Insect Biochemistry and Molecular Biology 38 (2008) 1158–1164

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Identification, genomic organization and expression pattern of glutathione *S*-transferase in the silkworm, *Bombyx mori*

Quanyou Yu [a,c], Cheng Lu [a,*], Bin Li [a], Shoumin Fang [b], Weidong Zuo [a], Fangyin Dai [a], Ze Zhang [a,c], Zhonghuai Xiang [a]

[a] *The Key Sericultural Laboratory of Agricultural Ministry, College of Biotechnology, Southwest University, Chongqing 400715, China*
[b] *College of Life Science, China West Normal University, Nanchong 637002, China*
[c] *The Institute of Agricultural and Life Sciences, Chongqing University, Chongqing 400030, China*

ARTICLE INFO

Article history:
Received 24 October 2007
Received in revised form 7 August 2008
Accepted 8 August 2008

Keywords:
Glutathione *S*-transferases
Bombyx mori
Genomic organization
Expression pattern

ABSTRACT

Glutathione *S*-transferases (GSTs) are a multifunctional supergene family and some play an important role in insecticide resistance. We have identified 23 putative cytosolic GSTs by searching the new assembly of the *Bombyx mori* genome sequence. Phylogenetic analyses on the amino acid sequences reveal that 21 of the *B. mori* GSTs fall into six classes represented in other insects, the other two being unclassified. The majority of the silkworm GSTs belong to the Delta, Epsilon, and Omega classes. Most members of each class are tandemly arranged in the genome, except for the Epsilon GSTs. Expressed sequence tags (ESTs) corresponding to 19 of the 23 GSTs were found in available databases. Furthermore RT-PCR experiments detected expression of all the GSTs in multiple tissues on day 3 of fifth instar larvae. Surprisingly, we found little or no expression of most Delta and Epsilon GSTs in the fat body, which is thought to be the main detoxification organ. This may explain the sensitivity of the silkworm to certain insecticides. Our data provide some insights into the evolution of the *B. mori* GST family and the functions of individual GST enzymes.

1. Introduction

Glutathione *S*-transferases (GSTs, EC2.5.1.18) are a superfamily of multifunctional enzymes found in almost all living organisms. These enzymes catalyze the nucleophilic attack of the tripeptide glutathione (GSH) on electrophilic centers of toxic compounds, including insecticides, arene oxides, quinones, and α,β-unsaturated carbonyl compounds (Motoyama and Dauterman, 1980; Clark and Shamaan, 1984; Vontas et al., 2001; Hayes et al., 2005). GSTs can serve as maleylacetoacetate isomerases and thiol transferases (Board et al., 1997, 2000). GSTs can also contain se-independent GSH peroxidase activity, which plays an important role in protecting against oxidative injury (Singh et al., 2001). In addition, GSTs have non-catalytic functions, mainly binding hydrophobic compounds such as drugs, hormones, and other metabolites (Hayes and Pulford, 1995).

GSTs may be microsomal, mitochondrial or cytosolic. Microsomal GSTs lie in a different superfamily of proteins known as the MAPEG superfamily (membrane-associated proteins in eicosanoid and glutathione metabolism) (Jakobsson et al., 1999). The mitochondrial GSTs are located in mammalian mitochondria and peroxisomes, but are not found in insects (Lander et al., 2004; Morel et al., 2004). Most insect GSTs belong to the cytosolic group. Based on amino acid sequence similarities and immunological relationships, seven classes of cytosolic GST have been recognized in mammalian species, designated Alpha, Mu, Pi, Omega, Sigma, Theta, and Zeta (Sheehan et al., 2001). Besides the ubiquitous Omega, Sigma, Theta, and Zeta classes, the insect specific Delta and Epsilon classes, as well as unclassified GSTs, have been identified in insects (Chelvanayagam et al., 2001; Ranson et al., 2001, 2002; Ding et al., 2003).

