



## OPEN

SUBJECT AREAS:  
PHYLOGENETICS  
GENE EXPRESSIONReceived  
28 November 2014Accepted  
4 February 2015Published  
10 March 2015Correspondence and  
requests for materials  
should be addressed to  
M.Y. (msyou@iae.fjau.  
edu.cn)

# Characterization and expression of the cytochrome P450 gene family in diamondback moth, *Plutella xylostella* (L.)

Liying Yu<sup>1,2,3</sup>, Weiqi Tang<sup>1,2</sup>, Weiyi He<sup>1,3</sup>, Xiaoli Ma<sup>1,3</sup>, Liette Vasseur<sup>1,4</sup>, Simon W. Baxter<sup>1,5</sup>, Guang Yang<sup>1,3</sup>, Shiguo Huang<sup>1,3</sup>, Fengqin Song<sup>1,2,3</sup> & Minsheng You<sup>1,3</sup>

<sup>1</sup>Institute of Applied Ecology, Fujian Agriculture and Forestry University, Fuzhou 350002, China, <sup>2</sup>Faculty of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China, <sup>3</sup>Key Laboratory of Integrated Pest Management for Fujian-Taiwan Crops, Ministry of Agriculture, Fuzhou 350002, China, <sup>4</sup>Department of Biological Sciences, Brock University, St. Catharines, Ontario, Canada, <sup>5</sup>School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, Australia.

Cytochrome P450 monooxygenases are present in almost all organisms and can play vital roles in hormone regulation, metabolism of xenobiotics and in biosynthesis or inactivation of endogenous compounds. In the present study, a genome-wide approach was used to identify and analyze the P450 gene family of diamondback moth, *Plutella xylostella*, a destructive worldwide pest of cruciferous crops. We identified 85 putative cytochrome P450 genes from the *P. xylostella* genome, including 84 functional genes and 1 pseudogene. These genes were classified into 26 families and 52 subfamilies. A phylogenetic tree constructed with three additional insect species shows extensive gene expansions of *P. xylostella* P450 genes from clans 3 and 4. Gene expression of cytochrome P450s was quantified across multiple developmental stages (egg, larva, pupa and adult) and tissues (head and midgut) using *P. xylostella* strains susceptible or resistant to insecticides chlorpyrifos and fiprinol. Expression of the lepidopteran specific CYP367s predominantly occurred in head tissue suggesting a role in either olfaction or detoxification. CYP340s with abundant transposable elements and relatively high expression in the midgut probably contribute to the detoxification of insecticides or plant toxins in *P. xylostella*. This study will facilitate future functional studies of the *P. xylostella* P450s in detoxification.

The cytochrome P450 monooxygenases (P450s) are a large and complex gene superfamily of heme-thiolate proteins that are found in almost all living organisms<sup>1</sup>. Insect P450s can be divided into four clans: the CYP2, CYP3, CYP4 and mitochondrial P450s, however, the number of genes among insect species can vary considerably (48 in *Apis mellifera* and 164 in *Aedes aegypti*)<sup>3</sup>. P450s play many functional roles in insect growth and development, including biosynthesis of hormones as well as inactivation and metabolism of xenobiotic compounds such as pesticides<sup>4–6</sup>.

Mechanisms of insecticide resistance are commonly associated with one or more detoxification genes, including P450s, esterases and glutathione S-transferases<sup>7–10</sup>. For example, overexpression of CYP6BQ9 in *Tribolium castaneum* brain tissue caused the majority of deltamethrin resistance, possibly through metabolizing the insecticide prior to it binding with the voltage-gated sodium channel receptor<sup>11</sup> and overexpression of CYP6G1 in *Drosophila melanogaster*, caused resistance to DDT and imidacloprid<sup>12</sup>. Pyrethroid resistance has also been documented through P450 duplications<sup>13</sup> and through the formation of a chimeric P450 gene, generated by unequal crossing-over between two parental genes<sup>14</sup>.

The diamondback moth (DBM), *Plutella xylostella* (L.) (Lepidoptera: Plutellidae), is a pest of Brassicaceae plants with a worldwide distribution. It causes damage to economically important crops, such as canola and cruciferous vegetables, and can undertake seasonal migration and reproduce rapidly (with a life as short as 14 days). DBM has the ability to quickly develop resistance to insecticides<sup>15–17</sup>, and populations have developed resistance to all classes of insecticide, including the bio-insecticide Bt produced by the bacteria *Bacillus thuringiensis*<sup>18</sup>.

Several studies have characterized individual DBM P450 genes, including CYP9G2<sup>19</sup>, CYP6BG1<sup>20</sup> and CYP6BF1<sup>21</sup>, and examined their association with insecticide resistance. Using next generation sequencing, Lin *et al.*<sup>22</sup> compared differential gene expression between three DBM strains with low, moderate and high levels of resistance to chlorantraniliprole. Compared to a susceptible reference strain, 189 unigenes were overexpressed, 19 of which were P450 metabolic detoxification enzymes. They suggested that these P450 genes play an important



role in resistance, as they were differentially expressed in all three resistant strains, while alterations in expression of target site ryanodine receptors were down-regulated in highly resistant strains. Understanding the mechanisms involved in insect resistance development is challenging as the metabolic pathway appears to contain several enzyme subfamilies.

With the rapid development of advanced next generation DNA sequencing technologies, genome-wide investigation of the P450 superfamily in some insect species has improved the identification of the genes responsible for resistance, their expression patterns and biological implications, and our understanding of their phylogenetic evolution<sup>2,23</sup>. From these analyses, we have been able to know that *Bombyx mori* has 84 CYP450s that are classified into 26 families and 47 subfamilies<sup>2</sup>. As well, 90 P450s have been identified in *D. melanogaster*<sup>24</sup>, 204 P450s in *Culex quinquefasciatus*<sup>23</sup> and 112 P450s in *Anopheles gambiae*<sup>2</sup>.

Transcriptomic analysis of DBM previously<sup>25</sup> identified 137 putative P450 gene fragments, from a set of 84,570 unigenes. In this study, we characterized a complete set of full length DBM P450s sequences (hereafter *PxP450s*) using genomic<sup>26</sup> and transcriptomic<sup>27</sup> datasets, and analyzed their phylogenetic relationships with other insects. We also investigated the differential expression of DBM P450 genes in function of the biology at its various life stages and among different strains susceptible or resistant to insecticides. These results provide a basis for further genetic studies on the biological function of P450s in DBM, particularly in relation to insecticide resistance.

