Genome-Wide Analysis of the Glutathione *S*-Transferase Gene Family in *Capsella rubella*: Identification, Expression, and Biochemical Functions

Gang He[1,2†], Chao-Nan Guan[3†], Qiang-Xin Chen[1], Xiao-Jun Gou[2], Wei Liu[2], Qing-Yin Zeng[1] and Ting Lan[1]*

[1] Functional Genomics and Protein Evolution Group, State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China, [2] The Key Laboratory of Medicinal and Edible Plants Resources Development of Sichuan Education Commission, Chengdu University, Chengdu, China, [3] College of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing, China

Extensive subfunctionalization might explain why so many genes have been maintained after gene duplication, which provides the engine for gene family expansion. However, it is still a particular challenge to trace the evolutionary dynamics and features of functional divergences in a supergene family over the course of evolution. In this study, we identified 49 Glutathione *S*-transferase (GST) genes from the *Capsella rubella*, a close relative of *Arabidopsis thaliana* and a member of the mustard family. *Capsella* GSTs can be categorized into eight classes, with tau and phi GSTs being the most numerous. The expansion of the two classes mainly occurs through tandem gene duplication, which results in tandem-arrayed gene clusters on chromosomes. By integrating phylogenetic analysis, expression patterns, and biochemical functions of *Capsella* and *Arabidopsis* GSTs, functional divergence, both in gene expression and enzymatic properties, were clearly observed in paralogous gene pairs in *Capsella* (even the most recent duplicates), and orthologous GSTs in *Arabidopsis/Capsella*. This study provides functional evidence for the expansion and organization of a large gene family in closely related species.

Keywords: gene family, gene duplication, genome, enzyme activity, functional divergence

INTRODUCTION

Glutathione *S*-transferases (GSTs; EC 2.5.1.18) are multifunctional proteins encoded by a large gene family that is found in most organisms. As classical phase II detoxification enzymes, GSTs mainly catalyze the conjugation of reduced glutathione (GSH) with a wide variety of reactive electrophiles (Hayes et al., 2005). In plants, GST proteins are involved in several crucial physiological and developmental processes, including xenobiotic (e.g., herbicides) detoxification, signal transduction, isomerization, and protection against oxidative damages, UV radiation, and heavy metal toxins (Dixon et al., 2010; Cummins et al., 2011). Based on amino acid sequence similarity and gene organization, plant GSTs have been categorized into eight classes: phi, tau, theta, zeta, lambda, dehydroascorbate reductase (DHAR), tetrachlorohydroquinone dehalogenase (TCHQD) and the

class containing the γ-subunit of the eukaryotic translation elongation factor 1B (EF1Bγ) (Oakley, 2005; Lan et al., 2009; Dixon and Edwards, 2010a). We recently identified two new GST classes (hemerythrin and iota) in non-vascular plants (Liu et al., 2013). Among the ten GST classes, phi, tau, lambda, and DHAR GSTs are considered unique to plants (Frova, 2006).

In plants, tau and phi class GSTs are the most numerous and play important roles in detoxification of xenobiotics (Frova, 2003). Overexpression of tau or phi GSTs in plants can increase tolerance to oxidation, herbicides, salinity, and chilling (Roxas et al., 1997; Cummins et al., 1999; Karavangeli et al., 2005; Benekos et al., 2010; Sharma et al., 2014). These proteins also participate in non-catalytic functions, e.g., binding/transport and signaling (Marrs, 1996; Lieberherr et al., 2003; Kitamura et al., 2004). Lambda and DHAR GSTs do not exhibit activity toward xenobiotics but are considered to be involved in redox and thiol transfer reactions (Dixon et al., 2002a; Dixon and Edwards, 2010b). DHAR GSTs have key functions not only in the ascorbate-GSH recycling reaction but also in stress resistance (Kwon et al., 2003; Chen and Gallie, 2006; Ushimaru et al., 2006). Recent studies demonstrated that some stress-inducible lambda GSTs could selectively bind flavonols and serve as antioxidants (Dixon and Edwards, 2010b; Dixon et al., 2011). The theta and zeta GSTs have counterparts in the mammalian system and function mainly as GSH-dependent peroxidases and isomerases (Thom et al., 2001; Basantani and Srivastava, 2007). GSTs in EF1Bγ class contain two domains: a typical GST domain and an EF1Bγ domain. The GST domain of EF1Bγ class GSTs functions as GSH peroxidases (Vickers et al., 2004).

Capsella rubella is from the same family as *Arabidopsis thaliana*. *C. rubella* is a model species widely used for studying natural variation in adaptive traits, such as flowering time (Guo et al., 2012). This species is also a good model for understanding the evolution of self-fertilization (Guo et al., 2009). In *Arabidopsis*, the haploid set consists of five chromosomes, whereas its close relative *C. rubella* has $n = 8$ chromosomes (Boivin et al., 2004). The progenitors of the lineage leading to *A. thaliana* and *C. rubella* diverged approximately 10 million years ago (Acarkan et al., 2000; Koch and Kiefer, 2005). The *C. rubella* genome has been completely sequenced (Slotte et al., 2013), thus facilitating the understanding of the evolutionary relationship between *C. rubella* and its relative *A. thaliana* from the gene family level. In this study, we performed genome-wide annotation of the GST gene family of *C. rubella*. Through phylogenetic analysis with expression and functional assays, we provided detailed characterization of the organization, gene expression pattern, and enzymatic properties of the GST members. Extensive functional divergence was observed among members within tandem-arrayed GST clusters and between paralogous gene pairs. Through comparative analyses of this family in *C. rubella* and *A. thaliana*, we examined the lineage-specific loss/gain events, and divergences in expression and substrate specificity in the orthologous GSTs. The genome-wide, multifaceted approach we employed provides new insights into the process of gene family evolution between closely related species.

MATERIALS AND METHODS

Gene Identification and Nomenclature

To identify putative GST members in *C. rubella*, we performed TBLASTN searches with default algorithm parameters in the *Capsella* genome database, version 1.0[1], using 55 GST protein sequences of *Arabidopsis* (Dixon and Edwards, 2010a), 81 of *populus* (Lan et al., 2009), and 575 of other plants, animals, fungi, and bacteria (Supplementary Table S5) as queries. These 575 full-length GSTs represent 36 GST sub-families defined by the NCBI Conserved Domain Database (CDD; Marchler-Bauer et al., 2011). All potential candidates identified were examined using the Pfam[2] and CDD[3] database to confirm the presence of typical GST N- and C-terminal domains in their protein structures. Preliminary classification of GST genes into subfamilies was performed using phylogenetic analysis. The proteins, which clustered with soluble cytosolic GSTs, have an ancient monophyletic origin (Dixon and Edwards, 2010a). They were used in subsequent analyses. Next, *Capsella* GSTs were amplified from genomic DNA and mRNA from mixed tissues of *C. rubella*, cloned into the pGEM-T Easy Vector (Promega), and sequenced in both directions to verify the gene sequences. The primers used for gene amplification are listed in Supplementary Table S2. Complete manual curation of the gene sequences and structures based on expressed sequence tag (EST) databases and experimental support was further performed to rectify incorrect start codon predictions, splicing errors, missed or extra exons, and incorrectly predicted pseudogenes. For genes that went undetected by PCR (5 out of 49 in this study), their gene structures were assumed to be identical to those of their closest phylogenetic relatives. This approach was adapted from other studies (Meyers et al., 2003).

The nomenclature for *Capsella* GSTs follows the system suggested by Dixon et al. (2002b) for plant GSTs. A univocal name was assigned to each *Capsella* GST gene consisting of two italic letters *Cr* denoting the source organism, the family name (e.g., *CrGSTU, CrGSTF, CrGSTT, CrGSTZ, CrGSTL, CrTCHQD, CrDHAR,* and *CrEF1Bγ* corresponding to tau, phi, theta, zeta, lambda, TCHQD, DHAR, and EF1Bγ classes, respectively) and a progressive number for each gene (e.g., *CrGSTU1*).

Phylogenetic Analyses

Full-length amino acid sequences were aligned using MUSCLE software[4] and adjusted manually with BioEdit (Hall, 1999). Phylogenetic analysis was performed using the maximum-likelihood (ML) method in PHYML software (Guindon and Gascuel, 2003) with the Jones, Taylor, and Thornton (JTT) amino acid substitution model. GRX2 protein from *Escherichia coli* was chosen as an out-group during phylogenetic analysis of the *Capsella* GST family, as cytosolic GSTs are thought to be derived from the GRX2 (Holm et al., 2006). For phylogenetic analysis of each GST class, members of the sister class were used as an

---

[1]https://phytozome.jgi.doe.gov/pz/portal.html
[2]http://pfam.xfam.org/search
[3]http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
[4]http://www.drive5.com/muscle/

out-group. One-thousand bootstrap replicates were conducted to obtain confidence support.

Expression of GST Genes in *Capsella* Tissues

The expression patterns of *Capsella* GST members during growth under normal conditions were examined by reverse transcription PCR (RT-PCR). Seeds of *C. rubella* were germinated on agar plates (Murashige and Skoog, 1962) and vernalized at 4°C for 4 days. Then, the seeds were grown in growth chambers under normal conditions (14 h light/10 h dark cycle) at a temperature of 25°C/22°C (day/night). Seedling plants were transplanted to soil for 2 weeks and harvested for RT-PCR analysis. We isolated total RNA from rossette leaves, roots, and hypocotyl tissues of each plant and dry seeds using an Aurum Total RNA kit (Bio-Rad Laboratories). Total RNA was treated with RNase-free DNase I (Promega) and reverse transcribed into cDNA using a TaKaRa RNA PCR kit (AMV), version 3.0. Forty-nine specific primer pairs were designed (Supplementary Table S3). The *actin* gene (Carubv10013961m.g) was used as an internal control. PCR conditions were optimized to consist of an initial denaturation step of 3 min at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 60°C and 30 s at 72°C, with a final extension of 5 min at 72°C. PCR products from each sample were analyzed on 1% agarose gel and were validated by DNA sequencing. Independent biological triplicates were used in all of the RT-PCR analyses.

