the **genetics**society

ARTICLE

Horizontally acquired cysteine synthase genes undergo functional divergence in lepidopteran herbivores

Yinghui Li[1] · Yanyan Zhou[1] · Wenhui Jing[1] · Shiliang Xu[1] · Yue Jin[1] · Yusong Xu[1] · Huabing Wang [1]

Abstract
Horizontal gene transfer (HGT) plays an important role in evolutionary processes as organisms adapt to their environments, and now cases of gene duplication after HGT in eukaryotes are emerging at an increasing rate. However, the fate and roles of the duplicated genes over time in eukaryotes remain unclear. Here we conducted a comprehensive analysis of the evolution of cysteine synthase (CYS) in lepidopteran insects. Our results indicate that HGT-derived *CYS* genes are widespread and have undergone duplication following horizontal transfer in many lepidopteran insects. Moreover, lepidopteran CYS proteins not only have β-cyanoalanine synthase activity but also possess cysteine synthase activity that is involved in sulfur amino acid biosynthesis. Duplicated *CYS* genes show marked divergence in gene expression patterns and enzymatic properties, suggesting that they probably have undergone subfunctionalization and/or neofunctionalization in Lepidoptera. The gene transfer of *CYS* genes and subsequent duplication appears to have facilitated the adaptation of lepidopteran insects to different diets and promoted their ecological diversification. Overall, this study provides valuable insights into the ecological and evolutionary contributions of *CYS* in lepidopteran insects.

Introduction

Horizontal gene transfer (HGT), a process in which the exchange of genes occurs between different species, is common among prokaryotes and recognized as one of the major forces in prokaryotic genome evolution (Ochman et al. 2000; Koonin et al. 2001; Pallen and Wren 2007; Polz et al. 2013). Compared with the frequency and level of HGT in prokaryotes, HGT events among eukaryotes are relatively rare (Boto 2014; Suzuki et al. 2015). An accumulation of evidence suggests that some horizontally transferred genes are truly expressed and have novel functions in eukaryotes (Jiggins and Hurst 2011; Acuna et al. 2012; Fan et al. 2020).

Recently, reports of HGT have become increasingly frequent in insects and nematodes. HGT may play an important role in the adaptive evolution of these organisms (Acuna et al. 2012; Wybouw et al. 2016; Husnik and McCutcheon 2018).

Lepidoptera (butterflies and moths) is the second-largest order of insects. Increasingly, incidences of HGT from bacteria and fungi to Lepidoptera are being reported. The majority of HTGs to Lepidoptera have likely provided physiological functions related to nutritional metabolism and detoxification; these functions have in turn facilitated the adaptation of lepidopteran insects to specific plant hosts (Li et al. 2011; Sun et al. 2013). For example, *chitinase* genes acquired through HGT can degrade chitin during molting in lepidopteran insects (Daimon et al. 2005). In addition, *Bombyx mori* β-fructofuranosidase gained through horizontal transfer is related to tolerance to sugar-mimic alkaloids contained in mulberry latex (Daimon et al. 2008; Gan et al. 2018). Consistent with prokaryotes, transferred genes in eukaryotes, including plant-parasitic nematodes (Danchin et al. 2010), aphids (Nikoh et al. 2010; Novakova and Moran 2012), and lepidopterans (Li et al. 2011; Sun et al. 2013), are more frequently duplicated than endogenetic. For example, among 14 genes acquired by HGT in lepidopteran insects, more than half have

✉ Huabing Wang
wanghb@zju.edu.cn

[1] College of Animal Sciences, Zhejiang University, Hangzhou, China

undergone gene duplication (Sun et al. 2013). Paralogous genes can evolve in several ways following duplication events. The majority of duplicate gene copies become nonfunctionalized; however, both paralogs survive in some circumstances, presumably due to neofunctionalization or subfunctionalization (Lee and Irish 2011). Therefore, studying the subsequent fate of these duplicate gene pairs can contribute to understand the role of post-HGT gene duplication in the evolution of eukaryotes.

Cysteine synthase (CYS) is a crucial enzyme involved in sulfur amino acid biosynthesis and exhibit dual functions in plants, bacteria (Feldman-Salit et al. 2009; Lai et al. 2009; Bogicevic et al. 2012; Yi et al. 2012), and nematodes (Budde and Roth 2011). Cysteine synthase are responsible for cysteine synthase (CYS) activity that catalyzes the reaction between H_2S and O-acetylserine (OAS) to generate cysteine, and β-cyanoalanine synthase (CAS) activity that catalyzes the conversion of hydrogen cyanide (HCN) into β-cyanoalanine. A recent study suggested that mites and lepidopteran insects possessed HGT-derived *CYS* genes and the recombinant mite CYS had dual functions (Wybouw et al. 2014). The recombinant CYS enzymes have CAS activity in eight lepidopteran species including *pierids* that feed on cyanogenic host plants and others species with varying diets; these enzymes are essential to lepidopterans being able to thrive on cyanogenic plants (Van Ohlen et al. 2016; Herfurth et al. 2017). However, previous studies were unable to clearly detect CYS activity in lepidopteran insects. Interestingly, *CYS* genes have also been identified in other lepidopteran insects that are not normally feed on cyanogenic plants, including *B. mori* and *Manduca sexta* (Li et al. 2011; Zhu et al. 2011; Sun et al. 2013; Wybouw et al. 2014). Moreover, the duplication of *CYS* genes has been observed in some lepidopteran species that feed on cyanogenic host plants such as *Heliconius melpomene* (Arias et al. 2016), *Spodoptera litura*, and *Pieris rapae* (Sun et al. 2013; Van Ohlen et al. 2016). Nevertheless, the diverse fates of *CYS* duplicates formed by post-HGT duplication remain to be elucidated.

In the present study, we examined the evolutionary fate of *CYS* genes by reconstructing their evolutionary history in lepidopteran insects. In addition, we characterized the functional divergence of duplicated *CYS* genes by examining the gene expression and enzymatic properties. Our results suggest that gene duplication and subsequent functional diversification of *CYS* genes could potentially facilitate the adaptation of lepidopteran insects to their plant hosts. Thus, our study provides valuable insights into the evolution and functional diversification of *CYS* genes in Lepidoptera; specifically, *CYS* genes acquired by HGT have apparent functional and ecological importance in lepidopteran insects.

Materials and methods

Insects material

The N4 strain of *B. mori* were reared on fresh mulberry leaves at 25 °C under standard conditions (14-h light/10-h dark cycle and 70% relative humidity). Larvae of *Spodoptera frugiperda* were reared on fresh corn in a conditioned insect rearing room. *Bombyx mori* and *S. frugiperda* samples collected from different developmental stages, were dissected in phosphate buffered saline (PBS). All samples were stored at −80 °C until used for analysis.

Phylogenetic analysis and conserved synteny analysis of *CYS* genes

To identify more potential CYS paralogs, we downloaded the Hidden Markov Model (HMM) file for the PLP-dependent enzymes superfamily (PALP) from the Pfam database (http://pfam.xfam.org/). Using HMM of PALP domain (PF00291), hmmsearch program from the HMMER package was performed (*e*-value <10e − 5) to identify the putative CYS paralogs on 4 representative lepidopteran protein dataset (*S. frugiperda*, *B. mori*, *P. rapae*, and *Amyelois transitella*) (Eddy 2011). CYS paralogs were then used as queries to perform BLASTp and tBLASTn searches in Lepbase (http://lepbase.org/) and National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/gene). Phylogenetic tree was generated based on the protein sequence alignments in Molecular Evolutionary Genetics Analysis (MEGA, version 7.0) using neighbor-joining (NJ) method under JTT model (Edgar 2004; Kumar et al. 2016) (Fig. S1). To further investigate the evolutionary relationships of HGT-derived *CYS* genes, homologs of *CYS* were identified by BLASTp and tBLASTn searches in Lepbase and NCBI with BmorCYS (XP_004932948.1), PxylCYS (XP_011554718.1), MspCYS (EIZ83823.1), and TurtCYS (XP_015786551.1) used as queries (Wybouw et al. 2014). The species names, protein names, and accession numbers were presented in Table S1. All protein sequences were first aligned in the MEGA7 program using MUSCLE with the default settings (Edgar 2004; Kumar et al. 2016). We then inspected for regions of high-quality alignment and manually checked for misalignments. The best amino acid substitution model selected by ProtTest was LG + G (Darriba et al. 2011). Phylogenetic trees were generated based on the protein sequence alignments in MEGA7 using Maximum Likelihood method (ML) under LG + G protein model and NJ method under JTT model. The final trees were visualized using the Interactive Tree of Life (iTOL) online tool (http://itol.embl.de).

The syntenic relationships of *CYS* genes and their flanking genes in lepidopteran and mite genomes were

manually analyzed in NCBI genome database using *Prap-CYS1* as the query gene. We analyzed genes flanking *CYS* gene and manually checked the results by performing reciprocal BLAST searches.

PCR amplification to exclude the possibility of bacterial contamination

Genomic DNA was extracted using DNA Isolation Mini Kit (Vazyme, Nanjing, China). Primer pairs were designed to amplify a 9400 bp-long genomic fragment of *BmorCYS* and its downstream *esterase FE4-like* (*FE4*) gene as well as their intergenic region on the *B. mori* genome. Briefly, the 20 μl reaction volume comprised 10 μl PCR Super Mix (Vazyme, Nanjing, China), 0.8 μl of each 10 μmol/L primer, and 1 μl template DNA. Amplification was performed using the following conditions: pre-denaturation at 94 °C for 3 min, followed by 35 cycles each with 94 °C for 10 s, 68 °C for 10 min. The PCR run ended with a post-extension step at 68 °C for 7 min.

