식물생명공학회지 제35권 제4호
J Plant Biotechnol
Vol. 35, No. 4, 265-274 (2008)

Genetic mapping and sequence analysis of Phi class Glutathione S-transferases (*BrGSTFs*) candidates from *Brassica rapa*

Tae-Ho Park[1,2,3], Mina Jin[1], Sang-Choon Lee[1], Joon Ki Hong[1], Jung Sun Kim[1], Jin A Kim[1], Soo-Jin Kwon[1], Yun-Xiang Zang[1], Young-Doo Park[2], and Beom-Seok Park[1*]

[1]*Genomics Division, National Academy of Agricultural Science, Suwon 441-707, Korea*

[2]*Department of Horticultural Biotechnology, College of Life Sciences, Kyung Hee University, Yongin 446-701, Korea*

[3]*Microbial Safety Division, National Academy of Agricultural Science, Suwon 441-707, Korea*

ABSTRACT Glutathione S-transferases (*GSTs*) are multifunctional proteins encoded by a large gene family divided into Phi, Tau, Theta, Zeta, Lambda and DHAR classes on the basis of sequence identity. The Phi(F) and Tau(U) classes are plant-specific and ubiquitous. Their roles have been defined as herbicide detoxification and responses to biotic and abiotic stresses. Fifty-two members of the *GST* super-family were identified in the *Arabidopsis thaliana* genome, 13 members of which belong to the Phi class of *GSTs* (*AtGSTFs*). Based on the sequence similarities of *AtGSTFs*, 11 BAC clones were identified from *Brassica rapa*. Seven unique sequences of ORFs designated the Phi class candidates of *GST* derived from *B. rapa* (*BrGSTFs*) were detected from these 11 BAC clones by blast search and sequence alignment. Some of *BrGSTFs* were present in the same BAC clones indicating that *BrGSTFs* could also be clustered as usual in plant. They were mapped on *B. rapa* linkage group 2, 3, 9 and 10 and their nucleotide and amino acid sequences were highly similar to those of *AtGSTFs*. In addition, *in silico* analysis of *BrGSTFs* using Korea Brassica Genome Project 24K oligochip and microarray database for cold, salt and drought stresses revealed 15 unigenes to be highly similar to *AtGSTFs* and six of these were identical to one of *BrGSTFs* identified in the BAC clones indicating their expression. The sequences of *BrGSTFs* and unigenes identified in this study will facilitate further studies to apply *GST* genes to medical and agriculture purposes.

Introduction

Glutathione S-transferases (*GSTs*) are a superfamily of multifunctional proteins distributed in highly diverse organisms such as bacteria, plants, animals and human. In plants, functions of *GSTs* are associated with a wide range of biotic and abiotic stresses including herbicide, organic pollutants, natural toxins and diseases with regards to detoxification and environmental safety (Hatton et al. 1996; Frova 2003). In addition to these functions, *GSTs* catalyze the reaction with glutathione and isothiocyanate (ITC) including sulforaphane (1-isothiocyanato-4-methylsulfinyl butane) to form ITC-glutathione in cruciferous vegetables (Rea 1999). Sulforaphane is one of the major functional ITCs which is known as anticancer compound and derived from the 4-methylsulfinybutyl glucosinolate that accumulates in Brassicaceae crops, especially Broccoli (*B. oleracea*) (Ambrosone et al. 2004; Gasper et al. 2005).

Typically *GSTs* are encoded by large gene families. Plant *GSTs* consist of six classes: Tau (U), Phi (F), Theta (T), Lambda (L), Zeta (Z) and DHAR classified based on sequence identity, gene organization and active site residues in the

*Corresponding author Tel 031-299-1671 Fax 031-299-1672
E-mail: pbeom@rad.go.kr

proteins. The fully sequenced *A. thaliana* and rice (*Oryza sativa*) genome revealed insight of *GSTs* in a dicot and a monocot species. In total, 52 *Arabidopsis GSTs* (AtGSTs) including 28 Tau, 13 Phi, 3 Theta, 2 Lambda, 2 Zeta and 4 DHAR were identified (Dixon et al. 2002a, 2002b; Wagner et al. 2002) and the analysis by an *in silico* approach led to the identification of 61 rice *GSTs* (*OsGSTs*) including 39 Tau, 16 Phi, 2 Theta, 1 Lambda and 3 Zeta (Soranzo et al. 2004). The classification of all groups of *GSTs* in both *Arabidopsis* and rice were well-studied by analyzing sequence similarity (Dixon et al. 2002b; Wagner et al. 2002; Frova 2003). Of all classes of *GSTs*, the Phi and Tau classes are plant specific (Dixon et al. 2002b).

As a part of Brassica Genome Sequencing Project, we have database of sequenced BAC and ESTs clones, Korea Brassica Genome Project 24K oligochip (KBGP-24K) and microarray database (http://www.brassica-rapa.org/BGP/NC_brgp.jsp; Lee et al. 2008). In this study, we tried specifically to identify Phi class *GSTs* from Chinese cabbage (*B. rapa*) (*BrGSTFs*) using our database. We mapped the identified *BrGSTFs* and compared the sequences of *BrGSTFs* with those of *AtGSTFs*. In addition, *in silico* analysis of gene expression was performed using KBGP-24K and microarray database.