## Results and discussion

**Identification and characteristics of *PxP450s*.** A total of 7,180 cytochrome P450 protein sequences from 24 insect species were downloaded from the Cytochrome P450 Homepage<sup>28</sup> and 18,701 predicted DBM peptide-coding sequences (CDS) were downloaded from DBM-DB<sup>29</sup>. Using the local BLAST and HMMER programs, followed by the FGENESH (+), 85 putative *PxP450s* were identified, of which 84 genes appeared to be functional and one was a probable

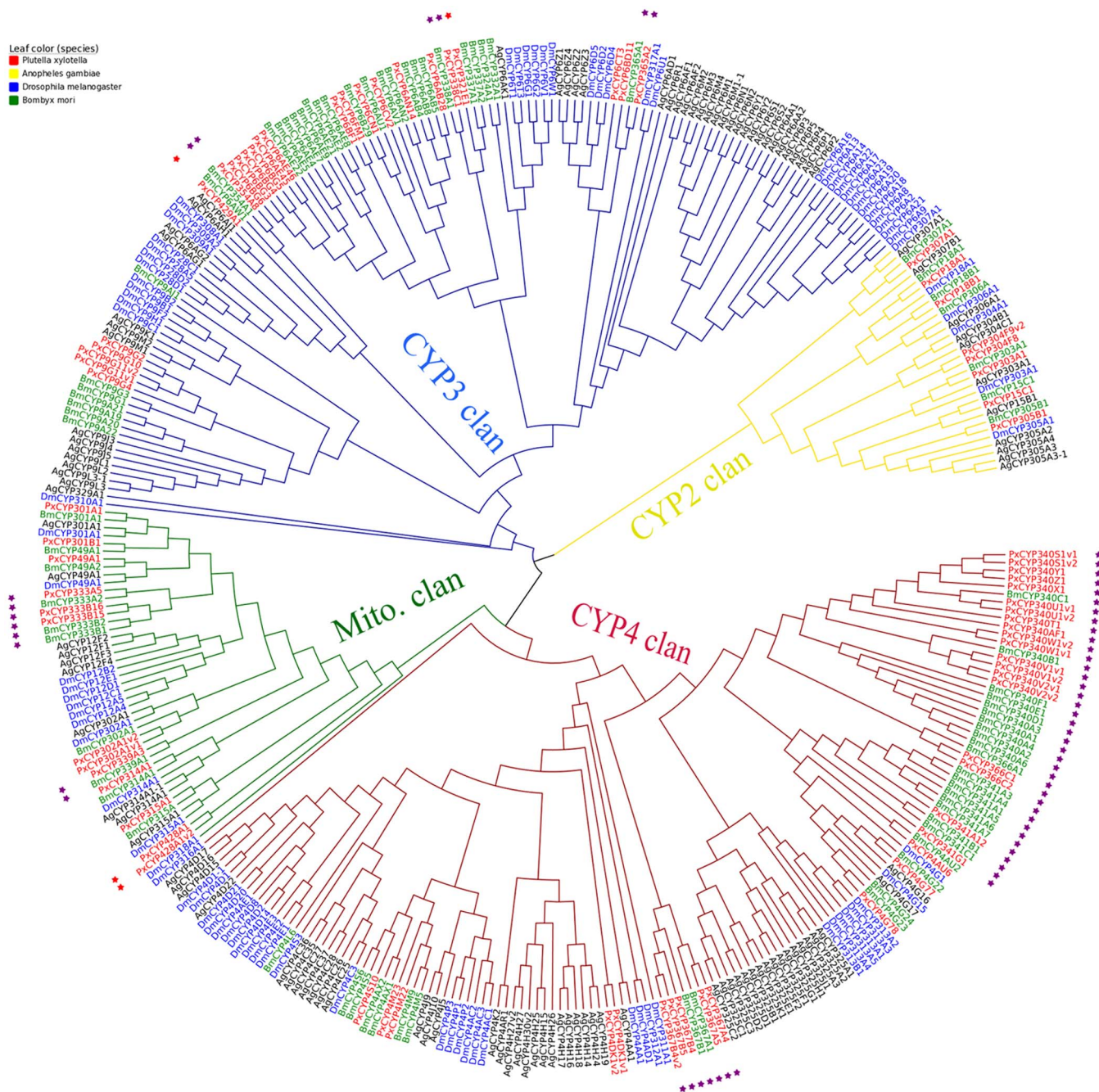
pseudogene (Table S1). According to the standard nomenclature, the 85 *PxP450s* were divided into 26 families and 52 subfamilies (Table 1). The largest families included CYP340 with 16 genes and CYP6 with 14 genes. Our results support previous estimates by Jouraku *et al.*<sup>25</sup> who predict that the number of the DBM P450 genes should be between 61 and 137. Interestingly, in the closely related lepidopteran species to DBM, *B. mori*, 84 genes are also found in 26 families and 47 subfamilies<sup>2</sup>.

The average length of *PxP450s* with complete open read frames was 504 amino acids (aa), which is consistent with the average gene length in insects<sup>2</sup>. Multiple-sequence alignments indicated that *PxP450* proteins had five basic and relatively conserved motifs that were arranged from the N-terminal to C-terminal as follows: helix C, helix I, helix K, PERF and heme-binding motifs (Figure S1). The 85 *PxP450s* were located on 60 scaffolds; 42 of them were found being individually located on a single scaffold, while the rest were on 18 scaffolds with each containing at least two *PxP450s*. A total of 40 genes from 18 clusters were predicted to be tandem duplicates, and they accounted for approximately 47% of all 85 *PxP450s* (Figure S2). Scaffold 130 had the largest *PxP450* cluster with 4 genes, which belong to the CYP340 family. Similarly, the largest CYP340 family cluster in *B. mori* is also located on a single chromosome<sup>2</sup>. P450 gene families evolved through duplication and diversification<sup>30</sup>, and it has been suggested that clusters of the P450 genes, in species such as *Tetrahymena thermophile*, arose through localized tandem duplications<sup>31</sup>. Complete chromosomal scaffolding of the DBM genome will ultimately enable analysis of *PxP450* tandem duplication events, providing insights into the evolution and diversification of this gene family.

**Phylogenetic analyses.** A neighbor-joining phylogeny of the P450 superfamily was constructed using *P. xylostella* and three other insect species (*A. gambiae*, *D. melanogaster* and *B. mori*) to identify gene orthologs (Figure 1). The phylogenetic tree resolved the four expected P450 groups including the CYP2, CYP3, CYP4 and mitochondrial clans.

**Table 1 | Family-based numbers of the cytochrome P450 genes and corresponding transposable elements (TEs) identified from the *P. xylostella* genome**

P450 clan	Family	Subfamily	No. of genes	Total No. of TEs	Average No. of TEs per gene	
CYP4	CYP340	AF, S, T, U, V, W, X, Y, Z	16	1282	80.1	
	CYP4	AU, DK, G, M, S	8	350	43.8	
	CYP367	A, B	8	401	50.1	
	CYP366	C	2	97	48.5	
	CYP341	A, G	2	79	39.5	
CYP3	CYP6	AB, AE, AN, BD, BF, BG, CN, CT, CV, FM	14	655	46.8	
	CYP9	G	5	271	54.2	
	CYP365	A	3	128	42.7	
	CYP354	A	1	28	28.0	
	CYP321	E	1	43	43.0	
	CYP429	A	1	30	30.0	
	CYP338	C	1	39	39.0	
	CYP2	CYP304	F	4	244	61.0
CYP2	CYP18	A, B	2	28	14.0	
	CYP15	C	1	25	25.0	
	CYP303	A	1	59	59.0	
	CYP307	A	1	58	58.0	
	CYP305	B	1	49	49.0	
	Mitochondrial	CYP333	A, B	3	59	19.7
		CYP301	A, B	2	124	62.0
CYP428		A	2	128	64.0	
CYP302		A	2	74	37.0	
CYP314		A	1	52	52.0	
CYP49		A	1	25	25.0	
CYP339		A	1	36	36.0	
Total	CYP315	A	1	38	38.0	
	26	52	85	4402	52.0	



**Figure 1 | Phylogenetic relationship of the P450s of *P. xylostella* and three other insect species.** The phylogenetic tree was divided into four P450 clans, each represented by a branch color. Red, green, blue and yellow branches represent clans of the CYP4, mitochondrial, CYP3 and CYP4, respectively. The four leaf colors are used to distinguish the four species: red for *P. xylostella*, yellow for *A. gambiae*, blue for *D. melanogaster* and green for *B. mori*. P450s in the four species of insects are named as DmCYP, BmCYP, PxCYP and AgCYP, with the first two letters representing the acronym of their scientific names. The outermost stars in purple show the P450s that are unique in Lepidoptera (*B. mori* and *P. xylostella*), while those in red are DBM-specific.

