

Identification of stage-specific larval camouflage associated genes in the swallowtail butterfly, *Papilio xuthus*

Ryo Futahashi · Haruhiko Fujiwara

Received: 16 April 2008 / Accepted: 29 July 2008 / Published online: 20 August 2008
© Springer-Verlag 2008

Abstract The swallowtail butterfly *Papilio xuthus* changes its larval body pattern dramatically during the fourth ecdysis. The penultimate (fourth) instar larvae mimic bird droppings that have many tubercle structures on their surface, whereas the final (fifth) instar larvae have a green camouflage color. To identify the genes involved in stage-specific larval mimicry markings, we compared the epidermal messenger ribonucleic acid expression between the third and fourth molts of *P. xuthus* using a complementary deoxyribonucleic acid subtraction method. After analyzing 2,072 clones from two subtractive libraries, we obtained 31 and 64 candidate genes for final (Fsg) or penultimate instar-specific genes (Psg), respectively. The expression pattern of each gene was examined by reverse transcriptase polymerase chain reaction and in situ hybridization. Among Fsgs, *Fsg02* and *Fsg20* were expressed in the presumptive green region only during the fourth molt, suggesting that these two genes were correlated with green coloration in the final instar. Among Psgs, 11 cuticular protein genes were expressed specifically in the tubercle structures during the third molt. These genes are likely to be involved in the formation of the unique tubercle structures observed in the juvenile instar. We found that genes with similar expression patterns do not necessarily share the same protein motifs and vice versa. This study provides novel

molecular markers and insights into the molecular mechanisms of the larval color pattern and body shape.

Keywords *Papilio xuthus* · Butterfly · Larval markings · Cuticle · Color formation

Introduction

One of the most attractive themes in modern biology is the pursuit of molecular mechanisms that underlie the phenotypic diversity upon which adaptive strategies are based. Insect body marking patterns have successfully diversified to adapt to their environments. Switches in larval coloration and body markings in the Lepidoptera are well known (Scoble 1992). The swallowtail butterfly, *Papilio xuthus*, changes its color pattern dramatically during the molt to the final instar (fourth molt, Fig. 1a). Juvenile instars of *P. xuthus* mimic bird droppings with many tubercle structures on their surface (Fig. 1a, arrowheads), whereas the fifth (last) instar shows a cryptic pattern. Recently, we found that this pattern switch is regulated by a juvenile hormone (JH, Futahashi and Fujiwara 2008). JH titer decreases during the fourth larval instar, and treatment with a JH analog (JHA) prevents this pattern switch. The expression levels or patterns of the genes involved in determining the larval body pattern should change during the fourth molt. We previously found that five genes involved in melanin synthesis—*tyrosine hydroxylase*, *dopa decarboxylase*, *ebony*, *yellow*, and *GTP cyclohydrolase I*—are differentially expressed between the third and fourth molts in accordance with the cuticular pigment pattern of the subsequent instar (Futahashi and Fujiwara 2005, 2006, 2007). Using complementary deoxyribonucleic acid (cDNA) subtraction methods, we cloned the hard cuticular protein genes associated with

Communicated by: P. Simpson

Electronic supplementary material The online version of this article (doi:10.1007/s00427-008-0243-y) contains supplementary material, which is available to authorized users.

R. Futahashi · H. Fujiwara (✉)
Department of Integrated Biosciences,
Graduate School of Frontier Sciences, The University of Tokyo,
Bioscience Bldg. 501,
Kashiwa, Chiba 277-8562, Japan
e-mail: haruh@k.u-tokyo.ac.jp

the specific tubercle structures (*HCP1* and *HCP2*) and the bilin-binding protein gene (*BBP*) that is only expressed at the final molt (Futahashi and Fujiwara 2008).

In this paper, we reported the expression pattern of 98 genes identified by cDNA subtraction. The spatial expression patterns of obviously differentially expressed genes were analyzed by whole-mount in situ hybridization and compared with stage-specific body markings. We identified 13 genes associated with green coloration or tubercle structures, three of which (*Fsg02*, *Psg12*, and *Psg14*) were also reported by Futahashi and Fujiwara (2008).

Materials and methods

Experimental animals and developmental staging

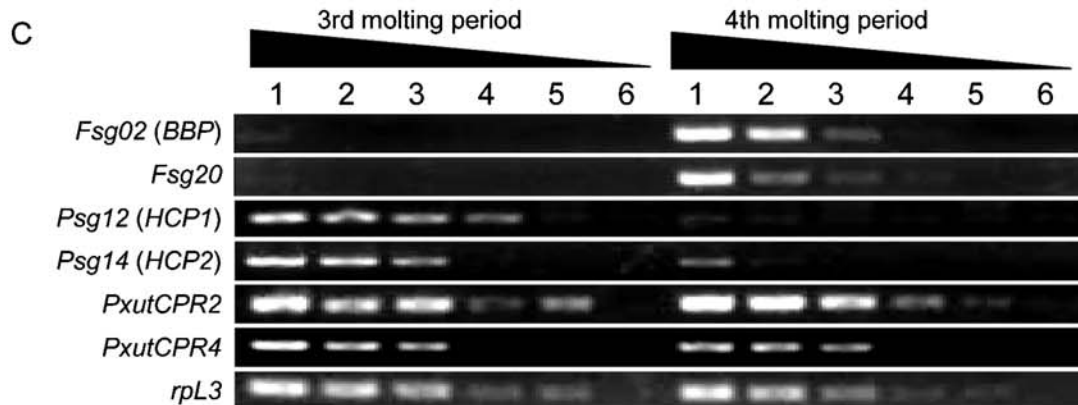
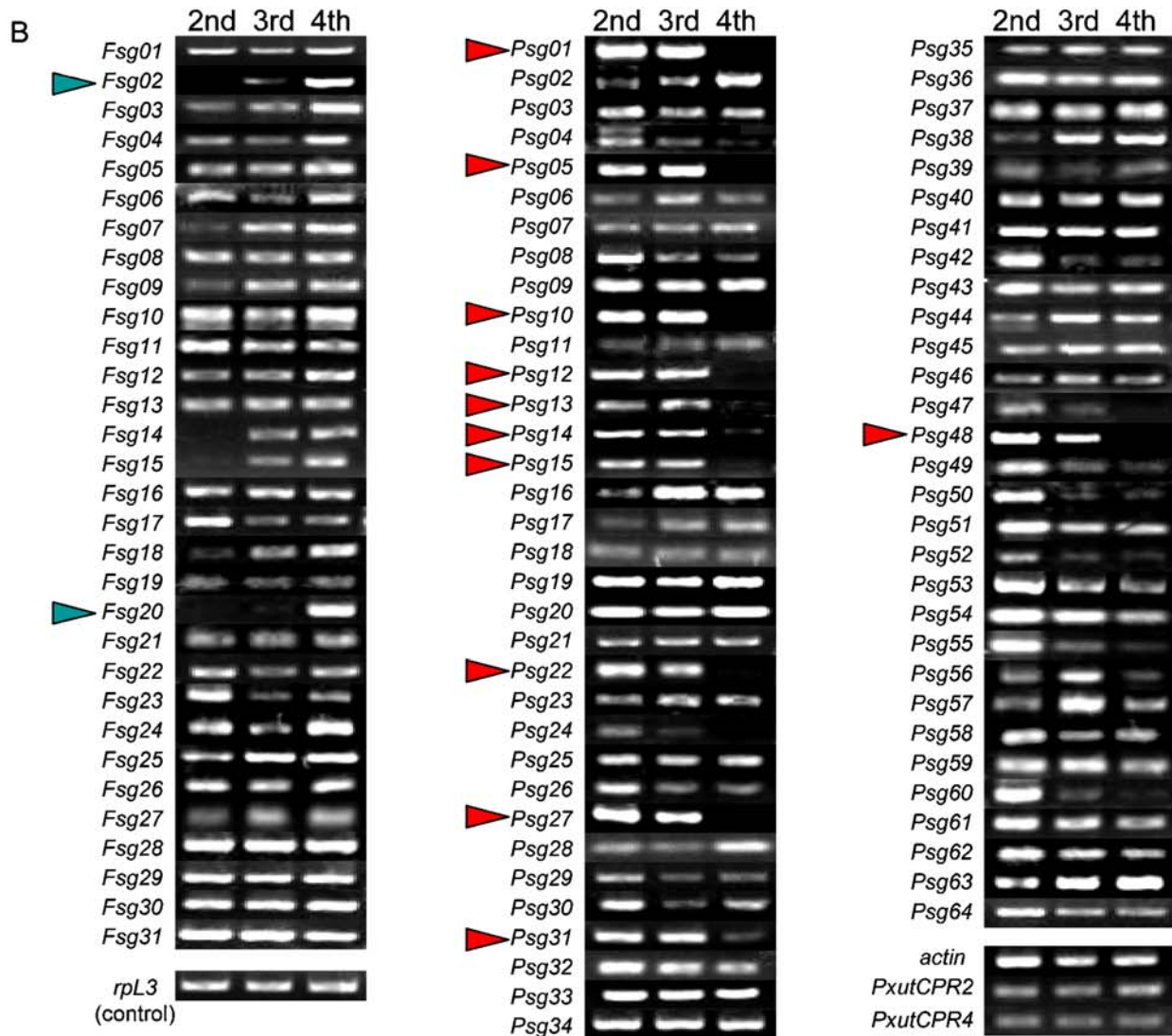
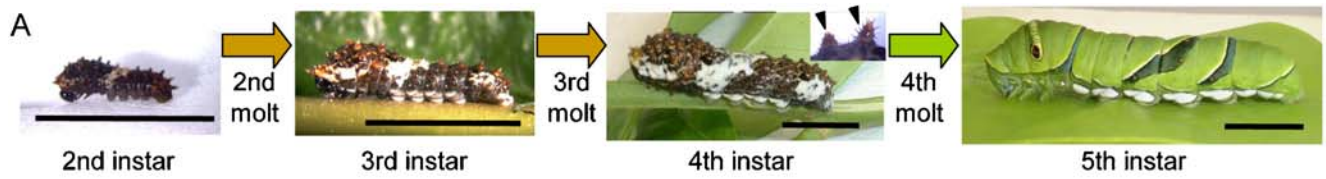
P. xuthus was purchased from Eiko-Kagaku (Osaka, Japan), kindly provided by Dr. Akira Yamanaka (Yamaguchi University, Japan), or collected from the field. Larvae were reared on leaves of *Zanthoxylum ailanthoides* (Rutaceae) at 25°C under long-day conditions (16 h light/8 h dark). The duration of the fourth instar (the third ecdysis to head capsule slippage, HCS) was approximately 4 days, and the duration of the fourth molt (pharate fifth instar) was approximately 21 h. The staging of molting period was based on the time when HCS occurred.