Materials and methods

Search for Phi class of *Brassica rapa* Glutathione S-transferases (*BrGSTFs*)

Initially, search for the Phi class of *A. thaliana* Glutathione S-transferases (*AtGSTFs*) was performed using public genome database. Sequences of *AtGSTFs* were obtained from The *Arabidopsis* Information Resource (TAIR) database (http://www.Arabidopsis.org). In order to search for sequences of the Phi class of *B. rapa* Glutathione S-transferases (*BrGSTFs*), the sequenced BAC clones originated from three different large-insert BAC libraries, KBrH (*Hin*dIII), KBrB (*Bam*HI) and KBrS (*Sau*3AI) constructed for full sequencing of *B. rapa* ssp. *pekinensis* cv. Chiifu (http://www.brassica-rapa.org/BGP/; Park et al. 2005; Yang et al. 2005) were used. 'tblastx' that searches translated nucleotide database (the sequenced BAC clones) using a translated nucleotide query (*AtGSTFs*) was performed at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and our own database.

Genetic positioning of *BrGSTs*

The BAC clones harboring *BrGSTFs* were genetically localized on two mapping populations, JWF3p (Kim et al. 2006) and VCS3-DH (Jin et al. unpublished). The mapping positions of *BrGSTFs* were compared with those of *AtGSTFs* confirmed at TAIR database (http://www.Arabidopsis.org).

Sequence comparison between *AtGSTFs* and *BrGSTFs*

The sequences of the selected BAC clones were analyzed using web-based gene prediction programs: GeneMark (http://opal.biology.gatech.edu/GeneMark/; Lukashin and Borodovsky 1998) and Genescan (http://genes.mit.edu/GENSCAN.html; Burge and Karlin 1997). The sequences of *BrGSTFs* were dissected from the BAC clones and compared with those of *AtGSTFs*. Sequence alignment and phylogenetic analysis were performed using AlignX in Vector NTI suite 9 (Invitrogen, Carlsbad). Sequence colinearity was also investigated using a web-based software, PipMaker (Schwartz et al. 2000).

Digital northern analysis of *BrGSTFs*

Digital northern analysis of *BrGSTFs* was conducted to have an idea of *BrGSTFs* expression *in silico* using KBGP- 24K and microarray database for cold, salt and drought stresses (http://www.brassica-rapa.org/BGP/NC_brgp.jsp). Initially, 23,939 unigenes presented in KBGP-24K were blasted to the *Arabidopsis* sequences and all unigenes were assigned to the loci of *Arabidopsis* (Lee et al. 2008). Therefore the loci of *AtGSTFs* which are similar to *BrGSTFs* and candidate unigenes could be selected. These selected unigenes were analyzed in the microarray database generated for cold, salt and drought stresses. The expression patterns of the candidate *BrGSTFs* were generated using an MeV software (http://www.tm2.org/mev.html).

Results

Collection of *AtGSTFs* and *BrGSTFs* sequences

To select BAC clones harboring the candidate *BrGSTFs*, 13 *AtGSTFs* (from *AtGSTF2* to *AtGSTF14*) were collected from TAIR database (http://www.Arabidopsis.org) (Table 1). The sequences of *AtGSTFs* were blasted to the sequenced BAC clones derived from *B. rapa* ssp. *pekinensis* cv. Chiifu using 'tblastx' at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and our own database. 11 BAC clones containing full length of open reading frames (ORFs) which were highly similar with any of *AtGSTFs*' sequences and their accession number registered in the NCBI database (http://www.ncbi.nlm.nih.gov/) were selected as shown in Table 2. Seven, two and two BAC clones were correspondent to *AtGSTF3*, *AtGSTF7* and *AtGSTF12*, respectively. All selected BAC clones were annotated using GeneMark (http://opal.biology.gatech.edu/GeneMark/; Lukashin and Borodovsky 1998) and Genescan (http://genes.mit.edu/GENSCANhtml; Burge and Karlin 1997) and predicted *BrGSTF* genes were dissected from the BAC clones for further analysis. Six and five BAC clones contain two and one copies of *GSTFs*, respectively. Of these, KBrB018D16, KBrH007D13 and KBrH 029J18 presented two identical copies highly similar to *AtGSTF3*, which were designated *BrGSTF3*_c1 & *BrGSTF3*_c2. Other three BAC clones for *AtGSTF3* (KBrB027P21, KBrB 048D13 and KBrH129I15) also consist of two identical copies designated *BrGSTF3*_c3 and *BrGSTF3*_c4. One more BAC clone (KBrH028K09) has one copy for *AtGSTF3* that was designated *BrGSTF3*_c5. For *AtGSTF7*, two BAC clones (KBrB030F10 and KBrH061D04) contain one identical copy each, and then it was designated *BrGSTF7*_c. For *AtGSTF12*, two BAC clones (KBrB012M04 and KBrH015C19) also contain one identical copy each, and then it was designated *BrGSTF12*_c.