Expression analysis by quantitative real-time PCR (qRT-PCR)

To obtain more detailed expression data of *CYS* genes in *B. mori* and *S. frugiperda*, we analyzed various tissues from different development stages: the third day of the last larval instar (L5D3), the eighth day after pupation (P8), and the first day of moth (A0). RNA was isolated with TRIzol reagent (Takara, Dalian, China) and quantified using a Nanodrop spectrophotometer (Thermo Fisher, USA). cDNA was synthesized using Primer Script RT Reagent Kit (Takara, Dalian, China).

qRT-PCRs were performed with specific primers as described previously (Gao et al. 2020). The cycling parameters were used as follows: 95 °C for 1 min, followed by 35 cycles at 95 °C for 15 s and 60 °C for 30 s. We analyzed three independent biological samples and calculated relative *CYS* expression levels against *B. mori* ribosomal protein L49 gene (*BmorRpl49*) or *S. frugiperda* glyceraldehyde 3-phosphate dehydrogenase gene (*SfruGAPDH*) using the $2^{-\Delta\Delta CT}$ method. The data were presented as means ± standard error of the mean. Statistical analysis between groups were performed using the one-way analysis of variance (ANOVA) on SPSS 20.0 software. Different letters over the bars showed that the expression levels between groups are statistically significant ($P \leq 0.05$).

Protein purification and determination of recombinant CYS proteins

BmorCYS, *SfruCYS1*, and *SfruCYS2* were respectively cloned into the corresponding site of a pCold vector (Takara, Dalian, China). Primers were listed in Table S2. *Escherichia coli* carrying pCold-*BmorCYS*, pCold-*Sfru-CYS1*, and pCold-*SfruCYS2* were cultured in Luria–Bertani medium at 37 °C. The recombinant proteins were induced by isopropyl-β-D-thiogalactoside until the absorbance at 600 nm reaches approximately 0.6. After incubation for 15 h at 18 °C, 220 rpm, the cells were harvested by centrifugation at 6000 × g, 4 °C for 5 min. The pellets were suspended in cold PBS, followed by sonication. The lysates were centrifuged at 10,000 × g for 30 min. After centrifugation, the supernatant was collected and purified as previously described (Zhou et al. 2019). Briefly, the supernatant was loaded onto a Ni-nitrilotriacetic acid (NTA) column. The column was washed with washing buffer consisting of 0.5 M NaCl, 20 mM sodium phosphate, and 20 mM imidazole, pH 7.4. Then, the CYS proteins were eluted with a stepwise gradient of imidazole (up to 500 mM) in washing buffer. The eluate was dialyzed against 20 mM sodium phosphate buffer (pH 7.4). Subsequently, recombinant BmorCYS, SfruCYS1, and SfruCYS2 were stored at 4 °C or −80 °C until use.

Preparation of antiserum and western blotting analysis

Purified proteins were used to generate rabbit polyclonal antibody according to previous study (Lee et al. 2016). Western blotting analysis was performed as previously described (Dai et al. 2019). Proteins were extracted from seven different tissues of L5D3 using RIPA lysis buffer (FDbio Science, China) and then separated by 10% SDS-PAGE before being transferred to polyvinylidene fluoride membranes (Millipore, USA). After blocking with 5% nonfat milk for 2 h, the membranes were incubated overnight with polyclonal rabbit anti-CYS immunoglobulin G. Subsequently, the membranes were incubated with a florescence-labeled secondary antibody (DingGuo ChangSheng Biotechnology Co., Ltd., China) for 2 h. The protein level of tubulin was used as a loading control. Western blotting signals were detected using an ECL Plus Kit (FDbio Science, China).

Identifying the subcellular localization of CYS proteins

A subcellular localization assay was performed as previously described (Zhou et al. 2019). In brief, *BmorCYS*, *SfruCYS1*, and *SfruCYS2* ORFs were cloned into a pIZ/V5-mCherry vector using specific primers (Table S2). Positive plasmids were transfected into Sf9 cells using Lipo8000™ Transfection Reagent (Beyotime, China). After 24 h of incubation, the nuclei were stained with Hochest 33258. The cells were then washed with PBS and
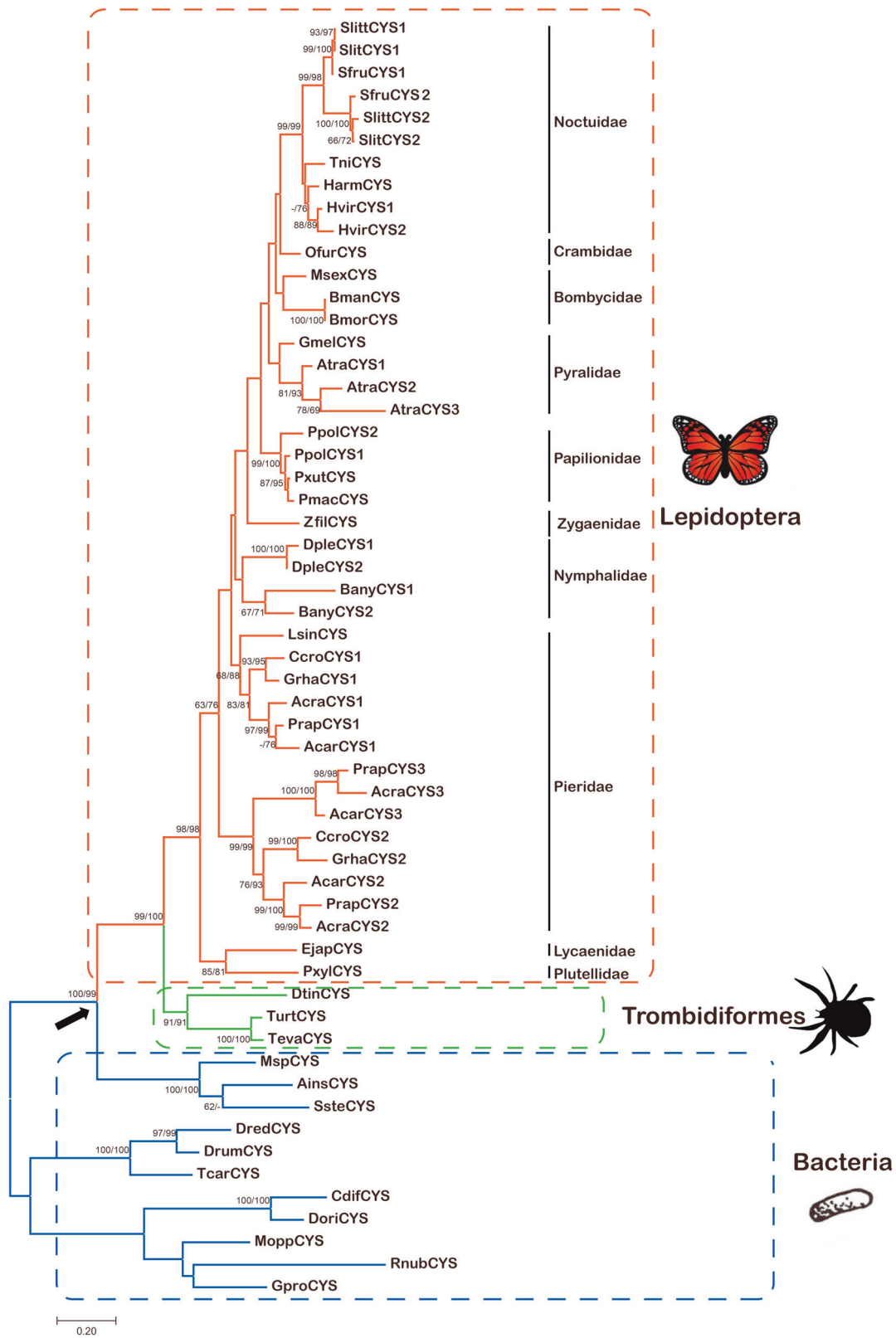
Fig. 1 Phylogenetic analysis of CYSs. The CYS sequences from bacteria, mites, and Lepidoptera are marked with different colors. Bootstrap values over 60% are indicated above the nodes, with the number on the right indicating NJ and that on the left indicating ML. Moreover, the bootstrap value at the branching point of arthropod and bacterial CYS homologs is 100/99 as indicated by the black arrow.

observed using a Zeiss LSM780 laser scanning confocal microscope.

Enzyme assays

Recombinant CYS proteins were incubated at 30 °C for 10 min in the reaction buffer containing 500 μM pyridoxal-5′-phosphate before measuring enzyme activity. The CYS activity assays were performed based on the method of Lunn (Lunn et al. 1990), with 1 μg of purified recombinant BmorCYS, SfruCYS1, or SfruCYS2 protein used per respective reaction. Standard substrate concentrations were 10 mM and 5 mM for OAS and sodium sulfide, respectively. The reaction product cysteine was quantified using the method of Gaitonde (Gaitonde 1967; Wybouw et al. 2014). The CAS activity assay were performed based on the method of Hendrickson (Hendrickson and Conn 1969). The standard reaction was performed with 500 μl of 12 mM cysteine, 500 μl of 12 mM potassium cyanide (KCN), and 1 μg recombinant CYS protein. These reactions were performed in a 1.5 mL microcentrifuge tubes with sealed caps. Cyanide solution was prepared by adding solid KCN immediately before use. After incubation at 30 °C for 10 min, CAS activity was quantified by measuring products H_2S with adding 250 μl 30 mM $FeCl_3$, and 250 μl 20 mM N, N-dimethyl-p-phenylendiamindihydrochloride. The absorbance was determined by spectrophotometrically at 650 nm.

