Article

CLONING, CHARACTERIZATION AND EXPRESSION OF TWO GLUTATHIONE *S*-TRANSFERASE cDNAS IN THE SPRUCE BUDWORM, *Choristoneura fumiferana*

Yufen Huang
Guangdong Provincial Key Lab for Plant Development, School of Life Sciences, South China Normal University, Guangzhou, China

Peter J. Krell
Department of Molecular and Cellular Biology, University of Guelph, Ontario, Canada

Tim Ladd
Great Lakes Forestry Centre, Canadian Forest Service, Sault Ste. Marie, Ontario, Canada

Qili Feng and Sichun Zheng
Guangdong Provincial Key Lab for Plant Development, School of Life Sciences, South China Normal University, Guangzhou, China

Two Choristoneura fumiferana glutathione S-transferase cDNAs (CfGSTs4 and CfGSTd5) were cloned from a cDNA library constructed using mRNA from the midgut cell line, CF-203. These cDNAs encoded two structurally different proteins with a predicted molecular mass of 23 and 24 kDa, respectively. Amino acid sequence analysis indicates that CfGSTs4 and CfGSTd5 contained Sigma and Delta GST domain, respectively. CfGSTs4 cDNA was expressed as a recombinant protein with the same molecular mass as predicted. Semi-quantitative reverse-transcription PCR analyses indicated that both of these genes were

Grant sponsor: Canadian Biotechnology Strategy Fund; Grant sponsor: Genome Canada through the Ontario Genomics Institute; Grant sponsor: South China Normal University.
Correspondence to: Sichun Zheng, Guangdong Provincial Key Lab of Biotechnology for Plant Development, School of Life Sciences, South China Normal University, Guangzhou, China, 510631.
E-mail: sczheng62@126.com

ARCHIVES OF INSECT BIOCHEMISTRY AND PHYSIOLOGY, Vol. 70, No. 1, 44–56 (2009)
Published online in Wiley InterScience (www.interscience.wiley.com).
© 2008 Wiley Periodicals, Inc. DOI: 10.1002/arch.20272

expressed in the epidermis, fat body, and midgut of the 6th instar larvae, as well as CF-203 cells. CfGSTs4 was highly and almost constantly expressed in all tissues during the 6th instar stage. There were higher levels of CfGSTs4 protein in the midgut and fat body than in the epidermis. CfGSTd5 was expressed in the fat body when the insects underwent pupal molting and was constantly expressed in the epidermis and midgut during 6th instar development. CfGSTs4 expression was not affected by ecdysone agonist tebufenozide (RH5992), whereas CfGSTd5 expression was slightly suppressed by the compound. *© 2008 Wiley Periodicals, Inc.*

Keywords: glutathione *S*-transferase; detoxification; resistance; insecticide

INTRODUCTION

Glutathione *S*-transferases are a superfamily of proteins involved in detoxification of xenobiotics, protection from oxidative damage, and intracellular transport of hormones, endogenous metabolites, and exogenous chemicals. GSTs can conjugate reduced glutathione on the thiol of cysteine to various electrophiles and bind to a variety of hydrophobic substances. In insects, interests have been focused on GST roles in enzymic detoxification of insecticides (Motoyama and Dauterman, 1980; Clark, 1989).

Based on sequence similarity, immunological characters, kinetic properties, and tertiary structure, mammal GSTs have been clustered into eight groups, Alpha, Mu, Pi, Sigma, Theta, Delta, Kappa, which are found in rat mitochondrial matrix (Pemble et al., 1996), and microsomal subunits (Jakobsson et al., 1997; Sheehan et al., 2001). In insects, seven groups of GSTs have been classified: Delta, Epsilon, Omega, Theta, Sigma, Zeta, and microsomal group (Enayati et al., 2005). GSTs in the Delta and Epsilon groups are present in only insect species, while GSTs in other groups are also present in other species. In addition, Delta and Epsilon GST transcripts are more abundant than other transcript classes in the species such as *Anopheles gambiae* and *Drosophila melanogaster.* Gene duplication, genetic rearrangement, and alternative splicing may contribute to the insect GSTs diversity. The presence of a wide range of GST isoforms implicates multifunctions of the GSTs superfamily in ligand binding for detoxification, removal of reactive oxygen species under oxidative stress, and catalytic activity in metabolic pathways not associated with detoxification. We previously identified three structurally different GST cDNAs (*Cf*GST, *Cf*GST-2, and *Cf*GST-3) from the spruce budworm, *Choristoneura furmiferana,* one of the most widely distributed destructive forest insect pests in North America, and characterized their nucleotide sequences, protein activity, localization, and developmental and stress-induced expression patterns (Feng et al., 1999, 2001; Zheng et al., 2007). *Cf*GST belongs to the Sigma group, while *Cf*GST-2 and *Cf*GST-3 are members of the Delta group, although the functions of all of these three GSTs have not been identified experimentally.

In this report, we describe another two cDNAs that encoded two structurally different GST proteins from the previous three. We analyzed their nucleotide sequence and the spatial and temporal expression patterns in the epidermis, fat body, and midgut during the development of 6th instar larvae. We suggest that these different genes might play various roles during development of the insect.

MATERIALS AND METHODS

Experimental Insects and Treatments

Third instar larvae of the spruce budworm (*Choristoneura fumiferana* Clem., Lepidoptera:Tortricidae) were maintained on artificial diet (McMorran, 1965) at 22°C, 70% relative humidity, and a photoperiod of 12-h light and 12-h dark cycle, and reared until they reached the pupal stage in the laboratory.

The nonsteroidal ecdysone agonist, tebufenozide (RH5992), was applied to two-day-old 5th instar larvae. One microliter of 70 ng/μl RH5992 in imbibing solution (1% sucrose and 0.1% red dye no. 28) was applied on a small diet plug fed to 5th instar larvae (Retnakaran et al., 1997). After consumption of the RH5992-containing diet, the larvae were transferred onto regular diet. The insects were collected at 6, 12, 24, and 48 h post-treatment for RNA isolation. The controls were treated with the imbibing solution alone.

Cell Culture

The *C. fumiferana* midgut cell line FPMI-CF-203 (CF-203, Sohi et al., 1993) was cultured at 28°C in 25-ml flasks with SF900 medium supplemented with 10% fetal bovine serum (Life Technologies, Inc.). Cells were seeded at the initial concentration of 1.5×10^5 cells/ml in 25-ml flasks. The cells were harvested after 2-day culture for RNA isolation.

