博士論文

論文題目

Molecular genetic studies on the formation of diverse larval pigmentation patterns caused by the *p* multiple alleles in the silkworm.

（カイコ*p*複対立遺伝子座によって生じる多様な幼虫斑紋形成に関する分子遺伝学的研究）

氏　名　　依田　真一

Molecular genetic studies on the formation of diverse larval pigmentation patterns caused by the *p* multiple alleles in the silkworm.

Shinichi Yoda

Department of Integrated Biosciences

Graduate School of Frontier Sciences

University of Tokyo

December, 2014

Acknowledgements

I would like to express my deepest gratitude to my supervisor Professor Haruhiko Fujiwara, for his kind guidance, support and advice throughout the study. I also express my great appreciation to Drs. Tetsuya Kojima and Junichi Yamaguchi (University of Tokyo), for providing me with the scientific background and generous advice throughout the study. I am grateful to Dr. Yusuke Kondo, Ms. Kyoko Chagi and Ms. Kazumi Arita (University of Tokyo) for technical support of TALEN mRNA microinjection or injection of siRNA into eggs. Thanks are due to Drs. Kimiko Yamamoto and Kazuei Mita (National Institute of Agrobiological Sciences) for their assistance on the linkage analysis. Thanks are also due to Drs. Shoji Kawamura (University of Tokyo) and Ryo Futahashi (National Institute of Advanced Industrial Science and Technology) for their helpful advice for molecular phylogenetic analysis. I am grateful to Drs. Yutaka Banno (Kyusyu University) and Takaaki Daimon (National Institute of Agrobiological Sciences) for kindly providing silkworm strains and TALEN constructions, respectively. Finally, I would like to express my great gratitude to my family for their encouragements and various supports throughout the study.

December, 2014

Shinichi Yoda

Contents

Abstract **1**

Introduction **2**

Materials and Methods **10**

Silkworm strains 10
Positional cloning of the p^S locus 10
RT-PCR and RACE 13
Phylogenetic analysis 14
Comparative genomics using mVISTA plots 15
Quantitative RT-PCR 16
Plasmid construction for *piggyBac*-based transgene expression 16
Transgene expression 18
Conforcal microscopy analysis 18
siRNA for gene knockdown 19
Scanning electron microscopy analysis 20
TALEN-mediated mosaic analysis 20
Injection of siRNA into eggs 22

Results **23**

Linkage analyses of *B. mori p* alleles 23
Structural features and phylogenetic analysis of the *3234/5* gene 24
Pigmentation-associated expression of *apt-like* 27
Functional analyses of *apt-like* for larval pigmentation 28
Six predicted genes in p^S are not involved in pigmentation 31
Isoform expression and functional domains of *apt-like* 32
Embryonic and tissue expressions of *apt-like* 34
Regulatory gene network between *apt-like* and *Wnt1* signaling 35

Discussion **38**

Biological function of *apt* in *D. melanogaster* 38
Putative biological function of *apt-like* in Lepidoptera 39
apt-like induces black pigmentation in a cell-autonomous manner 40
apt-like regulates multiple melanin-related genes in a module manner 41
Pleiotropic effect of *apt-like* in the larval integment 43
A putative molecular mechanism for producing $+^{p}$ phenotype 45
The molecular basis for generating multiple *p* phenotypes 46

References **50**

Figures **63**

Tables **90**

Appendix 1: References in Table 1 94

Appendix 2: Accession codes 95

Abstract

Genetic polymorphisms underlie the convergent and divergent evolution of various phenotypes. Diverse color patterns on caterpillars, which are ecologically important, are good models for understanding the molecular backgrounds of phenotypic diversity. Here, I show that a single evolutionarily conserved gene *apontic-like* (*apt-like*) encoding for a putative transcription factor accounts for the silkworm *p* locus, which causes at least 15 different larval markings involved in branch-like markings and eye-spot formation. The expression of *apt-like* and melanin synthesis genes are up-regulated in association with pigmented areas of marking mutants *Striped* (p^S), *normal* ($+^p$), *Moricaud* (p^M) and *Black* (p^B) but not in the non-marking allele *plain* (*p*). Functional analyses by ectopic expression, RNAi and TALEN, demonstrate that *apt-like* causes melanin pigmentation in a cell-autonomous manner. These results suggest that variation in *p* alleles is caused by the differential expression of the gene, *apt-like*, which induces targeted elevation of gene expressions in the melanin synthesis pathway. Furthermore, gene regulatory network analyses show that phenotypic differences between $+^p$ and *p* may be generated by their differential responses to the *Wnt1* pathway. In the $+^p$ larva, *apt-like* may be downstream of the *Wnt1* pathway, but not in the *p* larva probably because the *cis*-regulatory elements of *apt-like* is mutated and does not respond to the *Wnt1* pathway.

Introduction

One of the fundamental issues in evolutionary biology concerns the origin of morphological novelty, such as which of the genes contribute to novel traits and what kinds of genetic changes underlie phenotypic variations. One way of solving this problem is to study polymorphic species, with multiple variants (or morphs) coexisting within a single population.