To determine the optimum pH for CYS activity, 20 μl of reaction solution (containing purified CYS protein, 10 mM OAS, 5 mM sodium sulfide, and 20 mM Britton–Robinson's wide range buffer of pH 3–13) was incubated at 30 °C for 15 min. The product cysteine was determined as described above (Lunn et al. 1990). Various concentrations of OAS, sodium sulfide, and KCN were used to determine the kinetic parameters of the recombinant CYS proteins. Substrate affinity constants were determined by one-site-specific binding with the Hill slope using GraphPad Prism7 software. Each kinetic parameter, such as V_{max} and K_m, for OAS, sodium sulfide and KCN were derived using the Michaelis–Menten equation. GraphPad Prism7 was used to conduct statistical analysis. All experiments were repeated three times.

Results

Duplication of CYS genes occurred independently after the divergence of many lepidopteran species

To identify more potential CYS paralogs, we performed HMM and BLAST searches. The phylogenetic tree was then constructed using MEGA7. The phylogenetic trees indicated that CYS paralogs were clustered into three distinct, highly supported groups: Group I Cystathionine β-synthase, Group II Cysteine synthase, and Group III L-threonine ammonia lyase (Fig. S1). These are all members of pyridoxal-5′-phosphate (PLP) dependent enzyme family. 16 lepidopteran CYS homologs found within Group II tended to cluster with mite and bacterial sequences, suggesting that these CYS homologs within Group II might have been acquired from bacteria via horizontal gene transfer.

To further identify HGT-derived CYS homologs, we performed BLASTp and tBLASTn searches in the Lepbase and NCBI databases. In total, 57 genes encoding putative CYS were identified in lepidopteran, mite, and bacterial genomes. We found multiple putative CYS copies in several lepidopteran genomes from which 42 lepidopteran CYS proteins contained typical CBS-like domain and one lepidopteran CYS contained CysK domain, suggesting that they were members of the cysteine synthase family (Fig. S2). In addition, with the exception of M. sexta, all CYS genes lacked introns among 26 lepidopteran species (Fig. S2). Moreover, lepidopteran CYS genes had similar exon–intron structures to previously characterized members of the CYS gene family in mites and bacteria. To exclude the possibility that lepidopteran CYS genes were derived from contaminating bacterial sequences, we examined its position in the B. mori genome. The sequencing results showed that CYS genes were located on B. mori genome at downstream of FE4 gene, suggesting the CYS genes were not bacterial contamination (Figs. S3 and S4).

To further investigate the evolutionary relationships of HGT-derived CYS, we conducted an amino acid-based alignment of representative lepidopteran, mite, and bacterial CYS homologs and then constructed phylogenetic trees using the ML and NJ methods. ML and NJ methods yielded similar topologies and the phylogenetic tree reconstructed by ML was shown in Fig. 1. Phylogenetic analysis showed that arthropod CYS sequences form a highly supported cluster (bootstrap = 100/99) with a group of bacteria as their closest relatives; this suggests that the ancestors of the mite and lepidopteran species included in the phylogenetic tree acquired a CYS gene from bacteria via HGT. Despite some internal nodes being weakly supported within the lepidopteran CYS clade, most nodes of duplication events were supported with high confidence in our analysis (>80%; Fig. 1); thus, multiple homologous copies of CYS apparently exist in the studied lepidopteran species. In addition, the CYS homologs of several lepidopterans were clustered together by species, suggesting that duplication of CYS genes occurred independently within lepidopteran species. For instance, CYS1 and CYS2 paralogs of Heliothis virescens and A. transitella were reciprocally monophyletic. Moreover, CYS1 and CYS2 of Spodoptera clustered together within the species, suggesting that these proteins were derived from duplication before Spodoptera
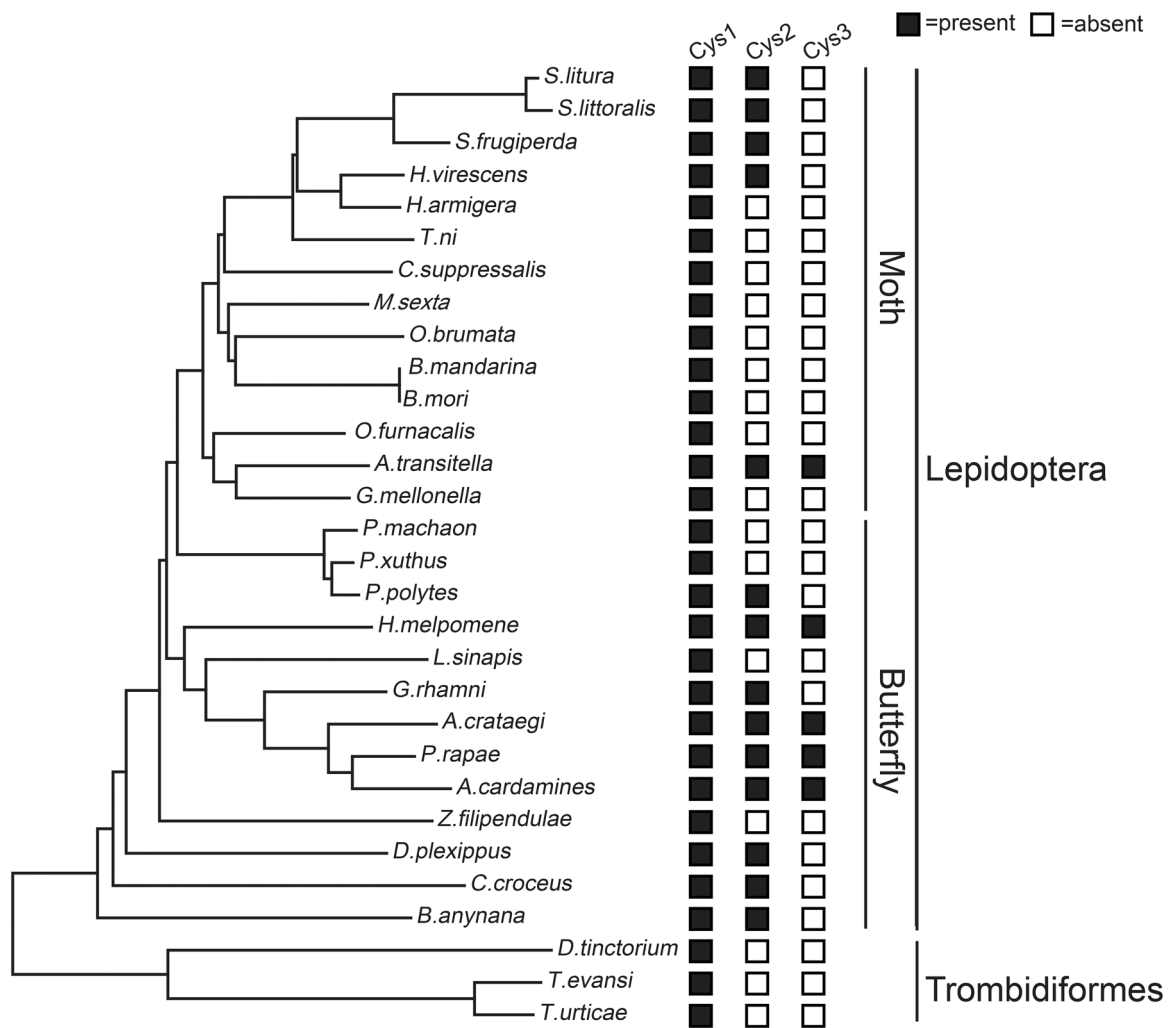
Fig. 2 **Identification of *CYS* copy number in mites and Lepidoptera.** The presence (black box) or absence (white box) of each *CYS* paralogs is detected by BLASTp and tBLASTn. This analysis indicated that there are multiple copies of *CYS* in many lepidopteran insects.

speciation. In *D. plexippus* and *Papilio polytes*, CYS1 and CYS2 paralogs were reciprocally monophyletic with high confidence. CYS paralogs of *Pierid* species were clustered into three groups, showing that gene duplication may have occurred in a common ancestor of these species. We also found that *CYS* duplication occurred in some polyphagous insects (such as *P. rapae*, *Spodoptera littoralis*, and *S. frugiperda*) that feed on cyanogenic plants (Fig. 2). Taken together, these results indicate that HGT-derived *CYS* genes are widespread and have undergone further duplications in many lepidopteran insects.

Genomic organization and synteny of *CYS* genes

To understand the evolutionary history of the *CYS* gene family, we investigated *CYS* genes and their adjacent genes in the genome of *P. rapae*, which had three *CYS*

paralogs in distinct genomic loci (Fig. 3B). *PrapCYS1* was located downstream of the *FE4* gene, whereas *PrapCYS2* and *PrapCYS3* were found in a distinct syntenic location.