Cloning and Sequencing of GST cDNAs

Partial sequences of two putative GST cDNAs were isolated from an expressed sequence tag (EST) project, in which a cDNA library was constructed in the Uni-ZAP XR vector by using the ZAP cDNA Gigapack II Gold Cloning Kit (Stratagene, La Jolla, CA) with mRNA isolated from the CF-203 cell line. The partial sequences were then assembled into full-length cDNA sequences by using the CAP3 program (Huang and Maddan, 1999) and confirmed by sequencing the longest target cDNA clones selected from the cDNA library. DNA sequencing was conducted using ABI 377 capillary automatic sequencer.

Sequence Analysis

Annotation, comparison, and alignment of sequences were performed using the National Center for Biotechnology Information BLAST search services (Altschul et al., 1990) and Clustal Alignment Program (Higgins and Sharp, 1988) of DNASTAR (DNASTAR, Inc., Madison, WI).

RNA Isolation and Reverse Transcription Polymerase Chain Reaction (RT-PCR)

Total RNA was isolated from larval tissues or CF-203 cells using the guanidinium thiocyanate method (Chomyczynski and Sacchi, 1987). For RT-PCR, mRNA was amplified using Superscript First Strand Synthesis System for RT-PCR of Invitrogen according to the manufacturer's instruction. The primers for amplifying *Cf*GSTs4 were: forward primer: 5′-GCTCGGACTTCGTTTGACCAACA-3′, reverse primer: 5′-GTTCGTGACAAAGTCCTCCGCTAA-3′; the primers for amplifying *Cf*GSTd5 were: forward primer: 5′-CTGGACCTGTGTGACTTCAAGTTC-3′, reverse primer: 5′-ACGGAAAAACATTGTTCAGTTATAC-3′. Template DNA was denatured at 94°C for

3 min, followed by 28 cycles of 94°C for 45 s, 60°C for 45 s, and 72°C for 90 s for each cycle. PCR products were separated in 0.9% agarose gels and stained with ethidium bromide.

Protein Expression and SDS-PAGE

Target cDNAs were cloned into the pET-22b(+) expression vector (Life Technologies, Burlington, Canada) in fusion with a 6xHis tag at the C-terminal end. *E. coli* cells (BL21-Blue) were transformed with the recombinant plasmid DNAs (pPROEX-*Cf*GST4 and pPROEX-*Cf*GSTs5). Recombinant protein expression was induced by adding isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM.

Protein samples were denatured at 100°C for 3 min in an equal volume of 2 × protein loading buffer (0.1 M Tris buffer, pH 6.8, 4% SDS, 0.2% β-mercaptoethanol, 40% glycerol, and 0.002% bromophenol blue). SDS-PAGE was performed in 12% acrylamide gels in tris-glycine-SDS buffer (10 mM Tris, 50 mM glycine, 0.1% SDS, pH 8.0). The gel was stained with Coomassie Blue R-250. The molecular mass of the proteins was calculated using the GeneTools program from Syngene (Synoptics Ltd., Cambridge, UK).

Antibody Production and Western Blotting

The recombinant GST proteins were excised from SDS-PAGE gels. Polyclonal antiserum was made in rabbits by three boost injections, each with 300-ng proteins in Freund's adjuvant. Antiserum was collected after the boost injections. Pre-immune serum collected from the same rabbit prior to immunization was used as a control.

For Western blotting analysis, proteins were transferred from the acrylamide gel to a nitrocellulose membrane. The membrane was blocked with 3% BSA in 1 × PBS buffer for 30 min at room temperature, and then incubated with the GST antibodies (1:500) at room temperature for 1 h. Goat anti- rabbit IgG conjugated with alkaline phosphatase was used as the secondary antibody at a dilution of 1:2,000. Nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate were used as substrates for color development.

