) | (ER) |
| PtGRXD2 | (CA) | n.d. | n.d. | n.d. | n.d. | (CRDC) | (ER) |
| PtGRXD3 | (CA) | n.d. | n.d. | n.d. | n.d. | (CQDC) | (ER) |

**Figure 6. Phylogenetic tree and enzymatic activities (A), active site motifs (B), and subcellular localizations (C) of *Populus* GRX proteins.** In (A), α, β, γ, and δ class GRXs are represented by black, blue, green, and purple letters, respectively. Numbers at each node in the phylogenetic tree represent bootstrap values. The pseudogene is indicated by asterisk. Suffixes attached to each GRX label: C, successfully cloned; A, purified GRX assayed; U, recombinant protein were not expressed; -, analysis not performed. Values shown are mean ± SD, as calculated from at least three replicates. n.d., no activity detected. In (B), the active site motif of each GRX protein is shown. In (C), CY, cytosol; NC, nucleus; ER, endoplasmic reticulum; MT, mitochondria; CP, chloroplast. Dash indicates that analysis is not performed.

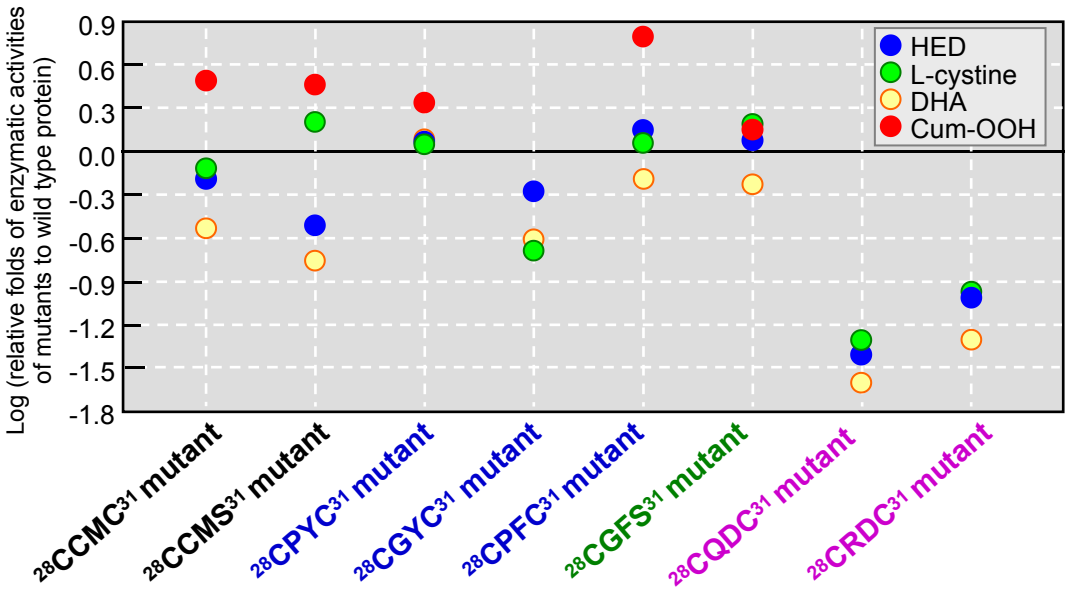



**Figure 7. Changes in enzymatic activities induced by mutations at the active site motif of PtGRXB4 protein.** The enzymatic activities are means as calculated from at least three replicates. The enzymatic activity of PtGRXB4 toward each substrate is set as baseline for comparison with the mutants specified on the x axis.