Gene expression profiles of the *Capsella* GSTs were compared with expression data from *Arabidopsis* ecotype *Columbia-0* (*Col-0*; Schmid et al., 2005) available through the *Arabidopsis* eFP browser at BAR (Winter et al., 2007). The eFP browser was set to the developmental map, with absolute expression values for gene expression. In this study, genes with values below 20 units were considered to be not expressed (Winter et al., 2007). The microarray data sets used in this study include leaves at rosette stage (ATGE_89_A, ATGE_89_B and ATGE_89_C), roots at rosette stage (ATGE_9_A, ATGE_9_B and ATGE_9_C), hypocotyls at seedling stage (ATGE_2_A, ATGE_2_B and ATGE_2_C), and dry seeds (RIKEN-NAKABAYASHI1A and RIKEN-NAKABAYASHI1B).

Putative Promoter Sequence Analysis

Gene promoter sequences were extracted 1000 pb upstream of the transcriptional start site of each *Capsella* GST. Plant CARE database[5] was used to find putative *cis*–elements among the promoter sequences. Divergence between upstream sequences of each paralogous gene pairs was measured by the GATA program (Nix and Eisen, 2005), with window size set as seven and lower cutoff score 12 bit.

Expression and Purification of Recombinant *Capsella* GST Proteins

To investigate the enzymatic functions of *C. rubella* GST proteins, 24 tau, 11 phi, three DHAR, and three zeta GSTs were selected for protein expression analysis and purification. The primers used to

[5]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

construct the GST expression vectors are listed in Supplementary Table S4. The products were subcloned into pET-30a expression vectors (Novagen) to obtain a 6×His-tag at the N-terminus. The resulting plasmids, pET-30a/GSTs, were transformed into *E. coli* BL21 (DE3) and verified by sequencing. The transformed *E. coli* cells were cultured at 37°C and grown until the optical density (A_{600}) reached 0.5. A final concentration of 0.1 mM isopropyl-β-D-thiogalactopyranoside was added to each culture, and the cultures were incubated at 37°C or 20°C overnight. The cells were harvested by centrifugation (10,000 g, 3 min, 4°C), resuspended in binding buffer (20 mM sodium phosphate, 0.5 M NaCl, and 20 mM imidazole, pH 7.4), and disrupted by cold sonication. The resulting homogenate was subjected to centrifugation (10,000 × g, 10 min, 4°C) and the supernatant was loaded onto a Ni Sepharose High Performance column (GE Healthcare Bio-Sciences) that had been pre-equilibrated with binding buffer. The GST proteins that bound to the Ni Sepharose High Performance column were eluted with elution buffer (20 mM sodium phosphate, 0.5 M NaCl, and 0.5 M imidazole, pH 7.4). The particulate material, a small portion of the supernatant and the purified proteins were analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) consisting of a 10% separating gel and a 5% stacking gel.

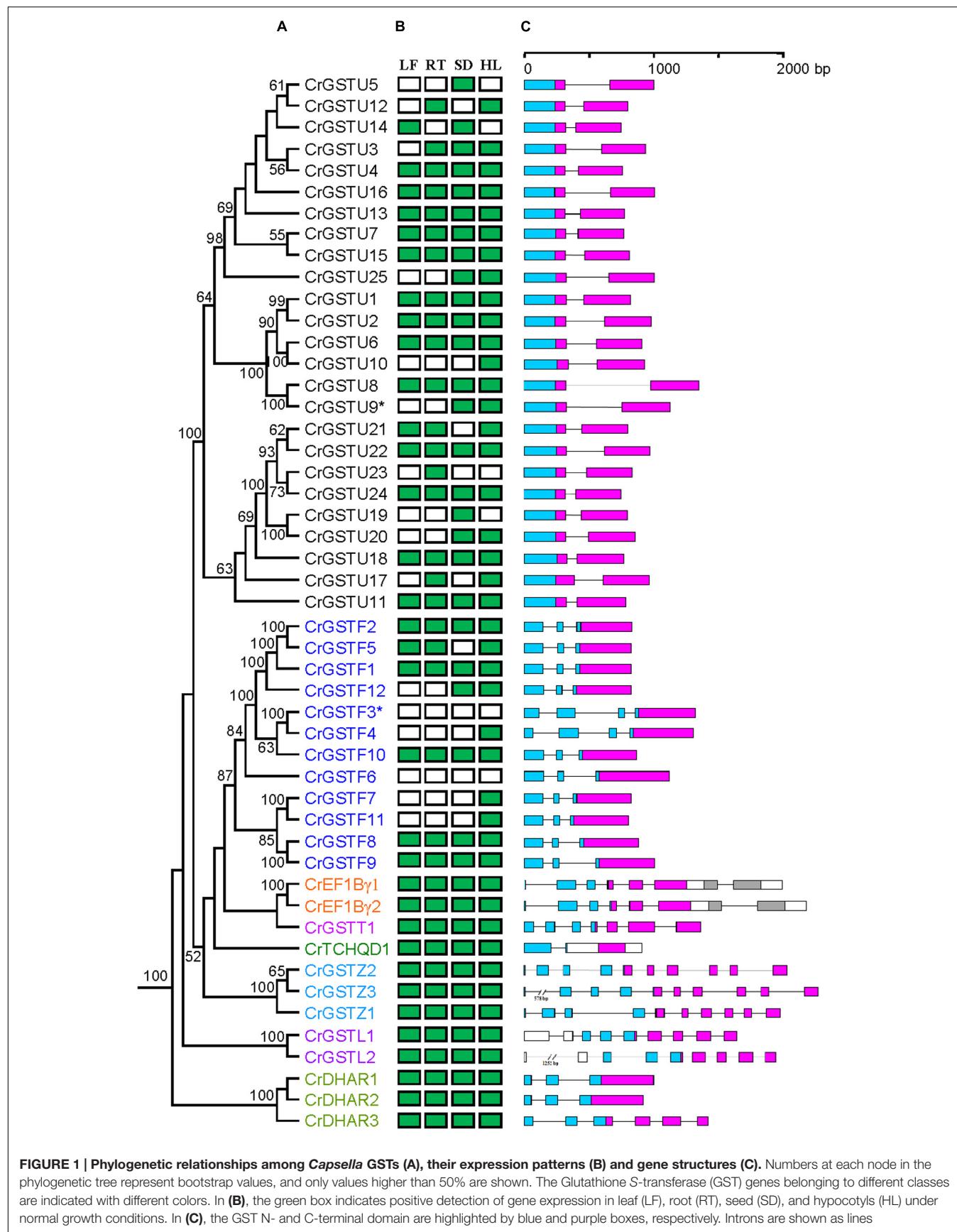## Enzyme Assays of *Capsella* GST Proteins

The enzyme activity of plant GSTs was measured using the following six substrates: 1-chloro-2,4-dinitrobenzene (CDNB) and 4-nitrobenzyl chloride (NBC), as described by Habig et al. (1974); 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole (NBD-Cl), as described by Ricci et al. (1994); and cumene hydroperoxide (Cum-OOH), dehydroascorbate (DHA), and diphenyl ethers (Fluorodifen), as described by Edwards and Dixon (2005). All assays were carried out at 25°C. Protein concentrations were determined by measuring the absorbance at 280 nm.
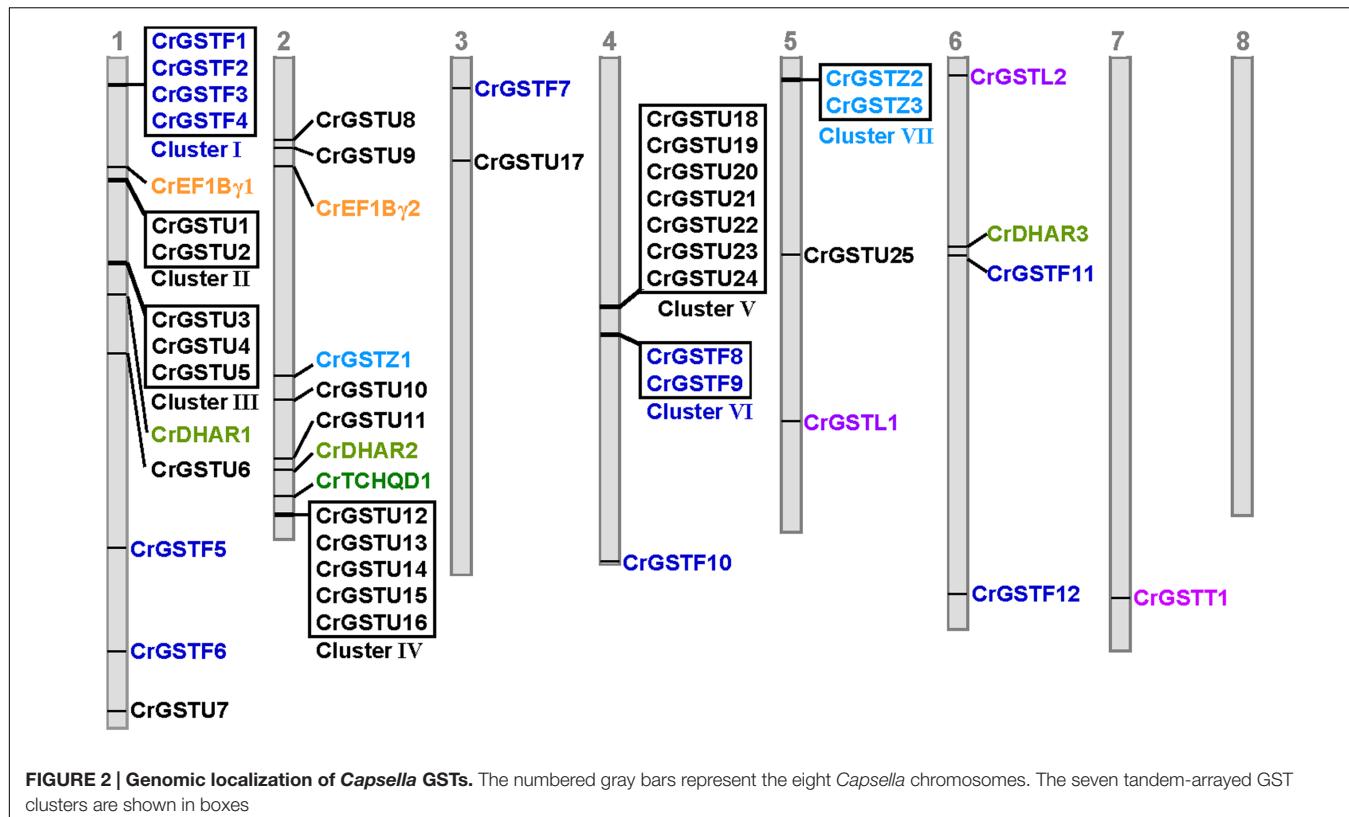
RESULTS