RESULTS

Cloning of CfGSTs4 and CfGSTd5

In an expressed sequence tags (ESTs) project, we identified nine ESTs (out of 1,776 ESTs) that coded for different GST proteins. Among these ESTs, three ESTs were assembled into a transcript named *Cf*GSTs4 and two were assembled into another one named *Cf*GSTd5 (Fig. 1), while another four ESTs coded two other GSTs, *Cf*GST-2 and *Cf*GST-3, respectively (Zheng et al., 2007). These two transcripts showed high similarities to the insect GSTs from different species, but differed from the previously reported three *C. fumiferana* GSTs (Feng et al., 1999; Zheng et al., 2007). To further characterize these two GSTs (*Cf*GSTs4 and *Cf*GSTd5), we cloned the full-length sequences of these GST cDNAs. *Cf*GSTs4 was 1,128 nucleotides in length and encoded a 204–amino acid protein with a predicted molecular mass of 23 kDa and a pI of 5.46 (Fig. 1A). *Cf*GSTd5 was 1,003 nucleotides in length, encoding a 217–amino acid protein with a predicted molecular mass of 24 kDa and a pI of 7.17 (Fig. 1B). There

```
A
1       GGCACGAGCTCGGACTTCGTTTGACCAACAGCTAACCTAAACAACCATGCCTAACATCGTGTTCAACTACTTCCCCGTC
                                       M  P  N  I  V  F  N  Y  F  P  V        11
80      AAGGCTCTGGGCGAGCCGAGCCGTCTGCTGCTGGCTTACGGCGGAGAGGGCTTCGAGGACAAGCGCATTTCGCCGGACAAC
         K  A  L  G  E  P  S  R  L  L  L  A  Y  G  G  E  G  F  E  D  K  R  I  S  P  D  N     38
161     TGGCCTGCCTACAAGCCTAGTACTCCGTTCGGCCAGATGCCCGTGATGGATATCGACGGCAAGCAATACGCTCAGAGCTAC
      W  P  A  Y  K  P  S  T  P  F  G  Q  M  P  V  M  D  I  D  G  K  Q  Y  A  Q  S  Y        65
242     GCCATCGCTCGCTACCTCGGCCGCAAGTACGGCCTGGCCGGCGCCGACATCGAAGAAGACTTCCAGATAGACCAGAACGTC
         A  I  A  R  Y  L  G  R  K  Y  G  L  A  G  A  D  I  E  E  D  F  Q  I  D  Q  N  V     92
323     GAGTACCTTAACGACATTCGTGCCAAGGCCGCTAGCGTCCATTACGAGCCTGACGCGGATTACAAGGATAAGAAGCAAGCC
         E  Y  L  N  D  I  R  A  K  A  A  S  V  H  Y  E  P  D  A  D  Y  K  D  K  K  Q  A    119
304     GAAAACAGCAAGAACGTGTATCCTGCAATGCTGGAGAAGCTGGATGCTATTCTCAAGGCCAATAACGGACACCTTGCGTGC
         E  N  S  K  N  V  Y  P  A  M  L  E  K  L  D  A  I  L  K  A  N  N  G  H  L  A  C    146
485     GGCAAGCTGACCTGGGGCGATTTCGTCTTCGCCGGCATCTACGCCTATCTTAAGACCATGCTGCAGACCCCAGATCTGGAC
         G  K  L  T  W  G  D  F  V  F  A  G  I  Y  A  Y  L  K  T  M  L  Q  T  P  D  L  D    173
566     GCGCAGTATCCTAGCTTCAAGAAACTGGAGGAGAAGGTCTACTCGCTGCCCAAGGTCAAGGCATGGGCTGACACCGCACCG
         A  Q  Y  P  S  F  K  K  L  E  E  K  V  Y  S  L  P  K  V  K  A  W  A  D  T  A  P    200
647     AAATGCGATTTTTAAGTTCATAATTTTAGCGGAGGACTTTGTCACGAACTGTTTTCTAATGCTGTGCCTGTTAAATTATCC
         K  C  D  F ***                                                                      204
728     TTATGGGTCCATTGCATGCTGAAGTTGTTTAAATGATAATTTAAATTTAAAATGTGATTGTATCTTATTTAACCGTTTATC
809     ACTTTCATTTTCAACTTAACTTAATAAAAAAAAGAAAATAAAAATAACGGAGGTTAGGGGTTCTTCTAATTTTTTCTTTCT
890     TGAAATTTGTGATATATTGAACTCTTTTCTAGAAGCGCAGCTACTACAGGAATTCACCATAGAACTACAATCTTAATCTAT
971     AAGCAACGGATAACTGTTTGTTATTTATTATTAAGATTGTGTATCTGCTATGCATGTAAGTGTTTATTCGCATAGGTTTTA
1052    AGTATATTAATAAAGTTGCAAGATGTATTTTTGAATGTAATAATACTATGTGTATCTGAAAAAAAAAAAAAAAAAAA

B
1          CGGCACGAGGAAAGAGTATAAGTACCTACCTATGTATCCATCCTTTCACTTTGCTCGTGCTAGTTATCCGTCCTGATCG
80       CTCTCACGTTTTTGTGTAACTGGACCTGTGTGACTTCAAGTTCTCTTTATAGTTAGTGTTCAATTTGCAATGTCTCCTAAA
                                                                           M  S  P  K      4
161      ATAGTTCTTTATCACTTCCCAATAAGCGGACCTTGTCGAGGCGCTCTACTTGCGGCCAGAGCAATCGATGTTCCGGTGGAA
      I  V  L  Y  H  F  P  I  S  G  P  C  R  G  A  L  L  A  A  R  A  I  D  V  P  V  E    31
242      ATTAAAATAGTTGATTTATTTAAAAAAGAACAGCTTAAAGAAGACTTCTTAAAAATTAATCCTCAGCACTGTGTACCGACT
           I  K  I  V  D  L  F  K  K  E  Q  L  K  E  D  F  L  K  I  N  P  Q  H  C  V  P  T    58
323      TTGGACGACGATGGCTTTGTTATTTGGGAGAGCCGCGCGATCGCTTGCTATCTGGCAGATAAGTATGGAAAGGATGATGAA
           L  D  D  D  G  F  V  I  W  E  S  R  A  I  A  C  Y  L  A  D  K  Y  G  K  D  D  E    85
404      CTTTATCCAAAAGATTTGAAACGTCGCGCCATTGTAAATCAGAGATTGCATTTCGATAGCTCGTCTCTGTACGTGAAAATT
           L  Y  P  K  D  L  K  R  R  A  I  V  N  Q  R  L  H  F  D  S  S  S  L  Y  V  K  I   112
485      CGAGCTATTTGTTTCCCAATTCTTTTTCAAGGAGTGACAGAAATTAAAGAGCCTTTAAAGAATGACCTTAATGTCACCCTT
           R  A  I  C  F  P  I  L  F  Q  G  V  T  E  I  K  E  P  L  K  N  D  L [N  V  T] L   139
566      GGATTTTTGGAGGAATTCTTAAAGAGCAGCAAGTGGGTGGCAGGGGACAACCCTACTATAGCAGATACATCCATCTATGCA
           G  F  L  E  E  F  L  K  S  S  K  W  V  A  G  D [N  P  T] I  A  D  T  S  I  Y  A   166
647       TCATTGTCAAGCATTTTGGCTGTGGGCTGGGACATTTCATTATTCCCTAACATCCAGCGTTGGATCAAACTTTGCAGTACA
           S  L  S  S  I  L  A  V  G  W  D  I  S  L  F  P  N  I  Q  R  W  I  K  L  C  S  T   193
728       CTGCCTGGCTATGCTGAAAATGAACAAGGAGCTAAAGAGTTTGGAGAAGCAGTGAAAAAAAATCTCAAGAGTTAATATAAA
           L  P  G  Y  A  E  N  E  Q  G  A  K  E  F  G  E  A  V  K  K  N  L  K  S ***         217
809      ATAGTAGTTATTGCTTAACAGATGTTATCTTATATGTATAACTGAACAATGTTTTTCCGTATAATATAATAAATATACAAT
890      ATTATTACATATAAAACTTCAATACTAATTGCTGTTCATTTCTAACTCTCCTCTTAATTACAACTTCTGACTTATATCCAT
971       AATTGGACATAATTAAAAAAAAAAAAAAAAAAAA
```

Figure 1. Nucleotide and deduced amino acid sequences of *Cf*GSTs4 (**A**) and *Cf*GSTd5 (**B**) cDNAs cloned from *C. fumiferana*. The numbers on the left are for nucleotide sequences while the numbers on the right are for amino acid sequences. The stop codon TAA is underlined with triple stars. The putative polyadenylation signals are double underlined. The primer sequences for RT-PCR are underlined. The putative *O*-glycosylation sites are shaded and the putative *N*-glycosylation site is boxed. The Genbank accession numbers for these sequences are EU912434 for *Cf*GSTs4 (A) and EU912435 for *Cf*GSTd5 (B).