Table 1 List of the *AtGSTFs* and the number of unigenes selected from *B. rapa* database. This table was partially adopted from Wagner et al. (2002)

Name	At locus ID	Number of unigenes[a]	Old name[b]	Reference
AtGSTF2	At4g02520	1	*GST2*	Zhou & Goldsborough 1993
			PM24.1	Zettl et al. 1994
AtGSTF3	At2g02930	4	*GST16*	unpublished
AtGSTF4	At1g02950	1	*GST31*	unpublished
AtGSTF5	At1g02940	0	-	unpublished
AtGSTF6	At1g02930	0	ERD11	Kiyosue et al. 1993
			GST1	unpublished
			AtGSTF3	Edwards et al. 2000
AtGSTF7	At1g02920	1	*GST11*	Yang et al. 1998
			AtGSTF8	Edwards et al. 2000
AtGSTF8	At2g47730	2	*GST6*	Chen et al. 1996
			AtGSTF5	Edwards et al. 2000
AtGSTF9	At2g30860	3	GLUTTR	unpublished
			AtGSTF7	Edwards et al. 2000
AtGSTF10	At2g30870	1	ERD13	Kiyosue et al. 1993
			AtGSTF4	Edwards et al. 2000
AtGSTF11	At3g03190	0	*AtGSTF6*	Edwards et al. 2000
AtGSTF12	At5g17220	1	*GST26*	unpublished
AtGSTF13	At3g67260	0	-	unpublished
AtGSTF14	At1g49860	1	-	unpublished

[a] Number of unigenes detected in microarray database generated for cold, salt and drought stresses
[b] A few names have been changed due to the new *AtGST* nomenclature suggested by Edwards et al. (2000)

Table 2 List of the selected BAC clones and candidate *BrGSTFs*

Name of *AtGST*	BAC clones	Accession number	Copy number	Name of *BrGST*
AtGSTF3	KBrB018D16	AC232454	2	*BrGSTF3_c1* & *BrGSTF3_c2*
	KBrH007D13	AC232397	2	*BrGSTF3_c1* & *BrGSTF3_c2*
	KBrH029J18	AC232400	2	*BrGSTF3_c1* & *BrGSTF3_c2*
	KBrB027P21	AC232474	2	*BrGSTF3_c3* & *BrGSTF3_c4*
	KBrB048D13	AC232395	2	*BrGSTF3_c3* & *BrGSTF3_c4*
	KBrH129I15	AC232570	2	*BrGSTF3_c3* & *BrGSTF3_c4*
	KBrH028K09	AC232399	1	*BrGSTF3_c5*
AtGSTF7	KBrB030F10	AC189302	1	*BrGSTF7_c*
	KBrH061D04	AC232401	1	*BrGSTF7_c*
AtGSTF12	KBrB012M04	AC232394	1	*BrGSTF12_c*
	KBrH015C19	AC189602	1	*BrGSTF12_c*

The list of the selected BAC clones with corresponding *AtGSTFs*, copy number and designated names of *BrGSTFs* are presented in Table 2.

Genetic positioning of *BrGSTs*

From the selected BAC clones, seven candidate *BrGSTFs* were genetically positioned on two different mapping populations and relative physical positions are shown in Figure 1B. The map positions of all *AtGSTFs* were also indicated on five *Arabidopsis* chromosomes (Figure 1A). *BrGSTF3_c1* and *BrGSTF3_c2* are located on the telomeric long arm of *B. rapa* linkage group R9. *BrGSTF3_c3* and *BrGSTF3_c4* are located on the long arm of R3. *BrGSTF3_c5* and *BrGSTF12_c* are positioned on the telomeric long and short arm of R2, respectively and *BrGSTF7_c* on the telomeric long arm of R10.

Sequence comparison of *GSTFs* from *Arabidopsis* and *Brassica rapa*

Phylogenetic tree was constructed with amino acid sequences of 13 *AtGSTFs* and seven candidate *BrGSTFs* using the neighbor-joining method (Saitou and Nei 1987) in Vector NTI suite 9 (Invitrogen, Carlsbad) as shown in Figure 2. It was built on a matrix of distances between all pairs of sequences to be analyzed. These distances are related to the degree of divergence between the sequences. In the phylogenetic tree, *BrGSTF12_c* is closest to *AtGSTF12*, *BrGSTF7_c* to *AtGSTF6* or *AtGSTF7* and *BrGSTF3_c1-c5* to *AtGSTF2* or *AtGSTF3*.

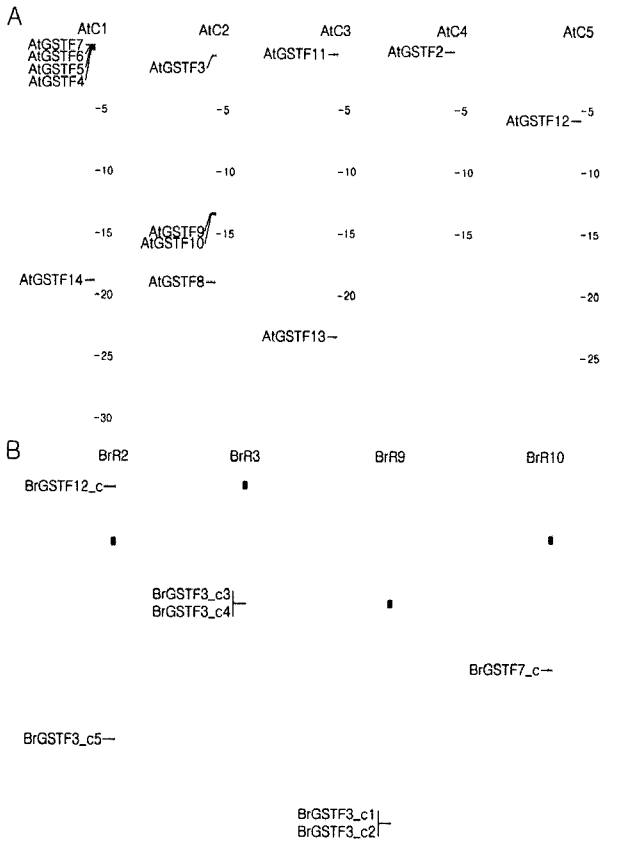

Figure 1. Chromosomal distribution of *GSTFs*. A: *Arabidopsis*. The length (Mbp) of each chromosome is indicated on the right of the gray bars. The positions of all *AtGSTFs* were adopted from TAIR database (http://www.Arabidopsis.org). B: *B. rapa*. The positions of centromeres are represented by the black boxes.