cDNA subtraction

A whole dorsal integument from the thoracic 2 segment to the abdominal 7 segment was dissected from a larva. After the fat body and muscle attached to the epidermis were carefully removed, total epidermal ribonucleic acid (RNA) was isolated using TRI reagent (Sigma, St. Louis, MO, USA) according to the manufacturer's instructions. Three or four specimens were used every 2 h from 0 to 20 after HCS of each molt (total 36–40 specimens at each molt). Poly (A)+RNA samples ($\approx 2 \mu\text{g}$) were prepared using Micro-Fast Track messenger RNA (mRNA) isolation kit (Clontech, San Jose, CA, USA) and subjected to cDNA subtraction by using the polymerase chain reaction (PCR)-select cDNA subtraction kit (Clontech) according to the manufacturer's instructions. We screened for final instar-specific genes (Fsg) and penultimate instar-specific genes (Psg). Fsgs were isolated by subtracting the epidermal cDNA population of the third molt from the epidermal cDNA population of the fourth molt, and Psgs were isolated by subtracting the epidermal cDNA population of the fourth molt from the epidermal cDNA population of the third molt (Fig. 1a). In the initial subtraction

Fig. 1 Larval body pattern of swallowtail butterfly *Papilio xuthus* and expression analysis of 98 genes identified by cDNA subtraction. **a** Larval pattern of the second, third, and fourth instar larvae. Lateral views of the third, fourth, and fifth instar larva are shown. *Inset* shows the tubercle structures (*arrowheads*) on the surface. *Scale bars*=1 cm. **b** Expression pattern of candidate genes by RT-PCR between the second, third, and fourth molts. The gene for the *ribosomal protein L3* (*rpL3*) was used as control. Three genes, *actin*, *PxutCPR2*, and *PxutCPR4*, were obtained from both subtractive libraries. *Green arrowheads* indicate the genes strongly expressed at the fourth molt, and *red arrowheads* indicate the genes weakly expressed at the fourth molt. **c** Relative abundance of mRNAs for cuticular protein genes was examined by semiquantitative RT-PCR. The dilutions for 1–6 thus range from none to 3,125-fold

screening, 40 randomly selected clones were sequenced. We selected the three redundant cDNA fragments as *Fsg01–Fsg03* in the fifth (final)-instar library and as *Psg01–Psg03* in the fourth (penultimate)-instar library. To avoid repeatedly sequencing clones corresponding to these genes and thus improve the chances of isolating low-abundance clones, we added specific primers for *Fsg01–Fsg03* or *Psg01–Psg03* cDNAs (Supplementary Table 1) to the T7 and SP6 primers used for subsequent subcloning (Supplementary Fig. 1a). When the cloned inserts were analyzed by PCR and electrophoresis, the redundant clones produced multiple bands because it is amplified by T7 and SP6 primers as well as redundant gene-specific primers, whereas the other clones produced single bands (Supplementary Fig. 1b). We therefore sequenced only the single-band clones by using a 3100 DNA sequencer (Applied Biosystems, Foster City, CA, USA). Among 1,044 randomly selected colonies from the final instar-specific library and 1,028 from the penultimate instar-specific library, 252 and 554 colonies were sequenced, respectively (including the 40 colonies analyzed initially, Supplementary Fig. 1). We also checked the eight multiple-band colonies of each library and confirmed that these colonies contain inserts of any one of *Fsg01–Fsg03* or *Psg01–Psg03*, respectively. The redundancy rates were $\{(988-514)/988\} \times 100 = 48\%$ and $\{(1,004-212)/1,004\} \times 100 = 79\%$, respectively. We obtained a sequence of ~ 650 nucleotides in average. To confirm the sequence of each cDNA, we performed sequencing from both ends. To identify the genes from which these partial sequences were derived, cDNA fragments were subjected to 5' and 3' rapid amplification of cDNA ends technique, and their nucleotide sequences were determined. Clustering was performed manually by collecting identical clones using a criterion of greater than 95% identity in a sequence greater than 100 nucleotides. We also referred to BLAST results for clustering. After clustering, 806 clones from both subtractive libraries were classified into 98 independent nonredundant contiguous cDNAs.



Expression analysis by RT-PCR

To analyze the transcription patterns of candidate stage-specific genes during molt, total epidermal RNA was isolated from larvae every 4 h during a molting period using TRI reagent. RNA (2 µg) was reverse-transcribed with random primer (N6) using a First-Strand cDNA Synthesis Kit (Amersham, Sunnyvale, CA, USA). Mixtures of the different timed cDNA samples (every 4-h cDNA samples of each stage) were used as the templates for the second, third, and fourth molts in initial screening and semiquantitative reverse transcriptase (RT) PCR, and each of the 4-h cDNA sample, which was not the same as used for mixture samples, was used as the templates to confirm stage specificity. After activating *Taq* DNA polymerase at 94°C for 2 min, PCR was performed for 30 cycles of 94°C for 30 s, 52–60°C for 30 s, and 72°C for 60 s. The primers for each gene are shown in Supplementary Table 1. The range of product length is 80–655 bp. The gene for the *ribosomal protein L3* (*rpL3*, Futahashi and Fujiwara 2006) that is expressed constitutively in the cell was used as an internal control for normalization of equal sample loading. For semiquantitative RT-PCR, RT products were subjected to fivefold serial dilutions followed by PCR with primer pairs for each gene simultaneously.

in situ hybridization

A whole dorsal integument from the thoracic 2 segment to the abdominal 7 segment was dissected from a larva. After the fat body and muscle attached to the epidermis were carefully removed, larval epidermis was fixed immediately in 4% paraformaldehyde in phosphate-buffered saline. In situ hybridization was performed as described in Futahashi and Fujiwara (2005). Digoxigenin-labeled antisense RNA probes were used, and the color reaction was performed at room temperature in 100 mM Tris-HCl, 100 mM NaCl, and 50 mM MgCl₂ (pH 9.5) containing 3.5 µl/ml 5-bromo-4-chloro-3-indolyl-phosphate, 4-toluidine salt, and 4.5 µl/ml nitroblue tetrazolium chloride. Digoxigenin-labeled sense-strand probes were used as negative controls. Template of the RNA probe for each gene was prepared using the same primers as RT-PCR analysis, and the probe lengths were described in Supplementary Table 1. The range of product length is 150–455 bp. Six to 24 pieces of integument were used in each probe.

Phylogenetic analysis

Sequences were aligned using Clustal_X (Thompson et al. 1997). Phylogenetic trees were constructed by the neighbor-joining method with the MEGA4 program (Tamura et al. 2007). The confidence of the various phylogenetic lineages

was assessed by the bootstrap analysis. We compared several related genes with the phylogenetic analysis. The following sequences were used to create the diagram (see Fig. 2): Fsg02, *P. xuthus* Fsg02 (AB264632); Aa-JF23, *Atrophaneura alcinous* JF23 (BAG32140); Ms-INSa, *Manduca sexta* insecticyanin-a (P00305); Ms-INSb, *M. sexta* insecticyanin-b (Q00630); Bm-bombyrin, *Bombyx mori* bombyrin (NP_001036872); Pb-BBP, *Pieris brassicae* bilin-binding protein (P09464); Pr-BP1, *Pieris rapae* bilin-binding protein 1; Pr-BP2, *P. rapae* bilin-binding protein 2. Amino acid sequences of Pr-BP1 and Pr-BP2 were referred to Kim et al. (2006).

The culturing of juvenile instar epidermis

To investigate the effects of JH on the gene expression, we dissected a whole dorsal integument from the thoracic 2 segment to the abdominal 7 segment of fourth instar epidermis containing tubercle regions just after the third ecdysis and cultured them with or without JHA (methoprene) in 0.5 ml Grace's medium at 25°C. To alter JH conditions, we precultured epidermis with 1 µg/ml methoprene (+JH) or without methoprene (-JH) for 24 h (Fig. 7a). After the JH treatment, we applied 20E to the medium for 16 h. The culture medium was then replaced with medium lacking 20E for 24 h (Fig. 7a). This 20E treatment mimics an ecdysone pulse during the fourth molt and is necessary for the induction of the fourth-molt-specific genes (Hiruma and Riddiford 1990). We compared mRNA expression during the fourth molt by RT-PCR.