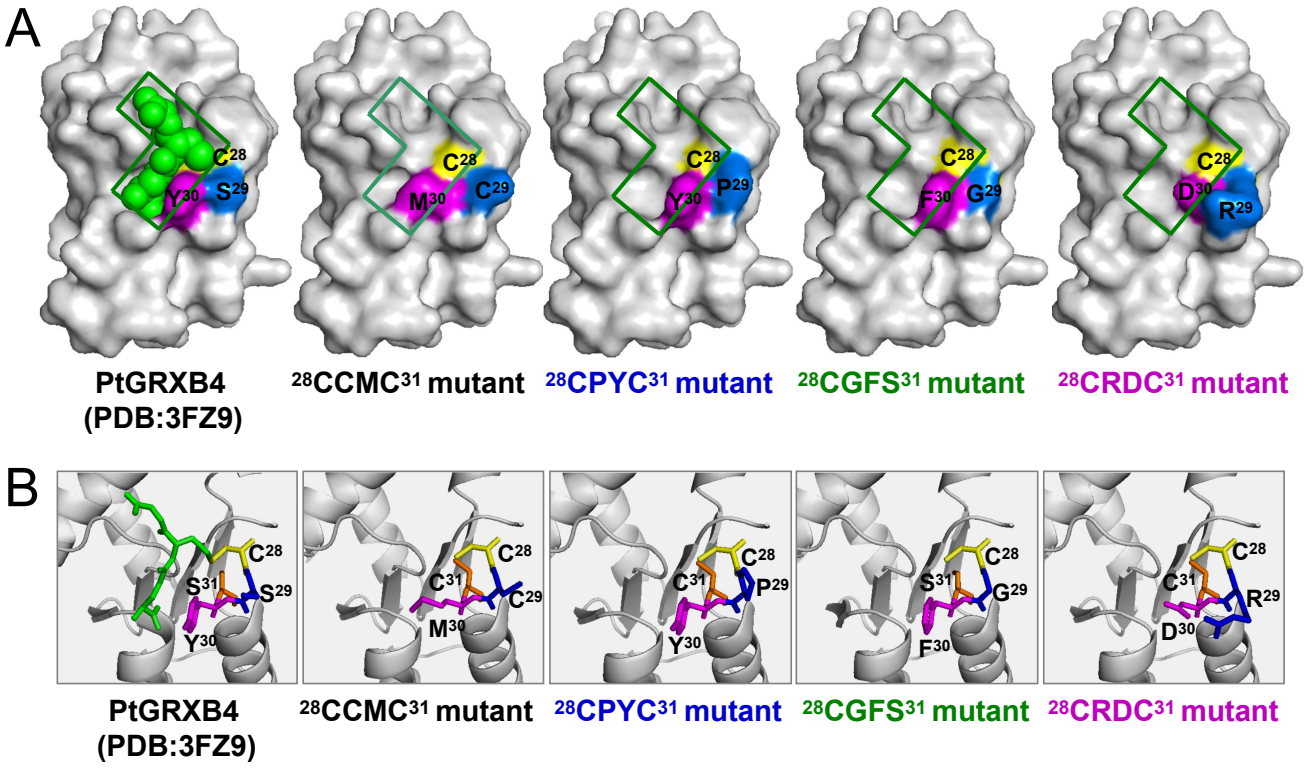



**Figure 8. GRX domain structures (A) and side chains of the amino acids in the active site motifs (B) of PtGRXB4 and its mutants.** In (A), glutathione (GSH) is shown as green balls. The GSH-binding pocket is indicated by a green L-shaped box. In (B), amino acids in the active site motif of both PtGRXB4 and its mutants are shown as sticks; GSH is shown as green sticks.

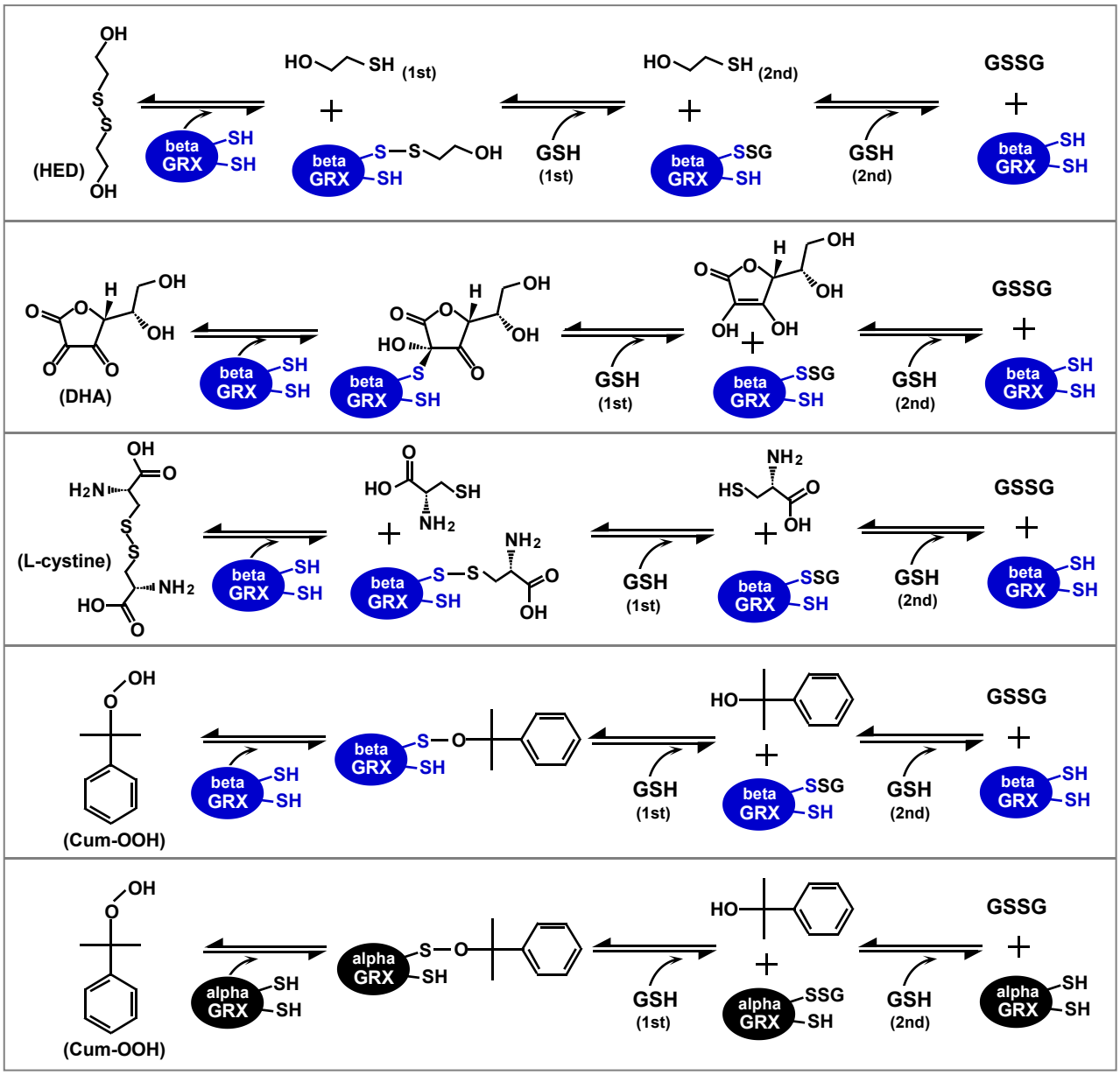



**Figure 9. Mechanistic models of reactions catalyzed by GRXs.** Alpha and beta GRXs are shown in black and blue ovals, respectively.