Identification of the GST Genes from the *C. rubella* Genome

Forty-nine full-length genes encoding putative cytosolic GST proteins were identified in the *C. rubella* genome (Supplementary Table S1). Among these 49 genes, two genes (*CrGSTF3* and *CrGSTU9*) were considered to be putative pseudogenes because one contained a frame shift disrupting the coding region and the other contained a premature stop codon. After revising the frame shifts by deleting two nucleotides or removing the stop codon, these two full-length sequences were included in the phylogenetic and gene expression analyses. Based on a conserved domain analysis, these 49 GST candidates were divided into eight classes. The tau and phi GSTs were most numerous, with 25 and 12 copies, respectively. The DHAR and zeta classes each contained three members. Both the lambda and EF1Bγ classes were each represented by two members, and the theta and TCHQD classes only had one member each.

FIGURE 1 | Phylogenetic relationships among *Capsella* GSTs (A), their expression patterns (B) and gene structures (C). Numbers at each node in the phylogenetic tree represent bootstrap values, and only values higher than 50% are shown. The Glutathione *S*-transferase (GST) genes belonging to different classes are indicated with different colors. In **(B)**, the green box indicates positive detection of gene expression in leaf (LF), root (RT), seed (SD), and hypocotyls (HL) under normal growth conditions. In **(C)**, the GST N- and C-terminal domain are highlighted by blue and purple boxes, respectively. Introns are shown as lines

FIGURE 2 | Genomic localization of *Capsella* GSTs. The numbered gray bars represent the eight *Capsella* chromosomes. The seven tandem-arrayed GST clusters are shown in boxes

Conserved gene structures were identified within each GST class. All 25 tau GST genes contained a one-intron/two-exon structure (**Figure 1C**). The intron positions were highly conserved among tau GST genes, and their lengths ranged from 70 to 642 bp. Ten of 12 phi GSTs had a two-intron/three-exon structure with a highly conserved first intron position. *CrGSTF3* and *CrGSTF4* contained a three-intron/four-exon structure. For the EF1Bγ, lambda, and zeta classes, the exon-intron architectures were conserved within each class, with each including 7, 9, and 10 exons, respectively.

Genomic Organization of the *Capsella* GST Gene Family

The genomic locations of 49 full-length GSTs were assigned to all of the *Capsella* chromosomes except for chromosome 8 (**Figure 2**). The distribution of the GST genes among the chromosomes was obviously heterogeneous. Seven clusters (clusters I, II, III, IV, V, VI, and VII) with relatively high densities of GSTs were discovered on four chromosomes. In total, 51% of *Capsella* GST genes were organized in tandem repeats, indicating that tandem duplications significantly contributed to the expansion of the *Capsella* GST gene family.

Among the seven clusters, cluster V was the largest, consisting of seven tau GSTs in a 25-kb region on chromosome 4. These seven GST genes were oriented in the same direction. Cluster IV contained five tau GSTs that were clustered in a 10-kb region on chromosome 2. Three clusters (cluster I, II, and III) were located on chromosome 1. Cluster I harbored four phi GSTs

arranged head-to-head in tandem in an 8-kb region. Clusters II and III contained two and three tau GSTs, respectively. Cluster VI contained two phi GST genes, and cluster VII contained two zeta GSTs.

Phylogenetic Relationship of *Capsella* and *Arabidopsis* GST Gene Families

To investigate the lineage-specific expansion of GST genes in the *Capsella* genome, we performed a phylogenetic analysis of all GSTs from *Capsella* and *Arabidopsis*. *Capsella* and *Arabidopsis* had 25 and 28 tau GST genes, respectively. There were at least 27 ancestral GST genes in the most recent common ancestor (MRCA) of *Capsella* and *Arabidopsis* (**Figures 3** and **4**). After the split, both *Capsella* and *Arabidopsis* gained one gene. However, *Capsella* had lost three genes, resulting in fewer tau GST genes in *Capsella* compared with *Arabidopsis*.

Capsella and *Arabidopsis* had 12 and 13 phi GST genes, respectively. There were at least 12 ancestral GST genes in the MRCA of *Capsella* and *Arabidopsis* (**Figures 3** and **5**). After the split, *Capsella* and *Arabidopsis* gained one and two genes, respectively. Additionally, *Capsella* and *Arabidopsis* each lost one gene. Thus, *Capsella* contains one less phi GST gene than *Arabidopsis*.

For DHAR and lambda GSTs, there were at least four and three ancestral GST genes, respectively, in the MRCA of *Capsella* and *Arabidopsis* (**Figures 3** and **5**). After the split, *Capsella* lost a DHAR and lambda GST; however, the *Arabidopsis* genome did not lose any DHAR or lambda GST genes.
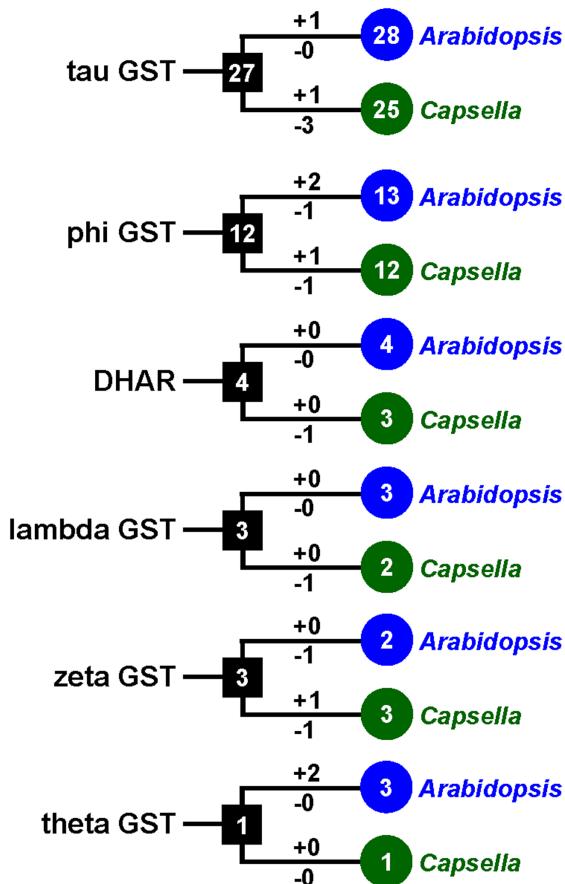
FIGURE 3 | The copy number changes of *Capsella* and *Arabidopsis* GSTs. Numbers in circles and rectangles represent the numbers of GST genes in extant and ancestral species, respectively. Numbers on branches with plus and minus symbols represent the numbers of gene gains and losses, respectively.

For zeta GST, the MRCA of *Capsella* and *Arabidopsis* had at least three ancestral zeta GSTs (**Figures 3** and **5**). After the split, the *Capsella* genome gained and lost one gene, and thus, *Capsella* still contains three zeta GSTs. *Arabidopsis* did not gain new copies. On the contrary, one zeta GST gene was lost, thus, the *Arabidopsis* genome contains two zeta GSTs.

For theta class GSTs, at least one ancestral GST gene was noted in the MRCA of the two species (**Figure 3**). Two duplication events in theta class GSTs were found in the *Arabidopsis* lineage after the split from *Capsella* (**Figure 5**). For TCHQD and EF1B class GSTs, we did not identify any gene gain or loss events after the split of these two species (**Figure 5**).

Expression Patterns of the *Capsella* GST Gene Family

We examined the tissue-specific expression patterns of all 49 *Capsella* GSTs in four tissues, including leaves, roots, seeds, and hypocotyl zones using RT-PCR (**Figure 1B**). The expression patterns of the six minor class GSTs (*CrEF1Bγ1* and *2*, *CrGSTT1*,

CrTCHQD1, *CrGSTZ1*, *2* and *3*, *CrGSTL1* and *2*, and *CrDHAR1*, *2* and *3*) were homogenous, as these GSTs were expressed in all tissues examined in this study. Expression divergences were observed among the tau and phi class GSTs (**Figure 1B**). Of the 25 tau GSTs, 13 (*CrGSTU1*, *2*, *4*, *6*, *7*, *8*, *11*, *13*, *15*, *16*, *18*, *22*, and *24*) were expressed in all tissues examined, and 12 genes (*CrGSTU3*, *5*, *9*, *10*, *12*, *14*, *17*, *19*, *20*, *21*, *23*, and *25*) were selectively expressed. *CrGSTU23* was exclusively expressed in root tissues. *CrGSTU5* and *CrGSTU19* were exclusively expressed in seed tissues, whereas *CrGSTU10* was only noted in hypocotyl zones. For the 12 phi GSTs, five (*CrGSTF1*, *2*, *8*, *9*, and *10*) were expressed in all the tissues examined, and five (*CrGSTF4*, *5*, *7*, *11*, and *12*) were selectively expressed. *CrGSTF3* and *CrGSTF6* were not expressed in any of the tissues examined (**Figure 1B**), which suggests that these two genes might exhibit loss of gene function by pseudogenization.