was no potential *N*-glycosylation site in the deduced amino acid sequence of *Cf*GSTs4, whereas two potential *N*-glycosylation sites (NVT at 136 and NPT at 156, Fig. 1B) were predicted in *Cf*GSTd5 by using NetNGlyc 1.0 software (http://www.cbs.dtu.dk/services/NetNGlyc). One potential *O*-β-glycosylation site in *Cf*GSTs4 (Ser45) and two potential *O*-β-glycosylation sites in *Cf*GSTd5 (Ser163 and Ser217) were predicted by using the YinOYang program (http://www.cbs.dtu.dk/services/YinOYang). Functional domain analysis indicated that *Cf*GSTs4 contained GST-Sigma-like domains, while *Cf*GSTd5 contained GST-Delta-Epsilon-like domains at the *N*- and *C*-terminal ends of the sequences. Therefore, they were classified into the Sigma and Delta groups, respectively. According to the GST nomenclature system suggested by Chelvanayagam et al. (2001), these two *C. fumiferana* GSTs are named *Cf*GSTs4 and *Cf*GSTd5, respectively.

Comparison and Phylogenetic Analysis of the C. fumiferana GSTs With Other Insect GSTs

*Cf*GSTs4 and *Cf*GSTd5 sequences were aligned with *Cf*GST (Feng et al., 1999), which is a member of the Sigma GST group and re-named *Cf*GSTs1 correspondingly in this report, and *Cf*GST-2 and *Cf*GST-3 (Zheng et al., 2007), which are members of the Delta group and re-named *Cf*GSTd2 and *Cf*GSTd3 in this study (Fig. 2). Most of the deduced amino acid residues in these cDNAs were variable, while nine amino acids were conserved in all of these GSTs (Fig. 2A). Among the identified *C. fumiferana* GST cDNAs, the lowest percent amino acid identity was 7.9 between *Cf*GSTs1 and *Cf*GSTd3 and the highest identity was 43.3 between *Cf*GSTd5 and *Cf*GSTd3 (Fig. 2B). *Cf*GSTs4 and *Cf*GSTs1 appeared to be clustered together, while *Cf*GSTd5 and *Cf*GSTd3 appeared to be closer to each other in the phylogenetic tree analysis (Fig. 2C). Table 1 summarizes the differences among these five *C. fumiferana* GSTs.

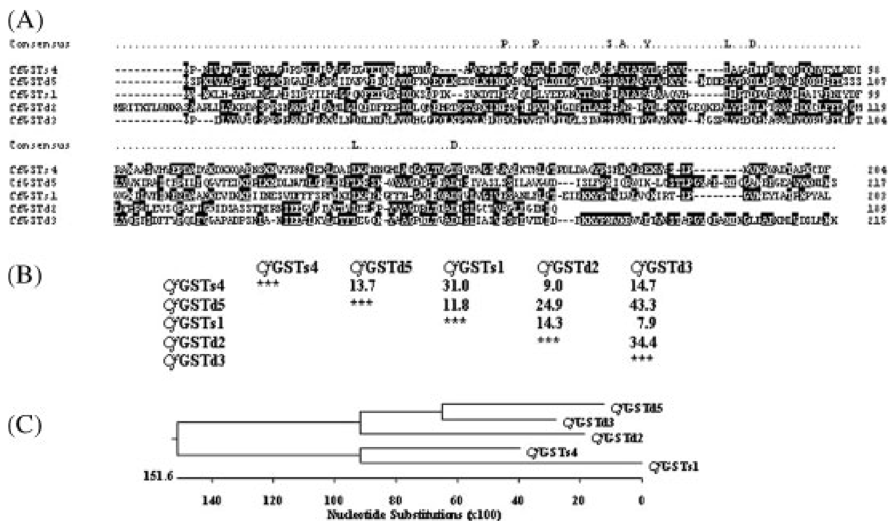

Figure 2. **A:** Amino acid sequence alignment of the five *C. fumiferana* GSTs: *Cf*GSTs1 (Feng et al., 1999), *Cf*GSTd2 and *Cf*GSTd3 (Zheng et al., 2007), *Cf*GSTs4 and *Cf*GSTd5 (this study). Amino acid residues that match each other among the sequences are blocked in black. When amino acid residues in all of the sequences match, they are shown in the consensus shown above the sequences, and otherwise are shown as "." **B:** Amino acid percent identities of the *C. fumiferana* GSTs. **C:** Phylogenetic tree of the *C. fumiferana* GSTs generated using the Clustal Alignment Program (Higgins and Sharp, 1988).

***Table 1. Comparison of the Five GST DNAs Isolated From** C. fumiferana*

	*Cf*GSTs1 (Feng et al., 1999)	NC*f*GSTd2 (Zheng et al., 2007)	*Cf*GSTd3 (Zheng et al., 2007)	*Cf*GSTs4 (this study)	*Cf*GSTd5 (this study)
Nucleotide length (nucleotide no.)	665	804	983	1,128	1,003
Deduced peptide length (amino acid no.)	203	189	215	204	217
Classification	Sigma	Delta	Delta	Sigma	Delta
Protein molecular mass (kDa)	23	21	24	23	24
pI	8.91	5.43	6.62	5.46	7.17
Putative O-glycosylation site	0	1	1	1	2
Putative N-glycosylation site	2	0	1	0	2
mRNA expression in the epidermis	+[a]	+++	+++	++	++
mRNA expression in the fat body	+++	++	+++	+++	+
mRNA expression in the midgut	no	++	+++	++	+
mRNA expression in the CF-203 cells	no	++	+++	+++	+
Responsiveness to RH5992	Yes	Not significant	Not significant	Not significant	Not significant

[a]Relative expression.