The CYP2 clan contained a single clade with fewer genes and less evolutionary differentiation than the other clans. The branching showed that the two lepidopterans, *B. mori* and *P. xylostella*, were most similar and evolutionally separated from *A. gambiae* and *D. melanogaster*. PxP450s from the CYP2 clan were largely orthologous to *B. mori* (Figure 1), including CYP307A1, CYP18A1, CYP18B1, CYP303A1, CYP15C1 and CYP305B1. The major difference between *B. mori* and *P. xylostella* in CYP2 clan was possible loss of the CYP304F paralogues or orthologues in the silkworm. The genes of CYP2 clan have generally been shown to be evolutionally conserved and play basic physiological functions in insects<sup>31</sup>. For

example, CYP303A1 had a single ortholog in all four species. In *D. melanogaster*, it is expressed in the sensory bristles, and plays an essential role in the development and structure of external sensory organs in response to vital mechanosensory and chemosensory stimuli<sup>32</sup>. In *D. melanogaster*, CYP307A1 encodes the Halloween family gene spook, which is required for the biosynthesis of 20-Hydroxyecdysone<sup>33</sup>, CYP18A1 is tightly correlated with ecdysteroid synthesis<sup>34</sup> and CYP15A1 is involved in juvenile hormone metabolic pathways<sup>5</sup>. As P450 gene clusters commonly have high levels of sequence identity and similar biological functions, the paralogous CYP2 genes in DBM may play similar roles.





In the mitochondrial clan, each species had a similar number of genes, and as expected, *PxP450s* were closely related to the lepidopteran *B. mori*, namely CYP301A1, CYP49A1, CYP333s, CYP302s, CYP314A1 and CYP315A1. Two basal *PxP450* genes, CYP428As (97% identity), were highly divergent compared with other mitochondrial P450s. CYP333s and CYP339s formed Lepidoptera-specific clusters (Figure 1). Ai *et al.*<sup>2</sup> proposed that CYP333s are probably related to xenobiotic metabolism. Two single-copy orthologs (1:1:1:1) were also found in this clan: CYP301A1 and CYP315A1. CYP301a1 is involved in the formation of the adult cuticle and ecdysone regulation in *D. melanogaster*<sup>35</sup>. Previous studies<sup>3</sup> have proposed that two types of mitochondrial P450s are present in insects: genes essential for physiological functions, such as the Halloween genes CYP302A1, CYP315A1 and CYP314A1 that produce the cyclic ecdysteroid pulses triggering moulting<sup>36</sup> and genes that have rapidly evolved taxon-specific paralogs P450s.

Genes in the CYP3 clan showed multiple expansions in DBM, resulting in species-specific clusters. New functions may have been gained in response to environmental changes<sup>37</sup>. Members of the CYP3 clan are involved in xenobiotic metabolism and insecticide resistance, and some are inducible by phenobarbital, pesticides and natural products<sup>3</sup>. Most *PxP450s* in the CYP3 clan had gene orthologs in *B. mori*, or gene expansions with close phylogenetic relationships. Two main clusters were formed: the CYP6s and CYP9s (Figure 1). In addition, the architecture of introns and exons showed the CYP6s and CYP9s clusters are relatively conserved with two exons split by one intron of phase 1 in the CYP6 subfamily, and 2-1-1-0-1-1-1-0-1 in CYP9 subfamilies (Figure S3). Such conserved gene architectures might result from a recent duplication of the genes and provide functions in response to environmental variations<sup>38</sup>. CYP6BGs, CYP6AEs and CYP9Gs formed small clusters in *P. xylostella*. Previous studies suggest that CYP6s are associated with insecticide resistance. CYP6Bs (CYP6B3, CYP6B4, CYP6B5) can promote the adaptation of *Papilio polyxenes* towards toxin-containing host plants<sup>16</sup>, and CYP6AB11 from *Amyelois transitella* efficiently metabolize several pesticides<sup>8</sup>. Although CYP9G2<sup>19</sup> and CYP6BF1<sup>21</sup> have been cloned in *P. xylostella*, their biological roles remain unknown. In contrast, overexpression of *PxCYP6BG1*<sup>20</sup> has been shown to play a role in resistance to the pyrethroid insecticide permethrin.

Gene expansions were also common within the CYP4 clan, particularly the cluster of CYP340s (Figure 1). Most *PxP450s* in this clan were also clustered with *B. mori*, including CYP4Ms, CYP367Bs, CYP4Gs, CYP341s and CYP340s. The *PxCYP4DKs* were phylogenetically separated from the other clusters, which seemed to have evolved at a relatively early stage. Members of the CYP4 clan are known to encode for constitutive and inducible isozymes linked to odorant and pheromone metabolism and are inducible metabolizers of xenobiotics<sup>3,39</sup>. CYP4M1 of *M. sexta* is induced by phenobarbital in the fat body and midguts, and CYP4M1 and CYP4M3 are induced by dietary clofibrate in the midguts<sup>40</sup>. CYP4G61 is thought to be a strong candidate for enhancing the pyriproxyfen resistance of the greenhouse whitefly, *Trialeurodes vaporariorum*<sup>41</sup>. CYP341A2 is preferentially expressed in the chemosensory organs and relates to the chemosensory reception for host plant recognition<sup>42</sup>.