Results

Subtraction screening of stage-specific epidermal cDNA clones

After analyzing 2,072 clones (including redundant clones) from two subtractive libraries, 98 independent nonredundant contiguous cDNAs were classified. Three cDNAs were found in both the penultimate- and final instar libraries. BLASTx searches for *Bombyx* cuticular protein genes (Futahashi et al. 2008) and NCBI database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) against all nonredundant sequences in GenBank showed that these three genes bore significant similarity to actin (accession no. CAA28818, $P < e^{-150}$), and two cuticular protein genes (BmorCPR2, $P = e^{-47}$ and BmorCPR4, $P = e^{-32}$) of the silkworm *B. mori*, respectively. We denoted these genes as *actin*, *PxutCPR2*, and *PxutCPR4*, respectively. Expression analysis with RT-PCR indicated that these genes are not differentially expressed (Fig. 1b,c), suggesting that their expression is ubiquitous and is not completely eliminated by the cDNA subtraction method. Because out of over 2,000 clones, only three genes were present in both

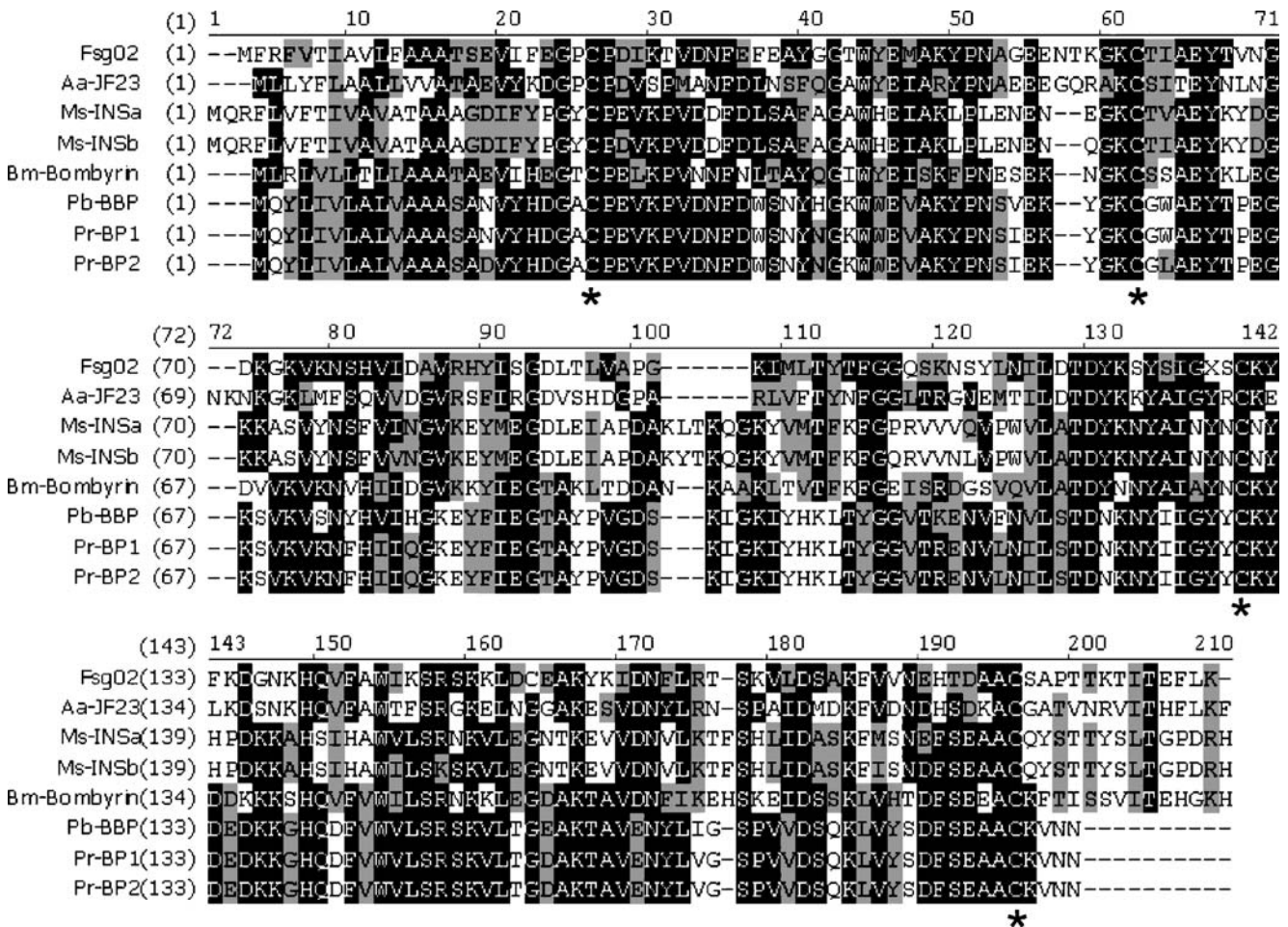


Fig. 2 Deduced amino acid sequence of Fsg02 protein and alignment of amino acid sequences of other related genes. The conserved four cysteine residues are denoted by asterisks. Conserved residues are

highlighted, and functionally or structurally similar residues are shadowed

(reciprocally subtracted) libraries, we thought that the subtraction step was successful. Excluding these three genes, 31 candidate Fsgs (*Fsg01–Fsg31*), and 64 candidate Psgs (*Psg01–Psg64*) were isolated (Tables 1 and 2). A BLASTx search for *Bombyx* cuticular protein genes (Futahashi et al. 2008) showed that 39 genes have sequence similarity (cutoff threshold E values: $P < e^{-10}$). Other than cuticular protein genes, BLASTx searches against all nonredundant sequences in GenBank showed that 49 genes display significant similarity with known proteins (cutoff threshold E values: $P < e^{-10}$), and a further gene, *Fsg25*, had weak similarity to known proteins ($P = 3e^{-9}$). Among the remaining six transcripts, four do not have long open reading frames (>200 bp), suggesting that these transcripts may be 3' untranslated regions. The deduced amino acid sequences of *Fsg20* and *Psg36* showed no significant similarity with known proteins (Tables 1 and 2). *Fsg02*, *Psg12*, and *Psg14* are also reported as *BBP*, *HCPI1*, and *HCP2* by Futahashi and Fujiwara (2008).

Stage-specific expression analysis among the second, third, and fourth molts in the epidermis

We first screened genes with stage-specific expression by RT-PCR. The expression of genes involved in the larval body markings is expected to be changed during the fourth molt. As shown in Fig. 1b, two Fsgs (*Fsg02*, *Fsg20*) were obviously expressed more strongly at the fourth molt than at the second or third molt (Fig. 1b, green arrowheads), and 11 Psgs (*Psg01*, *Psg05*, *Psg10*, *Psg12*, *Psg13*, *Psg14*, *Psg15*, *Psg22*, *Psg27*, *Psg31*, and *Psg48*) were obviously expressed more weakly at the fourth molt than at the second or third molt (Fig. 1b, red arrowheads). To confirm this stage specificity, the expression levels of these genes were also analyzed every 4 h during the third and fourth molts (Supplemental Fig. 2). We also confirmed that these genes were differentially expressed between the third and the fourth molts by semiquantitative RT-PCR (Fig. 1c). Although more quantitative analysis is needed to identify the stage specificity more

Table 1 Final instar specific cDNA clones from cDNA subtraction and match probabilities ($P < e^{-9}$) resulting from BLASTX search

Cluster name	clone numbers	Accession No.	Identified region	fragment length (bp)	Top hit ID (species)	E value (Blast X)	Accession No.
Fsg01	29	AB264631	full	1462	cuticular protein BmorCPT2 (Bombyx mori)	1.00E-125	BR000651
Fsg02	15	AB264632	full	784	Bombyrin (Bombyx mori)	4.00E-43	NP_001036872
Fsg03	8	AB264633	full	1100	C-type lectin (Anopheles gambiae)	6.00E-115	AAAB01008980
Fsg04	41	AB264634	full	659	cuticular protein BmorCPG6 (Bombyx mori)	4.00E-30	BR000427
Fsg05	29	AB264635	full	809	cuticular protein BmorCPG14 (Bombyx mori)	2.00E-87	BR000435
Fsg06	17	AB264636	full	1524	cuticular protein BmorCPG12 (Bombyx mori)	6.00E-178	BR000433
Fsg07	16	AB264637	full	1164	cuticular protein BmorCPT3 (Bombyx mori)	5.00E-144	BR000652
Fsg08	16	AB264638	full	1246	cuticular protein BmorCPG8 (Bombyx mori)	2.00E-107	BR000429
Fsg09	14	AB264639	full	1104	tropomyosin isoform 1 (Bombyx mori)	9.00E-146	NP_001103782
Fsg10	13	AB264640	full	1229	cuticular protein BmorCPG7 (Bombyx mori)	4.00E-94	BR000428
Fsg11	7	AB264641	full	3253	sarcalumenin (Aedes aegypti)	0.00E+00	EAT45792
Fsg12	6	AB264642	full	1480	cuticular protein BmorCPG24 (Bombyx mori)	0.00E+00	BR000445
Fsg13	5	AB264643	full	492	cuticular protein BmorCPR5 (Bombyx mori)	8.00E-25	BR000506
Fsg14	3	AB264644	full	611	cuticular protein BmorCPG15 (Bombyx mori)	3.00E-25	BR000436
Fsg15	3	AB264645	full	863	cuticular protein BmorCPG20 (Bombyx mori)	2.00E-61	BR000441
Fsg16	3	AB264646	full	606	cuticular protein BmorCPR3 (Bombyx mori)	4.00E-42	BR000504
Fsg17	3	AB264647	partial	1827	low homology		
Fsg18	2	AB264648	full	761	cuticular protein BmorCPH33 (Bombyx mori)	3.00E-82	BR000500
Fsg19	2	AB264649	partial	290	cuticular protein BmorCPG15 (Bombyx mori)	5.00E-32	BR000436
Fsg20	2	AB264650	full	1643	low homology (yellow related gene)		
Fsg21	1	AB264651	full	1026	cuticular protein BmorCPR17 (Bombyx mori)	3.00E-70	BR000518
Fsg22	1	AB264652	full	1084	Osiris 19 (Drosophila melanogaster)	6.00E-49	NP_649640
Fsg23	1	AB264653	partial	1965	organic cation transporter (Aedes aegypti)	1.00E-25	EAT38951
Fsg24	1	AB264654	full	2062	similar to CG12164 gene product (Tribolium castaneum)	2.00E-168	XP_970717
Fsg25	1	AB264655	full	919	similar to CG5597 gene product (Apis mellifera)	3.00E-09	XM_393193
Fsg26	1	AB264656	full	3506	elongase, putative (Aedes aegypti)	8.00E-118	EAT38554
Fsg27	1	AB264657	full	557	cuticular protein BmorCPG3 (Bombyx mori)	4.00E-47	BR000424
Fsg28	1	AB264658	full	1360	cuticular protein BmorCPR41 (Bombyx mori)	5.00E-37	BR000542
Fsg29	1	AB264659	full	1487	tropomyosin T (Bombyx mori)	0.00E+00	NP_001040221
Fsg30	1	AB264660	full	1126	tropomyosin 1 (Bombyx mori)	4.00E-149	NP_001040445
Fsg31	1	AB264661	full	6395	muscle myosin heavy chain (Drosophila melanogaster)	0.00E+00	NP_724004

Cuticular protein genes are denoted in color (red: RR-1, green: Tweedle, purple: glycine-rich, brown: others)

clearly, we used these results as an initial discovery tool. Other genes also show stage specificity. Expression of *Fsg07*, *Fsg09*, *Fsg14*, *Fsg15*, *Fsg18*, *Psg16*, and *Psg38* are stronger at the third and fourth molts than the second molt (Fig. 1b). Expression of *Psg24*, *Psg42*, *Psg47*, *Psg49*, *Psg50*, *Psg55*, and *Psg60* are weaker at the third and fourth molts than the second molt (Fig. 1b). However, the specificity did not correspond to the fourth molt, and we therefore exclude these genes from further study.