We also extended our syntenic analyses to other lepidopteran *CYS* genes. In the examined species, *CYS*2 genes were in the same syntenic location and shared the *FE4* gene as their 5′ neighbor gene (Fig. 3A). In addition, we also identified two *CYS* paralogs located tandemly on the same scaffold in several lepidopteran insects. The presence of many lepidopteran *CYS* genes in a shared syntenic location across species indicates that they were likely orthologous. Finally, we examined *CYS* genes in mites; they were not adjacent to *FE4*, *histidine-rich glycoprotein-like* (*HRG*), or *protein extra-macrochaetae* genes (*Emc*) (Fig. 3C), suggesting that the *CYS* locus of mites was not syntenic with that of lepidopteran insects.
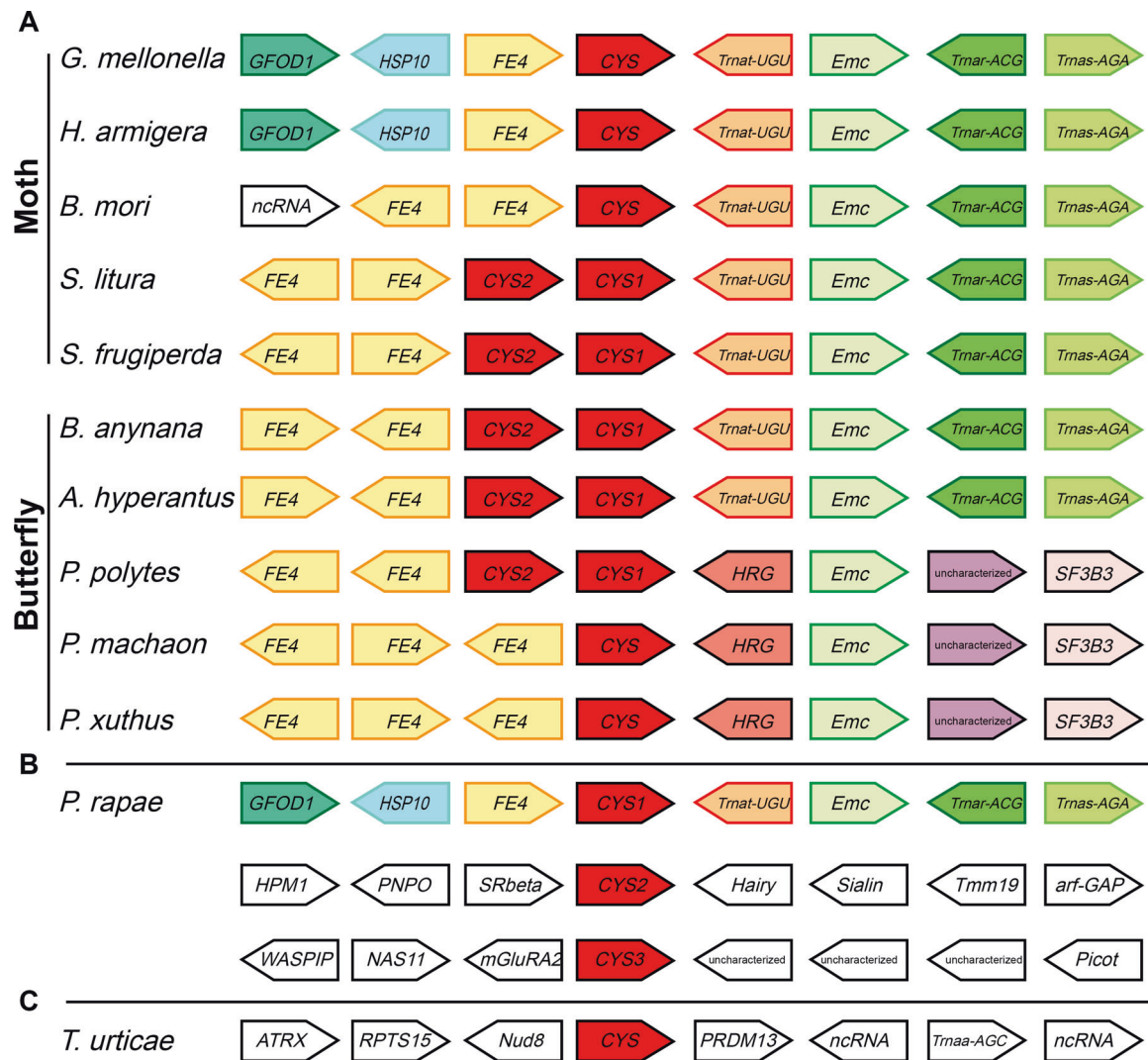
Fig. 3 Synteny of genes flanking *CYS* in mite and 11 lepidopteran chromosomes. Synteny was analyzed using previously published genome sequences from the NCBI database in (**A**) 10 lepidopteran insects, (**B**) *P. rapae*, and (**C**) mite. Genes and their positions are not drawn to scale and shown as the series of colored arrows. Homologs are represented by the same color, and genes with no homolog were indicated by white arrows. Arrows point in the sense direction of the gene. Insect *CYS*s are placed in the middle and colored in red.

Functionally important amino acid residues in cysteine synthase are conserved in lepidopteran orthologs

Enzymes that catalyze the formation of cysteine are highly conserved in plants and bacteria; they have a conserved lysine residue and two domains: a pyridoxal-5′-phosphate (PLP)-binding domain and an OAS-binding domain (Burkhard et al. 1998; Bonner et al. 2005). To evaluate the nature of lepidopteran *CYS* orthologs, we identified functionally important amino acid residues in lepidopteran CYS proteins based on previous studies (Bonner et al. 2005; Wybouw et al. 2014). Such residues, which were responsible for the molecular interaction between the PLP cofactor

(H182, G206, T207, T210, T296) and substrate OAS (T91, S92, G93, N94, Q95, Q172), were conserved among lepidopteran CYS proteins; the exception was A/S/G208, which had different residues in *Eumeta japonica*, *Plutella xylostella*, *Pieridae*, and *Spodoptera* species (Figs. S5 and S6). Of note, the functionally important amino acids of lepidopteran CYS proteins were conserved.

Expression pattern and subcellular localizations of CYS proteins in *B. mori* and *S. frugiperda*

The spatial and temporal expression data for a corresponding gene should be useful for determining its biological or physiological role (Doxey et al. 2007).
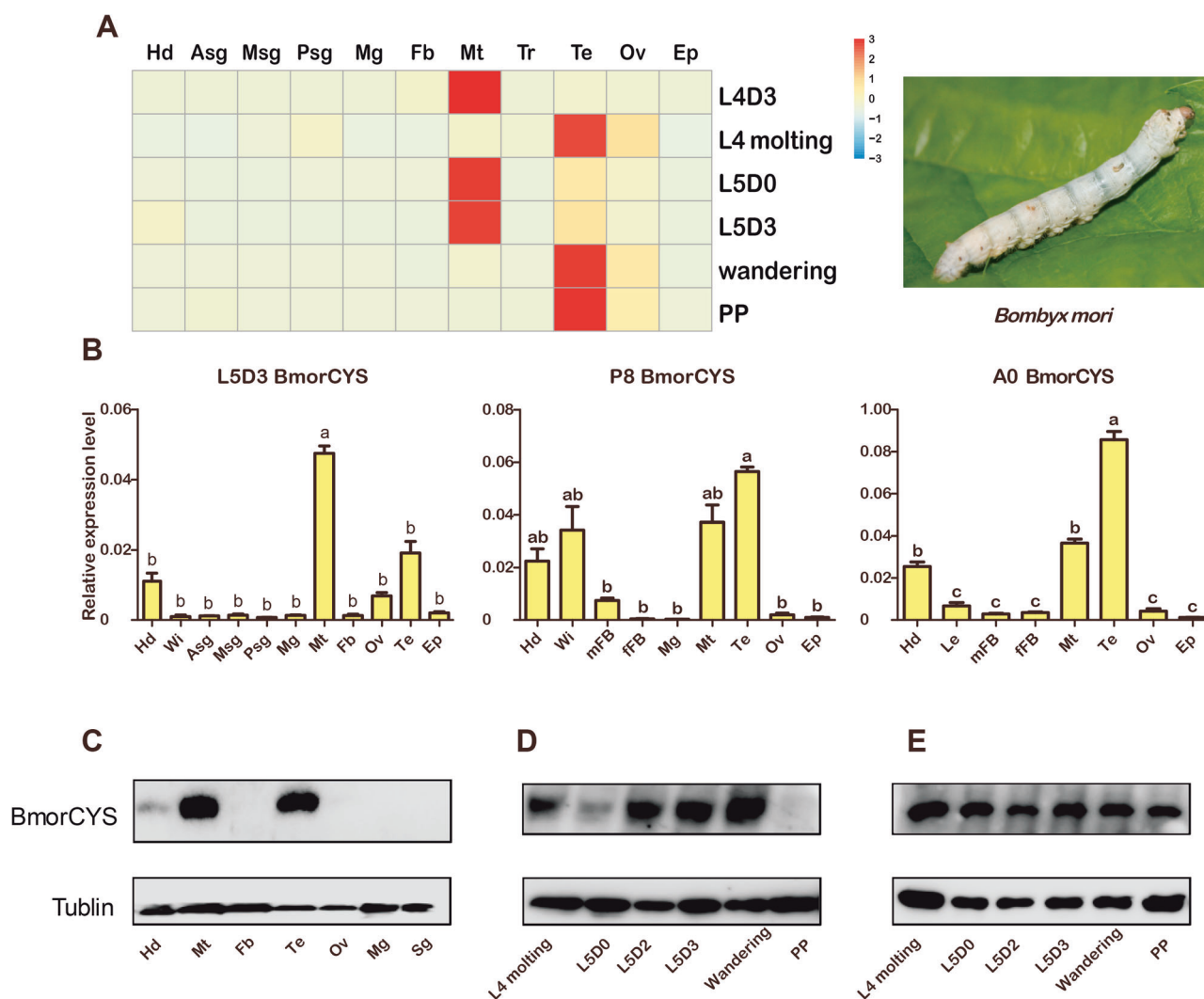
Fig. 4 Expression profiles of *CYS* in *Bombyx mori*. A Heat map for expression patterns of *CYS* gene in different stages and tissues of *B. mori*. **B** qRT-PCR assays of *BmorCYS*. **C** Protein expression analysis of BmorCYS in different tissues of L5D3 larvae. **D** Protein expression analysis of BmorCYS in Malpighian tubule. **E** Protein expression analysis of BmorCYS in testis. L4D3 day 3 of fourth-instar larvae, L4 molting fourth larval molting, L5D0 newly emerged fifth-instar larvae, L5D3 day 3 of fifth-instar larvae, PP pre-pupa, Ep epidermis, Mt Malpighian tubule, mFb male fat body, fFb female fat body, Ov ovary, Te testis, Mg midgut, Sg silk gland, Asg anterior silk gland, Msg middle silk gland, Psg posterior silk gland, Wi wing disc, Hd head, Tr Trachea.