The expression profiles of 23 tau and 8 phi orthologous GSTs in *Capsella* and *Arabidopsis* revealed a high degree of divergence (**Figure 6**). In total, 7 of the 31 orthologs displayed similar expression patterns, whereas the remaining 24 orthologs exhibited considerable expression divergence in some tissues. For example, *AtGSTU21* was not detected in any tissue of *Arabidopsis* (Goda et al., 2008; Kram et al., 2009), but its orthologous gene *CrGSTU14* was expressed in leaf and seed tissues (**Figures 1B** and **6**).

Potential regulatory motifs analysis using PlantCARE (Plant *cis*-acting regulatory element database) revealed a number of *cis*-elements in the promoter sequences of *Capsella* GST genes (Supplementary Table S6). These motifs were divided into at least eight functional categories, such as core promoters, ABA/abiotic stress, light, phytohormones, pathogen/elicitor/wound responsive elements as well as elements responsible for metabolism regulation, developmental stage, and organ specific expressions. The result showed considerable differences in the regulatory elements among the *Capsella* GSTs and within the subfamilies. Comparative analysis of upstream regions of close paralogs showed divergence, although there are conserved regions (Supplementary Figure S1), indicating that rapid divergence has occurred in the regulatory regions. Further experimental validation step is required to assess the changes in the *cis*-elements that are responsible for the expression diversity in GST genes.

Substrate Specificities and Activities of *Capsella* GSTs]

In the *Capsella* genome, the tau and phi GSTs are most numerous, with 25 and 12 copies, respectively. The DHAR and zeta classes each contain three members. Thus, in this study, we selected tau, phi, DHAR, and zeta GSTs to express and purify GST proteins. Except for two pseudogenes (CrGSTU9 and CrGSTF3), Forty-one genes were cloned into expression vector pET-30a. Twenty-five of the 41 recombinant proteins were expressed as soluble proteins in *E. coli*, whereas the other 16 were insoluble. To determine the enzyme activity and substrate specificity of the soluble proteins, six substrates were selected: CDNB, NBD-Cl, NBC, Cum-OOH, and DHA.
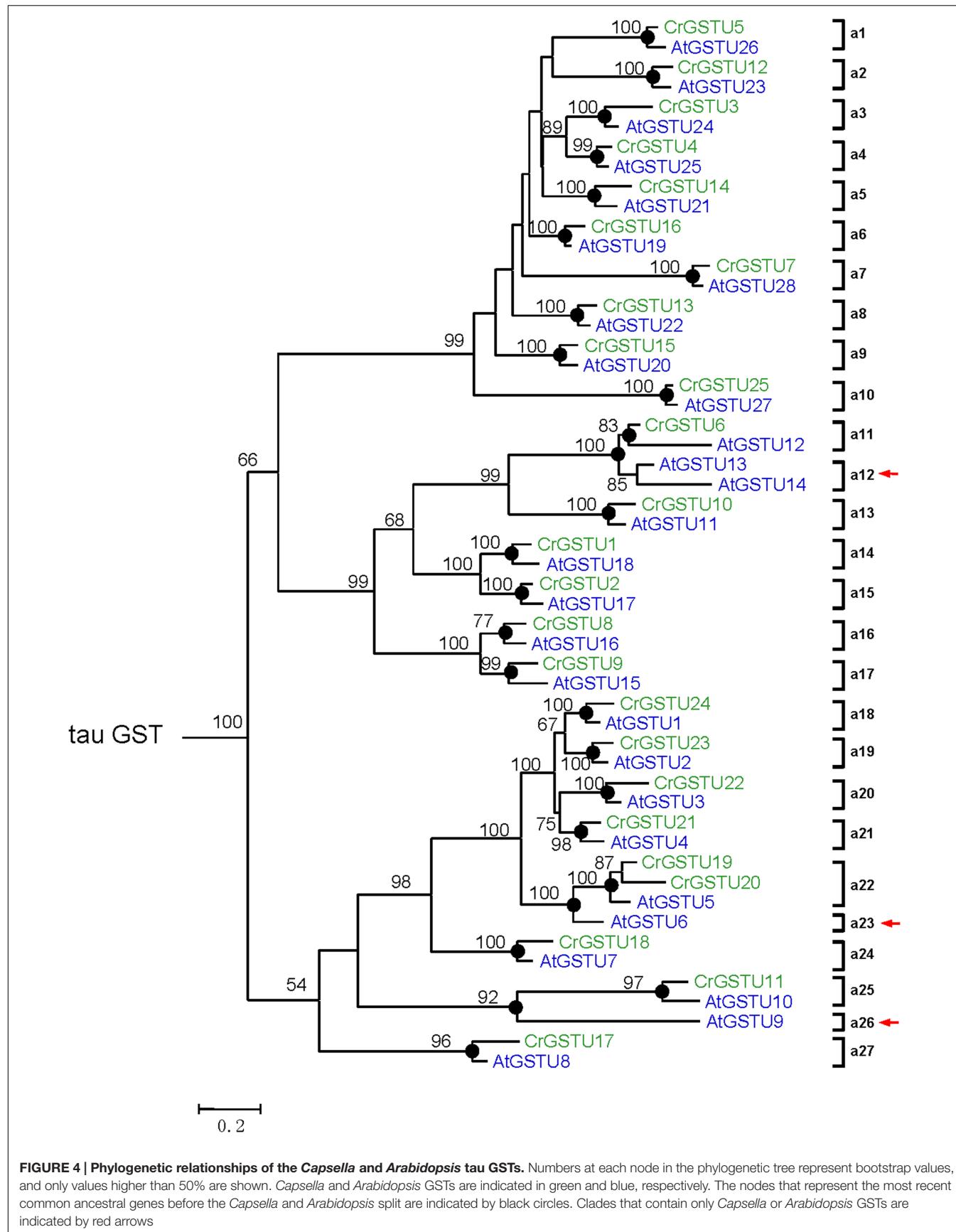
FIGURE 4 | Phylogenetic relationships of the *Capsella* and *Arabidopsis* tau GSTs. Numbers at each node in the phylogenetic tree represent bootstrap values, and only values higher than 50% are shown. *Capsella* and *Arabidopsis* GSTs are indicated in green and blue, respectively. The nodes that represent the most recent common ancestral genes before the *Capsella* and *Arabidopsis* split are indicated by black circles. Clades that contain only *Capsella* or *Arabidopsis* GSTs are indicated by red arrows