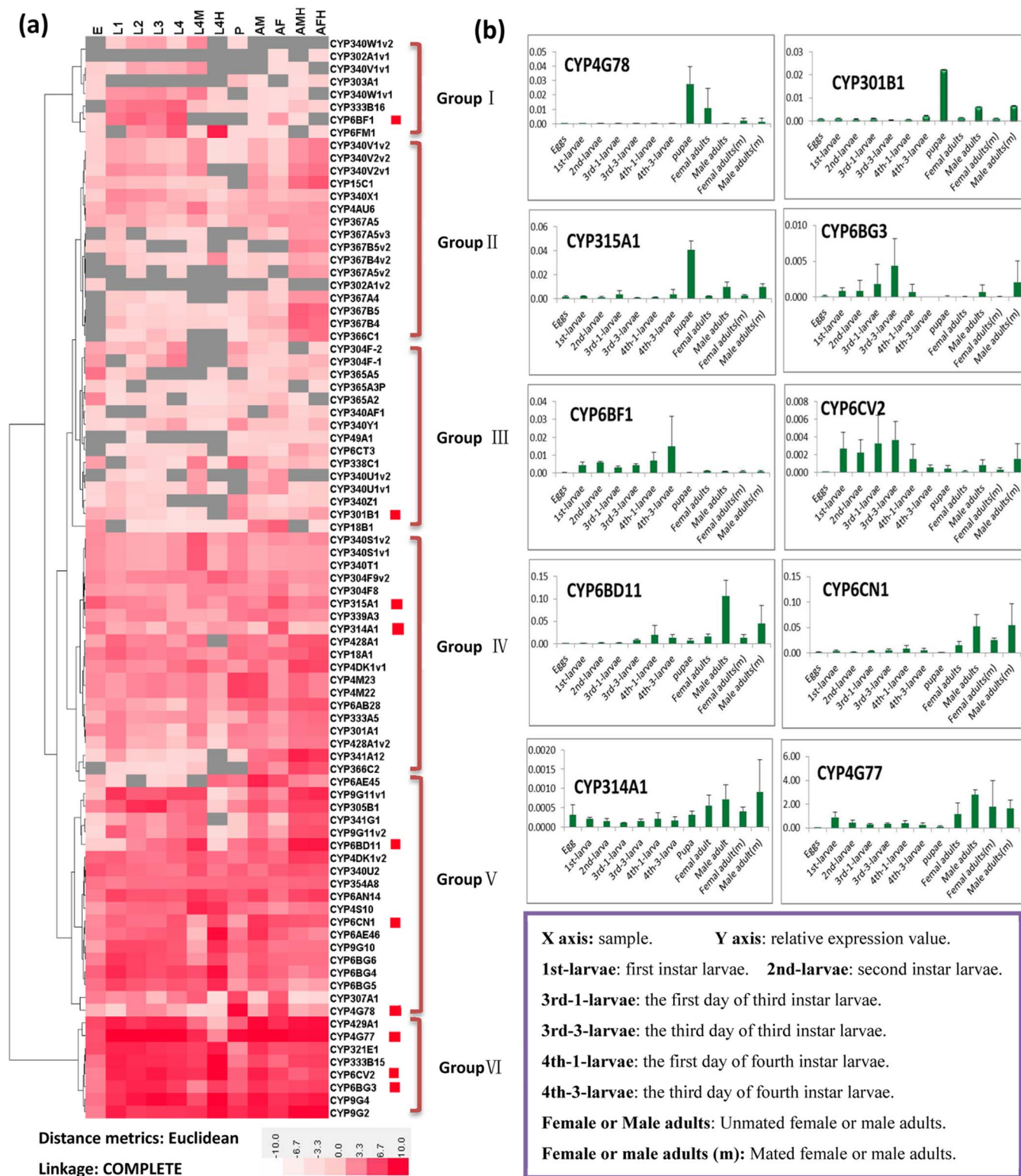
The general clustered *PxP450* and *BmP450* genes indicated their evolutionally close origins<sup>26</sup> and the possibility that P450s might be quite evolutionally conserved within Lepidoptera. The larger number of CYP3 and CYP4 genes in insects might suggest that they probably play broader functions required for their development or to deal with environmental stress (including exposure to insecticides and plant chemical defenses). Further studies would be required to test this hypothesis.

**Stage- and tissue-specific expression profiles of the *PxP450s*.** Expression profiles of 85 *PxP450* genes were generated using

transcriptome sequence data from multiple developmental stages of the susceptible reference strain (SS) including egg, larva, pupa and adult. Hierarchical clustering across different developmental stages was used to divide these genes into six distinct groups (Figure 2a). Group I consisted of 8 genes that showed higher expression in larval stages than any other stages. Foraging can be essential for larvae<sup>43</sup>, and expression of many diverse detoxification genes may help defend themselves from toxic chemicals produced by host plants. Nearly all of the genes in Group II had higher expression levels in heads than in other tissues or stages. Group III genes exhibited relatively high expression values in eggs and pupae. Especially, the P450 pseudogene, CYP365A3P, was expressed in almost all stages and tissues and have similar expression pattern compared to other two CYP365As (CYP365A2 and CYP365, Figure 2a). Transcribed pseudogenes are contributors to the transcriptional landscape and positioned to play potential roles in regulating cognate wild-type gene expression or function<sup>44</sup>. Group IV showed differential expression, with most genes highly expressed in midguts or heads. Midguts serve as the first line in the detoxification of xenobiotics, which are associated with mono-oxygenase activities<sup>45</sup>. For example, CYP340S1v1 and CYP340S1v2 were highly expressed in the midguts, and CYP428A1 and CYP18A1 showed higher expression in female and male heads. Group VI contained 18 genes that were expressed at relatively low levels in eggs compared to other developmental stages. Group VI consisted of the remaining 8 genes that had the highest average expression level. qRT-PCR validation of 10 genes with higher RPKM values, which mostly were consistent with the transcriptome sequence data, showed three preferential expression patterns: 1) CYP4G78, CYP301B1 and CYP315A1 in pupae, 2) CYP6BF1, CYP6CV2 and CYP6BG3 in larvae, and 3) CYP6BD11, CYP6CN1, CYP314A1 and CYP4G77 in adults, especially in males (Figure 2b, Figure S4). CYP312A1 in *D. melanogaster* has been shown to be expressed 82-times higher in adult males than in females, and is suggested to be involved in endogenous catalytic reactions<sup>46</sup>. Our results confirm that gene expression of P450s varies with developmental stages and sexes<sup>47</sup>. Although we have inferred some biological functions of the *PxP450s* based on gene expression and homology, the exact functional mechanisms remain to be elucidated.

**Strain-specific expression profiles of *PxP450s*.** To investigate the expression profiles of *PxP450* genes in three DBM strains, we analyzed gene expression pattern from the 3<sup>rd</sup> instar larvae using transcriptome data from the insecticide-susceptible genome reference strain, Fuzhou-S (SS), and two insecticide-resistant strains FR (fipronil-resistant, 273.5 fold) and CR (chlorpyrifos-resistant, 3417.2 fold).

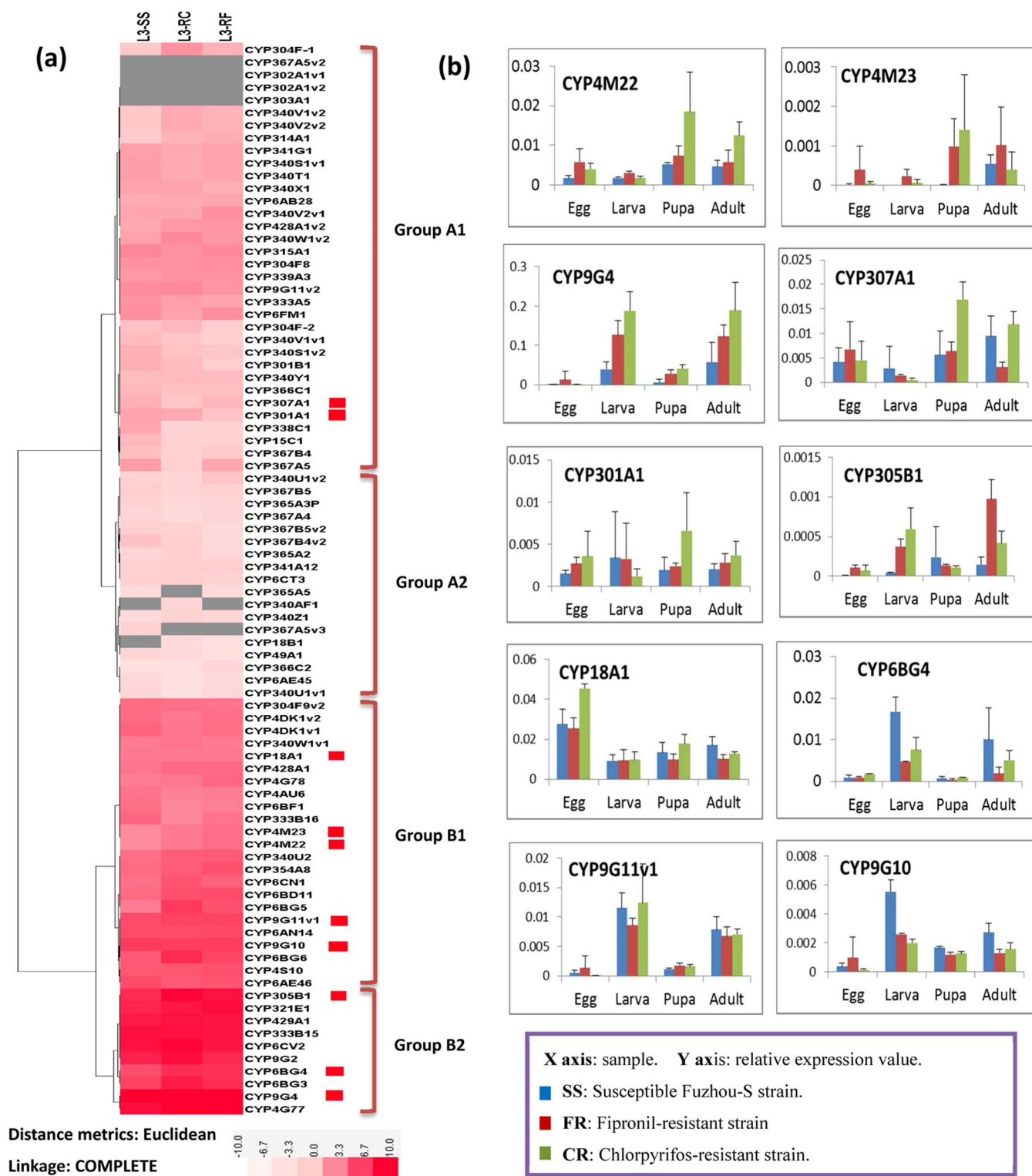
Based on the heatmap representing the expression profiles of the 85 *PxP450s* genes in these three strains (Figure 3a), two major groups could be defined: Group A had 33 genes with low expression (RPKM < 10.3, average RPKM = 2.3) in the 3<sup>rd</sup> instar larvae and Group B consisted of 52 highly expressed genes (RPKM > 8.6, average RPKM = 208.6). Each major group could be further classified into two subgroups. Subgroup A1 contained relatively more P450s members showed higher expression values (more than twice) in one or both insecticide-resistant strains relative to the susceptible strain, including CYP304F-1, CYP340V1v2, CYP340V2v2, CYP314A1, CYP340V2v1, CYP428A1v2 and CYP340W1v2 (Figure 3). Over expression of P450 genes commonly causes insecticide resistant phenotypes<sup>23</sup>. Subgroup A2 genes showed no expression differences between the three strains, suggesting little involvement in resistance to these two chemicals. Subgroup B1 genes were highly expressed with some genes, including CYP6BG5 and CYP6BG6 that were over-expressed in the insecticide-resistant strains. Quantitative RT-PCR was used to assess expression of 10 genes with higher RPKM values and variation between resistant and susceptible larvae. Multiple