Many Lepidoptera (butterflies and moths) show several color polymorphisms and have caught attention of evolutionary biologists for more than a century. In particular, the highly diverse color patterns on caterpillars are of great evolutionary interest because of their association with natural selection (Greene, 1989; Berenbaum, 1995; Futahashi & Fujiwara, 2008). Similar to the adult wings of butterflies and moths, their larvae, which are often preyed by other animals, also show diverse camouflage and warning color patterns (Janzen *et al*., 2010; Shirataki *et al*., 2010). For example, the swallowtail butterfly *Papilio xuthus* changes its larval body pattern dramatically during the molt to the final instar (fourth molt). In the early stadia of *P. xuthus*, the larvae have black and white markings, while the last (fifth) instar larvae have a green camouflage with a pair of large false eyespots on the metathorax (Futahashi & Fujiwara, 2008). The larvae of monarch butterfly *Danaus plexippus* are unpalatable to predators because of its ability to accumulate host-plant cardenolides and to display aposematic coloration (vivid yellow, black, and white stripes) (Berenbaum, 1995).

Both larvae and various adult moths are usually cryptically colored and frequently polymorphic (Hebert *et al*., 2004; Bond, 2007). The peppered moth *Biston betularia*, a classical genetic model of industrial melanism in 19th-century Britain, is one of the most widely recognized examples of a rapid adaptive response to environmental changes (Cook, 2003). The emerald moth *Nemoria arizonaria* is bivoltine, with distinct broods of caterpillars hatching in the spring and summer. Caterpillars of the spring brood develop into mimics of oak catkins, while caterpillars of the summer brood develop into mimics of oak twigs (Green, 1989).

Some lepidopteran species exhibit highly diverse color patterns not only among closely related species but also among the same species. The mocker swallowtail *Papilio dardanus*, an iconic example of Batesian mimicry (an undefended mimic evolves to look like a toxic model), can develop distinctly different patterns because of at least 10 gene alleles at a single locus, *H* (Nijhout, 2003; Clark *et al*., 2008). *P. dardanus* shows a female-limited polymorphic mimicry, whereas the males all look identical across this species (Nijhout, 2003). The majority of these morphs are Batesian mimics of several unpalatable species of Danaidae and Acraeidae butterflies (Poulton, 1924; Ford, 1936; Bernardi *et al*., 1985). The remaining morphs, although highly distinctive, do not resemble any other known species of butterfly, and they are not considered to be mimics. Laboratory crosses have shown that most of the phenotypic variation is determined by the *H* locus in which various alleles exhibit a dominance hierarchy

such that most of them are inherited without producing intermediate phenotypes (Ford, 1964; Nijhout, 2003). The *H* locus has been physically mapped through linkage mapping and single nucleotide polymorphism (SNP)-association analysis to a narrow region, including the transcription factor genes *engrailed* and *invected* (Clark *et al*., 2008; Timmermans *et al*., 2014). However, the gene responsible for divergence of wing phenotypes in this butterfly has not yet been functionally identified. *Heliconius* species, which are well known examples of Müllerian mimicry (toxic species gain mutual protection by sharing warning signals) also display phenotypic diversity of their wing color patterns (Joron *et al*., 2011; The *Heliconius* Genome Consortium, 2012). Among 43 species of *Heliconius* butterflies, there are hundreds of races with different color patterns (The *Heliconius* Genome Consortium, 2012). Geographic mosaics of multiple color-pattern races, such as those in *Heliconius melpomene*, converge to similar phenotypes in other sympatric *Heliconius* species, probably because their aposematic signals are locally monomorphic but differ between localities. *Heliconius* are unpalatable to vertebrate predators, and their warning color patterns enable them to share the cost of educating predators (Turner, 1981). Because of its dual role in mimicry and mate selection, divergence of wing patterns is also associated with speciation and adaptive radiation (Turner, 1981; Jiggins *et al*., 2001; The *Heliconius* Genome Consortium, 2012). The above examples demonstrate that color polymorphisms can contribute to both convergent evolution (genetically diverse organisms evolve to have similar

morphs) and divergent evolution (a single ancestor gives rise to a diverse group of organisms).

Recently, detailed analyses of their developmental processes have revealed how polymorphisms could have been acquired during their evolution. Some of the most convincing data for the genetic and evolutionary control of color patterns were obtained from different species or spontaneous mutants within several species of butterflies and fruit flies (Carroll *et al*., 1994; Brakefield *et al*., 1996; Rebeiz *et al*., 2009; Wittkopp *et al*., 2009; Nijhout, 2010; Werner *et al*., 2010; Reed *et al*., 2011; Kronforst *et al*., 2012; Martin *et al*., 2012; Saenko *et al*., 2012; Rogers *et al*., 2013). In those reports, various genetic changes that contributed to phenotypic differences within or between species were identified; however, the molecular mechanism by which a single locus (gene) can produce multiple phenotypic traits are largely unknown (Wittkopp *et al*., 2009). Because genetic polymorphisms underlie the convergent and divergent evolution of color pattern diversity (Joron *et al*., 2006), identifying the essential gene that controls this phenotypic diversity is of great interest.