Phylogenetic analysis of *C. fumiferana* GSTs with homologues from other insect species indicated that the *C. fumiferana* GSTs distributed to various groups (Fig. 3). *Cf*GSTs4 and *Cf*GSTs1, which contain GST-Sigma-like domains, were clustered into two separate groups, which include Sigma members of the other insect GSTs, while *Cf*GSTs4 was closer to lepidopteran Sigma GST members, such as GST-S of *Bombyx mori* (NP_001037077 and NP_001036994), *Platynota idaeusalis* (AAC34097), *Helicoverpa armigera* (ABU88426), *Plutella xylostella* (BAD26698), and *Manduca sexta* (P46429), while *Cf*GSTs1 was clustered together with dipteran Sigma GSTs, such as *Apis mellifera* (XP_624682) and *Solenopsis invicta* (ABA39530). *Cf*GSTd2, *Cf*GSTd3, and *Cf*GSTd5 contained the Delta domain and separated from the Sigma groups, with *Cf*GSTd2 and *Cf*GSTd3 being more similar to the lepidopteran Delta members, such as BmGST-D (NP_001036974) and PxGST-3 (AAC35245), and *Cf*GSTd5 being more similar to the dipteran Delta and Theta members, such as AaGST-T (*Aedes Aegypti*, XP_001654620) and PpGST (*Phlebotomus paptasi*, ABV44726). It appears that Sigma and Delta members of insect GSTs were evolutionally separated before the separation of lepidopteran and dipteran species, as both Sigma and Delta members could be found in these orders of insects.

Developmental Expression

*Cf*GSTs4 and *Cf*GSTd5 were constantly expressed in the epidermis of 6th instar larvae (Fig. 4A). Expression levels of these two genes in the epidermis appeared to be similar in most stages of 6th instar (up to 156 h post-ecdysis). *Cf*GSTs4 appeared to express constantly in all of the stages, whereas *Cf*GSTd5 had higher expression levels when the larvae were at the HCS and prepupal stages than in the intermolt stages.

In the fat body of 6th instar larvae, transcripts of both *Cf*GSTs4 and *Cf*GSTd5 were also detected (Fig. 4B). *Cf*GSTs4 appeared to express constantly and in higher levels than *Cf*GSTd5 at all times. There was no expression of *Cf*GSTd5 during the early time (intermolt) of the 6th instar stage and the expression level increased when the insects approached near to the molting to pupae stage. This expression pattern was similar to that of *Cf*GSTd2 (Zheng et al., 2007).

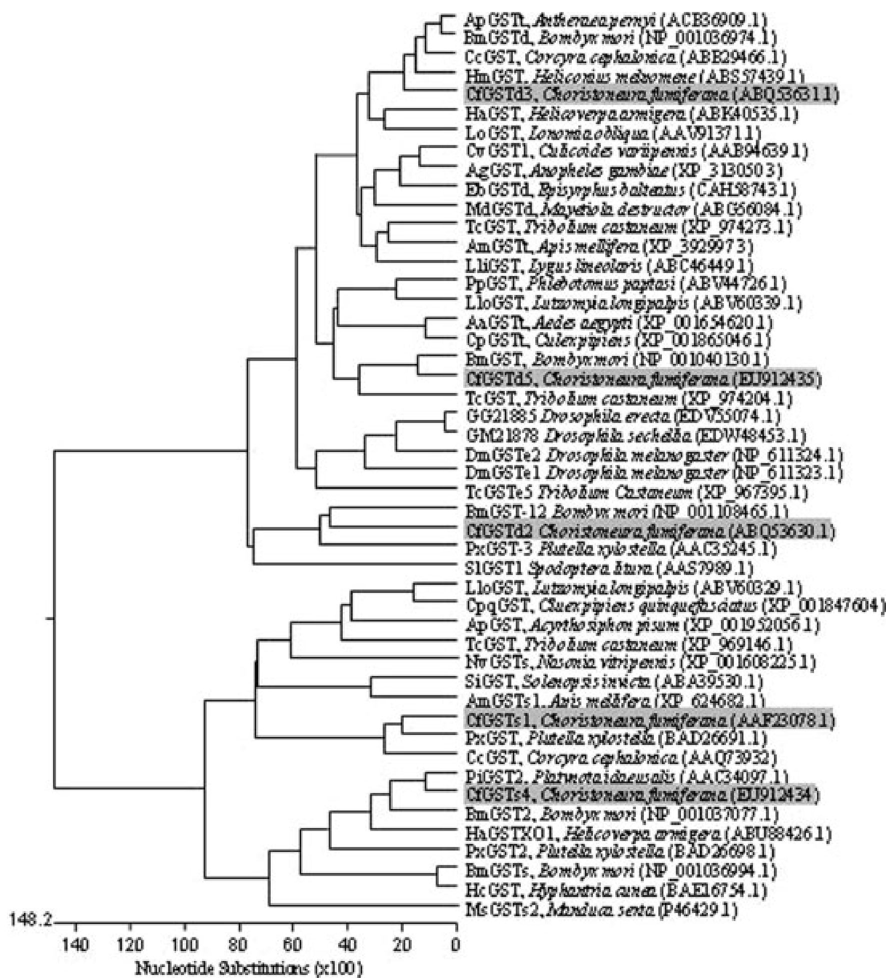

Figure 3. Phylogenetic tree analysis of *C. fumiferana* GSTs with other insect GSTs using the Clustal Alignment Program (Higgins and Sharp, 1988). This phylogenetic analysis was done with the five *C. fumiferana* GSTs and their top ten most homologous sequences in the GenBank.

*Cf*GSTs4 and *Cf*GSTd5 were also detected in the midgut of 6th instar larvae (Fig. 4C). *Cf*GSTs4 appeared to have higher expression levels than *Cf*GSTd5 in most of the stages, particularly during the intermolt.

The developmental expression patterns of these two GSTs indicated that (1) these GSTs were not specifically related to any particular tissues, instead they were present in all tissues; (2) *Cf*GSTd5, like *Cf*GSTd2 (Zheng et al., 2007), had higher expression levels in the fat body when the insects were about to molt; *Cf*GSTs4, like *Cf*GSTd3 (Zheng et al., 2007), was expressed in all of the tissues during the 6th instar larval development.

Transcripts of these two GST genes were also detected using reverse-transcription PCR in the CF-203 midgut cell line in vitro cultured without any treatment (data not shown). Much higher expression levels were detected for *Cf*GSTs4 than for *Cf*GSTd5, which had only a trace expression. This result confirmed the original identification of the transcripts in EST format in CF-203 cells.