**Figure 2 | Stage- and tissue-specific expression of the *P. xylostella* P450s.** (a) Hierarchical clustering of the 85 *PxP450s* based on RPKM values from the digital gene expression profiling (DGE). The samples are described as in Figure 2a. The log<sub>2</sub> RPKM values are presented with varying colors. The darker red represents higher expression values, and the gray in the heat map missing values. (b) qRT-PCR-based expression of 10 selected *P. xylostella* P450s across developmental stages of the susceptible strain.

tissues sources were used including egg, whole larvae, (3<sup>rd</sup> instar), pupae (day 2) and adult (male and female). CYP4M22 and CYP4M23 had higher expression levels in insecticide-resistant strains than in the susceptible strain, supporting the RNA-seq dataset (Figure 3b, Figure S5). The *PxP450* genes of subgroup B2 showed

relatively high levels of expression in all strains. qRT-PCR results indicated that CYP9G4 was up-regulated in insecticide-resistance strains, while CYP6BG4 down-regulated. The expressions of P450s are induced or inhibited by many xenobiotics and some endogenous compounds<sup>48</sup>. Decreased P450 expression is a homeostatic or adaptive



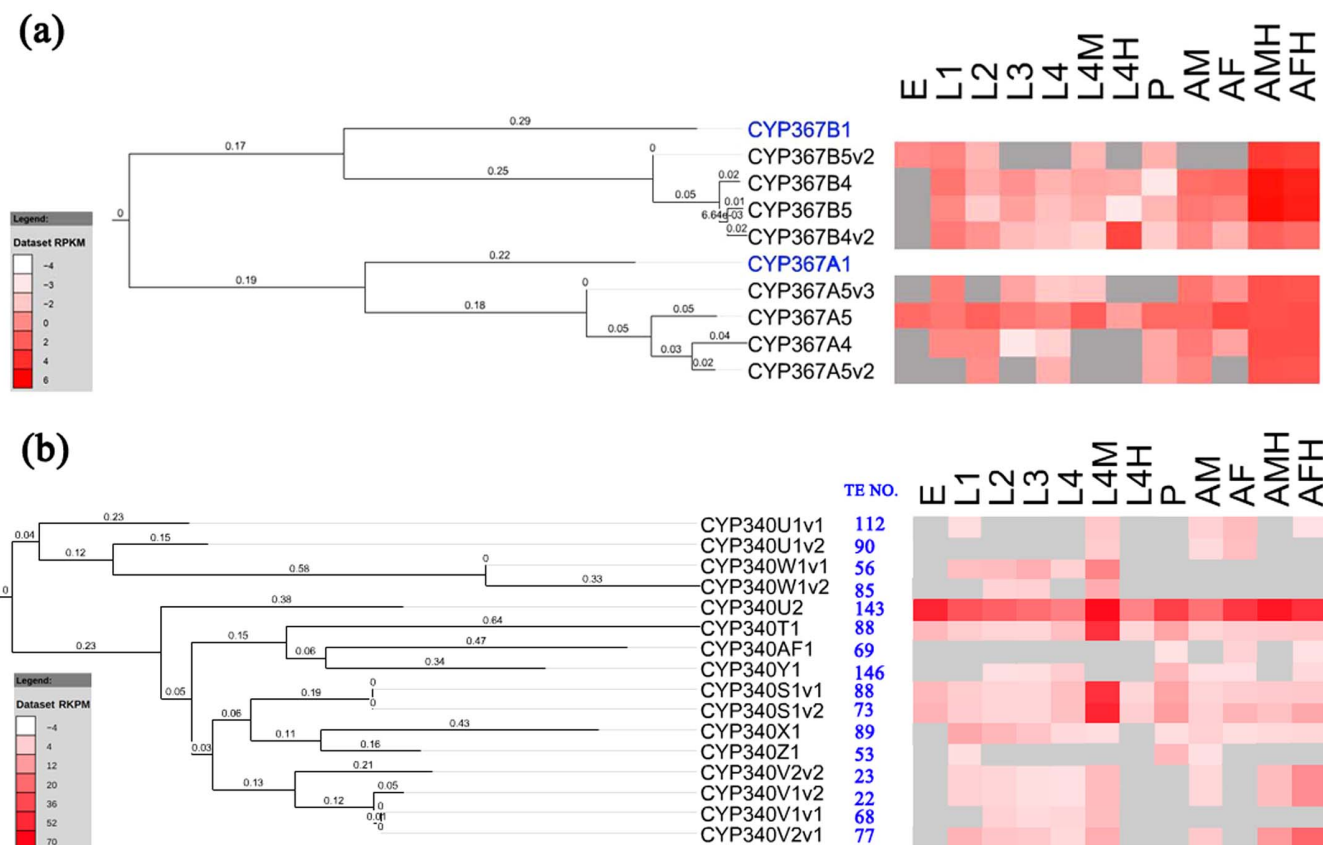
**Figure 3** | Strain-specific expression of the *P. xylostella* P450s. (a) Hierarchical clustering of the 85 *PxP450*s based on RPKM values from the transcriptome sequencing. The log<sub>2</sub> RPKM values are presented with varying colors. The darker red represents higher expression values, and the gray in the heat map missing values. L3-SS: 3<sup>rd</sup>-instar larva from the susceptible strain; L3-FR: 3<sup>rd</sup>-instar larva from the fipronil-resistant strain; and L3-CR: 3<sup>rd</sup>-instar larva from the chlorpyrifos-resistant strain. (b) qRT-PCR-based expression of 10 selected *P. xylostella* P450s across four development stages in three strains.

response, whereby genes are suppressed in response to pathophysiological signals and various exogenous or endogenous compounds<sup>47,49</sup>.