Bombyx mori was the first lepidopteran insect for which a genome was fully sequenced and has been extensively studied (Goldsmith et al. 2005). To explore the expression profiles of *CYS*, we investigated the expression of *BmorCYS* from SilkDB 3.0 (http://silkworm.swu.edu.cn/silkdb/) in various tissues and at different development stages (Fig. 4A). *BmorCYS* was expressed at relatively high levels in larval Malpighian tubule and testis. Transcriptional levels of *BmorCYS* were highest in the Malpighian tubule during feeding stages, whereas levels dropped dramatically during molting, wandering, and prepupal stages. In contrast, *BmorCYS* was expressed in the testis at all larval stages. To verify our findings and obtain more detailed expression data, we dissected larva, pupae, and adult tissues from *B. mori* and investigated the expression of *BmorCYS* (Fig. 4B). In general, qRT-PCR analysis suggested that *BmorCYS* was expressed at relatively high levels in the larval Malpighian tubule. Moreover, *BmorCYS* was expressed in pupal testis, wing, and Malpighian tubule, as well as adult testis, Malpighian tubule, and legs. These results suggest that *BmorCYS* shows diverse and dynamic expression patterns in multiple tissues.

We also investigated the expression of BmorCYS protein on the third day of fifth-instar larvae (Fig. 4C). Western blotting analysis revealed that BmorCYS was expressed in the Malpighian tubule and testis but not in the fat body, midgut, and silk gland, which was consistent with the mRNA expression level of *BmorCYS*. Moreover, BmorCYS protein was detected in the Malpighian tubule during

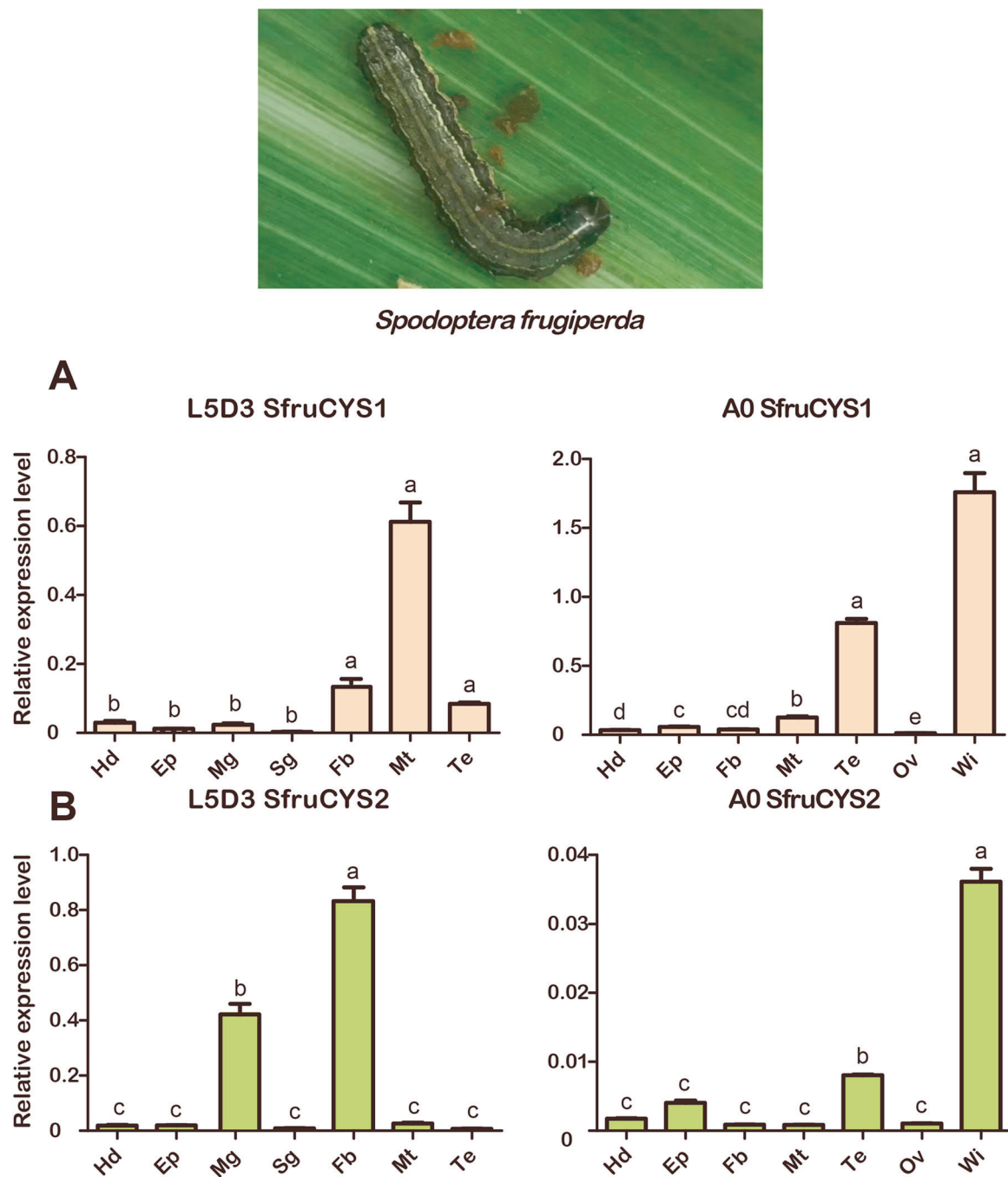**Fig. 5 Expression profiles of *CYS* in *Spodoptera frugiperda*.** qRT-PCR assays of *SfruCYS1* and *SfruCYS2*. Ep epidermis, Te testis, Ov ovary, Mt Malpighian tubule, Sg silk gland, Fb fat body, Wi wing disc, Mg midgut, Hd head.

molting, the newly emerged fifth-instar, day-2 and day-3 fifth-instar larval, and wandering stages (Fig. 4D). Bmor-CYS was expressed in the testis at all larval stages examined (Fig. 4E). Thus, dynamic expression analyses revealed that BmorCYS was stably expressed in the testis but variable in the Malpighian tubule.

Duplicated gene copies often differ in their expression patterns (Farre and Alba 2010). To test whether such

differences exist for *CYS* paralogous genes, we extended our expression analyses to another lepidopteran species, *S. frugiperda*, containing duplicate *CYS* genes and compared gene expression patterns of duplicate gene pairs. We found that two *SfruCYS* genes were significantly differentially expressed in larval tissues (Fig. 5). *Sfru-CYS*1 was highly expressed in the Malpighian tubule, fat body, and testis. *SfruCYS*2 was highly expressed in the

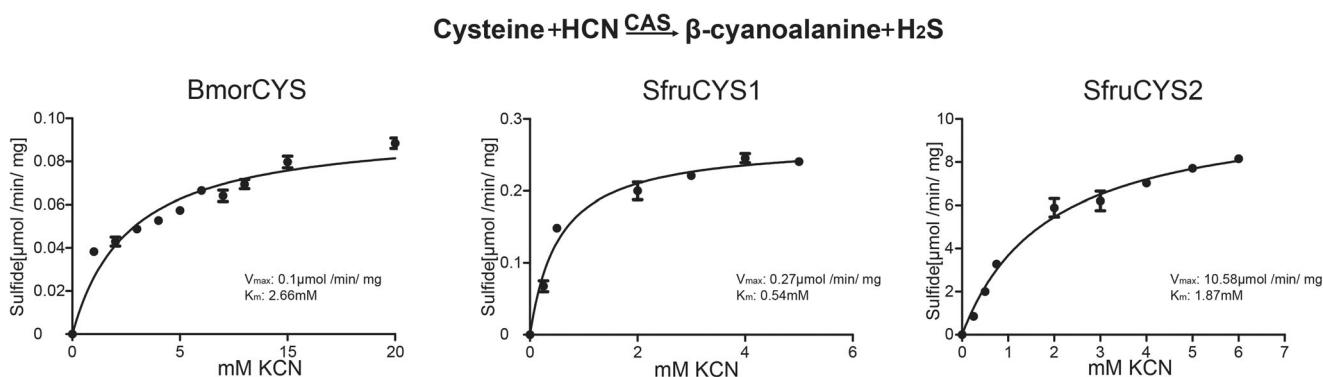**Cysteine +HCN $\xrightarrow{\text{CAS}}$ β-cyanoalanine+H2S**

Fig. 6 The CAS activity analysis of BmorCYS/SfruCYS1/SfruCYS2. Michaelis–Menten analysis of recombinant BmorCYS/SfruCYS1/SfruCYS2. Various concentrations of KCN was used to determine the kinetic parameters of the recombinant CYSs.

larval fat body and midgut, whereas low expression of *SfruCYS2* was observed in the Malpighian tubule, suggesting that the expression pattern of *CYS* genes had diverged in *S. frugiperda*. *SfruCYS1* and *SfruCYS2* were similarly expressed in the adult stage: expression levels were high in the wing and testis. Moreover, the expression of *SfruCYS1* was similar to that of *BmorCYS*, which had high expression levels in the Malpighian tubule during the feeding stage.