FIGURE 5 | Phylogenetic relationships of the *Capsella* and *Arabidopsis* phi and all six minor class GSTs. Numbers at each node in the phylogenetic tree represent bootstrap values, and only values higher than 50% are shown. *Capsella* and *Arabidopsis* GSTs are indicated in green and blue, respectively. The nodes that represent the most recent common ancestral genes before the *Capsella* and *Arabidopsis* split are indicated by black circles. Clades that contain only *Capsella* or *Arabidopsis* GSTs are indicated by red arrows
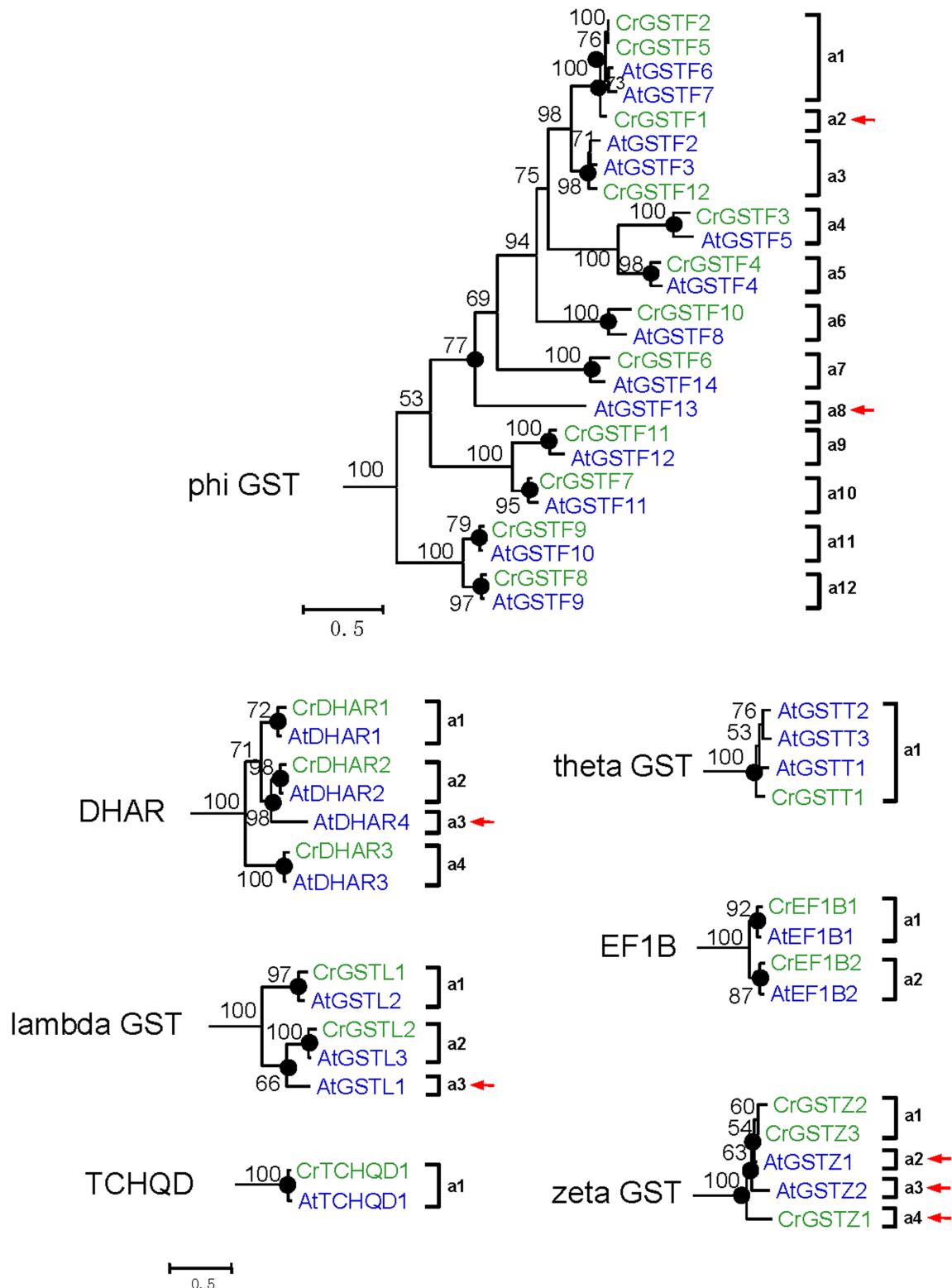
For the tau GSTs, all 14 proteins showed specific activity toward CDNB, 11 toward NBD-Cl, nine toward Cum-OOH, and seven toward NBC and fluorodifen (**Figure 7**). Among the 14 tau GSTs, two (CrGSTU2 and 4) had enzymatic activity toward all five of the substrates. Five proteins exhibited activity toward four substrates, and four toward three substrates. Among the tau GSTs, CrGSTU4 showed the highest enzymatic activity toward CDNB, CrGSTU16 toward NBD-Cl, CrGSTU7 toward NBC, CrGSTU21 toward Cum-OOH, and CrGSTU2 toward fluorodifen. For the 12 phi GSTs, only five proteins were expressed as soluble proteins in *E. coli*. Among these five proteins, CrGSTF10 exhibited boarder substrate spectra and enzymatic activities for four substrates. CrGSTF1 did not exhibit any activity toward any of the tested substrates. All three DHAR GSTs exhibited activity toward DHA. Compared with CrDHAR1 and 3, CrDHAR2 had noticeably reduced enzymatic activity toward DHA, whereas CrDHAR2 exhibited activity toward Cum-OOH.

Substantial variations in specific activities toward different substrates were noted among the members of tandem-arrayed GST clusters. For example, CrGSTU12, 15, and 16 belong to cluster IV. CrGSTU16 displayed a much broader substrate spectrum than did CrGSTU12 and 15. Although CrGSTU12 and 15 shared a similar substrate spectrum, their specific activities toward NBD-Cl varied 10-fold (**Figure 7**).

The enzymatic activities of orthologous GSTs in *Capsella* and *Arabidopsis* also displayed variations. We made a comparison of enzyme specificity toward CDNB and Cum-OOH between orthologous GSTs (Dixon et al., 2009). For example, CrGSTU7, CrGSTU12, CrGSTU15, and CrGSTU18 had enzymatic activity for CDNB but no activity for Cum-OOH, whereas their orthologs in *Arabidopsis* had enzymatic activity for both substrates (**Figures 6** and **7**).
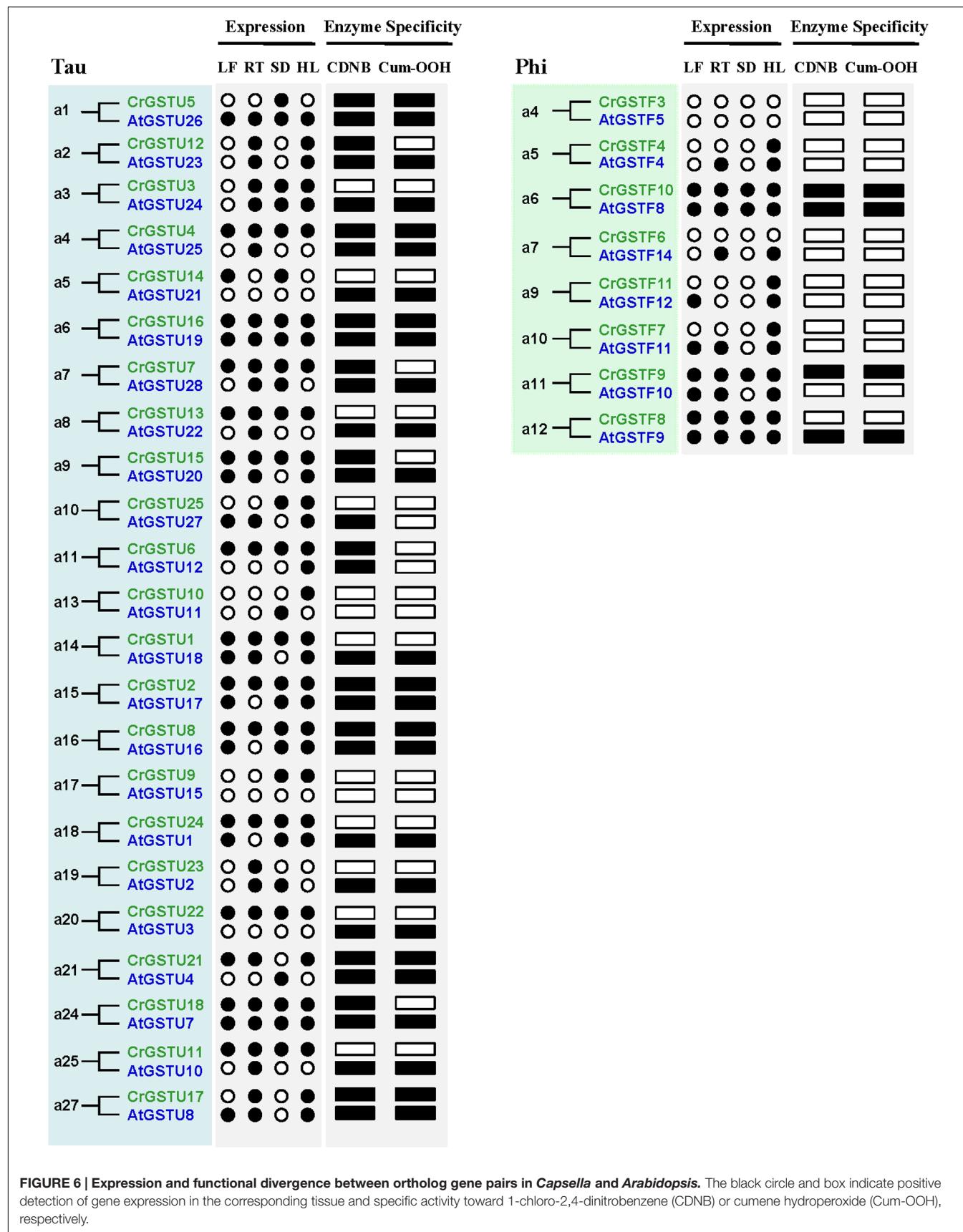
DISCUSSION

Functional divergence of duplicated genes is a major factor promoting their retention in the genome (Ohno, 1970; Zhang, 2003). Many theoretical models have been proposed to explain the mechanisms leading to the divergence include sub-functionalization, neo-functionalization, and dosage-balance model, etc (Ohno, 1970; Hughes, 1994; Force et al., 1999; Walsh, 2003; Moore and Purugganan, 2005; Veitia et al., 2008; Innan and Kondrashov, 2010). However, our understanding of the mechanisms driving the evolution of a large and functionally heterogeneous gene family is limited. Because to determine whether the duplicates have identical, similar, or different functions requires comprehensive examination of the functions of each gene product, while this approach is useful at a single gene level, genome-scale analyses of functional divergence of a supergene family are unfeasible. Our study combined bioinformatics and experimental approaches to explore the functional diversification of GST gene family at different levels of genomic organization: among subfamily classes, within tandem clusters, in paralogous and orthologous gene pairs.

Genome annotation identified 49 full-length GST genes from the *C. rubella* genome, which were divided into eight classes.

Extensive study has showed that tau and phi classes were the most abundant with wide interspecific variation in copy number in plants (Lan et al., 2009, 2013; Dixon and Edwards, 2010a; Jain et al., 2010; Liu et al., 2013, 2015). For instance, our study and previous studies showed that tau GSTs was not found in moss and green algae, whereas it's ubiquitous in tracheophytes (25–62 copies). Seventeen phi GSTs were found in rice while only two were represented in *S. moellendorffii*. However, other six classes remain comparatively small, with only 1–5 members. Comparison of copy numbers among the classes indicated that they might follow distinct evolutionary paths. Why did extensive expansion of tau and phi classes occur? A possible explanation is functional requirement. Tau and phi GSTs play an important role in the detoxification of xenobiotics and defense responses against both biotic and abiotic stresses (Loyall et al., 2000; Karavangeli et al., 2005; Benekos et al., 2010; Dixon et al., 2011; Jha et al., 2011; Cummins et al., 2013). Thus, the large scale expansion within tau and phi classes might provide defense against a broader range of xenobiotics and facilitated their tolerance to various environmental hazards. Our study exhibited extensive diversification in enzyme substrate specificity and transcript expression in tau and phi classes. This might further support the diversification in response to a set of changing substrates and regulatory properties.