According to the result of the phylogenetic tree, three groups of *GSTFs* were further investigated separately. The amino acid sequences of *GSTFs* in three groups were aligned

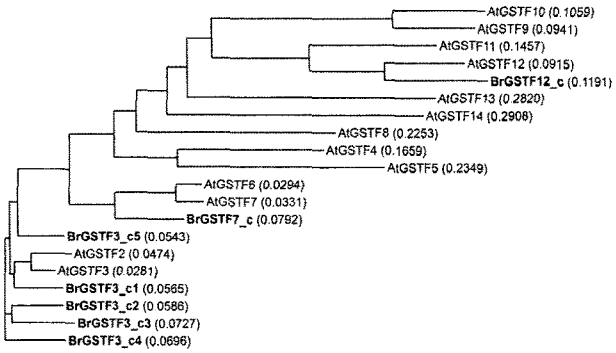

Figure 2. Phylogenetic tree of *GSTFs* identified in *A. thaliana* and *B. rapa*. The calculated distance values in parentheses following the name of *GSTFs* are displayed.

(Figure 3) and pairwise amino acid and nucleotide sequence similarity in each group was examined (Figure 4) using AlignX in Vector NTI suite 9 (Invitrogen, Carlsbad). The first group designated *GSTF3* includes *AtGSTF2*, *AtGSTF3* and *BrGSTF3*_c1-c5 (Figure 3A and Figure 4A), the second group designated *GSTF7* includes *AtGSTF6*, *AtGSTF7* and *BrGSTF7*_c (Figure 3B and Figure 4B) and the third group designated *GSTF*12 includes *AtGSTF12* and *BrGSTF12*_c (Figure 3C and Figure 4C). The amino acid sequences of all At/*BrGSTFs* in each group were highly conserved as shown in Figure 3 and it was confirmed that the candidate *BrGSTFs* were the most likely to be close to *AtGSTF3*, *AtGSTF7* and *AtGSTF12* in the group of *GSTF3*, *GSTF7* and *GSTF12*, respectively. In the *GSTF3* group, the pairwise similarity of the amino acid and nucleotide sequences between *AtGSTF3* and *BrGSTF3*_c or among *BrGSTF3*_c ranged from 78% to 90% and from 83% to 91%, respectively. In the *GSTF7* and *GSTF12* groups, they were 81% and 86% between *AtGSTF7* and *BrGSTF7*_c and 76% and 83% between *AtGSTF12* and *BrGSTF12*_c, respectively. Sequence similarity and colinearity within three groups were confirmed using PipMaker (results not shown). Similarity within *BrGSTFs* ranged from 51% to 91% at the amino acid level.

Digital northern analysis of *BrGSTFs*

KBGP-24K oligochip consisting of 23,929 unigenes and microarray database for cold, salt and drought stress (http://www.brassica-rapa.org/BGP/NC_brgp.jsp) were used for digital northern analysis of *BrGSTFs*. All the sequences of unigenes were assigned to the locus of *A. thaliana* based on the best matches at the sequence blast. Of 23,929 unigenes on the KBGP-24K oligochip, 15 were detected to be similar to one of nine *AtGSTFs* (Table 1).