To identify the subcellular locations of lepidopteran CYS proteins, pIZ-mCherry-BmorCYS, pIZ-mCherry-SfruCYS1, and pIZ-mCherry-SfruCYS2 were transfected into Sf9 cells, which encode transiently expressed mCherry-BmorCYS, mCherry-SfruCYS1, and mCherry-SfruCYS2 fusion proteins, respectively (Fig. S7). Strong signals from the three proteins were detected in the cytoplasm of Sf9 cells, indicating that *B. mori* and *S. frugiperda* CYS proteins had similar subcellular localizations in the cytoplasm.

Biochemical divergence of lepidopteran CYS proteins in CYS and CAS activity

In order to identify whether lepidopteran CYS have CYS activity and CAS activity, we expressed and purified recombinant BmorCYS, SfruCYS1, and SfruCYS2 proteins (Fig. S8). Subsequent biochemical assays suggested that three purified recombinant proteins had CAS and CYS activity (Figs. 6 and 7). The CYS/CAS ratio of the specificity constants was 2.75, 0.3, and 0.03 for BmorCYS, SfruCYS1, and SfruCYS2, respectively (Tables S3 and S4). This suggested that CYS activity of BmorCYS was higher than the CAS activity, whereas CAS activity was higher for SfruCYS2. The optimum pH of BmorCYS, SfruCYS1, and SfruCYS2 for cysteine synthesis activity was 9, 10, and 9, respectively. Moreover, the pH curve was broad for all CYS enzymes, with ~50% of the maximum activity occurring over pH 7–11 (Fig. 7B). This finding suggests that lepidopteran CYS proteins might have adapted to a highly alkaline pH environment.

To determine whether lepidopteran CYS enzymes are biochemically diverged, we compared previously published data (Van Ohlen et al. 2016; Herfurth et al. 2017) and the enzyme activity of BmorCYS and SfruCYSs. For CAS activity, K_m and V_{max} values of CYS3 for KCN were one to two orders of magnitude higher than CYS2 and CYS1 among homologs from the same species in *Pieridae*. K_m and V_{max} values of CYS2 for KCN were one order of magnitude higher than CYS1 among homologs from the same species in *Noctuidae* (Fig. S9). In addition, the CAS activity of BmorCYS was low. For CYS activity, we found no significant difference in K_m values between SfruCYS1 and SfruCYS2. However, SfruCYS2 had a higher V_{max} than SfruCYS1 (Table S3). Taken together, our results suggested lepidopteran CYS enzymes have undergone divergence in CAS and CYS activities.

Discussion

In this study, we provided a comprehensive analysis of the evolution of the *CYS* family in Lepidoptera. We showed that *CYS* genes underwent further duplications in many lepidopteran insects, and duplicated *CYS* genes showed marked divergence in gene expression patterns and enzymatic properties. Lepidopteran CYSs not only have β-cyanoalanine synthase activity but also possess cysteine synthase activity; these processes may facilitate the adaptation of lepidopteran insects to various diets and diverse habitats. Thus, our findings provide valuable insights into the function and fate of lepidopteran *CYS* genes.

CYS genes may have horizontally transferred to common ancestor of Lepidoptera (Wybouw et al. 2014). Interestingly, *CYS* gene is widespread and has been maintained as a functional gene in several lepidopteran insects. Recombinant CYS enzymes apparently have β-cyanoalanine synthase activity in several lepidopteran insects, which is essential to thrive on cyanogenic plants (Stauber et al. 2012;
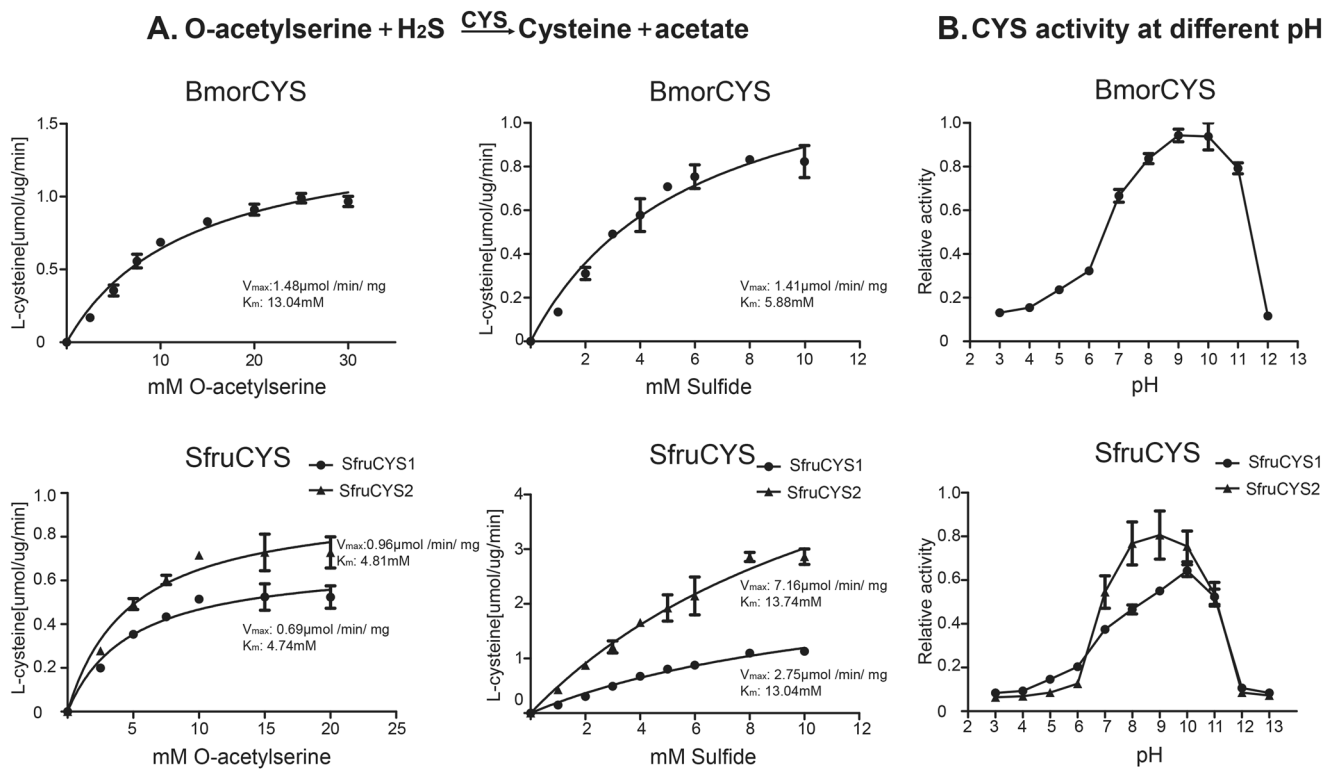
A. O-acetylserine + H_2S $\xrightarrow{CYS}$ Cysteine + acetate

B. CYS activity at different pH

Fig. 7 The CYS activity analysis of BmorCYS/SfruCYS1/SfruCYS2. **A** Michaelis–Menten analysis of recombinant BmorCYS/SfruCYS1/SfruCYS2. Various concentrations of OAS and sodium sulfide were used to determine the kinetic parameters of the recombinant CYSs. **B** pH profile of recombinant BmorCYS/SfruCYS1/SfruCYS2. The pH–activity relationship was determined using 10 mM OAS and 5 mM sodium sulfide as the substrate over a pH range of 3–13. Recombinant CYSs (1 μg) were incubated for 15 min at 30 °C. Cysteine was quantified by measuring the absorbance at 560 nm. These experiments were repeated three independent times.

Van Ohlen et al. 2016). In addition, *CYS* genes were observed in other lepidopteran insects that do not feed on cyanogenic host plants. It is unclear why these *CYS* genes have been retained in these organisms. We showed that lepidopteran CYS proteins not only have β-cyanoalanine synthase activities but also possess cysteine synthase activities. Both of these activities have also been reported in CYS proteins of plants, bacteria, and mites (Yamaguchi et al. 2000; Wada et al. 2004; Bogicevic et al. 2012; Wybouw et al. 2014). Moreover, CYS activity of BmorCYS was higher than the CAS activity, suggesting that lepidopteran CYSs are not only related to detoxification but also possess cysteine synthase activity that is involved in sulfur amino acid biosynthesis. Cysteine is the least abundant amino acid in plant leaves, which may be growth limiting for herbivorous insects (Barbehenn et al. 2013a; Barbehenn et al. 2013b). CYS activity may increase the levels of cysteine to promote insect's growth. Moreover, the synthesis of glutathione (GSH) is tightly regulated and highly dependent on the availability of cysteine (Aoyama et al. 2008). A large fraction (20%) of total cysteine content is used to produce GSH (Jeschke et al. 2016), which may exert an anti-oxidant effect as a metabolic precursor of GSH and play roles in detoxification of specific plant secondary

metabolites to facilitate organisms adaptation to host plants (Wadleigh and Yu 1988; Schramm et al. 2012). These might be one of the reasons why these *CYS* genes acquired from bacteria have been retained in these organisms that are not feeding on cyanogenic host plants. In addition, previous studies suggested that insects either obtain cysteine directly from their diet or synthesize it from methionine (Jeschke et al. 2016). Our results provide evidence for an alternative cysteine biosynthesis route in Lepidoptera: they can synthesize cysteine independently from methionine by CYS activity (Fig. 8).