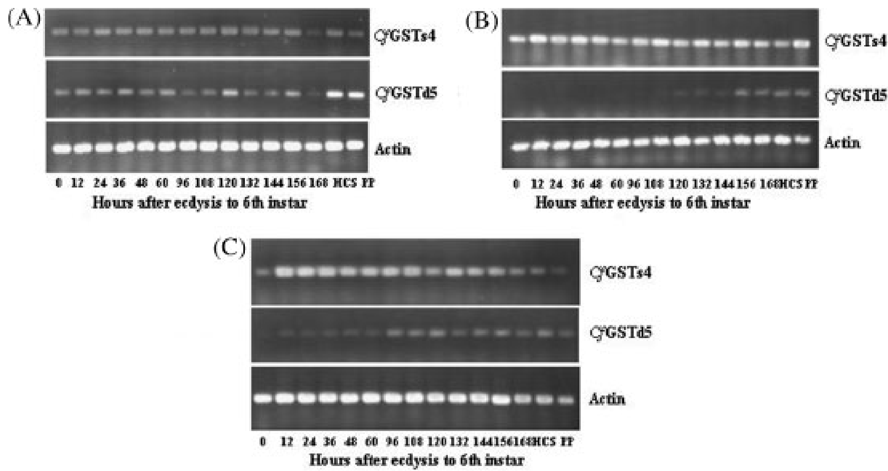

Figure 4. Semi-quantitative RT-PCR analysis of temporal and spatial expression of *Cf*GSTs4 and *Cf*GSTd5 in the epidermis (**A**), fat body (**B**), and midgut (**C**) of 6th instar larvae and pre-pupae. Actin expression is used as a control indicating the equal loading of RNA in the RT-PCR reactions. The expected sizes of the PCR products of *Cf*GSTs4 and *Cf*GSTd5 are 688 and 770 bp, respectively. HCS: head capsule slippage; PP: pre-pupa.

Induced Expression by RH5992

To examine whether or not the expression of these two genes was induced by RH5992, as reported for *Cf*GSTs1 (Feng et al., 2001), we tested the induced expression of *Cf*GSTs4 and *Cf*GSTd5 by applying this synthetic insecticide to 5th instar larvae. The results indicated that *Cf*GSTs4 and *Cf*GSTd5 appeared not to be inducible by RH5992 at the tested concentration of 70 ng/larva in 5th instar larvae (Fig. 5), whereas *Cf*GSTd5 appeared to be slightly but significantly suppressed by RH5992 at 6 h post-treatment (Fig. 5A,C).

In Vitro and In Vivo Protein Expression of CfGSTs4

Whereas in vitro expression of *Cf*GSTd5 was not successful, recombinant *Cf*GSTs4 was obtained through in vitro expression with pET-22b(+) vector in *E. coli* (Fig. 6A). The recombinant protein was produced in the form of in the form of soluble protein and and was approximately 23 kDa in size, the same as the predicted size based on the deduced amino acid sequence of *Cf*GSTs4. The protein was isolated and purified was isolated and purified by multiple centrifugations. Antibody against the purified recombinant protein was generated and showed a high specificity to *Cf*GSTs4 (Fig. 6B). In vivo expression of *Cf*GSTs4 in different tissues was examined by Western blotting analysis (Fig. 5C). Higher levels of the protein were present in the midgut and fat bodies than in the epidermis. This is consistent with the results of transcript expression by RT-PCR (Fig. 4).

DISCUSSION

Three GSTs were previously identified from *C. fumiferana* (Feng et al., 1999; Zheng et al., 2007). Those three *Cf*GSTs1, *Cf*GSTd2, and *Cf*GSTd3, are different in

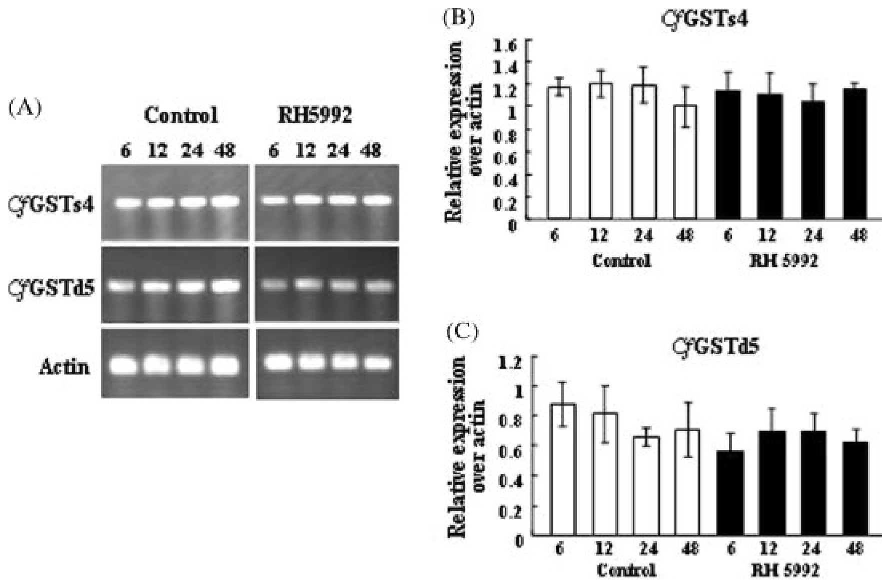

Figure 5. A: RT-PCR analysis of induced expression of *Cf*GSTs4 and *Cf*GSTd5 genes by the ecdysone agonist tebufenozide (RH5992) in 5th instar larvae. Total RNA was extracted from 5th instar larvae at 6, 12, 24, and 48 h post-treatment. Controls were mock-treated and samples were taken at the same times post–mock treatment. Actin expression was used as an internal reference for the loading of RNA in the RT-PCT reactions and non-responsiveness to RH5992. **B,C:** Semiquantitative analysis of three replicates of RT-PCR for *Cf*GSTs4 (B) and *Cf*GSTd5 (C). The PCR products were quantified using ImageQuant TL v2005 (Amersham Inc.). The relative expression of GSTs is calculated and normalized based on the levels of expression of actin.

nucleotide sequences, expression patterns, and responsiveness to the synthetic ecdysone agonist tebufenozide (RH5992). *Cf*GSTs1 is a member of the insect (Sigma) subfamily, while *Cf*GSTd2 and *Cf*GSTd3 belong to the insect (Delta) subfamily. In this study, we identified another two GSTs, *Cf*GSTs4 and *Cf*GSTd5, from the same species. Sequence comparison revealed that *Cf*GSTs4 is closest to *Cf*GSTs1, while *Cf*GSTd5 and *Cf*GSTd3 are more similar to each other. Thus, we have now identified three members (*Cf*GSTd2, *Cf*GSTd3, and *Cf*GSTd5) of the insect Delta subfamily and two members (*Cf*GSTs1 and *Cf*GSTs4) of the insect Sigma subfamily from the same species.