Insect GSTs play an important role in detoxifying insecticides. GST activities (toward model substrates 1-chloro-2,4-dinitro benzene (CDNB) and 1,2-dichloro-4-nitrobenzene (DCNB)) are significantly increased in OP (organophosphate) and pyrethroid resistant strains of the fall armyworm (Yu, 1992). *Plutella xylostella* GST3 exhibits relatively high activities toward DCNB and some OPs, suggesting that it might be related to OP resistance in the diamondback moth (Chiang and Sun, 1993; Ku et al., 1994; Huang et al.,

Abbreviations: CDNB, 1-chloro-2,4-dinitrobenzene; DCNB, 1,2-dichloro-4-nitrobenzene; OPs, organophosphate insecticides.

* Corresponding author. Tel.: +86 23 68250346; fax: +86 23 68251128.
E-mail address: lucheng@swu.edu.cn (C. Lu).

0965-1748/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2008.08.002

1998). The OP phoxim was widely used to control *Helicoverpa armigera* in China in the 1980s, and resistance to it has developed in Chinese populations (Wu et al., 1997). GST activity was also found to be significantly higher in phoxim-resistant *H. armigera*. These observations suggest a correlation between increased GST activity and insecticide resistance.

In this study, we have identified the silkworm cytosolic GSTs using the newly assembled 9× genome sequence (The International Silkworm Genome Sequencing Consortium, submitted for publication) and analyzed their genomic distribution and intron characteristics. We have searched available EST data for each silkworm GST to confirm active transcription and examined the expression patterns for all the GST genes in multiple tissues on day 3 fifth instar larvae by reverse transcription-polymerase chain reaction (RT-PCR). Our results provide preliminary insights into the evolution and functions of the silkworm GSTs.

2. Materials and methods

2.1. Identification of the Bombyx mori GST genes

To search for putative silkworm GST genes, GST protein sequences of *Drosophila melanogaster*, *Anopheles gambiae* and *Apis mellifera* were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/) and used as queries to perform TBLASTN searches against the silkworm 9× genome database (Altschul et al., 1997). If a piece of genomic sequence showed even weak sequence similarity to any query sequence, its flanking regions (1 kb or longer) were extracted. Genes within the extracted sequences were predicted using BGF software (Wang et al., 2005) and Fgenesh+ (http://www.softberry.com/). The silkworm GSTs were classified according to conventional nomenclature (Chelvanayagam et al., 2001). Their numbering within each class mainly reflects the order of their submission to GenBank.

2.2. EST collection

To search for evidence of transcription of individual GST genes, a BLASTN search was conducted against the silkworm EST database downloaded from GenBank. The putative coding sequences were used as queries. A 95% or greater identity and minimum cut-off *E*-value (e−20) were employed to discriminate between duplicated genes.

2.3. Phylogenetic analysis

As well as the *D. melanogaster*, *A. gambiae*, and *A. mellifera* GSTs, six lepidopteran GSTs with functional data were also included to inform the possible functions of the *B. mori* clades. Putative amino acid sequences of GSTs were aligned using Clustal X (Thompson et al., 1997). Positions that had a high percentage of gaps (>70%) were manually trimmed. A phylogenetic tree was reconstructed using the neighbor-joining method (Saitou and Nei, 1987) implemented in MEGA 4.0 (Tamura et al., 2007). In addition, a maximum likelihood (ML) tree was reconstructed using PHYLIP 3.65 (Felsenstein, 2005) with the Jones–Taylor–Thornton model and 100 bootstrap replicates. A maximum parsimony (MP) tree was also built by PAUP* 4.0b10 (Swofford, 2002) with heuristic searches with tree bisection reconnection (TBR) and 100 random-taxon-addition replicates.

2.4. RNA extraction and RT-PCR

Total RNA was extracted from the fat body using Trizol reagent (Invitrogen, USA) and RNA concentration was determined using a spectrophotometer (Gene Spec V: HITACHI, Japan). DNA within RNA samples was digested with RNase-free DNase I. The first strand of cDNA was synthesized using M-MLV Reverse Transcriptase following the manufacturer's instructions (Promega, USA).