The rapid expansion of GST gene family in plants is largely the result of the expansion of tau and phi classes. In the *C. rubella* genome, 17 of the 25 (68%) tau GSTs consisted of four clusters, and six of the 12 (50%) phi GSTs consisted of two clusters, indicating that tandem duplication considerably contributed to the expansion of tau and phi GSTs in the *C. rubella* genome. Previous studies also indicate that tandem duplication played important roles in the expansion of tau and phi GSTs in poplar, soybean, *Arabidopsis,* and rice genomes (Dixon et al., 2002b; Soranzo et al., 2004; Lan et al., 2009; Liu et al., 2015). Why have so many duplicate genes been retained for such a long time in the *C. rubella* genome? To investigate this question, we examined the seven tandem-arrayed clusters (Cluster I–VII). We found two categories of expression patterns. In the first, all the members in each cluster were expressed in all tissues. This pattern was observed in tau cluster II (*CrGSTU1/2*), phi cluster VI (*CrGSTF8/9*), and zeta cluster VII (*CrGSTZ2/3*; **Figure 1**). In the second category, found in tau cluster III (*CrGSTU3-5*), IV (*CrGSTU12-16*), V (*CrGSTU18-24*), and phi cluster I (*CrGSTF1-4*), some copies were expressed in all tissues, some had restricted tissue-specific expression or were not expressed in any tissue examined (**Figure 1B**). When enzyme assays were examined, no GST proteins in clusters showed identical enzymatic activities and specificities toward different substrates (**Figure 7**). Through this integrated approach, we found that rapid divergence has occurred in the regulatory regions of genes and in their biochemical properties within clusters, suggesting that partial sub-functionalization has indeed taken place. This seems to be an important factor promoting the duplicated GSTs' retention in the genome.

A major challenge in comparative genomics is to find sufficient functional differences between species. However, it remains challenging in *Arabidopsis* and other plants, partly due to

FIGURE 6 | Expression and functional divergence between ortholog gene pairs in *Capsella* and *Arabidopsis*. The black circle and box indicate positive detection of gene expression in the corresponding tissue and specific activity toward 1-chloro-2,4-dinitrobenzene (CDNB) or cumene hydroperoxide (Cum-OOH), respectively.

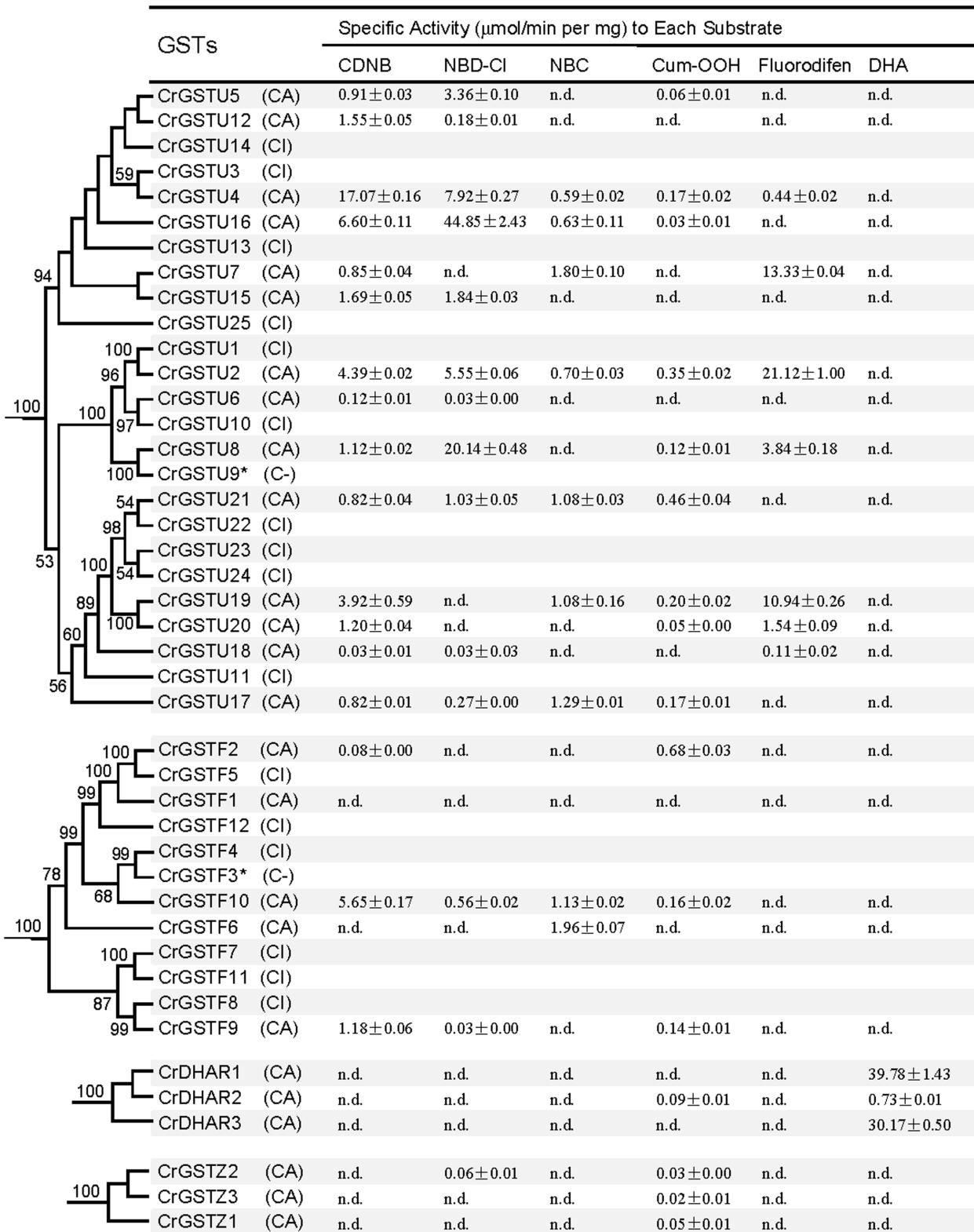| GSTs		Specific Activity (µmol/min per mg) to Each Substrate					
		CDNB	NBD-Cl	NBC	Cum-OOH	Fluorodifen	DHA
CrGSTU5	(CA)	0.91±0.03	3.36±0.10	n.d.	0.06±0.01	n.d.	n.d.
CrGSTU12	(CA)	1.55±0.05	0.18±0.01	n.d.	n.d.	n.d.	n.d.
CrGSTU14	(CI)						
CrGSTU3	(CI)						
CrGSTU4	(CA)	17.07±0.16	7.92±0.27	0.59±0.02	0.17±0.02	0.44±0.02	n.d.
CrGSTU16	(CA)	6.60±0.11	44.85±2.43	0.63±0.11	0.03±0.01	n.d.	n.d.
CrGSTU13	(CI)						
CrGSTU7	(CA)	0.85±0.04	n.d.	1.80±0.10	n.d.	13.33±0.04	n.d.
CrGSTU15	(CA)	1.69±0.05	1.84±0.03	n.d.	n.d.	n.d.	n.d.
CrGSTU25	(CI)						
CrGSTU1	(CI)						
CrGSTU2	(CA)	4.39±0.02	5.55±0.06	0.70±0.03	0.35±0.02	21.12±1.00	n.d.
CrGSTU6	(CA)	0.12±0.01	0.03±0.00	n.d.	n.d.	n.d.	n.d.
CrGSTU10	(CI)						
CrGSTU8	(CA)	1.12±0.02	20.14±0.48	n.d.	0.12±0.01	3.84±0.18	n.d.
CrGSTU9*	(C-)						
CrGSTU21	(CA)	0.82±0.04	1.03±0.05	1.08±0.03	0.46±0.04	n.d.	n.d.
CrGSTU22	(CI)						
CrGSTU23	(CI)						
CrGSTU24	(CI)						
CrGSTU19	(CA)	3.92±0.59	n.d.	1.08±0.16	0.20±0.02	10.94±0.26	n.d.
CrGSTU20	(CA)	1.20±0.04	n.d.	n.d.	0.05±0.00	1.54±0.09	n.d.
CrGSTU18	(CA)	0.03±0.01	0.03±0.03	n.d.	n.d.	0.11±0.02	n.d.
CrGSTU11	(CI)						
CrGSTU17	(CA)	0.82±0.01	0.27±0.00	1.29±0.01	0.17±0.01	n.d.	n.d.
CrGSTF2	(CA)	0.08±0.00	n.d.	n.d.	0.68±0.03	n.d.	n.d.
CrGSTF5	(CI)						
CrGSTF1	(CA)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
CrGSTF12	(CI)						
CrGSTF4	(CI)						
CrGSTF3*	(C-)						
CrGSTF10	(CA)	5.65±0.17	0.56±0.02	1.13±0.02	0.16±0.02	n.d.	n.d.
CrGSTF6	(CA)	n.d.	n.d.	1.96±0.07	n.d.	n.d.	n.d.
CrGSTF7	(CI)						
CrGSTF11	(CI)						
CrGSTF8	(CI)						
CrGSTF9	(CA)	1.18±0.06	0.03±0.00	n.d.	0.14±0.01	n.d.	n.d.
CrDHAR1	(CA)	n.d.	n.d.	n.d.	n.d.	n.d.	39.78±1.43
CrDHAR2	(CA)	n.d.	n.d.	n.d.	0.09±0.01	n.d.	0.73±0.01
CrDHAR3	(CA)	n.d.	n.d.	n.d.	n.d.	n.d.	30.17±0.50
CrGSTZ2	(CA)	n.d.	0.06±0.01	n.d.	0.03±0.00	n.d.	n.d.
CrGSTZ3	(CA)	n.d.	n.d.	n.d.	0.02±0.01	n.d.	n.d.
CrGSTZ1	(CA)	n.d.	n.d.	n.d.	0.05±0.01	n.d.	n.d.