Fsg02 is similar to blue pigment-binding proteins

Fsg02 has sequence similarity to the lipocalin family of butterflies and moths (Fig. 2), including JF23 of *A. alcinous* (48% identity), bombyrin of *B. mori* (43% identity), bilin-binding protein of *P. brassicae* (41% identity), two bilin-binding proteins of *P. rapae* (41% and 41% identity), and insecticyanin-b and insecticyanin-a of *M. sexta* (39% and 38% identity, respectively). Bilin-binding proteins of *P. rapae* was referred to Kim et al. (2008), and other genes were displayed significant sequence similarity by BLASTX search against all nonredundant sequences in GenBank. The four conserved cysteine residues of these proteins are also conserved in Fsg02 (Fig. 2, asterisks). All of them are binding proteins of lipophilic ligands. JF23 is a putative binding protein for lipophilic substances related to butterfly

oviposition (Tsuchihara et al. 2000), and bombyrin is a neural lipocalin, especially present in the central nervous system (Sakai et al. 2001). Importantly, both bilin-binding protein and insecticyanin are blue pigment-binding proteins, which are associated with larval greenish coloration (Riddiford et al. 1990; Schmidt and Skerra 1994; Kim et al. 2006). The strong similarity between the Fsg02 amino acid sequence and the sequences of these proteins, together with its specific expression at the fourth molt, suggests that Fsg02 is involved in the greenish coloration of *P. xuthus* larvae. As described above, Fsg20 shows no similarity to other sequences in the databases, and its homolog is not even found in the *B. mori* genome.

Fsg02 and Fsg20 expression correlates strongly with the green region of the final instar

We used whole-mount in situ hybridization to investigate the relationships between the distribution of the two Fsgs (*Fsg02*, *Fsg20*) mRNAs and larval green coloration. *Fsg02* was strongly expressed in the presumptive green region of the fifth instar larvae (Fig. 3a,b). In the final instar, *P. xuthus* larva represents overall green coloration, except for white and black markings. Green pigments were not observed in the black and white stripes at the edge of the V-shaped markings (Fig. 3a). *Fsg02* expression was not

Table 2 Penultimate instar specific cDNA clones from cDNA subtraction and match probabilities ($P < e^{-9}$) resulting from BLASTX search

Cluster name	clone numbers	Accession No.	Identified region	fragment length (bp)	Top hit ID (species)	E value (Blast X)	Accession No.
Peg01	23	AB264662	full	920	cuticular protein BmorCPR77 (Bombyx mori)	1.00E-75	BR000578
Peg02	13	AB264663	partial	2533	serine protease (Aedes aegypti)	2.00E-127	EAT42573
Peg03	4	AB264664	full	2503	CG1673 gene product (Drosophila melanogaster)	3.00E-135	NP_572884
Peg04	40	AB264665	full	1550	glycoside hydrolases (Aedes aegypti)	6.00E-149	EAT38910
Peg05	38	AB264666	full	643	cuticular protein BmorCPR76 (Bombyx mori)	9.00E-68	BR000577
Peg06	34	AB264667	full	741	cuticular protein BmorCPR6 (Bombyx mori)	6.00E-28	BR000461
Peg07	26	AB264668	partial	1707	chromator (Drosophila melanogaster)	5.00E-17	NP_649426
Peg08	25	AB264669	full	2520	hypothetical protein (Rattus norvegicus)	6.00E-11	XP_001068310
Peg09	24	AB264670	full	830	cuticular protein BmorCPR59 (Bombyx mori)	2.00E-60	BR000560
Peg10	24	AB264671	full	587	cuticular protein BmorCPR71 (Bombyx mori)	3.00E-72	BR000572
Peg11	24	AB264672	partial	772	low homology		
Peg12	21	AB264673	full	790	cuticular protein BmorCPFL4 (Bombyx mori)	5.00E-42	BR000421
Peg13	21	AB264674	full	758	cuticular protein BmorCPR71 (Bombyx mori)	2.00E-68	BR000572
Peg14	20	AB264675	full	729	cuticular protein BmorCPFL4 (Bombyx mori)	6.00E-36	BR000421
Peg15	19	AB264676	full	707	cuticular protein BmorCPR99 (Bombyx mori)	3.00E-72	BR000600
Peg16	14	AB264677	full	736	cuticular protein BmorCPR4 (Bombyx mori)	8.00E-49	BR000454
Peg17	13	AB264678	full	2145	DEAD box polypeptide 5 (Bombyx mori)	0.00E-00	NP_001037582
Peg18	13	AB264679	full	1467	pyruvate dehydrogenase (Aedes aegypti)	9.00E-144	EAT39862
Peg19	12	AB264680	partial	1551	sec71 (Drosophila melanogaster)	1.00E-84	NP_69675
Peg20	9	AB264681	full	970	ribosomal protein L23A (Bombyx mori)	1.00E-137	NP_001037229
Peg21	8	AB264682	full	4681	leucine-rich transmembrane protein (Aedes aegypti)	0.00E-00	EAT45504
Peg22	8	AB264683	full	712	cuticular protein BmorCPR97 (Bombyx mori)	3.00E-66	BR000598
Peg23	7	AB264684	full	3489	importin beta-3 (Aedes aegypti)	0.00E-00	EAT37896
Peg24	7	AB264685	partial	587	low homology		
Peg25	6	AB264686	full	1214	similar to CG6803 gene product (Tribolium castaneum)	7.00E-75	XP_975727
Peg26	6	AB264687	full	2026	lethal (2) 01810 (Drosophila melanogaster)	3.00E-128	NP_620115
Peg27	6	AB264688	full	672	cuticular protein BmorCPR67 (Bombyx mori)	5.00E-71	BR000568
Peg28	5	AB264689	partial	2433	Titin-like protein (Bombyx mori)	0.00E-00	AB090308
Peg29	4	AB264690	full	4480	Contactin (Drosophila melanogaster)	0.00E-00	NP_649461
Peg30	4	AB264691	partial	924	similar to CG10263 product (Tribolium castaneum)	2.00E-67	XP_971828
Peg31	4	AB264692	full	1370	cuticular protein BmorCPR54 (Bombyx mori)	2.00E-126	BR000555
Peg32	4	AB264693	full	754	cuticular protein BmorCPR29 (Bombyx mori)	2.00E-37	BR000450
Peg33	4	AB264694	full	1671	elongation factor 1 alpha (Bombyx mori)	0.00E-00	NP_001037510
Peg34	3	AB264695	full	616	cuticular protein BmorCPR59 (Bombyx mori)	3.00E-41	BR000560
Peg35	3	AB264696	full	1520	secreted protein acidic and rich in cysteine (Bombyx mori)	2.00E-149	NP_001040421
Peg36	3	AB264697	full	1535	low homology		
Peg37	3	AB264698	full	691	cuticular protein BmorCPR4 (Bombyx mori)	4.00E-37	BR000425
Peg38	2	AB264699	full	703	cuticular protein BmorCPR19 (Bombyx mori)	6.00E-54	BR000520
Peg39	2	AB264700	full	1695	cuticular protein BmorCPGI3 (Bombyx mori)	2.00E-25	BR000434
Peg40	2	AB264701	full	1393	squid 1 (Bombyx mori)	4.00E-28	NP_001037323
Peg41	2	AB264702	full	1969	very low-density lipoprotein receptor (Bombyx mori)	0.00E-00	ABF51235
Peg42	2	AB264703	full	1657	eukaryotic initiation factor 4B-2 (Bombyx mori)	2.00E-107	NP_001091833
Peg43	2	AB264704	full	2265	translation initiation factor 5A (Bombyx mori)	9.00E-89	NP_001037538
Peg44	1	AB264705	full	943	muscular protein 20 (Bombyx mori)	6.00E-95	NP_001040476
Peg45	1	AB264706	full	1052	3-dehydrocyclohexanone 3B-reductase (Trichoplusia ni)	1.00E-84	AF409102
Peg46	1	AB264707	full	1049	cuticular protein BmorCPR41 (Bombyx mori)	5.00E-55	BR000542
Peg47	1	AB264708	full	547	cuticular protein BmorCPR11 (Bombyx mori)	4.00E-41	BR000466
Peg48	1	AB264709	full	753	cuticular protein BmorCPR103 (Bombyx mori)	2.00E-81	BR000604
Peg49	1	AB264710	partial	1778	glycoside hydrolases (Aedes aegypti)	2.00E-146	EAT32750
Peg50	1	AB264711	full	3121	Translocase of outer membrane 34 (Drosophila melanogaster)	0.00E-00	NP_524796
Peg51	1	AB264712	full	2128	tetraspanin E (Bombyx mori)	2.00E-21	ABD36170
Peg52	1	AB264713	partial	1124	low homology		
Peg53	1	AB264714	full	4251	similar to CG13676 gene product (Tribolium castaneum)	2.00E-92	XP_972676
Peg54	1	AB264715	partial	2812	similar to CG17342 gene product (Tribolium castaneum)	1.00E-10	XP_970209
Peg55	1	AB264716	partial	2114	coracle protein, putative (Aedes aegypti)	7.00E-46	EAT40392
Peg56	1	AB264717	full	981	cuticular protein BmorCPFL4 (Bombyx mori)	3.00E-37	BR000421
Peg57	1	AB264718	full	1628	CG3590 gene product (Drosophila melanogaster)	0.00E-00	NP_650586
Peg58	1	AB264719	full	1693	DnaJ homolog subfamily A member 2 (Tribolium castaneum)	2.00E-155	XP_970724
Peg59	1	AB264720	full	1768	nucleosome assembly protein isoform 1 (Bombyx mori)	2.00E-148	ABF51222
Peg60	1	AB264721	full	1215	ubiquitin carboxyl-terminal hydrolase (Tribolium castaneum)	4.00E-107	XP_967733
Peg61	1	AB264722	partial	424	ras-related GTP-binding protein Rab11 (Bombyx mori)	1.00E-33	NP_001037618
Peg62	1	AB264723	full	938	proteasome alpha 3 subunit (Bombyx mori)	3.00E-133	NP_001040387
Peg63	1	AB264724	full	665	60S ribosomal protein L7/L12 precursor (Bombyx mori)	1.00E-65	NP_001040523
Peg64	1	AB220983	full	1848	phenylalanine hydroxylase (Papilio xuthus)	0.00E+00	AB220983