**Lepidoptera-specific P450s.** Multiple P450 clusters were identified that were specific to the Lepidoptera analyzed here, including the clan

4 CYP367s and CYP340s. A phylogeny of the CYP367s (Figure 4a) showed expansions of CYP367A and CYP367B in DBM, compared only one member in these subfamilies in *B. mori*<sup>2</sup>, suggesting that *P. xylostella* may have evolved some specific functions. More interestingly, all *PxP450*s showed adult head-specific expression,





**Figure 4** | Characterization of lepidoptera-specific P450s (CYP367s and CYP340s). (a) Phylogenetic tree of CYP367s of *B. mori* (in blue) and *P. xylostella* (in black) (left), and stage- and tissue-specific DGE profiles of *P. xylostella* (right). (b) Phylogenetic relationships of the *P. xylostella* CYP340s (left), and stage- and tissue-specific DGE profiles of *P. xylostella* (right). E: egg; L1: 1<sup>st</sup>-instar larva; L2: 2<sup>nd</sup>-instar larva; L3: 3<sup>rd</sup>-instar larva; L4: 4<sup>th</sup>-instar larva; L4M: midgut of the 4<sup>th</sup>-instar larva; L4H: head of the 4<sup>th</sup>-instar larva; P: pupa; AM: male adult; AF: female adult; AMH: head of male adult; and AFH: head of female adult. The RPKM values are presented with varying colors. The darker red represents higher expression values, and the gray in the heat map missing values.

and CYP367B4v2 tended to specifically express in heads of the 4<sup>th</sup>-instar larva and adult (Figure 4a). P450s specifically expressed in heads probably have specific functions, which is evidenced by fact that CYP6BQ9 predominantly expressed in *T. castaneum* brains confers deltamethrin resistance<sup>11</sup>. In insects, some P450s in the antennae play a role in xenobiotic metabolism or odorant processing, and the antenna is a key site for P450s mediating the metabolism of many exogenous and endogenous molecules<sup>50</sup>. Considering the facts that *PxCYP367s* were greatly expanded and mainly expressed in heads, we suggest that they may play special roles in detoxification or metabolic processing of environmental chemicals.

*PxCYP340s* constituted the largest family of genes in DBM with 16 genes and the largest number of transposable elements per gene (Table 1). CYP340Y1 and CYP340U2 have the most abundant transposable elements inside and in their 2-kb upstream regions (Figure 4b). The phylogenetic tree showed only two clusters, confirming their relative conserved gene intron and exon organization (Figure 4b, Figure S3). P450s with abundant transposable elements have been proposed to be involved in induced xenobiotic detoxification<sup>26</sup>. Nearly all *PxCYP367s* showed larva midgut-specific expression. (Figure 4b). In midguts of *M. sexta*, CYP4M1 and CYP4M3 are induced by dietary clofibrate<sup>40</sup>. Midguts serving as the first line in the detoxification of xenobiotics are associated with monooxygenase activities<sup>45</sup>. The CYP340 family with abundant transposable elements and midgut-specific expression might contribute to P450-mediated detoxification in *P. xylostella*.

## Methods

**Identification and nomenclature of the *PxP450s*.** Two approaches were used to identify the putative *PxP450s*. The first method consisted at downloading from the Cytochrome P450 Homepage<sup>28</sup> (<http://drnelson.uthsc.edu/CytochromeP450.html>) known P450 amino acid sequences of 24 species and then, subjected the protein sequences of DBM OGSv1 (Office Gene set version 1) (<http://iae.fafu.edu.cn/DBM/>) to BLASTP analysis with a threshold of  $P < 10^{-7}$ . The targeted DBM genes that were similar to unknown proteins were verified using BLASTP to search against the NCBI and UniPort databases, and those with the best hit to the annotated P450s were considered to be possible P450s.

The second method used to identify P450 genes was carried out using the HMMER3 (<http://hmmer.janelia.org/>) software with the default parameters. The P450 domain (PF00067) was searched against the DBM OGSv1 using the *hmmsearch* program of HMMER3. Then, DBM proteins harboring at least one P450 domain were searched against the Pfam-A dataset that was downloaded from PFAM (<http://pfam.janelia.org/>) using the *hmmsearch* program and those best-matched domain of P450 were considered as the putative P450s.

We used the online FGENESH and FGENESH+ (<http://linux1.softberry.com/berry.phtml>) programs to find the entire length of the P450 genes that had no complete open read frames and predicted gene structure based on their whole-length sequences and most similar protein sequences. All the putative *PxP450s* were named by David R. Nelson to maintain consistency in the nomenclature.

**Sequence alignment and motif visualization.** The average length of cytochrome P450 genes was approximately 500 aa. Five conserved motifs exist in insect P450s: the helix C motif (WxxxR), the helix I motif (Gx[ED]T[TS]), the helix K motif (ExLR), the PERF motif (PxxFxP[ED]RE) and the heme-binding motif (PFxxGxRx[Gx])<sup>2</sup>. The fifth conserved motif can be used to easily recognize P450 family members. To visualize the conserved motifs, 77 functional *PxP450s* with more than 300 aa were aligned by ClustalX 2.0.11 using the default parameters, and the five conserved motifs were illustrated using WebLogo<sup>51</sup> (<http://weblogo.threeplusone.com/create.cgi>) based on the sequence alignment results.



**Determination of gene location and duplication.** Pxp450s were mapped on scaffolds using Mapchart<sup>32</sup>, and only scaffolds with at least two genes were displayed. Gene location information from the DBM genome (assembly version 2.0) was obtained from the DBM-DB (<http://iae.fafu.edu.cn/DBM/>)<sup>29</sup>, and genes separated by 0–5 genes were considered to be tandem duplicates<sup>53</sup>.

**Generation of gene architecture and phylogenetic trees.** Using the information from the published DBM genome (assembly version 2.0), a figure of the gene architecture was generated online (GSDS) (<http://gsds.cbi.pku.edu.cn/>). Phylogenetic trees were constructed using the MEGA5.10 based on the neighbor-joining method, including trees of 77 putative Pxp450s (amino acid residues >300 aa) and four insects (*P. xylostella*, *A. gambiae*, *D. melanogaster* and *B. mori*), CYP367s and CYP340s (regardless of the length of amino acid residues). Bootstrap analysis was performed using 1000 replicates, and the Newick output format was used to visualize the phylogenetic tree online with the EvolView<sup>54</sup> tool (<http://www.evolgenius.info/evolview.html#login>) or iTOL<sup>55</sup> (<http://itol.embl.de/>).

**Expression profile and quantitative RT-PCR analysis.** Expression profiling of insect strains, developmental stages and tissue types was generated by Cluster and Treeview software based on the DBM RNA-seq and DGE data, respectively. The RPKM values were log transformed, and genes were clustered in terms of their expression patterns using the similarity metric of Euclidean distance and clustering method of complete linkage.