In *M. sexta,* two isoforms (*Ms*GST1 and *Ms*GST2) were identified from the midgut (Snyder et al., 1995). Protein purification and isoelectric focusing analysis revealed that at least 6 different larval midgut GSTs were present in *M. sexta* (Snyder et al., 1995). Four subunits of the cytosolic GSTs were identified in *Orthosia gothica* fed on a native diet and a semisynthetic bean diet (Egaas et al., 1995). One of them is a member of GST1 and the other three are members of GST2. There are at least four active GSTs presenting in *Galleria mellonella* (Baker et al., 1994). Four isozymes were identified in the larval midgut of *Plutella xylostella* (Huang et al., 1998). In *Bombyx mori*, the seven GST genes are found to be in different chromosomes, though expression, protein distribution, and enzymic activity of these genes have not been reported. In other lepdipotera such as *Trichoplusia ni, Heliothis virescens, Annticarsia gemmatalis, Spodoptera frugiperda, Wiseana cervinata*, and *Heliothis zea*, at least one GST isoform has been identified by different approaches. Thus, multiple genes and gene products of GSTs

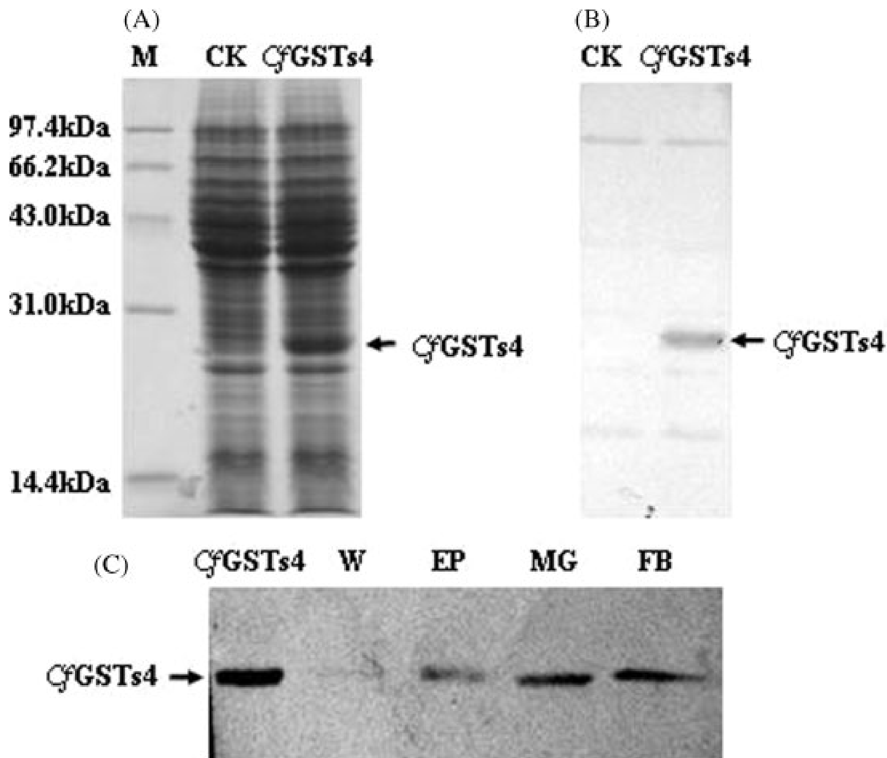

Figure 6. **A:** In vitro expression of *Cf*GSTs4 in *E. coli;* **B:** Western blot analysis of recombinant *Cf*GSTs4 protein using the *Cf*GSTs4 antibody; **C:** Western blot analysis of in vivo expression of *Cf*GSTs4 protein using the *Cf*GSTs4 antibody. Thirty micrograms of total proteins extracted from the tissues were loaded on each of the lanes. Lanes CK: control sample, the proteins isolated from the *E. coli* cells transformed with the pET-22b(+) expression vector alone; Lanes *Cf*GSTs4: the protein isolated from the *E. coli* cells transformed with the pET-22b(+) expression vector containing *Cf*GSTs4; W: whole larvae; EP: epidermis; MG: midgut; FB: fat body.

are present in different tissues of a species, which may reflect the multiple functions of various insect GSTs.

The five *C. fumiferana* GSTs that we cloned and identified appear to be different in nucleotide and protein sequences, as well as gene expression patterns (Table 1). Although *Cf*GSTs1 and *Cf*GSTs4 are apparently closer to each other in peptide structure, *Cf*GSTs1 has a basic pI while *Cf*GSTs4 has an acidic pI. Similarly, *Cf*GSTd2, *Cf*GSTd3, and *Cf*GSTd5 are more similar to each other but *Cf*GSTd5 has a basic pI while both *Cf*GSTd2 and *Cf*GSTd3 have an acidic pI. This may suggest that they play their roles under different physiological conditions for their optimal enzymic activity.

All five of these genes express in the epidermis and fat body of 6th instar larvae. Except for *Cf*GSTs1, which was not detected in the midgut, transcripts of the other four GSTs were found in the midgut and the in vitro cultured midgut cell line. However, immnohistochemistry analysis clearly demonstrated that the proteins of *Cf*GSTs1, *Cf*GSTd2, and *Cf*GSTd3 were mainly in the fat body (Feng et al., 1999; Zheng et al., 2007), although low levels of the proteins were also found in the epidermis. No or only a trace amount of proteins of these three genes was found in the

midgut. This is inconsistent with other reports on lepidopteran GSTs, in which the proteins were found in the midgut (Snyder et al., 1995; Huang et al., 1998). It is not clear whether or not the genes express into mRNA in the midgut and then the proteins are synthesized, stored, and/or play their roles in the fat body. On the other hand, *Cf*GSTs4 had higher levels of the transcript and protein in the fat body and midgut than in the epidermis, whereas *Cf*GSTd5 had higher levels in the epidermis than in the fat body and midgut. *Cf*GSTd5 was also not highly expressed in the CF-203 cell line and at early times post-ecdysis in the midgut and fat body. These results indicate that these GST genes have different expression patterns and, therefore, may play various physiological functions.