Cuticular protein genes are denoted in color (red RR-1, blue: RR-2, yellow: CPFL, purple: glycine-rich, brown: others)

detected in the two lines corresponding to these black (Fig. 3a, arrows) and white stripes (Fig. 3a, arrowheads). To confirm the relationship between *Fsg02* expression and the V-shaped markings, an integument sample was taken just before the fourth ecdysis (17 h after HCS), at which time the black stripe was already pigmented and used for in situ hybridization. One line of the *Fsg02*-nonexpressed region is identical to the black stripe of the V-shaped markings (Fig. 3a, arrows). During the third molt, the *Fsg02* signal was not detectable, except for a spot on the abdomen (Fig. 3c, arrows). These spots probably coincide with the pale blue spot of the fourth instar larvae (Fig. 3c, white arrows). Thus, as expected by its presumed blue pigment-binding function, *Fsg02* was specifically expressed in the presumptive green region of the fifth instar and the pale blue spot region of the third instar.

Like *Fsg02*, *Fsg20* was also expressed in the presumptive green region of the fifth instar larva (Fig. 3a,b). Unlike that of *Fsg02*, the *Fsg20* signal was not detected in one line of the V-shaped markings (Fig. 3a, arrows). Around eyespot markings, expression of *Fsg20* differed from that of *Fsg02*. The *Fsg02* signal was detected at the black eyespot region, but the *Fsg20* signal was not (Fig. 3b, arrows). Around the black eyespot region, there were some regions where the *Fsg20* signal was detected, but the *Fsg02* signal was not (Fig. 3b, arrowheads). The body coloration was yellowish at the area where the *Fsg20* signal alone was detected, suggesting that *Fsg20* is associated with yellowish coloration. During the third molt, the *Fsg20* signal was not detectable, even in the pale spot region (Fig. 3c). These results suggest that *Fsg20* expression is associated with the yellow or green region of the final instar.

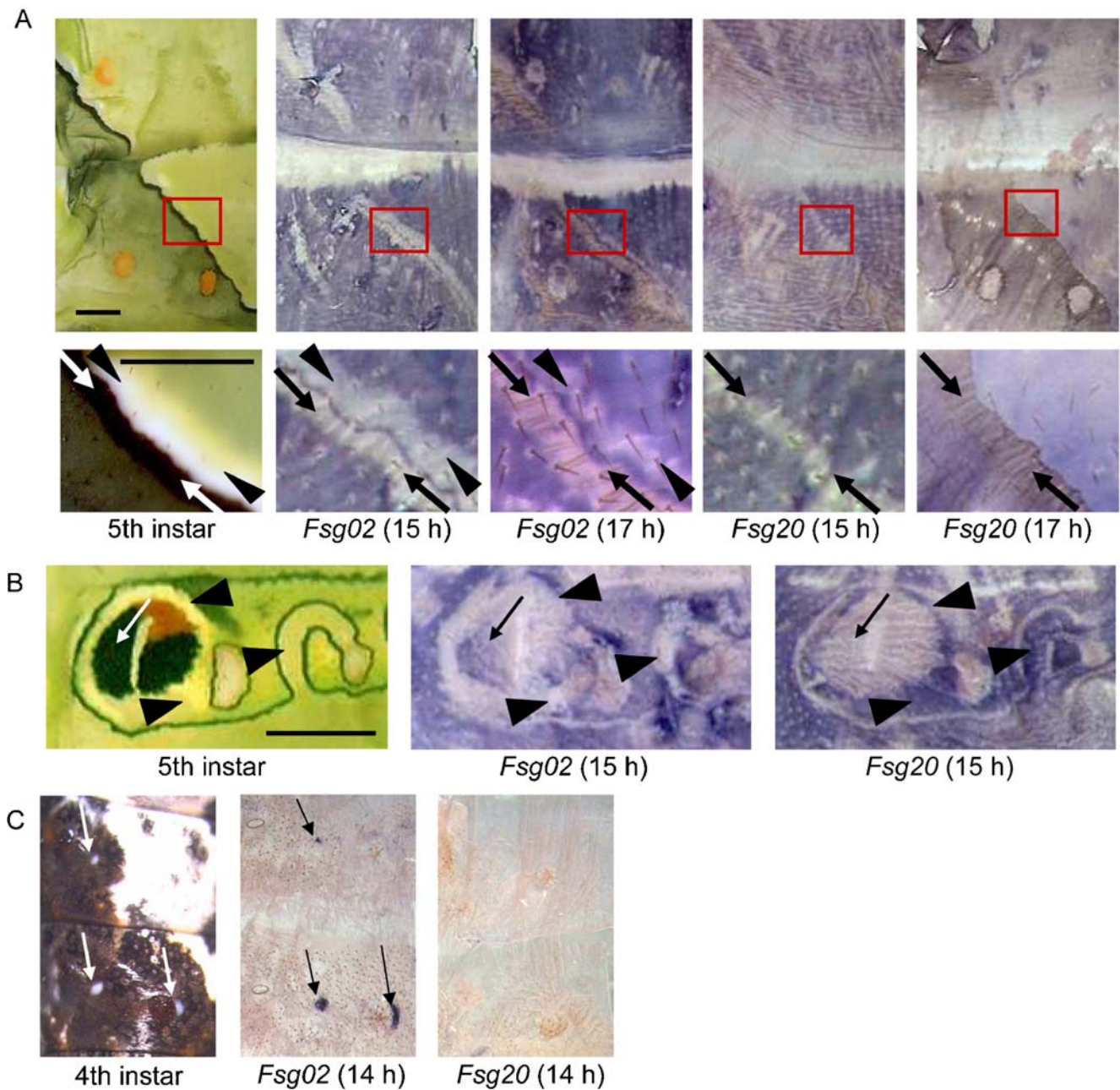


Fig. 3 Expression patterns of *Fsg02* and *Fsg20*. Expression patterns of *Fsg02* and *Fsg20* at 15 or 17 h after HCS during the molt to the fifth instar (**a**, **b**) and at 14 h after HCS during the molt to the fourth instar (**c**) are shown (abdominal 4–5 segments [**a**, **c**] and thoracic 3 segment [**b**]). Lower panels in **a** show the magnification of red boxes in the upper panels. The black region was already pigmented at 17 h

after HCS. Arrows and arrowheads in **a** indicate the black and white stripe at the edge of V-shaped markings, respectively. Arrows in **b** indicate the eyespot black region, and arrowheads in **b** indicate the region where *Fsg20* signal is detected but *Fsg02* signal is not. Arrows in **c** indicate the pale spot region. Scale bars=1 mm

Several cuticular proteins are associated with juvenile instar-specific tubercle structures

Between two subtractive libraries, we found 41 putative cuticular protein genes (17 Fsgs, 22 Psgs, *PxutCPR2*, and *PxutCPR4*, the latter two genes being found in both subtractive libraries, Tables 1 and 2). Each of them contains

a putative signal peptide sequence at the N terminus. Among these 41 cuticular proteins, 11 genes (*Psg01*, *Psg05*, *Psg10*, *Psg12*, *Psg13*, *Psg14*, *Psg15*, *Psg22*, *Psg27*, *Psg31*, and *Psg48*) were strongly expressed at the second and third molts and thought to be juvenile instar specific in the larval epidermis (Fig. 1b, red arrowheads). Juvenile instar larvae of *P. xuthus* have many tubercle structures on their surfaces

(Fig. 1a, inset, Fig. 4, arrowheads). Cuticular proteins are necessary for exoskeleton construction and are good candidates for regulators of juvenile instar-specific tubercle characteristics. Interestingly, all of these 11 young larva-specific Psgs were expressed specifically in the presumptive tubercle regions at 14 h after HCS during the third molt (Fig. 4, arrowheads). The expression pattern of each Psg gene was slightly different. In the small tubercle region, Psg01 and Psg10 signals were barely detected, Psg05 and Psg48 signals were faintly detected, whereas other Psgs were clearly expressed (Fig. 4, red arrowheads), suggesting that the composition of cuticular proteins was different due to the

tubercle size. We also analyzed spatial expression pattern 13 h after HCS and found similar expression patterns (data not shown). During the fourth molt, the expression of these genes was not detectable (Futahashi and Fujiwara 2008, and data not shown).