The susceptible strain and two insecticide (chlorpyrifos and fipronil)-resistant strains of DBM (were reared with radish seedlings at 25 ± 2°C. The median lethal concentrations (LC<sub>50</sub>) of chlorpyrifos- and fipronil-resistant strains were 574-fold (51,500.00 mg·L<sup>-1</sup> vs. 89.79 mg·L<sup>-1</sup>) and 72-fold (16.85 mg·L<sup>-1</sup> vs. 0.23 mg·L<sup>-1</sup>) higher than susceptible strain, respectively. Strain- and stage-specific DBM samples were collected and held at -20°C for RNA isolation and qRT-PCR analyses. The strain-specific samples included newly laid eggs, 3<sup>rd</sup>-instar larvae, pupae, and male or female adults of the three strains, while stage-specific samples involved eggs, 1<sup>st</sup>–4<sup>th</sup> instar larvae, pupae, and unmated or mated males and females of the susceptible strain. Total RNA was extracted from the samples using Trizol reagent according to the manufacturer's instructions (Invitrogen, USA). First-strand cDNAs were synthesized from the total RNA in a 20-μl reaction using the GoScript Reverse Transcription System following the manufacturer's instructions (Promega, USA). Each cDNA sample was diluted to 200 μl for qPCR. Gene-specific primers of 20 Pxp450 genes were designed using the Primer premier 5 software (Table S2), and the RIBP gene was used as an internal reference for gene analyses.

The qRT-PCR analyses were conducted using the Bio-RAD CFX96 Real-time PCR Detection System (Bio-Rad, USA) according to the manufacturer's instructions using the GoTaq qPCR Master Mix (Promega, USA). The PCR reaction system contained 12.5 μl of 2× real-time PCR Master Mix (containing SYBR Green I), 0.5 μl of each primer, 2 μl of the diluted cDNA samples and nuclease-free water in a final volume of 25 μl. The PCR thermal cycling conditions were as follows: 95°C for 3 min, followed by 44 cycles of 95°C for 15 s and 57°C for 35 s, and a final melt curve at 60°C for 5 s to 95°C at 0.5°C increments. Each treatment was technically and biologically conducted independently and in triplicate. The relative gene expression was calculated according to the delta Ct method of the system. To determine gene expressions and standard deviation in different lines, data were statistically analyzed by univariate ANOVA from IBM SPSS Statistics 19.

## Conclusion

With the recent advancement of genome sequencing technology, it has become possible to identify and investigate expression of complex gene families in non-model systems. In this study, we identified 85 putative P450s in *P. xylostella*. By analyzing evolutionary relationships with other insects and the expression patterns of the Pxp450s, we provided detailed information on their diversity and evolution. The Lepidoptera-specific clusters of CYP367s, which show abundant expression in adult heads, and CYP340s, that are associated with transposable element density, are excellent candidates for investigating detoxification of insecticides and plant toxins. Genomic characterization and gene expression analysis of the Pxp450s during development will facilitate future functional studies of this complex gene family in *P. xylostella*.

1. Werck-Reichhart, D. & Feyereisen, R. Cytochromes P450: a success story. *Genome Biol* **1**, Reviews3003 (2000).
2. Ai, J. *et al.* Genome-wide analysis of cytochrome P450 monooxygenase genes in the silkworm, *Bombyx mori*. *Gene* **480**, 42–50 (2011).
3. Feyereisen, R. Evolution of insect P450. *Biochem Soc Trans* **34**, 1252–1255 (2006).
4. Scott, J. G. Cytochromes P450 and insecticide resistance. *Insect Biochem Mol Biol* **29**, 757–777 (1999).
5. Iga, M. & Kataoka, H. Recent studies on insect hormone metabolic pathways mediated by cytochrome P450 enzymes. *Biol Pharm Bull* **35**, 838–843 (2012).

6. Bernhardt, R. Cytochromes P450 as versatile biocatalysts. *J Biotechnol* **124**, 128–145 (2006).
7. Stevenson, B. J. *et al.* Cytochrome P450 6M2 from the malaria vector *Anopheles gambiae* metabolizes pyrethroids: Sequential metabolism of deltamethrin revealed. *Insect Biochem Mol Biol* **41**, 492–502 (2011).
8. Niu, G. *et al.* A substrate-specific cytochrome P450 monooxygenase, CYP6AB11, from the polyphagous navel orangeworm (*Amyelois transitella*). *Insect Biochem Mol Biol* **41**, 244–253 (2011).
9. Matambo, T. S., Paine, M. J., Coetzee, M. & Koekemoer, L. L. Sequence characterization of cytochrome P450 CYP6P9 in pyrethroid resistant and susceptible *Anopheles funestus* (Diptera: Culicidae). *Genet Mol Res* **9**, 554–564 (2010).
10. Martinez-Paz, P., Morales, M., Martinez-Guitarte, J. L. & Morcillo, G. Characterization of a cytochrome P450 gene (CYP4G) and modulation under different exposures to xenobiotics (tributyltin, nonylphenol, bisphenol A) in *Chironomus riparius* aquatic larvae. *Comp Biochem Physiol C Toxicol Pharmacol* **155**, 333–343 (2012).
11. Zhu, F. *et al.* A brain-specific cytochrome P450 responsible for the majority of deltamethrin resistance in the QTC279 strain of *Tribolium castaneum*. *Proc Natl Acad Sci U S A* **107**, 8557–8562 (2010).
12. Daborn, P., Boundy, S., Yen, J., Pittendrigh, B. & French-Constant, R. DDT resistance in *Drosophila* correlates with Cyp6g1 over-expression and confers cross-resistance to the neonicotinoid imidacloprid. *Mol Genet Genomics* **266**, 556–563 (2001).
13. Wondji, C. S. *et al.* Two duplicated P450 genes are associated with pyrethroid resistance in *Anopheles funestus*, a major malaria vector. *Genome Res* **19**, 452–459 (2009).
14. Jousen, N. *et al.* Resistance of Australian *Helicoverpa armigera* to fenvalerate is due to the chimeric P450 enzyme CYP337B3. *Proc Natl Acad Sci U S A* **109**, 15206–15211 (2012).
15. Furlong, M. J., Wright, D. J. & Dossdall, L. M. Diamondback moth ecology and management: problems, progress, and prospects. *Annu Rev Entomol* **58**, 517–541 (2013).
16. Scott, J. G. & Wen, Z. Cytochromes P450 of insects: the tip of the iceberg. *Pest Manag Sci* **57**, 958–967 (2001).
17. Talekar, N. S. Biology, ecology, and management of the diamondback Moth. *Annu Rev Entomol* **38**, 275–301 (1993).
18. Tabashnik, B. E., Liu, Y. B., Finson, N., Masson, L. & Heckel, D. G. One gene in diamondback moth confers resistance to four *Bacillus thuringiensis* toxins. *Proc Natl Acad Sci U S A* **94**, 1640–1644 (1997).
19. Shen, B., Zhao, D., Qiao, C. & Lan, W. Cloning of CYP9G2 from the Diamondback Moth, *Plutella xylostella* (Lepidoptera: Yponomeutidae). *DNA Seq* **15**, 228–233 (2004).
20. Bautista, M. A., Miyata, T., Miura, K. & Tanaka, T. RNA interference-mediated knockdown of a cytochrome P450, CYP6BG1, from the diamondback moth, *Plutella xylostella*, reduces larval resistance to permethrin. *Insect Biochem Mol Biol* **39**, 38–46 (2009).
21. Li, H., Dai, H. & Wei, H. Molecular cloning and nucleotide sequence of CYP6BF1 from the diamondback moth, *Plutella xylostella*. *J Insect Sci* **5**, 45 (2005).
22. Lin, Q. *et al.* Transcriptome analysis of chlorantraniliprole resistance development in the diamondback moth *Plutella xylostella*. *PLoS One* **8**, e72314 (2013).
23. Yang, T. & Liu, N. Genome analysis of cytochrome P450s and their expression profiles in insecticide resistant mosquitoes, *Culex quinquefasciatus*. *PLoS One* **6**, e29418 (2011).
24. Tijet, N., Helvig, C. & Feyereisen, R. The cytochrome P450 gene superfamily in *Drosophila melanogaster*: annotation, intron-exon organization and phylogeny. *Gene* **262**, 189–198 (2001).
25. Jouraku, A. *et al.* KONAdbase: a genomic and transcriptomic database for the diamondback moth, *Plutella xylostella*. *BMC Genomics* **14**, 464 (2013).
26. You, M. *et al.* A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics* **45**, 220–225 (2013).
27. He, W. *et al.* Developmental and insecticide-resistant insights from the de novo assembled transcriptome of the diamondback moth, *Plutella xylostella*. *Genomics* **99**, 169–177 (2012).
28. Nelson, D. R. The cytochrome p450 homepage. *Hum Genomics* **4**, 59–65 (2009).
29. Tang, W. *et al.* DBM-DB: the diamondback moth genome database. *Database (Oxford)* **2014**, 87; DOI:10.1093/database/at08b7 (2014).
30. Gotoh, O. [Evolution and differentiation of P450 genes] *Cytochrome P450* [Omura, T., Ishimura, Y. & Fujii-Kuriyama, Y. (eds.)] [255–272] (VCH, New York, 1993).
31. Fu, C., Xiong, J. & Miao, W. Genome-wide identification and characterization of cytochrome P450 monooxygenase genes in the ciliate *Tetrahymena thermophila*. *BMC Genomics* **10**, 208 (2009).
32. Willingham, A. T. & Keil, T. A tissue specific cytochrome P450 required for the structure and function of *Drosophila* sensory organs. *Mech Dev* **121**, 1289–1297 (2004).
33. Rewitz, K. F., O'Connor, M. B. & Gilbert, L. I. Molecular evolution of the insect Halloween family of cytochrome P450s: phylogeny, gene organization and functional conservation. *Insect Biochem Mol Biol* **37**, 741–753 (2007).