RT-PCR primers were designed on the basis of the coding sequences and ESTs of the silkworm GSTs (Supplemental Table S1). The silkworm cytoplasmic actin *A3* gene (forward primer: 5′-AACACCCCGTCCTGCTCACTG-3′; reverse primer: 5′-GGGCGAGACGT GTGATTTCCT-3′) was used as an internal control. PCR amplification was performed in a total reaction volume of 25 μl, containing normalized cDNA, 15 pmol of each primer, 2 mM MgCl2, 0.25 mM dNTP, 1× buffer and 2.5 units of Taq DNA polymerase. PCRs were performed with the following cycles: initial denaturation at 95 °C for 5 min; followed by 25 cycles of 1 min at 95 °C, 30 s annealing (temperatures listed in Supplemental Table S1), 1 min extension (72 °C), and a final extension at 72 °C for 10 min. The amplification products were analyzed on 1.5% agarose gels, purified from the gel, and directly sequenced using an ABI 3100 automated sequencer.

3. Results

3.1. Classification and phylogeny of B. mori GSTs

Using the amino acid sequences of *A. gambiae*, *D. melanogaster*, and *A. mellifera* GSTs as queries, we identified 23 putatively cytosolic GST genes by a local TBLASTN search of the silkworm genome sequence (Table 1). To reveal the relationships among the GSTs, we reconstructed phylogenetic trees using NJ, ML, and MP methods. Because the topologies of the three resulting trees are similar in overall structure, we show only the NJ tree (Fig. 1). The higher-level structure of the phylogeny is well supported in terms of bootstrap scores and distinguishes all the major classes. The silkworm GSTs cover all six classes (Delta, Epsilon, Omega, Sigma, Theta, and Zeta) found in other insects (Ding et al., 2003), with two that could not be readily assigned within one of the known classes being designated as 'unclassified'. Interestingly, both the latter are most closely related to the Delta classes (Fig. 1). Most of the BmGSTs sit within

Table 1
Summary of the silkworm GSTs

Gene name	Length of putative proteins	Number of ESTs	Chromosome	Old name	Accession number
GSTd1	218	2	6		AJ006502[b]
GSTd2	216	54	6	*GSTt*	AB176691[b]
GSTd3	220	3	6	*GST3*	DQ355374[b]
GSTd4	–	1	6		BGIBMGA006538[a]
GSTe1	222	22	7	*GST1*	AY192575[b]
GSTe2	215	0	26	*GST5*	DQ355376[b]
GSTe3	217	2	UN	*GST10*	EF506488[b]
GSTe4	223	1	7	*GST9*	EF506489[b]
GSTe5	229	1	7	*GST11*	EU216542[b]
GSTe6	223	0	21	*GST12*	EU216543[b]
GSTe7	–	6	19	*GST14*	EU216545[b]
GSTe8	–	0	10		BGIBMGA006639[a]
GSTo1	254	19	11	*GSTo1*	DQ311183[b]
GSTo2	256	22	11	*GST6*	DQ355373[b]
GSTo3	240	12	24	*GSTo*	DQ443293[b]
GSTo4	247	0	11	*GST13*	EU216544[b]
GSTs1	206	9	3	*GST2*	AY297161[b]
GSTs2	204	33	3	*GSTs*	AB206971[b]
GSTt1	229	2	8	*GST8*	EF506487[b]
GSTz1	215	12	25	*GST4*	DQ355375[b]
GSTz2	216	3	15	*GSTz*	EF565386[b]
GSTu1	233	2	6	*GST7*	EF423869[b]
GSTu2	216	12	UN		DQ311182[b]

The dash represents incomplete coding sequences. UN represents unknown chromosome locations.

[a] Accession number of the silkworm 9× genome database (http://silkworm.swu.edu.cn/silkdb/).

[b] GenBank accession number.