Relationships between cuticular protein motif and expression pattern

Cuticular proteins have been categorized according to their conserved motifs. The type of each cuticular protein gene is denoted in color in Tables 1 and 2 (red: RR-1, blue: RR-2,

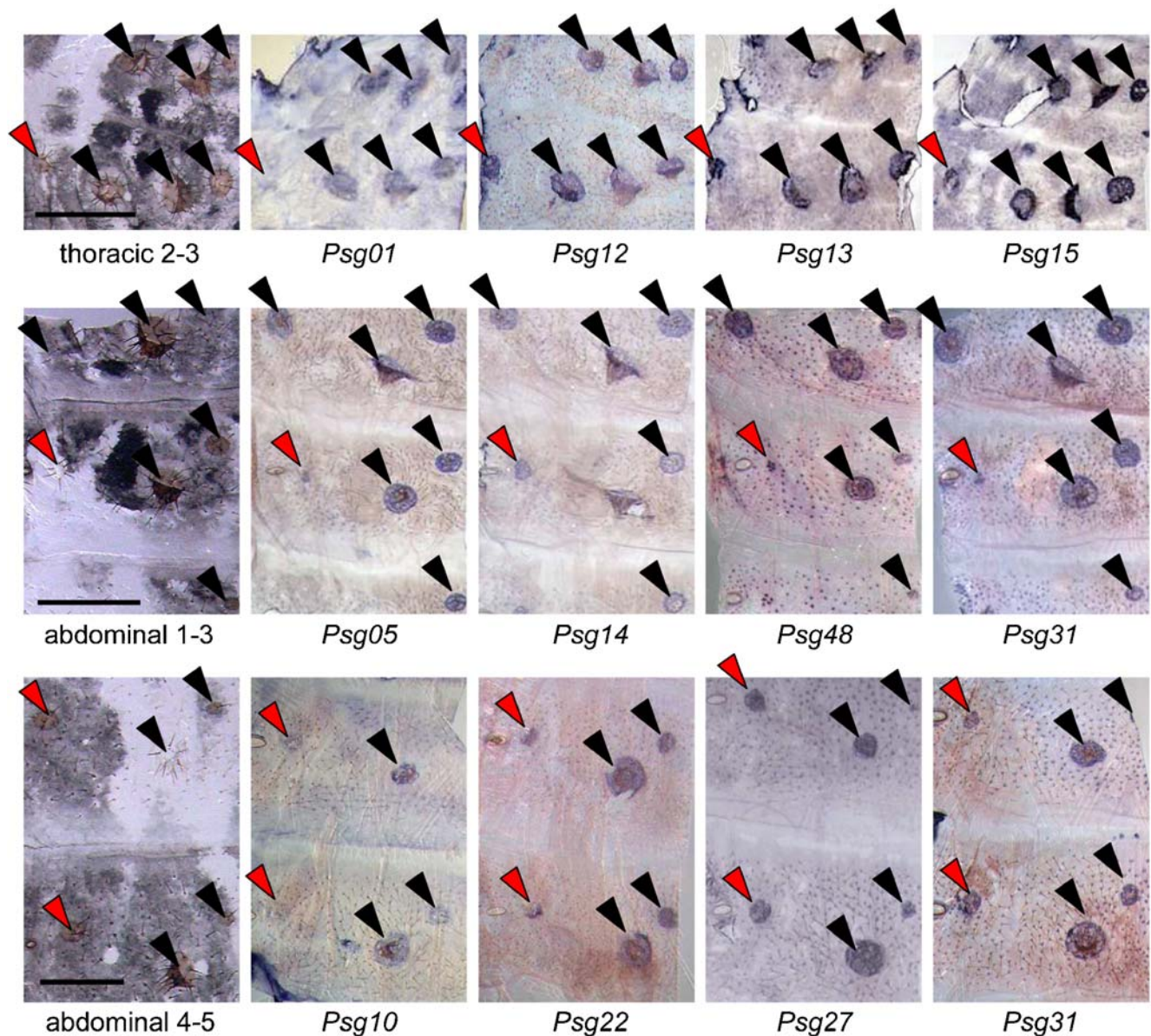


Fig. 4 Expression patterns of tubercle-associated hard cuticular protein genes at 14 h after HCS during the third molt. *Black arrowheads* indicate the position of large tubercle regions, whereas

red arrowheads indicate the position of small tubercle regions. The fourth instar patterns are also shown (*leftmost*). Scale bars=1 mm

green: Tweedle, yellow: cuticular protein with a 44-amino acid motif-like (CPFL), purple: glycine-rich, and brown: others). The most prevalent motif is the Rebers and Riddiford Consensus (R&R Consensus), which in an extended form is known to bind chitin (Rebers and Willis 2001; Togawa et al. 2004, Willis et al. 2005). Proteins with R&R Consensus can be split into three groups, RR-1, RR-2, and RR-3, according their amino acid sequences (Willis et al. 2005). It is generally claimed that RR-1 is characteristic of proteins in soft and flexible cuticles, and RR-2 proteins are associated with stiff and hard cuticles (Willis et al. 2005). Because the precise localization of individual cuticular protein mRNAs has been rarely reported, we analyzed the relationships between the cuticular protein motifs and their expression pattern. Nine genes (*Fsg13*, *Fsg16*, *Fsg21*, *Fsg28*, *Psg31*, *Psg38*, *Psg46*, *PxutCPR2*, and *PxutCPR4*) belong to RR1 protein, and ten genes (*Psg01*, *Psg05*, *Psg09*, *Psg10*, *Psg13*, *Psg15*, *Psg22*, *Psg27*, *Psg34*, and *Psg48*) belong to RR2 protein judged by a tool based on profile hidden Markov models on the cuticleDB website (<http://bioinformatics2.biol.uoa.gr/cuticleDB/index.jsp>; Supplementary Table 2). Most of these juvenile instar tubercle-specific proteins belong to RR-2 protein (Fig. 6); however, *Psg31* belongs to RR-1 protein (Supplementary Table 2).

Among RR-2 protein genes, *Psg09* and *Psg34* seem to be not stage specific (Fig. 1b). By whole-mount in situ hybridization, we found that *Psg09* was expressed broadly, whereas *Psg34* was expressed at the basal cells of hairs at both the third and fourth molts (Fig. 5). Except for *Psg31*, RR-1 protein genes were expressed at both the third and fourth molts (Fig. 1b). Interestingly, the expression of RR-1 protein gene *PxutCPR4* was weaker at the tubercle than the other region (Fig. 5, arrowheads). We also analyzed the spatial expression pattern of cuticular protein without R&R Consensus. The expression of the glycine-rich type of cuticular protein gene *Psg32* was quite similar to that of *PxutCPR4* (Fig. 5). As for CPFL proteins, there are three CPFL genes (*Psg12*, *Psg14*, and *Psg56*) in two subtractive libraries, and all three have sequence similarity with *BmorCPFL4* (Table 2). Although expression of *Psg12* and *Psg14* coincided with the tubercle region (Fig. 4), *Psg56* was broadly expressed in both the third and the fourth molts (Fig. 5). We found similar expression patterns 13 and 14 h after HCS during the third molt and 13, 14, and 15 h after HCS during the fourth molt (data not shown). Thus, the expression pattern of each cuticular protein genes was intricately regulated and was not entirely associated with the sequence similarity or specific motif. Phylogenetic analysis of cuticular proteins showed that tubercle-associated

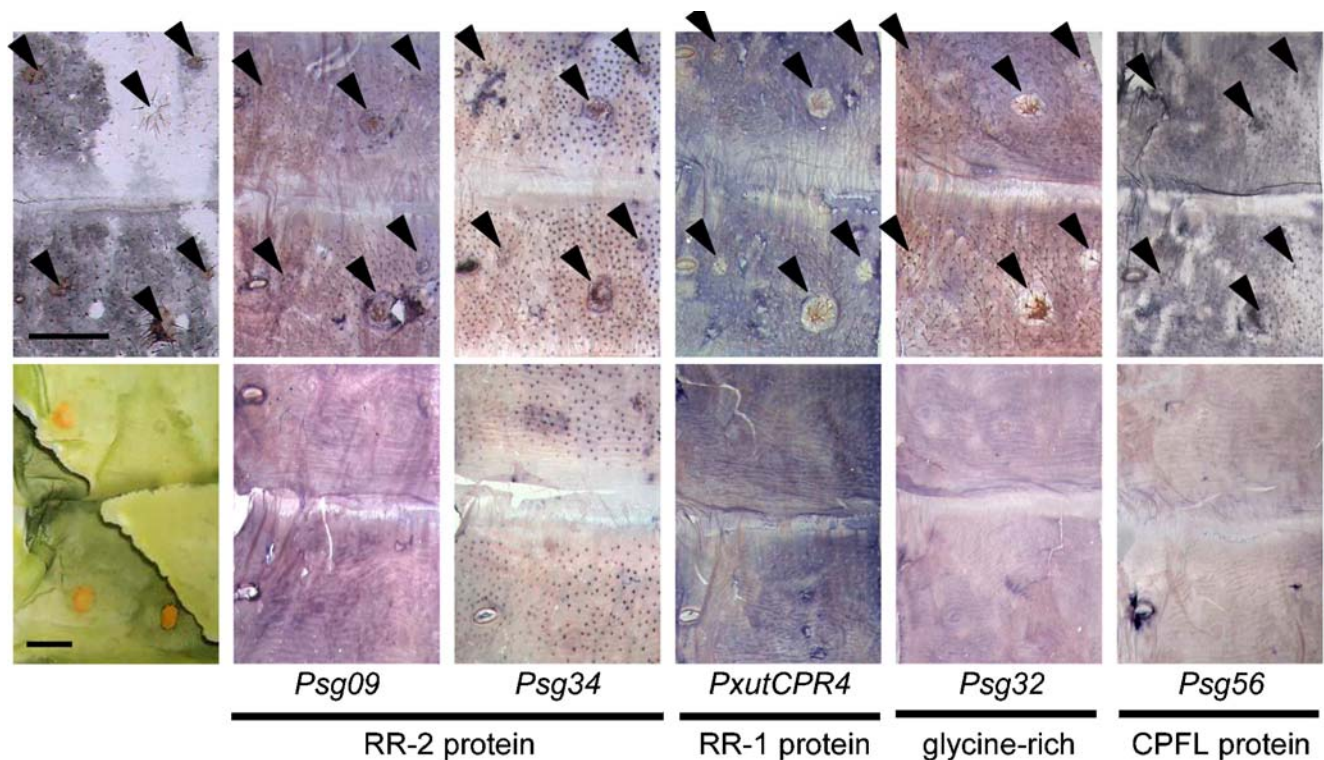


Fig. 5 Expression patterns of cuticular protein genes at 14 h at abdominal 4–5 segments after HCS during the third molt (*upper panels*) and at 15 h after HCS during the fourth molt (*lower panels*).