34. Bassett, M. H., McCarthy, J. L., Waterman, M. R. & Sliter, T. J. Sequence and developmental expression of Cyp18, a member of a new cytochrome P450 family from *Drosophila*. *Mol Cell Endocrinol* **131**, 39–49 (1997).
35. Sztal, T. *et al.* A cytochrome p450 conserved in insects is involved in cuticle formation. *PLoS One* **7**, e36544 (2012).
36. Rewitz, K. F. & Gilbert, L. I. Daphnia Halloween genes that encode cytochrome P450s mediating the synthesis of the arthropod molting hormone: evolutionary implications. *BMC Evol Biol* **8**, 60 (2008).
37. Gasch, A. P. *et al.* Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* **11**, 4241–4257 (2000).
38. Kondrashov, F. A. Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc Biol Sci* **279**, 5048–5057 (2012).
39. Simpson, A. E. The cytochrome P450 4 (CYP4) family. *Gen Pharmacol* **28**, 351–359 (1997).
40. Snyder, M. J., Stevens, J. L., Andersen, J. F. & Feyereisen, R. Expression of cytochrome P450 genes of the CYP4 family in midgut and fat body of the tobacco hornworm, *Manduca sexta*. *Arch Biochem Biophys* **321**, 13–20 (1995).
41. Karatolos, N. *et al.* Over-expression of a cytochrome P450 is associated with resistance to pyriproxyfen in the greenhouse whitefly *Trialeurodes vaporariorum*. *PLoS One* **7**, e31077 (2012).
42. Ono, H., Ozaki, K. & Yoshikawa, H. Identification of cytochrome P450 and glutathione-S-transferase genes preferentially expressed in chemosensory organs of the swallowtail butterfly, *Papilio xuthus* L. *Insect Biochem Mol Biol* **35**, 837–846 (2005).
43. Sokolowski, M. B., Kent, C. & Wong, J. *Drosophila* larval foraging behaviour: Developmental stages. *Anim Behav* **32**, 645–651 (1984).
44. Kalyana-Sundaram, S. *et al.* Expressed pseudogenes in the transcriptional landscape of human cancers. *Cell* **149**, 1622–1634 (2012).
45. Scott, J. G. [Insect cytochrome P450s, Thinking beyond detoxification] *Recent Advances in Insect Physiology, Toxicology and Molecular Biology* [Nannan, L. (ed.)] [117–124] (Kerala, India, 2008).
46. Kasai, S. & Tomita, T. Male specific expression of a cytochrome P450 (Cyp312a1) in *Drosophila melanogaster*. *Biochem Biophys Res Commun* **300**, 894–900 (2003).
47. Riddick, D. S. *et al.* Transcriptional suppression of cytochrome P450 genes by endogenous and exogenous chemicals. *Drug Metab Dispos* **32**, 367–375 (2004).
48. Hodgson. The significance of cytochrome in insect. *Insect Biochem* **13**(3), 237–246 (1983).
49. Cui, P. H., Lee, A. C., Zhou, F. & Murray, M. Impaired transactivation of the human CYP2J2 arachidonic acid epoxigenase gene in HepG2 cells subjected to nitrate stress. *Br J Pharmacol* **159**, 1440–1449 (2010).
50. Pottier, M. A. *et al.* Cytochrome P450s and cytochrome P450 reductase in the olfactory organ of the cotton leafworm *Spodoptera littoralis*. *Insect Mol Biol* **21**, 568–580 (2012).
51. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188–1190 (2004).
52. Voorrips, R. E. MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* **93**, 77–78 (2002).
53. Zhao, H., Ma, H., Yu, L., Wang, X. & Zhao, J. Genome-wide survey and expression analysis of amino acid transporter gene family in rice (*Oryza sativa* L.). *PLoS One* **7**, e49210 (2012).
54. Zhang, H., Gao, S., Lercher, M. J., Hu, S. & Chen, W. H. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res* **40**, W569–572 (2012).
55. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).

## Acknowledgments

We are grateful to Dr. David R. Nelson (the International Cytochrome Nomenclature Committee), for nomenclature of the DBM P450 genes, and to Dr. Xianchun Li for his advice on sequence analysis. We thank Yanchun You for her advice on primer design and the RT-PCR experiment, and Jianyu Li and Lishuang Zheng for DBM rearing and sampling. We appreciate Dr. Mark Goettel for his comments and suggestions on the manuscript revision. The work is supported by National Natural Science Foundation of China (No. 31320103922 and No. 31230061), and National Key Project of Fundamental Scientific Research (“973” Programs, No. 2011CB100404) in China. L. V. is supported by the Minjiang Scholar Fellowship, and S. B. is supported by the Jinshan Scholar Fellowship and the Australian Research Council.

## Author contributions

L.Y. carried out the bioinformatics analysis and the biological experiments, and drafted the manuscript. W.T., W.H., S.B., L.V., Y.G., S.H. and M.Y. contributed to the study design or manuscript revision. X.M. and F.Q. participated in the experiments. All authors read and approved the final manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Yu, L. *et al.* Characterization and expression of the cytochrome P450 gene family in diamondback moth, *Plutella xylostella* (L.). *Sci. Rep.* **5**, 8952; DOI:10.1038/srep08952 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>