FIGURE 7 | Specific activities of the *Capsella* tau, phi, dehydroascorbate reductase (DHAR), and zeta class GSTs toward six substrates (Mean ± SD obtained from at least three independent determinations). C, successfully cloned; A, purified GST protein assayed; I, recombinant protein totally insoluble; dash, analysis not performed; n.d., no activity detected.

technical limits and potential functional redundancy within the family (Sappl et al., 2009). We identified 23 tau and 8 phi orthologous GSTs in the two relatives, and most of the gene pairs exhibited variations at expression and biochemical level (**Figure 6**), indicating that their functions may have evolved after the split. For example, *AtGSTF12* showed high expression in senescent leaf and was demonstrated to be involved in flavonoid metabolism (Kitamura et al., 2004; Dixon and Edwards, 2010a). But its orthologous gene *CrGSTF11* was not detected in leaf tissue (**Figure 1B**). AtGSTU25 and AtGSTU28 have the highest activity in tau class when assayed with CDNB or Cum-OOH as substrates (Dixon et al., 2009), whereas their orthologs CrGSTU4 and CrGSTU7 have low activity for Cum-OOH and CDNB, respectively (**Figure 7**). *AtGSTU20* was showed to interact with *Far-red (FR) insensitive 219* (*FIN219*) in response to light and could regulate cell elongation and plant development (Chen et al., 2007). We detected some light responsive elements in the promoter of *CrGSTU15* (Supplementary Table S6), suggesting that *CrGSTU15* may also involve in light regulation. In addition, 5 of the 31 orthologs displayed similar patterns (**Figure 6**). *AtGSTU19* and *CrGSTU16* displayed similar expression and substrate spectrum. *AtGSTU19* showed tolerance to salt, drought, and methyl viologen stresses (Xu et al., 2015). Several *cis*-acting elements involved in abscisic acid, anaerobic, heat, low-temperature, drought, defense, and phytohormones responsiveness were indentified in *CrGSTU16* as well, suggesting that this gene may also be induced by several stimuli. However, *AtGSTF8* and *CrGSTF10* exhibited a different example: As an enzyme, *AtGSTF8* was the most active member in phi class when assayed with CDNB and Cum-OOH (Dixon et al., 2009). Its expression was strongly induced by salicylic acid and H_2O_2 in root tissue, and *ocs* element in the promoter region played an important role (Chen and Singh, 1999). Unlike *AtGSTF8*, *CrGSTF10* didn't contain *ocs*-like element and its specific activity toward Cum-OOH was low (Supplementary Table S6 and **Figure 7**). Protein subcellular relocalization is also considered as another form of functional divergence (Marques et al., 2008; Qian and Zhang, 2009; Wang et al., 2009). *AtGSTU12* is the only tau class GST localized entirely to the nucleus, containing a putative nuclear localization signal KKRKK (Takahashi et al., 1995). But we did not find this signal in *CrGSTU6*. All these results suggested that functional divergence previously had occurred in the two lineages after the split.

CONCLUSION

In this study, we characterized the complete set of GST gene family in *C. rubella* genome. By phylogenetic and functional analysis, we compared it to that in *Arabidopsis*. We examined the gene gain and loss events after the divergence of the relatives. Also, we evaluated the functional divergences of recently expanded GSTs and orthologs. Through these analyses, we were able to draw a picture illustrating how gene duplication and sub-functionalization influence the divergence, retention, and functions of GST genes in the *Capsella* genome. Furthermore, by extending genome-wide comparison analysis of GST gene family with more species in the Brassicaceae, the study will provide a comprehensive overview of the evolutionary history of a large gene family among lineages and mechanism of functional diversification and retention of duplicates.

AUTHOR CONTRIBUTIONS

TL and QZ designed the study. GH, CG, and QC performed the experiments. TL, XG, and WL analyzed the data. TL wrote the paper.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01325

REFERENCES

Acarkan, A., Rossberg, M., Koch, M., and Schmidt, R. (2000). Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*. *Plant J.* 23, 55–62. doi: 10.1046/j.1365-313x.2000.00790.x

Basantani, M., and Srivastava, A. (2007). Plant glutathione transferases - a decade falls short. *Can J. Bot.* 85, 443–456. doi: 10.1139/B07-033

Benekos, K., Kissoudis, C., Nianiou-Obeidat, I., Labrou, N., Madesis, P., Kalamaki, M., et al. (2010). Overexpression of a specific soybean GmGSTU4 isoenzyme improves diphenyl ether and chloroacetanilide herbicide tolerance of transgenic tobacco plants. *J. Biotechnol.* 150, 195–201. doi: 10.1016/j.jbiotec.2010.07.011

Boivin, K., Acarkan, A., Mbulu, R. S., Clarenz, O., and Schmidt, R. (2004). The *Arabidopsis* genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the *Arabidopsis* and *Capsella rubella* genomes. *Plant Physiol.* 135, 735–744. doi: 10.1104/pp.104.040030

Chen, I. C., Huang, I. C., Liu, M. J., Wang, Z. G., Chung, S. S., and Hsieh, H. L. (2007). Glutathione S-transferase interacting with far-red insensitive 219 is involved in phytochrome A-mediated signaling in *Arabidopsis*. *Plant Physiol.* 143, 1189–1202. doi: 10.1104/pp.106.094185

Chen, W., and Singh, K. B. (1999). The auxin, hydrogen peroxide and salicylic acid induced expression of the *Arabidopsis* GST6 promoter is mediated in part by an ocs element. *Plant J.* 19, 667–677. doi: 10.1046/j.1365-313x.1999.00560.x

Chen, Z., and Gallie, D. R. (2006). Dehydroascorbate reductase affects leaf growth, development, and function. *Plant Physiol.* 142, 775–787. doi: 10.1104/pp.106.085506

Cummins, I., Cole, D. J., and Edwards, R. (1999). A role for glutathione transferases functioning as glutathione peroxidases in resistance to multiple herbicides in black-grass. *Plant J.* 18, 285–292. doi: 10.1046/j.1365-313X.1999.00452.x

Cummins, I., Dixon, D. P., Freitag-Pohl, S., Skipsey, M., and Edwards, R. (2011). Multiple roles for plant glutathione transferases in xenobiotic detoxification. *Drug Metab. Rev.* 43, 266–280. doi: 10.3109/03602532.2011.552910

Cummins, I., Wortley, D. J., Sabbadin, F., He, Z., Coxon, C. R., Straker, H. E., et al. (2013). Key role for a glutathione transferase in multiple-herbicide resistance in grass weeds. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5812–5817. doi: 10.1073/pnas.1221179110

Dixon, D. P., Davis, B. G., and Edwards, R. (2002a). Functional divergence in the glutathione transferase superfamily in plants. Identification of two classes with putative functions in redox homeostasis in *Arabidopsis thaliana*. *J. Biol. Chem.* 277, 30859–30869.

Dixon, D. P., and Edwards, R. (2010a). Glutathione transferases. *Arabidopsis Book* 8:e0131. doi: 10.1199/tab.0131

Dixon, D. P., and Edwards, R. (2010b). Roles for stress-inducible lambda glutathione transferases in flavonoid metabolism in plants as identified by ligand fishing. *J. Biol. Chem.* 285, 36322–36329. doi: 10.1074/jbc.M110.164806

Dixon, D. P., Lapthorn, A., and Edwards, R. (2002b). Plant glutathione transferases. *Genome Biol.* 3, 1–10. doi: 10.1186/gb-2002-3-3-reviews3004

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. doi: 10.1093/jxb/ern365

Dixon, D. P., Skipsey, M., and Edwards, R. (2010). Roles for glutathione transferases in plant secondary metabolism. *Phytochemistry* 71, 338–350. doi: 10.1016/j.phytochem.2009.12.012

Dixon, D. P., Steel, P. G., and Edwards, R. (2011). Roles for glutathione transferases in antioxidant recycling. *Plant Signal. Behav.* 6, 1223–1227. doi: 10.4161/psb.6.8.16253

Edwards, R., and Dixon, D. P. (2005). Plant glutathione transferases. *Methods Enzymol.* 401, 169–186. doi: 10.1016/S0076-6879(05)01011-6

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545.

Frova, C. (2003). The plant glutathione transferase gene family: genomic structure, functions, expression and evolution. *Physiol. Plant.* 119, 469–479. doi: 10.1046/j.1399-3054.2003.00183.x

Frova, C. (2006). Glutathione transferases in the genomics era: new insights and perspectives. *Biomol. Eng.* 23, 149–169. doi: 10.1016/j.bioeng.2006.05.020

Goda, H., Sasaki, E., Akiyama, K., Maruyama-Nakashita, A., Nakabayashi, K., Li, W., et al. (2008). The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. *Plant J.* 55, 526–542. doi: 10.1111/j.0960-7412.2008.03510.x

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704. doi: 10.1080/10635150390235520

Guo, Y. L., Bechsgaard, J. S., Slotte, T., Neuffer, B., Lascoux, M., Weigel, D., et al. (2009). Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5246–5251. doi: 10.1073/pnas.0808012106

Guo, Y. L., Todesco, M., Hagmann, J., Das, S., and Weigel, D. (2012). Independent FLC Mutations as Causes of Flowering-Time Variation in *Arabidopsis thaliana* and *Capsella rubella*. *Genetics* 192, 729–739. doi: 10.1534/genetics.112.143958

Habig, W. H., Pabst, M. J., and Jakoby, W. B. (1974). Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. *J. Biol. Chem.* 249, 7130–7139.