Arrowheads indicate the tubercle positions of the fourth instar larva. The fourth and fifth instar patterns are also shown (*leftmost*). Type of cuticular proteins are shown *below*. Scale bars=1 mm

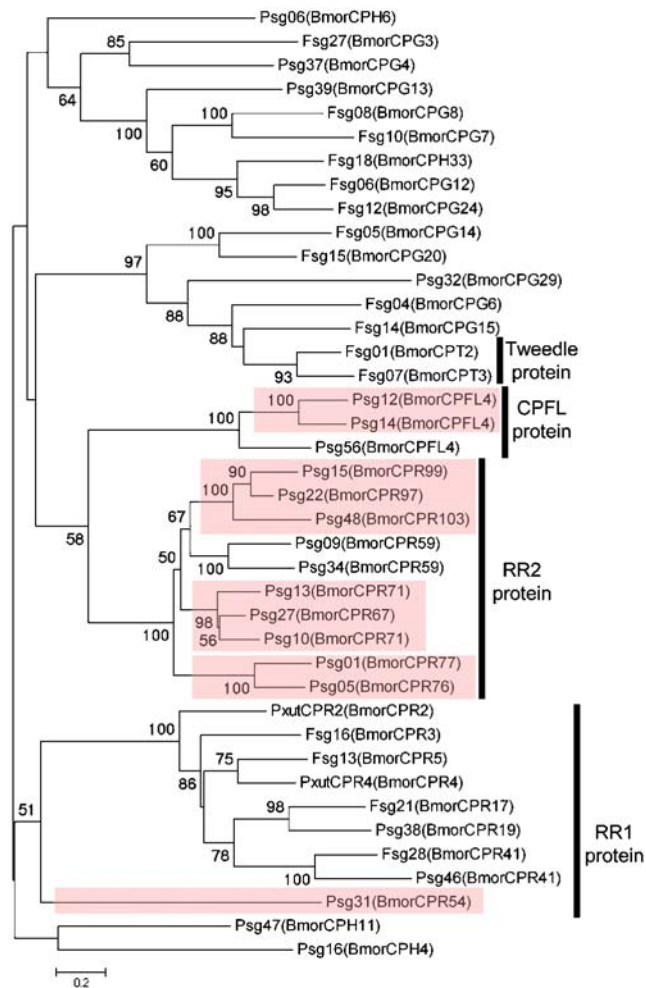


Fig. 6 Neighbor-joining tree of cuticular protein genes based on their amino acid sequences. The numbers at the tree edges represent the bootstrap values. The scale bar indicates the evolutionary distance between the groups. Red boxes indicate the tubercle-associated cuticular proteins of *P. xuthus* in this study

cuticular proteins do not form a single cluster (Fig. 6, red boxes) and suggests that association with tubercle development evolved independently.

Expression of stage-specific cuticular proteins is regulated by JH

We previously found that JH treatment induces *HCP1* (*Psg12*) and *HCP2* (*Psg14*) expression and suppresses *BBP* (*Fsg02*) expression in vivo (Futahashi and Fujiwara 2008). To investigate JH effect of expression of stage-specific Fsgs and Psgs in vitro, we cultured juvenile instars' epidermis with or without JHA (methoprene, Fig. 7a). The expression of the young larva-specific Psg mRNAs was suppressed in the absence of JHA, whereas

the expression of *PxutCPR2* and *PxutCPR4* mRNAs was not (Fig. 7b, -JH). However, JHA treatment of the epidermis induced the expression of young larva-specific Psg mRNAs during the fourth molt, caused by an ecdysone pulse (Fig. 7b, +JH), which is consistent with the results shown in Fig. 1b. These results suggest that JH regulates the expression of the young larva-specific Psg genes, and high levels of JH in the early phase of the fourth instar cause juvenile-type gene expression. We did not observe *Fsg02* and *Fsg20* expression even in the absence of JHA (data not shown), suggesting that the cultured condition was insufficient or that another factor was necessary for the induction of these two Fsgs.

Discussion

Expression analysis of the genes isolated by cDNA subtraction revealed that *Fsg02* and *Fsg20* are strongly expressed in the presumptive green regions only during the

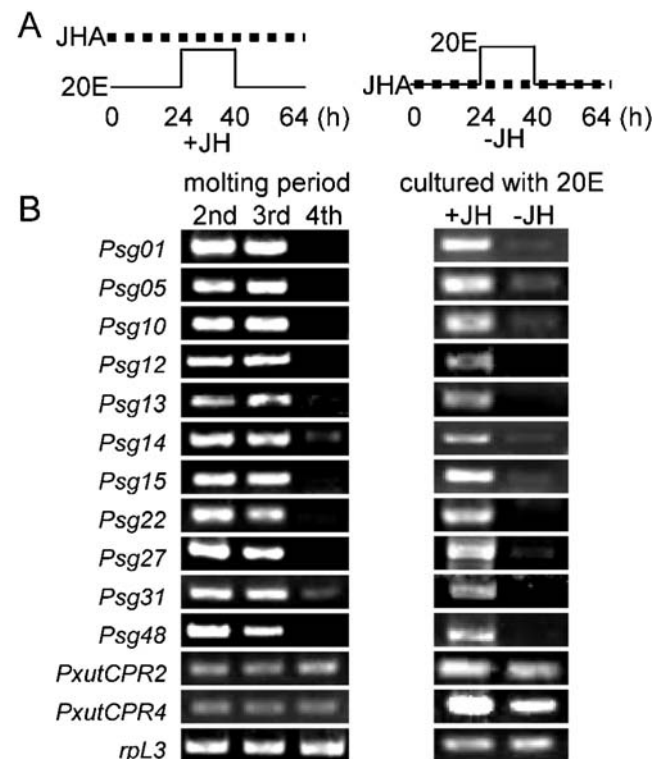


Fig. 7 a Time course of hormonal treatment in epidermis culture. See “Materials and methods” for detail. b Expression analysis of cuticular protein genes by RT-PCR. Larval epidermis during the second, third, and fourth molt (left) and epidermis cultured with JHA or without JHA (right) were used as the template. Culture condition is shown in a. The gene for the ribosomal protein L3 (*rpl3*) that is expressed constitutively in the cell was used as the control

molt to the final instar (Fig. 1b,c). Also, several cuticular protein genes are strongly expressed in the tubercle region during the early larval molting period but not during the final larval molt (Fig. 1b). The expression of these genes coincides with either the final instar-specific green coloration or the complex juvenile instar-specific exoskeletal structure (Figs. 3 and 4). Our results provide novel molecular markers and insights into the molecular mechanisms of the larval color pattern and the body shape.

Genes involved in green coloration

The green coloration of butterfly larvae is due in part to the presence of blue and yellow pigments. Blue pigments are thought to function in conjunction with yellow carotenoids as camouflage (Riley et al. 1984), and they are usually associated with proteins in pigment–protein complexes (Law and Wells 1989). However, both blue and yellow pigment-binding protein has been identified only in a few species. Similarity analysis of the amino acid sequence of Fsg02, together with its region-specific expression at the fourth molt, suggests that Fsg02 may be equivalent to a *P. xuthus* bilin-binding protein (or insecticyanin).

Our present results also show that there are some differences among Fsg02, *M. sexta* insecticyanin and *P. brassicae* bilin-binding protein. The first difference of Fsg02, *M. sexta* insecticyanin and *P. brassicae* bilin-binding protein is its expression pattern. Fsg02 and *M. sexta* insecticyanin is strongly expressed mainly in the epidermis (Li and Riddiford 1992, 1994), whereas *P. brassicae* bilin-binding protein is expressed not in the epidermis but in the fat body and gonads (Schmidt and Skerra 1994). Fsg02 is strongly expressed only at the final larval molt (Fig. 1b,c), whereas *M. sexta* insecticyanin is strongly expressed during all the early larval instars (Li and Riddiford 1992), and *P. brassicae* bilin-binding protein is expressed from the third ecdysis to adult (Schmidt and Skerra 1994). The second difference is its sequence similarity. Notably, these three genes do not form a single cluster in the phylogenetic analysis. The highest degree of identity was observed between Fsg02 and JF23 of *A. alcinous*, a putative binding protein for lipophilic substances related to butterfly oviposition (Tsuchihara et al. 2000). Because *A. alcinous* larva does not exhibit green coloration, JF23 is unlikely to be associated with the production of larval coloration. Although the sequence similarity among Fsg02, *M. sexta* insecticyanin, and *P. brassicae* bilin-binding protein is rather weak (Fig. 2), the three-dimensional structures of *M. sexta* insecticyanin are similar to that of *P. brassicae* bilin-binding protein (Holden et al. 1987; Huber et al. 1987; Schmidt and Skerra 1994), suggesting that these three proteins bind the same known ligand, biliverdin IX γ (Goodman et al. 1985; Huber et al. 1987). These findings suggest that these three proteins evolved independently, to

serve a common physiological role in each organism. The comparison of blue pigment-binding protein among more species may clarify the evolution of both pigment-binding protein and larval green coloration.