CYS gene has undergone further duplications in many lepidopteran insects. 2-3 *CYS* paralogs were found in the butterflies *D. plexippus*, *H. melpomene*, and *P. rapae* (Sun et al. 2013). However, the number of *CYS* paralogs has yet to be identified in other lepidopteran genomes. Of 27 lepidopteran species examined, we found that 14 species have 2-3 copies of *CYS* genes. Interestingly, we found that *CYS* duplication occurred in some polyphagous insects (such as *P. rapae, S. littoralis,* and *S. frugiperda*) that feed on cyanogenic plants, which may be due to challenges in the dietary conditions required to metabolize plant secondary metabolites. The large number genes found in many insect lineages perhaps reflect their enormous ecological success
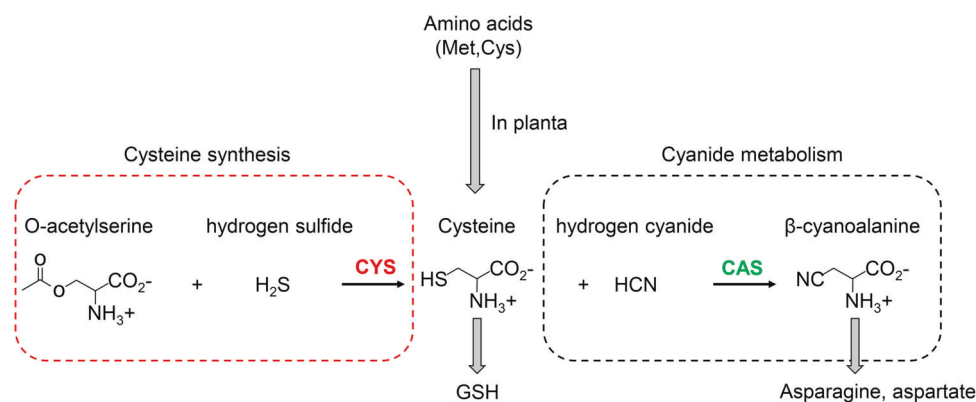
Fig. 8 Summary of the synthesis and metabolism of cysteine in lepidopteran larvae. Insects can obtain cysteine directly from their diet or via synthesis from methionine. Our results provide an alternative cysteine biosynthesis route in Lepidoptera: synthesis of cysteine from OAS and sodium sulfide by CYS activity (marked by a red box). Other reactions involving the lepidopteran *CYS* gene are marked by a black box. CAS detoxifies cyanide by incorporating it into cysteine to form β-cyanoalanine, which is further metabolized into asparagine and aspartate. In addition, cysteine can be used to produce GSH, which plays a role in detoxifying specific plant secondary metabolites.

and diversity (Helmkampf et al. 2015; Rane et al. 2016). For instance, *P. rapae* is a specialist herbivore that feeds on cyanogenic plants, whereas *Spodoptera* species, such as *S. littoralis* and *S. frugiperda*, are major polyphagous pests that infest >80 economically important plant species from 40 different families (Brown and Dewhurst 2009; Baloch et al. 2020). These insect species are exposed to very different plant secondary metabolites, and the expansion of *CYS* genes may allow them to handle the broad range of secondary metabolites (Schramm et al. 2012).

Over the course of genome evolution, functional divergence is the most likely explanation for the retention of duplicate genes. Divergence in expression patterns has been reported for various functional duplicated genes among multicellular species (Gu et al. 2002; Wagner 2002; Gu et al. 2005; Ganko et al. 2007; Farre and Alba 2010; Liu et al. 2015; Leite et al. 2018; Zhou et al. 2019). In the present study, pronounced divergence between the spatial expression patterns of *S. frugiperda* CYS paralogs was observed, suggesting that the *CYS* paralogs have undergone subfunctionalization or neofunctionalization. Previous studies demonstrated that the transcript levels of *PrapCYS1* were highest in gut tissue, which was likely the major site of cyanide liberation upon ingestion of cyanogenic plants. *PrapCYS1* has β-cyanoalanine synthase activity and plays an important role in cyanide detoxification (Van Ohlen et al. 2016). Our results show that *SfruCYS2* was highly expressed in the larval midgut and the recombinant SfruCYS2 also showed higher CAS enzyme activity. In contrast, *SfruCYS1* and *BmorCYS* were highly expressed in the Malpighian tubule, but showed relatively low expression in the midgut. These results suggest that BmorCYS and SfruCYS1 were likely involved in diuresis, detoxification, and metabolism in larval Malpighian tubule. Early expression divergence between duplicated genes may be a phenomenon that

reduces the chances that one paralog becomes a pseudogene (Force et al. 1999; Lynch and Force 2000). In addition, lepidopteran *CYS* duplicate genes showed different enzymatic activities. For example, the kinetic values of *Pieridae* CYS3 and *Noctuidae* CYS2 for KCN was higher than that of other CYSs from the same species in *Pieridae* and *Noctuidae*, indicating divergence in their biochemical properties.

The evolutionary mechanisms responsible for the retention and subsequent functional divergence of duplicate genes after HGT events are poorly understood. In this study, we reconstructed the evolutionary history of HGT-derived *CYS* genes in lepidopteran genomes, supporting hypothesis that HGT events might have occurred in mite ancestor and lepidopteran ancestor, respectively (Wybouw et al. 2014). By examining the gene expression and the enzymatic properties, we revealed the evolutionary and functional dynamics of *CYS* genes. Lepidopteran *CYS* genes were differentially expressed in a tissue-specific manner, which presumably accelerates functional diversification to optimally regulate cysteine biosynthesis and cyanide detoxification. Our investigation provides valuable insights into the function and fate of lepidopteran *CYS* genes. This study also provides an example of how gene transfer and gene duplication can contribute to the adaptation and ecological success of insects.

Data availability

qRT-PCR data is available from the Dryad Digital Repository https://doi.org/10.5061/dryad.pzgmsbckj.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Acuna R, Padilla BE, Florez-Ramos CP, Rubio JD, Herrera JC, Benavides P et al. (2012) Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. Proc Natl Acad Sci USA 109(11):4197–4202

Aoyama K, Watabe M, Nakaki T (2008) Regulation of neuronal glutathione synthesis. J Pharm Sci 108(3):227–238

Arias M, Meichanetzoglou A, Elias M, Rosser N, de-Silva DL, Nay B et al. (2016) Variation in cyanogenic compounds concentration within a *Heliconius* butterfly community: does mimicry explain everything? BMC Evol Biol 16(1):272

Baloch MN, Fan JY, Haseeb M, Zhang RZ (2020) Mapping potential distribution of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in central Asia. Insects 11(3):172

Barbehenn RV, Kochmanski J, Menachem B, Poirier LM (2013a) Allocation of cysteine for glutathione production in caterpillars with different antioxidant defense strategies: a comparison of *Lymantria dispar* and *Malacosoma disstria*. Arch Insect Biochem 84(2):90–103

Barbehenn RV, Niewiadomski J, Kochmanski J (2013b) Importance of protein quality versus quantity in alternative host plants for a leaf-feeding insect. Oecologia 173(1):1–12

Bogicevic B, Berthoud H, Portmann R, Meile L, Irmler S (2012) CysK from *Lactobacillus casei* encodes a protein with *O*-acetylserine sulfhydrylase and cysteine desulfurization activity. Appl Microbiol Biot 94(5):1209–1220

Bonner ER, Cahoon RE, Knapke SM, Jez JM (2005) Molecular basis of cysteine biosynthesis in plants: structural and functional analysis of *O*-acetylserine sulfhydrylase from *Arabidopsis thaliana*. J Biol Chem 280(46):38803–38813

Boto L (2014) Horizontal gene transfer in the acquisition of novel traits by metazoans. Proc Biol Sci 281(1777):20131834

Brown ES, Dewhurst CF (2009) The genus *Spodoptera* (Lepidoptera, Noctuidae) in Africa and the Near East. Bull Entomological Res 65(2):221–262

Budde MW, Roth MB (2011) The response of *Caenorhabditis elegans* to hydrogen sulfide and hydrogen cyanide. Genetics 189(2):521–532

Burkhard P, Rao GS, Hohenester E, Schnackerz KD, Cook PF, Jansonius JN (1998) Three-dimensional structure of *O*-acetylserine sulfhydrylase from *Salmonella typhimurium*. J Mol Biol 283(1):121–133

Dai X, Li R, Li X, Liang Y, Gao Y, Xu Y et al. (2019) Gene duplication and subsequent functional diversification of sucrose hydrolase in *Papilio xuthus*. Insect Mol Biol 28(6):862–872

Daimon T, Katsuma S, Iwanaga M, Kang WK, Shimada T (2005) The BmChi-h gene, a bacterial-type chitinase gene of *Bombyx mori*, encodes a functional exochitinase that plays a role in the chitin degradation during the molting process. Insect Biochem Mol Biol 35(10):1112–1123