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

Hayes, J. D., Flanagan, J. U., and Jowsey, I. R. (2005). Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* 45, 51–88. doi: 10.1146/annurev.pharmtox.45.120403.095857

Holm, P. J., Bhakat, P., Jegerschold, C., Gyobu, N., Mitsuoka, K., Fujiyoshi, Y., et al. (2006). Structural basis for detoxification and oxidative stress protection in membranes. *J. Mol. Biol.* 360, 934–945. doi: 10.1016/j.jmb.2006.05.056

Hughes, A. L. (1994). The evolution of functionally novel proteins after gene duplication. *Proc. Biol. Sci.* 256, 119–124. doi: 10.1098/rspb.1994.0058

Innan, H., and Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* 11, 97–108. doi: 10.1038/nrg2689

Jain, M., Ghanashyam, C., and Bhattacharjee, A. (2010). Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione S-transferase genes during development and stress responses. *BMC Genomics* 11:73. doi: 10.1186/1471-2164-11-73

Jha, B., Sharma, A., and Mishra, A. (2011). Expression of SbGSTU (tau class glutathione S-transferase) gene isolated from *Salicornia* brachiata in tobacco for salt tolerance. *Mol. Biol. Rep.* 38, 4823–4832. doi: 10.1007/s11033-010-0625-x

Karavangeli, M., Labrou, N. E., Clonis, Y. D., and Tsaftaris, A. (2005). Development of transgenic tobacco plants overexpressing maize glutathione S-transferase I for chloroacetanilide herbicides phytoremediation. *Biomol. Eng.* 22, 121–128. doi: 10.1016/j.bioeng.2005.03.001

Kitamura, S., Shikazono, N., and Tanaka, A. (2004). TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. *Plant J.* 37, 104–114. doi: 10.1046/j.1365-313X.2003.01943.x

Koch, M. A., and Kiefer, M. (2005). Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of three diploid species– *Capsella rubella*, *Arabidopsis lyrata* subsp. petraea, and A. thaliana. *Am. J. Bot.* 92, 761–767. doi: 10.3732/ajb.92.4.761

Kram, B. W., Xu, W. W., and Carter, C. J. (2009). Uncovering the *Arabidopsis thaliana* nectary transcriptome: investigation of differential gene expression in floral nectarious tissues. *BMC Plant Biol.* 9:92. doi: 10.1186/1471-2229-9-92

Kwon, S. Y., Choi, S. M., Ahn, Y. O., Lee, H. S., Lee, H. B., Park, Y. M., et al. (2003). Enhanced stress-tolerance of transgenic tobacco plants expressing a human dehydroascorbate reductase gene. *J. Plant Physiol.* 160, 347–353. doi: 10.1078/0176-1617-00926

Lan, T., Wang, X. R., and Zeng, Q. Y. (2013). Structural and functional evolution of positively selected sites in pine glutathione S-transferase enzyme family. *J. Biol. Chem.* 288, 24441–24451. doi: 10.1074/jbc.M113.456863

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. doi: 10.1105/tpc.109.070219

Lieberherr, D., Wagner, U., Dubuis, P. H., Metraux, J. P., and Mauch, F. (2003). The rapid induction of glutathione S-transferases AtGSTF2 and AtGSTF6 by avirulent *Pseudomonas syringae* is the result of combined salicylic acid and ethylene signaling. *Plant Cell Physiol.* 44, 750–757. doi: 10.1093/pcp/pcg093

Liu, H. J., Tang, Z. X., Han, X. M., Yang, Z. L., Zhang, F. M., Yang, H. L., et al. (2015). Divergence in enzymatic activities in the soybean gst supergene family provides new insight into the evolutionary dynamics of whole-genome duplicates. *Mol. Biol. Evol.* 32, 2844–2859. doi: 10.1093/molbev/msv156

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. doi: 10.1104/pp.112.205815

Loyall, L., Uchida, K., Braun, S., Furuya, M., and Frohnmeyer, H. (2000). Glutathione and a UV light-induced glutathione S-transferase are involved in signaling to chalcone synthase in cell cultures. *Plant Cell* 12, 1939–1950. doi: 10.2307/3871204

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39, D225–D229. doi: 10.1093/nar/gkq1189

Marques, A. C., Vinckenbosch, N., Brawand, D., and Kaessmann, H. (2008). Functional diversification of duplicate genes through subcellular adaptation of encoded proteins. *Genome Biol.* 9, R54. doi: 10.1186/gb-2008-9-3-r54

Marrs, K. A. (1996). The functions and regulation of glutathione s-transferases in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 127–158. doi: 10.1146/annurev.arplant.47.1.127

Meyers, B. C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R. W. (2003). Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* 15, 809–834. doi: 10.1105/tpc.009308

Moore, R. C., and Purugganan, M. D. (2005). The evolutionary dynamics of plant duplicate genes. *Curr. Opin. Plant Biol.* 8, 122–128. doi: 10.1016/j.pbi.2004.12.001

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Nix, D. A., and Eisen, M. B. (2005). GATA: a graphic alignment tool for comparative sequence analysis. *BMC Bioinformatics* 6:9. doi: 10.1186/1471-2105-6-9

Oakley, A. J. (2005). Glutathione transferases: new functions. *Curr. Opin. Struct. Biol.* 15, 716–723. doi: 10.1016/j.sbi.2005.10.005

Ohno, S. (1970). *Evolution by gene duplication*. Heidelberg: Springer-Verlag.

Qian, W., and Zhang, J. (2009). Protein subcellular relocalization in the evolution of yeast singleton and duplicate genes. *Genome Biol. Evol.* 1, 198–204. doi: 10.1093/gbe/evp021

Ricci, G., Caccuri, A. M., Lo Bello, M., Pastore, A., Piemonte, F., and Federici, G. (1994). Colorimetric and fluorometric assays of glutathione transferase based on 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole. *Anal. Biochem.* 218, 463–465. doi: 10.1006/abio.1994.1209

Roxas, V. P., Smith, R. K. Jr., Allen, E. R., and Allen, R. D. (1997). Overexpression of glutathione S-transferase/glutathione peroxidase enhances the growth of transgenic tobacco seedlings during stress. *Nat. Biotechnol.* 15, 988–991. doi: 10.1038/nbt1097-988

Sappl, P. G., Carroll, A. J., Clifton, R., Lister, R., Whelan, J., Harvey Millar, A., et al. (2009). The *Arabidopsis* glutathione transferase gene family displays complex stress regulation and co-silencing multiple genes results in altered metabolic sensitivity to oxidative stress. *Plant J.* 58, 53–68. doi: 10.1111/j.1365-313X.2008.03761.x

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., et al. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* 37, 501–506. doi: 10.1038/ng1543

Sharma, R., Sahoo, A., Devendran, R., and Jain, M. (2014). Over-expression of a rice tau class glutathione s-transferase gene improves tolerance to salinity and oxidative stresses in *Arabidopsis*. *PLoS ONE* 9:e92900. doi: 10.1371/journal.pone.0092900

Slotte, T., Hazzouri, K. M., Agren, J. A., Koenig, D., Maumus, F., Guo, Y. L., et al. (2013). The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. *Nat. Genet.* 45, 831–835. doi: 10.1038/ng.2669

Soranzo, N., Sari Gorla, M., Mizzi, L., De Toma, G., and Frova, C. (2004). Organisation and structural evolution of the rice glutathione S-transferase gene family. *Mol. Genet. Genomics* 271, 511–521. doi: 10.1007/s00438-004-1006-8

Takahashi, Y., Hasezawa, S., Kusaba, M., and Nagata, T. (1995). Expression of the auxin-regulated parA gene in transgenic tobacco and nuclear localization of its gene products. *Planta* 196, 111–117. doi: 10.1007/BF00193224

Thom, R., Dixon, D. P., Edwards, R., Cole, D. J., and Lapthorn, A. J. (2001). The structure of a zeta class glutathione S-transferase from *Arabidopsis thaliana*: characterisation of a GST with novel active-site architecture and a putative role in tyrosine catabolism. *J. Mol. Biol.* 308, 949–962. doi: 10.1006/jmbi.2001.4638

Ushimaru, T., Nakagawa, T., Fujioka, Y., Daicho, K., Naito, M., Yamauchi, Y., et al. (2006). Transgenic *Arabidopsis* plants expressing the rice dehydroascorbate reductase gene are resistant to salt stress. *J. Plant Physiol.* 163, 1179–1184. doi: 10.1016/j.jplph.2005.10.002

Veitia, R. A., Bottani, S., and Birchler, J. A. (2008). Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. *Trends Genet.* 24, 390–397. doi: 10.1016/j.tig.2008.05.005

Vickers, T. J., Wyllie, S., and Fairlamb, A. H. (2004). Leishmania major elongation factor 1B complex has trypanothione S-transferase and peroxidase activity. *J. Biol. Chem.* 279, 49003–49009. doi: 10.1074/jbc.M311039200

Walsh, B. (2003). Population-genetic models of the fates of duplicate genes. *Genetica* 118, 279–294. doi: 10.1023/A:1024194802441

Wang, X., Huang, Y., Lavrov, D. V., and Gu, X. (2009). Comparative study of human mitochondrial proteome reveals extensive protein subcellular relocalization after gene duplications. *BMC Evol. Biol.* 9:275. doi: 10.1186/1471-2148-9-275

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2:e718. doi: 10.1371/journal.pone.0000718

Xu, J., Tian, Y. S., Xing, X. J., Peng, R. H., Zhu, B., Gao, J. J., et al. (2015). Over-expression of AtGSTU19 provides tolerance to salt, drought and methyl viologen stresses in *Arabidopsis*. *Physiol. Plant.* 156, 164–175. doi: 10.1111/ppl.12347

Zhang, J. Z. (2003). Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298. doi: 10.1016/s0169-5347(03)00033-8