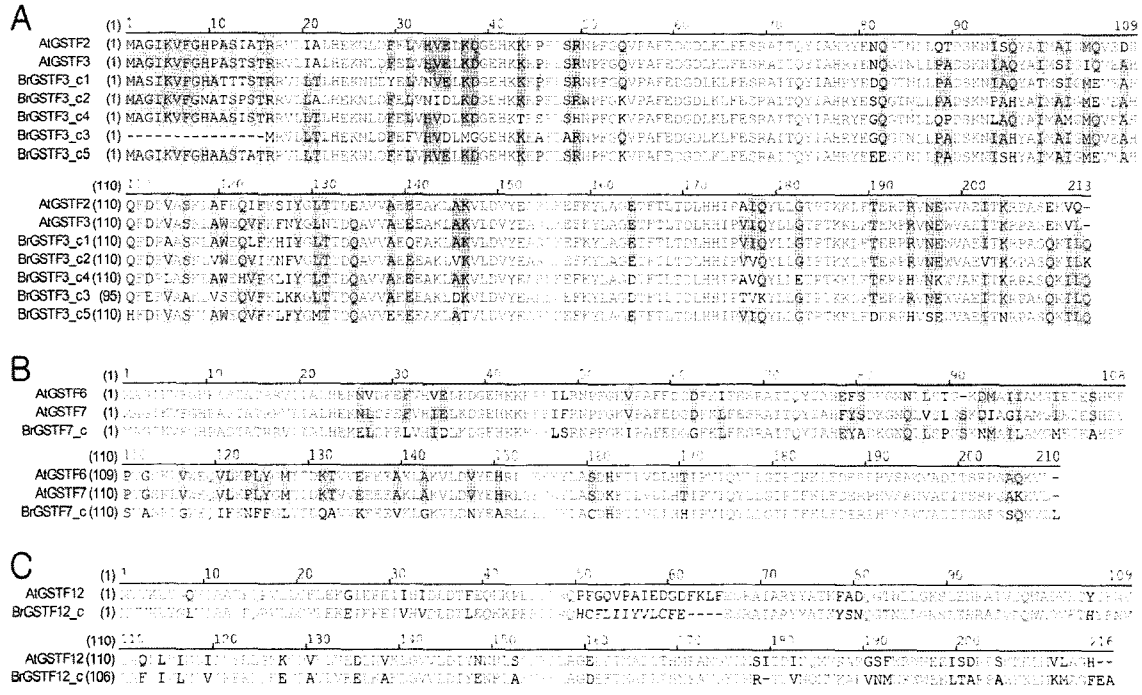

Figure 3. Amino acid sequence alignments of three different groups of *GSTFs*. A: *GSTF3*, B: *GSTF7*, and C: *GSTF12*.

A

Pairwise % similarity of amino acid sequences

		A	B	C	D	E	F	G
AtGSTF2	A	-	92	86	84	77	86	86
AtGSTF3	B	94	-	90	87	79	86	88
BrGSTF3_c1	C	89	90	-	88	79	86	88
BrGSTF3_c2	D	88	89	91	-	81	85	87
BrGSTF3_c3	E	83	83	84	84	-	80	86
BrGSTF3_c4	F	90	90	90	90	86	-	78
BrGSTF3_c5	G	90	90	91	90	83	91	-

Pairwise % similarity of nucleotide sequences

B

Pairwise % similarity of amino acid sequences

		A	B	C
AtGSTF6	A	-	93	81
AtGSTF7	B	94	-	81
BrGSTF6_c	C	85	86	-

Pairwise % similarity of nucleotide sequences

C

Pairwise % similarity of amino acid sequences

		A	B
AtGSTF12	A	-	76
BrGSTF12_c	B	83	-

Pairwise % similarity of nucleotide sequences

Figure 4. Pairwise amino acid and nucleotide sequence similarities in three different groups of *GSTFs*. A: *GSTF3*, B: *GSTF7*, and C: *GSTF12*.

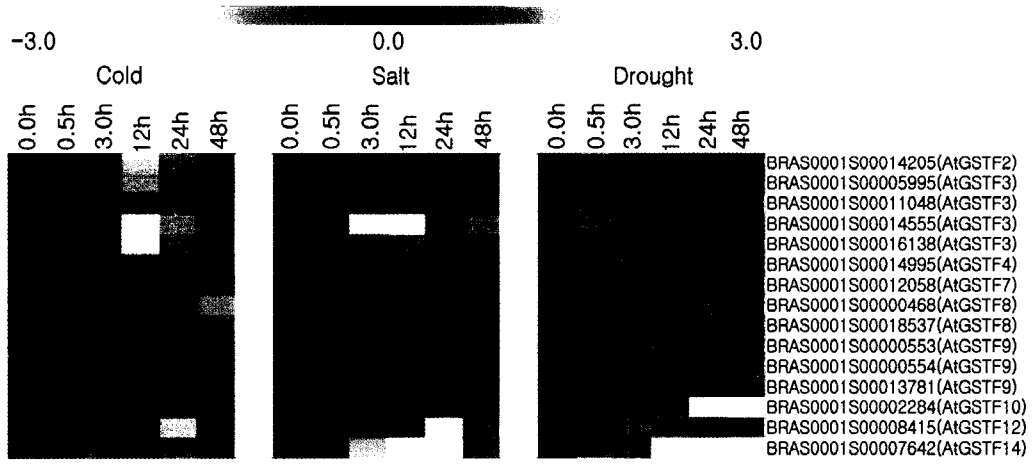

Figure 5. Hierarchical cluster display of 15 selected *BrGSTFs* candidates. In the microarray analysis, cold (4℃), salt (250 mM NaCl) and drought (air-dry) stresses for various times (0.5 to 48 hours) were applied (Lee et al. 2008). The codes of *B. rapa* unigenes and *AtGSTFs* related to *BrGSTFs* are presented on the right of the cluster. The color scale bar shown above the cluster indicates the maximum and minimum brightness values that represent expression ratio in $\log_2$.

The selected 15 unigenes were investigated in the results of microarray analyses of *B. rapa* plants treated with abiotic stresses including cold, salt and drought. Their expression patterns are presented in Figure 5. The color scale indicated the fold-change of genes converted from perfect match (PM) values of the six probes which were designed in the sense direction (5'-UTR to 3'-UTR) (Lee et al. 2008). Six of the selected 15 unigenes, BRAS0001S00005995 (*AtGSTF3*), BRAS0001S00011048 (*AtGSTF3*), BRAS0001S00014555 (*AtGSTF3*), BRAS0001S00016138 (*AtGSTF3*), BRAS0001S0001205 (*AtGSTF7*) and BRAS0001S00008415 (*AtGSTF12*) were compared with those of candidates *BrGSTF3*, *BrGSTF7* and *BrGSTF12* because they were selected and comparable in both *B. rapa* BAC clones and unigenes. BRAS0001S00005995 (*AtGSTF3*), BRAS0001S00011048 (*AtGSTF3*), BRAS0001S00014555

(*AtGSTF3*), BRAS0001S00016138 (*AtGSTF3*) and BRAS0001S00008415 (*AtGSTF12*) were exactly identical to *BrGSTF3*_c1, *BrGSTF3*_c2, *BrGSTF3*_c3, *BrGSTF3*_c4 and *BrGSTF12*_c, respectively (result not shown).