Tebufenozide is a synthetic non-steroidal agonist of the insect molting hormone, 20-hydroxyecdysone (20E). But unlike 20E, tebufenozide is selectively toxic to lepidopteran larvae (Dhadialla et al., 1998). Different metabolic fate, ecdysone receptor binding affinity (Dhadialla et al., 1998), and exclusion (Sundaram et al., 1998) of tebufenozide have been suggested to be associated with the selective toxicity. Whether or not different insect GSTs contribute the tebufenozide resistance is not clear. The results of this and previous studies indicate that *C. fumiferana* GSTs show different responsiveness to tebufenozide. *Cf*GSTs1 is significantly inducible by tebufenozide (Feng et al., 2001). At the tested concentration of 10^{-6} M, the expression of *Cf*GSTd2 seemed to be induced by tebufenozide before 12 h post-treatment in 5th instar larvae, followed by a decrease, while the *Cf*GSTd3 expression was increased by tebufenozide after 12 h post-treatment (Zheng et al., 2007). *Cf*GSTs4 and *Cf*GSTd5 appeared not to be inducible by tebufenozide at the tested concentration and instead CfGSTd5 was even slightly suppressed by tebufenozide at 6 h post-treatment. Because they are not induced by tebufenozide but are constantly expressed during larval development, it would be interesting to know why these genes are highly expressed in different tissues and in vitro cells if their major function is not detoxification.

ACKNOWLEDGMENTS

This research was supported in part by Genome Canada through the Ontario Genomics Institute and a start-up fund of South China Normal University.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

Baker WL, Clark AG, Faulds G, Nielsen JS. 1994. Multiple glutathione *S*-transferases in *Galleria mellonella*: their detection with fluorogenic substrates. Insect Biochem Mol Biol 24:301–307.

Chelvanayagam G, Parker MW, Board PG. 2001. Fly fishing for GSTs: a unified nomenclature for mammalian and insect glutathione transferases. Chem Biol Interact 133:256–260.

Chomyczynski P, Sacchi N. 1987. Methods of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem162:156–159.

Clark AG. 1989. The comparative enzymology of the glutathione *S*-transferases from non-vertebrate organisms. Comp Biochem Physiol 92B:419–446.

Dhadialla TS, Carlson GR, Le DP. 1998. New insecticides with ecdysteroidal and juvenile hormone activity. Annu Rev Entomol 43:545–569.

Egaas E, Sandvik M, Svendsen NO, Skaare JU. 1995. The separation and identification of glutathione *S*-transferase subunits from *Orthosia gothica*. Insect Biochem Mol Biol 25:783–788.

Enayati AA, Ranson H, Hemingway J. 2005. Insect glutathione transferases and insecticide resistance. Insect Mol Biol 14:3–8.

Feng QL, Davey KG, Pang AS, Primavera M, Ladd TR, Zheng SC, Sohi SS, Retnakaran A, Palli. SR. 1999. Glutathione *S*-transferase from the spruce budworm, *Choristoneura fumiferana*: identification, characterization, localization, cDNA cloning, and expression. Insect Biochem Mol Biol 29:779–793.

Feng Q, Davey KG, S D Pang A, Ladd TR, Retnakaran A, Tomkins BL, Zheng S, Palli SR. 2001. Developmental expression and stress induction of glutathione *S*-transferase in the spruce budworm, *Choristoneura fumiferana*. J Insect Physiol 47:1–10.

Higgins DG, Sharp PM. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244.

Huang HS, Hu NT, Yao YE, Wu CY, Chiang SW, Sun CN. 1998. Molecular cloning and heterologous expression of a glutathione *S*-transferase involved in insecticide resistance from the diamondback moth, *Plutella xylostella*. Insect Biochem Mol Biol 28:651–658.

Huang X, Maddan A. 1999. CAP3: a DNA sequence assembly program. Genome Res 9: 868–877.

Jakobsson PJ, Mancini JA, Riendeau D, Ford-Hutchinson AW. 1997. Identification and characterization of a novel microsomal enzyme with glutathione-dependent transferase and peroxidase activities. J Biol Chem 272:22934–22939.

McMorran A. 1965. A synthetic diet for the spruce budworm, *Choristoneura fumiferana* (Lepidoptera: Tortricidae). Can Entomol 97:58–62.

Motoyama N, DautermanWC. 1980. Glutathione *S*-transferases: their role in the metabolism of organophosphorus insecticides. Rev Biochem Toxic 2:49–69.

Pemble SE, Wardie AF, Taylar JB. 1996. Glutathione *S*-transferase class Kappa: characterization by the cloning of rat mitochondrial GST and identification of a human homologue. Biochem J 319:749–754.

Retnakaran A, Smith LFR, Tomkins WL, Primavera MJ, Palli SR, Payne NJ, Jobin L. 1997. Effect of RH-5992, a nonsteroidal ecdysone agonist, on the spruce budworm, *Choristoneura fumiferana* (Lepidoptera: Tortricidae): Laboratory, greenhouse, and ground spray trials. Can Entomol 129:871–885.

Sheehan D, Meade G, Foley VM, Dowd CA. 2001. Structure, function and evolution of glutathione transferases: implication for classification of non-mammalian members of an ancient enzyme superfamily. Biochem J. 360:1–16.

Snyder MJ, Walding JK, Feyereisen R. 1995. Glutathione *S*-transferases from larval *Manduca sexta* midgut: sequence of two cDNAs and enzyme induction. Insect Biochem Mol Biol 25:455–465.

Sohi, SS, Lalouette W, MacDonld JA, Gringorten JL, Budau CB. 1993. Establishment of continuous midgut cell lines of spruce budworm (Lepidoptera:Tortricidae). In Vitro Cell Dev Biol 29A:56A.

Sundaram M, Palli SR, Krell PJ, Sohi SS, Dhadialla TS, Retnakaran A. 1998. Basis for selective action of a synthetic molting hormone agonist, RH-5992 on lepidopteran insects. Insect Biochem Mol Biol 28:693–704.

Zheng SC, Deng HM, Ladd T, Tomkins BL, Krell PJ, Feng QL. 2007. Cloning and characterization of two glutathione *S*-transferase cDNAs in the spruce budworm, *Choristoneura fumiferana*. Arch Insect Biochem Physiol 66:146–157.