Expression analysis suggests that Fsg20 is also associated with green coloration (Fig. 3). Since the areas where Fsg20 was expressed but Fsg02 does not acquire a yellowish coloration (Fig. 3b, arrowheads), one possibility is that Fsg20 is associated with yellow pigment binding. Larvae of the closely related species *Papilio machaon* have many yellow spots in the black marking. Expression analysis of Fsg20-like genes among the green larvae of other species may clarify the relationship between Fsg20 and coloration.

The body markings and structure of the larva of *P. xuthus* change dramatically during the final molt; therefore, the epidermis is likely to display a unique gene expression pattern during the fourth molt. However, in a search for final instar-specific genes among the more than 1,000 clones isolated, only Fsg02 and Fsg20 are highly expressed at the final molt, suggesting that only a few genes can change body coloration dramatically.

Intricate regulation of the spatial expression pattern of each cuticular protein gene

The cuticle structures between the flexible (expansion) and rigid (nonexpansion) cuticle in *M. sexta* larva are different (Wolfgang and Riddiford, 1981). Several reports have suggested that RR-1 protein is associated with a flexible cuticle and RR-2 protein is associated with a rigid cuticle (Cox and Willis 1985; Lampe and Willis 1994, Gu and Willis 2003; Willis et al. 2005). However, precise spatial localization of each cuticular protein gene was largely unknown. We newly revealed the spatial expression pattern of many cuticular proteins and found that the spatial expression pattern of each cuticular protein did not entirely correlate with their motif (Fig. 6). The expression of 11 cuticular protein genes coincided with the tubercle structures, suggesting that these proteins are associated with complex structure. The combination of cuticular protein genes at the tubercle region is intricately regulated. Some cuticular proteins, such as *PxutCPR4* and *Psg32*, are less expressed at the tubercle regions (Fig. 5), suggesting that the expression pattern change of flexible cuticular protein is also necessary for the construction of tubercle structures. We found that the expression patterns of two RR-2 protein genes are quite different from other RR-2 proteins and that expression patterns are different even among the closely related CPFL protein genes (Fig. 5). Our findings indicate that it is difficult to predict the expression pattern of each cuticular protein gene by sequence similarity alone (Figs. 5 and 6). Although it is needed to study physical differences of the cuticle between the tubercle and nontubercle regions,

comparison of the spatial expression pattern is necessary for understanding each cuticular protein function. Juvenile instar larva of *P. xuthus* is a good example to analyze, because it is easy to recognize both rigid (tubercle) and flexible regions.

In *M. sexta*, JH alters the combinations of cuticular protein genes expressed. Although the JH regulation in the epidermis has been reported between final larval instar and pupa (Riddiford et al. 2003), no differential regulation of gene expression by hormones has been reported between the third and fourth molts. We found that the expression of tubercle-associated cuticular protein genes is regulated by JH both in vivo (Futahashi and Fujiwara 2008) and in vitro (Fig. 7). The expression of these cuticular protein genes is induced by 20E in the presence of JH during a larval molt. Then, during the final larval molt, the lower level of JH prevents this induction. Cuticular proteins are good representative proteins with which to study the effects of JH on gene expression. In this study, we focused on genes whose expression patterns were obviously different at the fourth molt (Fig. 1b). A more quantitative comparison of the remaining genes may clarify the differences in gene expression between juvenile and final instar larvae.

Acknowledgments This work was supported by grants from the Ministry of Education, Science, and Culture (MEXT) of Japan, the Program for Promotion of Basic Research Activities for Innovative Bioscience (PROBRAIN), and Research Fellowship of Japan Society for the Promotion of Science for Young Scientists. We are grateful to Mrs. Rumi Fujioka for technical support of RT-PCR analysis and to Dr. Akira Yamanaka for kindly providing many larvae of *P. xuthus*. We would like to thank Dr. Lynn M. Riddiford for many helpful comments on the manuscript.

References

- Cox DL, Willis JH (1985) The cuticular proteins of *Hyalophora cecropia* from different anatomical regions and metamorphic stages. *Insect Biochem* 15:349–362
- Futahashi R, Fujiwara H (2005) Melanin-synthesis enzymes coregulate stage-specific larval cuticular markings in the swallowtail butterfly, *Papilio xuthus*. *Dev Genes Evol* 215:519–529
- Futahashi R, Fujiwara H (2006) Expression of one isoform of GTP cyclohydrolase I coincides with the larval black markings of the swallowtail butterfly, *Papilio xuthus*. *Insect Biochem Mol Biol* 36:63–70
- Futahashi R, Fujiwara H (2007) Regulation of 20-hydroxyecdysone on the larval pigmentation and the expression of melanin synthesis enzymes and *yellow* gene of the swallowtail butterfly, *Papilio xuthus*. *Insect Biochem Mol Biol* 37:855–864
- Futahashi R, Fujiwara H (2008) Juvenile hormone regulates butterfly larval pattern switches. *Science* 319:1061
- Futahashi R, Okamoto S, Kawasaki H, Zhong YS, Iwanaga M, Mita K, Fujiwara H (2008) Genome-wide identification of cuticular protein genes in the silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* (in press). doi:10.1016/j.ibmb.2008.05.007
- Goodman WG, Adams B, Trost JT (1985) Purification and characterization of a biliverdin-associated protein from the hemolymph of *Manduca sexta*. *Biochemistry* 24:1168–1175
- Gu S, Willis JH (2003) Distribution of cuticular protein mRNAs in silk moth integument and imaginal discs. *Insect Biochem Mol Biol* 33:1177–1188
- Hiruma K, Riddiford LM (1990) Regulation of dopa decarboxylase gene expression in the larval epidermis of the tobacco hornworm by 20-hydroxyecdysone and juvenile hormone. *Dev Biol* 138:214–224
- Holden HM, Rypniewski WR, Law JH, Rayment I (1987) The molecular structure of insecticyanin from the tobacco hornworm *Manduca sexta* L. at 2.6 Å resolution. *EMBO J* 6:1565–1570
- Huber R, Schneider M, Mayr I, Muller R, Deutzmann R, Suter F, Zuber H, Falk H, Kayser H (1987) Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2.0 Å resolution. *J Mol Biol* 198:499–513
- Kim HJ, Yun CY, Han YS, Lee IH, Kang YJ, Jin BR, Seo SJ (2006) cDNA sequences of two biliproteins, BP1 and BP2, from the cabbage white butterfly, *Pieris rapae* and their tissue- and stage-specific accumulation. *Insect Biochem Mol Biol* 36:54–62
- Lampe DJ, Willis JH (1994) Characterization of a cDNA and gene encoding a cuticle protein from rigid cuticles of the giant silkworm, *Hyalophora cecropia*. *Insect Biochem Mol Biol* 24:419–435
- Law JH, Wells MA (1989) Insects as Biochemical Models. *J Biol Chem* 264:16335–16338
- Li W, Riddiford LM (1992) Two distinct genes encode two major isoelectric forms of insecticyanin in the tobacco hornworm, *Manduca sexta*. *Eur J Biochem* 205:491–499
- Li WC, Riddiford LM (1994) The two duplicated insecticyanin genes, ins-a and ins-b are differentially expressed in the tobacco hornworm, *Manduca sexta*. *Nucleic Acids Res* 22:2945–2950
- Rebers JE, Willis JH (2001) A conserved domain in arthropod cuticular proteins binds chitin. *Insect Biochem Mol Biol* 31:1083–1093
- Riddiford LM, Palli SR, Hiruma K, Li W, Green J, Hice RH, Wolfgang WJ, Webb BA (1990) Developmental expression, synthesis, and secretion of insecticyanin by the epidermis of the tobacco hornworm, *Manduca sexta*. *Arch Insect Biochem Physiol* 14:171–190
- Riddiford LM, Hiruma K, Zhou X, Nelson CA (2003) Insights into the molecular basis of the hormonal control of molting and metamorphosis from *Manduca sexta* and *Drosophila melanogaster*. *Insect Biochem Mol Biol*, 33:1327–1338
- Riley CT, Barbeau BK, Keim PS, Kezdy FJ, Heinrikson RL, Law JH (1984) The covalent protein structure of insecticyanin, a blue biliprotein from the hemolymph of the tobacco hornworm. *J Biol Chem* 259:3159–3165
- Sakai M, Wu C, Suzuki KJ (2001) Nucleotide and deduced amino acid sequences of a cDNA encoding a lipocalin protein in the central nervous system of *Bombyx mori*. *J Insect Biotech Sericol* 70:105–111
- Schmidt FS, Skerra A (1994) The bilin-binding protein of *Pieris brassicae*. cDNA sequence and regulation of expression reveal distinct features of this insect pigment protein. *Eur J Biochem* 219:855–863
- Scoble MJ (1992) *The Lepidoptera*. Oxford University Press, New York
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882

- Togawa T, Nakato H, Izumi S (2004) Analysis for the chitin recognition mechanism of cuticle proteins from the soft cuticle of the silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* 34:1059–1067
- Tsuchihara K, Ueno K, Yamanaka A, Isono K, Endo K, Nishida R, Yoshihara K, Tokunaga F (2000) A putative binding protein for lipophilic substances related to butterfly oviposition. *FEBS Lett* 478:299–303
- Willis JH, Iconomidou VA, Smith RF, Hamodraskis SJ (2005) Cuticular proteins. In: Gilbert LI, Iatrou K, Gill SS (eds) *Comprehensive molecular insect science*. vol. 4. Elsevier, Oxford, UK, pp 79–110
- Wolfgang WJ, Riddiford LM (1981) Cuticular morphogenesis during continuous growth of the final instar larva of a moth. *Tissue Cell* 13:757–772