Discussion

The completion of several genome sequencing projects has led to gene annotation and construction of physical maps and provided the genome distribution of *GST* genes in various plants. In *Arabidopsis*, the *GST* gene family of the identified 52 members was divided into five distinct classes was identified (Dixon et al. 2002b; Wagner et al. 2002). The Phi and Tau classes are plant specific with multifunctional proteins encoded by a large gene family (Dixon et al. 2002b; Nutricati et al. 2006). Because the roles of these genes are related to agriculturally and medically important traits such as herbicide detoxification, responses to biotic and abiotic stress and accumulation of ITC (Hatton et al. 1996; Frova 2003; Ambrosone et al. 2004; Gasper et al. 2005), researches on plant *GSTs* are getting increased. In *Arabidopsis*, many *GSTs* have been characterized (DeRidder et al. 2002; ; Wagner et al. 2002; DeRidder and Goldsbrough 2006; Nutricati et al. 2006) and several *GSTs* have also been identified and characterized in rice (*Oryza sative*; Cho et al. 2007), wheat (*Triticum aestivum*; Thom et al. 2002), maize (*Zea mays*; Neuefeind et al. 1997a; 1997b; Farkas et al. 2007), and tall fescue (*Festuca arundinacea*; Buono et al. 2007). They were mostly related to detoxifying activities of herbicides. Especially in *Brassica* species, *GSTs* are of interest and importance regarding anticancer function related to glucosinolate (sulforaphane) mainly detected in Broccoli (*B. oleracea*), and many researches therefore have been focused on it during last a few years (Ambrosone et al. 2004; Gasper et al. 2005; Zhang et al. 2006). However, little has been reported in *B. rapa* (Chinese cabbage). Therefore, we tried to identify *GSTFs in silico* derived from *B. rapa* in this study as a part of Brassica Genome Sequencing Project.

Eleven BAC clones were selected from the sequenced BAC clones by comparison with *AtGSTFs* (Table 2). Blast search and sequence alignment resulted in designating seven unique *BrGSTFs* whose map positions were identified (Figure 1B). As *AtGSTs* are usually known to be clustered in the genome as a result of gene duplication giving rise to sequence diversification (Dixon et al. 2002b; Frova 2003, 2006), two sets of duplicated *BrGSTFs* were detected in the same BAC clones and mapped on the chromosomes R3 and R9. On the map of *A. thaliana*, four and two *AtGSTFs* are located on the same regions of chromosome 1 and 2, respectively (Figure 1A) and expanding to other classes of *AtGSTs*, the clustering of *GST* genes is much clearer. Of 48 *AtGST* genes, excluding the DHAR genes which are not precisely positioned, the non-random distribution of *AtGSTs* along the chromosomes was reported (Lin et al. 1999). Only 14 *AtGST* genes are solely present, but the remaining *GSTs* showed tight class-specific clusters containing two to seven members. For instance, seven clusters with two to five *AtGSTs* are present on chromosome 1 which is the richest (in total 23 *AtGSTs*). The densest cluster is present on chromosome 2 where seven *AtGSTUs* lay tandomly in a 14 Kb segment (Lin et al. 1999). In rice, the analysis of the chromosomes *in silico* after the release of the whole genome sequences also revealed clustering of *GSTs* of the same class, the most obvious of which are a large cluster of 24 Tau genes on chromosome 10 and nine Phi genes on chromosome 1 (Dixon et al. 2002b; Yuan et al. 2002; Soranzo et al. 2004; Frova 2003, 2006). The presence of clusters of *GSTs* in the genomes of rice and *Arabidopsis* suggests it to be a common feature within this gene family and we also found BAC clones containing two members of *GSTFs* in this study and five members of *GSTUs* in our contemporary study.

The *AtGST* family shows considerable sequence divergence with less than 25% identity at the amino acid level between classes (Wagner et al. 2002). Within the *AtGSTF* and *BrGSTF* it ranged from 33% to 95% and from 51% to 91%, respectively. By comparison of amino acid sequences between *BrGSTFs* and *AtGSTFs*, high percentage of affinities and colinearities was detected within sub-classes of *GSTFs* (Figure 2 and 3). This result could be derived from the facts that *Arabidopsis* and *B. rapa* belong to the same family of Brassicaceae and the Brassica genome has been known to be

evolutionarily triplicated from the *Arabidopsis* genome, resulting sequence/gene divergence (O'Neill and Bancroft 2000; Rana et al. 2004). Similar result was observed in the *FLC* gene (Yang et al. 2006) and they suggested that the number of genes in the *Brassica* genome is increased approximately 1.7-fold compared with that of the *Arabidopsis* genome. Therefore the number of *BrGSTFs* could be estimated to be 22 although we identified the limited number of *BrGSTFs* due to the limited information of the sequenced BAC clones in this study. Moreover, the number of unigenes found using KBGP-24K which might cover approximately 50% of estimated gene in *B. rapa* was 15, slightly higher than expected. *In silico* analysis of *BrGSTFs* and unigenes proposes at least six *BrGSTF* genes to be expressed.