Daimon T, Taguchi T, Meng Y, Katsuma S, Mita K, Shimada T (2008) Beta-fructofuranosidase genes of the silkworm, *Bombyx mori* - Insights into enzymatic adaptation of *B. mori* to toxic alkaloids in mulberry latex. J Biol Chem 283(22):15271–15279

Danchin EG, Rosso MN, Vieira P, de Almeida-Engler J, Coutinho PM, Henrissat B et al. (2010) Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. Proc Natl Acad Sci USA 107(41):17651–17656

Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27(8):1164–1165

Doxey AC, Yaish MWF, Moffatt BA, Griffith M, McConkey BJ (2007) Functional divergence in the *Arabidopsis* beta-1,3-glucanase gene family inferred by phylogenetic reconstruction of expression states. Mol Biol Evol 24(4):1045–1055

Eddy SR (2011) Accelerated profile HMM searches. PLoS Comput Biol 7(10):e1002195

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. Bmc Bioinforma 5:1–19

Fan X, Qiu H, Han W, Wang Y, Xu D, Zhang X et al. (2020) Phytoplankton pangenome reveals extensive prokaryotic horizontal gene transfer of diverse functions. Sci Adv 6(18):eaba0111

Farre D, Alba MM (2010) Heterogeneous patterns of gene-expression diversification in mammalian gene duplicates. Mol Biol Evol 27(2):325–335

Feldman-Salit A, Wirtz M, Hell R, Wade RC (2009) A mechanistic model of the cysteine synthase complex. J Mol Biol 386(1):37–59

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerate mutations. Genetics 151(4):1531–1545

Gaitonde MK (1967) A spectrophotometric method for direct determination of cysteine in presence of other naturally occurring amino acids. Biochem J 104(2):627–633

Gan Q, Zhang XW, Zhang DB, Shi L, Zhou Y, Sun TT et al. (2018) BmSUC1 is essential for glycometabolism modulation in the silkworm, *Bombyx mori*. BBA-Gene Regul Mech 1861(6):543–553

Ganko EW, Meyers BC, Vision TJ (2007) Divergence in expression between duplicated genes in *Arabidopsis*. Mol Biol Evol 24(10):2298–2309

Gao Y, Liu YC, Jia SZ, Liang YT, Tang Y, Xu YS et al. (2020) Imaginal disc growth factor maintains cuticle structure and controls melanization in the spot pattern formation of *Bombyx mori*. PLoS Genet 16(9):e1008980

Goldsmith MR, Shimada T, Abe H (2005) The genetics and genomics of the silkworm, *Bombyx mori*. Annu Rev Entomol 50:71–100

Gu X, Zhang ZQ, Huang W (2005) Rapid evolution of expression and regulatory divergences after yeast gene duplication. Proc Natl Acad Sci USA 102(3):707–712

Gu ZL, Nicolae D, Lu HHS, Li WH (2002) Rapid divergence in expression between duplicate genes inferred from microarray data. Trends Genet 18(12):609–613

Helmkampf M, Cash E, Gadau J (2015) Evolution of the insect desaturase gene family with an emphasis on social Hymenoptera. Mol Biol Evol 32(2):456–471

Hendrickson HR, Conn EE (1969) Cyanide metabolism in higher plants. IV. Purification and properties of the beta-cyanolanine synthase of blue lupine. J Biol Chem 244(10):2632–2640

Herfurth AM, van Ohlen M, Wittstock U (2017) Beta-cyanoalanine synthases and their possible role in *Pierid* host plant adaptation. Insects 8(2):62

Husnik F, McCutcheon JP (2018) Functional horizontal gene transfer from bacteria to eukaryotes. Nat Rev Microbiol 16(2):67–79

Jeschke V, Gershenzon J, Vassao DG (2016) A mode of action of glucosinolate-derived isothiocyanates: detoxification depletes glutathione and cysteine levels with ramifications on protein metabolism in *Spodoptera littoralis*. Insect Biochem Molec 71:37–48

Jiggins FM, Hurst GD (2011) Microbiology. Rapid insect evolution by symbiont transfer. Science 332(6026):185–186

Koonin EV, Makarova KS, Aravind L (2001) Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol 55:709–742

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874

Lai KW, Yau CP, Tse YC, Jiang LW, Yip WK (2009) Heterologous expression analyses of rice OsCAS in *Arabidopsis* and in yeast provide evidence for its roles in cyanide detoxification rather than in cysteine synthesis in vivo. J Exp Bot 60(3):993–1008

Lee BS, Huang JS, Jayathilaka LP, Lee J, Gupta S (2016) Antibody production with synthetic peptides. Methods Mol Biol 1474:25–47

Lee HL, Irish VF (2011) Gene duplication and loss in a MADS box gene transcription factor circuit. Mol Biol Evol 28(12): 3367–3380

Leite DJ, Baudouin-Gonzalez L, Iwasaki-Yokozawa S, Lozano-Fernandez J, Turetzek N, Akiyama-Oda Y et al. (2018) Homeobox gene duplication and divergence in arachnids. Mol Biol Evol 35(9):2240–2253

Li ZW, Shen YH, Xiang ZH, Zhang Z (2011) Pathogen-origin horizontally transferred genes contribute to the evolution of lepidopteran insects. Bmc Evolut Biol 11(1):356

Liu HJ, Tang ZX, Han XM, Yang ZL, Zhang FM, Yang HL 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(11): 2844–2859

Lunn JE, Droux M, Martin J, Douce R (1990) Localization of atp sulfurylase and O-acetylserine(Thiol)lyase in spinach leaves. Plant Physiol 94(3):1345–1352

Lynch M, Force A (2000) The probability of duplicate gene preservation by subfunctionalization. Genetics 154(1):459–473

Nikoh N, McCutcheon JP, Kudo T, Miyagishima S, Moran NA, Nakabachi A (2010) Bacterial genes in the aphid genome: absence of functional gene transfer from *Buchnera* to its host. Plos Genet 6(2):e1000827

Novakova E, Moran NA (2012) Diversification of genes for carotenoid biosynthesis in aphids following an ancient transfer from a fungus. Mol Biol Evol 29(1):313–323

Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405(6784):299–304

Pallen MJ, Wren BW (2007) Bacterial pathogenomics. Nature 449 (7164):835–842

Polz MF, Alm EJ, Hanage WP (2013) Horizontal gene transfer and the evolution of bacterial and archaeal population structure. Trends Genet 29(3):170–175

Rane RV, Walsh TK, Pearce SL, Jermiin LS, Gordon KH, Richards S et al. (2016) Are feeding preferences and insecticide resistance associated with the size of detoxifying enzyme families in insect herbivores? Curr Opin Insect Sci 13:70–76

Schramm K, Vassao DG, Reichelt M, Gershenzon J, Wittstock U (2012) Metabolism of glucosinolate-derived isothiocyanates to glutathione conjugates in generalist lepidopteran herbivores. Insect Biochem Molec 42(3):174–182

Stauber EJ, Kuczka P, van Ohlen M, Vogt B, Janowitz T, Piotrowski M et al. (2012) Turning the 'mustard oil bomb' into a 'cyanide bomb': aromatic glucosinolate metabolism in a specialist insect herbivore. Plos One 7(4):e35545

Sun BF, Xiao JH, He SM, Liu L, Murphy RW, Huang DW (2013) Multiple ancient horizontal gene transfers and duplications in lepidopteran species. Insect Mol Biol 22(1):72–87

Suzuki K, Moriguchi K, Yamamoto S (2015) Horizontal DNA transfer from bacteria to eukaryotes and a lesson from experimental transfers. Res Microbiol 166(10):753–763

Van Ohlen M, Herfurth AM, Kerbstadt H, Wittstock U (2016) Cyanide detoxification in an insect herbivore: molecular identification of beta-cyanoalanine synthases from *Pieris rapae*. Insect Biochem Molec 70:99–110

Wada M, Awano N, Yamazawa H, Takagi H, Nakamori S (2004) Purification and characterization of O-acetylserine sulfhydrylase of *Corynebacterium glutamicum*. Biosci Biotech Bioch 68(7): 1581–1583

Wadleigh RW, Yu SJ (1988) Detoxification of isothiocyanate allelochemicals by glutathione transferase in three lepidopterous species. J Chem Ecol 14(4):1279–1288

Wagner A (2002) Asymmetric functional divergence of duplicate genes in yeast. Mol Biol Evol 19(10):1760–1768

Wybouw N, Dermauw W, Tirry L, Stevens C, Grbic M, Feyereisen R et al. (2014) A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. Elife 3: e02365

Wybouw N, Pauchet Y, Heckel DG, Van Leeuwen T (2016) Horizontal gene transfer contributes to the evolution of arthropod herbivory. Genome Biol Evol 8(6):1785–1801

Yamaguchi Y, Nakamura T, Kusano T, Sano H (2000) Three arabidopsis genes encoding proteins with differential activities for cysteine synthase and beta-cyanoalanine synthase. Plant Cell Physiol 41(4):465–476

Yi H, Juergens M, Jez JM (2012) Structure of soybean beta-cyanoalanine synthase and the molecular basis for cyanide detoxification in plants. Plant Cell 24(6):2696–2706

Zhou YY, Li XT, Katsuma S, Xu YS, Shi LG, Shimada T et al. (2019) Duplication and diversification of trehalase confers evolutionary advantages on lepidopteran insects. Mol Ecol 28(24):5282–5298

Zhu B, Lou MM, Xie GL, Zhang GQ, Zhou XP, Li B et al. (2011) Horizontal gene transfer in silkworm, *Bombyx mori*. Bmc Genomics 12:248