Chinese cabbage is one of major vegetables in Korea, but the content of sulforaphane, a kind of ITC which is of interest due to its anticancer activity is relatively much lower than that in broccoli. Therefore recently many researches have been focused on generating Chinese cabbages containing high amount of various ITCs including sulforaphane through bio-technology approaches. A knockout of *BrGSTFs* we identified in this study will be one of the methods to achieve it and additionally *BrGSTs* can be used for further studies on herbicide detoxification and responses to biotic and abiotic stresses, which are of agriculturally interest in the cultivation and breeding of Chinese cabbage.

Acknowledgements

This work was supported by grants from National Institute of Agricultural Biotechnology (Code #04-1-12-2-3) and the BioGreen21 Programs (Code #20050301034438 and Code #20070301034037), the Rural Development Administration, Republic of Korea.

References

Ambrosone CB, McCann SE, Freudenheim JL, Marchall JR, Zhang Y and Shields PG (2004) Breast cancer risk in premenopausal women is inversely associated with consumption of broccoli, a source of isothiocyanates, but is not modified by *GST* genotype. J Nutr 134: 1134-1138

Buono DD, Scarponi L and Espen L (2007) Glutathione S-transferases in *Festuca arundinacea*: identification, characterization and inducible by safeners benoxacor. Phytochemistry 68: 2614-2624

Burge CB and Karlin S (1997) Prediction of complete gene structure in human genomic DNA. J Mol Biol 268: 78-94

Chen WQ, Chao G and Singh KB (1996) The promoter of a H_2O_2-inducible, *Arabidopsis* glutathione S-transferase gene contains closely linked OBF- and OBP1-binding sites. Plant J 10: 955-966

Cho H-Y, Lee HJ and Kong K-H (2007) A Phi class glutathione S-transferase from *Oryza sativa* (*OsGSTF5*): molecular cloning, expression and biochemical characteristics. J Biochem Mol Biol 40: 511-516

DeRidder BP and Goldsbrough PB (2006) Organ-specific expression of glutathione S-transferases and the efficacy of herbicide safeners in *Arabidopsis*. Plant Physiol 140: 167-175

DeRidder BP, Dixon DP, Beussman DJ, Edwards R and Goldsbrough PB (2002) Induction of glutathione S-transferases in *Arabidopsis* by herbicide safeners. Plant Physiol 130: 1497-1505

Dixon DP, Davies BG and Edwards E (2002a) Functional divergence in the glutathione transferase superfamily in plant. J Biol Chem 277: 30859-30869

Dixon DP, Lapthorn A and Edwards R (2002b) Plant glutathione transferases. Genome Biol 3: reviews 3004.1 -3004.10

Edwards R, Dixon DP and Walbot V (2000) Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. Trends Plant Sci 5: 193-198

Farkas M, Berry JO and Aga DS (2007) Chlortetracycline detoxification in maize via induction of glutathione S-transferases after antibiotic stress. Environ Sci Technol 15: 1450-1456

Frova C (2003) The plant glutathione transferase gene family: genomic structure, functions, expression and evolution. Physiol Plant 46: 469-479

Forva C (2006) Glutathione transferases in the genomics era: new insights and perspectives. Biomol Eng 23: 149-169

Gasper AV, Al-janobi A, Smith JA, Bacon JR, Fortun P, Atherton C, Taylor M, Hawkey CJ, Barrett DA and Mithen RF (2005) Glutathione S-transferase M1 polymorphism and metabolism of sulforaphane from standard and high-glucosinolate broccoli. Am J Clin Nutr 82: 1283-1291

Hatton PJ, Dixon D, Cole DJ and Edwards R (1996) Glutathione transferase activity and herbicide selectivity

in maize and associated weed species. Pestic Sci 46: 267-275

Kim JS, Chung TY, King GJ, Jin M, Yang T-J, Jin Y-M, Kim H-I and Park B-S (2006) A sequence-tagged linkage map of *Brassica rapa*. Genetics 174: 29-39

Kiyosue T, Yamaguchi-Shinozaki K and Shinozaki K (1993) Characterization of two cDNAs (ERD11) and (ERD13) for dehydration-inducible genes that encode putative glutathione S-transferases in *Arabidopsis thaliana* L. FEBS Lett 335: 189-192

Lee S-C, Lim M-H, Kim JA, Lee S-I, Kim JS, Jin M, Kwon S-J, Mun J-H, Kim Y-K, Kim HU, Hur Y and Park B-S (2008) Transcriptome analysis in *Brassica rapa* under the abiotic stresses using Brassica 24K oligo microarray. Mol Cells 26: in press

Lin X, Kaul S, Rounsley S, Shea TP, Benito MI, Town CD, Fujii CY, Mason T, Bowman CL, Barnstead M, Feldblyum TV, Buell CR, Ketchum KA, Lee J, Ronning CM, Koo HL, Moffat KS, Cronin LA, Shen M, Pai G, van Aken S, Umayam L, Tallon LJ, Gill JE, Adams MD, Carrera AJ, Creasy TH, Goodman HM, Somerville CR, Copenhavor GP, Preuss D, Nierman WC, White O, Eisen JA, Salzberg SL, Fraser CM and Venter JC (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402: 761-768

Lukashin AV and Borodovsky M (1998) GeneMark.Hmm: new solutions for gene finding. Nucl Acids Res 26: 1107-1115

Neuefeind T, Huber R, Dasenbrock H, Prade L and Bieseler B (1997a) Crystal structure of herbicide-detoxifying maize glutathione S-transferase-I in complex with lactoylglutathione: evidence for an induced-fit mechanism. J Mol Biol 274: 446-453

Neuefeind T, Huber R, Reinemer P, Knäblein J, Prade L, Mann K and Bieseler B (1997b) Cloning, sequencing, crystalization and X-ray structure of glutathione S-transferase-III from *Zea mays* var. *mutin*: a leading enzyme in detoxification of maize herbicides. J Mol Biol 274: 577-587

Nutricati E, Miceli A, Blando F and De Bellis L (2006) Characterization of two *Arabidopsis thaliana* glutathione S-transferases. Plant Cell Rep 25: 997-1005

O'Neill CM and Bancroft I (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. Plant J 23: 233-243

Park J, Koo DH, Hong CP, Lee SJ, Jeon JW, Lee SH, Yun PY, Park BS, Kim HR, Bang JW, Plaha P, Bancroft I and Lim YP (2005) Physical mapping and microsynteny of *Brassica rapa* ssp. *pekinensis* genome corresponding to a 222 kbp gene-rich region of *Arabidopsis* chromosome 4 and partially duplicated on chromosome 5. Mol Genet Genomics 274: 579-588

Rana D, van den Boogaart T, O'Neill CM, Hynes L, Bent E, Macpherson L, Park JY, Lim YP and Bancroft I (2004) Conservation of the microstructure of genome segments in *Brassica rapa* and its diploid relatives. Plant J 40: 725-733

Rea PA (1999) MRP subfamily ABC transporters from plants and yeast. J Exp Bot 50: 895-913

Reinemer P, Prade L, Hof P, Neuefeind T, Huber R, Zettl R, Palme K, Schell J, Koelln I, Bartunik HD and Bieseler B (1996) Three-dimensional structure of glutathione S-transferase from *Arabidopsis thaliana* at 2.2 Å resolution: structural characterization of herbicide-conjugation plant glutathione S-transferases and a novel active site architecture. J Mol Biol 255: 289-309

Saitou N and Nei M (1987) The neighbor-joining method: a new method for reconstruction Phylogenetic trees. Mol Biol Evol 4: 406-425

Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer G, Bouck J, Gibbs R, Hardison R and Miller W (2000) PipMaker-a web server for aligning two genomic DNA sequences. Genome Res 10: 577-586

Soranzo N, Sari Gorla M, Mizzi L, De Toma G and Frova C (2004) Organization and structural evolution of the rice glutathione S-transferase gene family. Mol Gen Genomics 271: 511-521

Thom R, Cummins I, Dixon DP, Edwards R, Cole DJ and Lapthorn AJ (2002) Structure of a Tau class glutathione S-transferase from wheat active in herbicide detoxification. Biochemistry 41: 7008-7020

Wagner U, Edwards R, Dixon DP and Mauch F (2002) Probing the diversity of the *Arabidopsis* glutathione S-transferase gene family. Plant Mol Biol 49: 515-532

Yang K-Y, Kim E-Y, Kim C-S, Guh J-O, Kim K-C and Cho B-H (1998) Characterization of a glutathione S-transferase gene *AtGST1* in *Arabidopsis thaliana*. Plant Cell Rep 17: 700-704

Yang T-J, Kim J-S, Lim K-B, Kwon S-J, Kim J-A, Jin M, Park JY, Lim M-H, Kim H-I, Kim SH, Lim YP and Park B-S (2005) The Korea Brassica Genome Project: a glimpse of the Brassica genome based on comparative genome analysis with *Arabidopsis*. Comp Funct Genomics 6: 138-146

Yang T-J, Kim J-S, Kwon S-J, Lim K-B, Choi B-S, Kim J-A, Jin M, Park JY, Lim M-H, Kim H-I, Lim YP, Kang JJ, Hong J-H, Kim C-B, Bhak J, Bancroft I and Park B-S (2006) Sequence-level analysis of the diploidization process in the triplicated *FLOWERING LOCUS* C region of *Brassica rapa*. Plant Cell 18: 1339-1347

Yuan Q, Hill J, Hsiao J, Moffat K, Ouyang S, Cheng Z, Jiang J and Buell CR (2002) Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion. Mol Genet Genomics 267: 713-720

Zettl R, Schell J and Palme K (1994) Photoaffinity labeling of *Arabidopsis thaliana* plasma membrane vesicles by 5-azido-[7-^{3}H]indole-3-acetic acid: identification of a glutathione S-transferase. Proc Natl Acad Sci USA 91: 689-693

Zhang Y, Munday R, Jobson HE, Munday CM, Lister C, Wilson P, Fahey JW and Mhawech-Fauceglia P (2006) Induction of *GST* and NQO1 in cultured bladder cells and in the urinary bladders of rats by an extract of broccoli (*Brassica oleracea*) sprouts. J Agric Food Chem 13: 9370-9376

Zhou J and Goldsbrough PB (1993) An *Arabidopsis* gene with homology to glutathione S-transferases is regulated by ethylene. Plant Mol Biol 22: 517-523

(Received September 23, 2008; Accepted November 25, 2008)