Among Lepidoptera, the silkworm *Bombyx mori* is the most suitable for identifying the gene responsible for color patterns because there are numerous pigmentation pattern mutants available, particularly for its larval stages (Banno *et al*., 2005), and its genome sequence and a high-density linkage map are available (International Silkworm Genome Consortium, 2008; Yamamoto *et al*., 2008). Molecular genetic studies have identified the genes responsible for several color

pattern mutants in the silkworm (Liu *et al*., 2010; Zhan *et al*., 2010; Wei *et al*., 2013). Turner (1984) introduced 4 different mutations of this silkworm, *Moricaud* (p^M), *Zebra* (*Ze*), *Multilunar* (*L*) and *quail* (*q*), which are good models for investigating caterpillar mimicry with various patterns. Previous work found that the melanin synthesis genes *yellow* and *ebony* were responsible for the silkworm color mutants *chocolate* (*ch*) and *sooty* (*so*) loci, respectively (Futahashi *et al*., 2008). In addition, recent study found that *L* resulted from the ectopic expression of *Wnt1* in the epidermis (Yamaguchi *et al*., 2013). These studies demonstrated that silkworm mutants are useful resources for understanding color pattern formation.

Larval color patterns primarily depend on the nature and distribution of pigments in the epidermal cells and epicuticle. The larval body color in the silkworm is mainly determined by the abundance of melanin in the cuticle (Ohashi *et al*., 1983). In lepidopteran larvae, melanin granules contribute to thermoregulation, UV resistance, cuticle strength, desiccation tolerance, and body color (crypsis, aposematism, mimicry) (True, 2003; Hu *et al*., 2013; Shamim *et al*., 2014). Melanin is produced by the epidermal cells via the melanin synthesis pathway. Genes responsible for melanin synthesis have been widely studied in many insects, particularly in *Drosophila melanogaster* (True, 2003; Wittkopp & Beldade, 2009). Biosynthesis of melanin begins with the conversion of tyrosine to DOPA and then to dopamine by tyrosine hydroxylase (encoded by *TH* gene) and dopa decarboxylase (*DDC* gene), respectively (Wright, 1987).

Dopamine is the primary precursor of insect melanin. Dopamine is incorporated into cuticular premelanin granules that contain two secreted proteins encoded by *yellow* and *Laccase 2*. Finally, the pigment precursor molecule is converted into melanin (Shamin *et al*., 2014). Following a different route, *N*-*β*-alanyldopamine (NBAD) synthetase (the product of the *ebony* gene) can use dopamine to produce NBAD, which is the precursor of yellow sclerotin (Wittkopp *et al*., 2002). NBAD can also be converted back to dopamine by NBAD hydrolase (encoded by the gene *tan*) (Wittkopp *et al*., 2003; True *et al*., 2005). Dopamine-*N*-acetyl transferases (NATs) convert dopamine to *N*-acetyl-dopamine, a precursor of colorless sclerotin (Brodbeck *et al*., 1998). In the silkworm and several swallowtail butterflies, the cuticular markings in the larvae are also associated with the genes of the melanin synthesis pathway, such as *TH*, *DDC*, *yellow*, *Laccase2*, *ebony*, and *tan*. These genes are strictly regulated in a spatiotemporal manner (Futahashi & Fujiwara, 2005, 2007; Okamoto, 2005; Futahashi *et al*., 2008, 2010; Shirataki *et al*., 2010; Yu *et al*., 2011).

Silkworms were considered to be handed down from China to Japan in the 2nd century (Tazima, 1993). The normal silkworm (denoted by $+^{p}$) pattern comprises three types of spots: (i) eye spots on the second thoracic segment, (ii) crescents on the second abdominal segment and (iii) star spots on the fifth abdominal segment (Fig. 1a, b; $+^{p}$) (Tazima, 1978). These markings are primarily controlled by one allele ($+^{p}$) of the *p* locus, which was shown to be located at 3.0 cM of Linkage Group 2 by classical chromosomal mapping (Banno *et al*., 2005).

This *p* locus comprises at least 15 alleles (Table 1), which are suitable for studying the molecular background underlying phenotypic polymorphisms.

The genetics of larval color pattern determination in the *p* locus was studied from the 1900s (Table 1). Among *p* alleles, *normal pattern* ($+^p$) and *plain* (*p*) are the most widespread markings among silkworm varieties (Fig. 1a). In *p* larvae, no markings are observed compared with $+^p$ larvae (Fig. 1a, b; *p*). The larval markings of the wild silk moth *B. mandarina* resemble those of *Moricaud* (p^M) of *B. mori*; in addition, p^M markings of *B. mori* consist of dark grayish brown lines and dots (Tazima, 1978). Figure 1c shows the branch mimicking of the wild silkworm with noticeable eye spot markings on the second thoracic segment. These eye spots are also observed on $+^p$ larvae (Fig. 1b). For *Striped* (p^S) larvae, the body surface is solid black, except for the posterior margin of each segment. Several other alleles are known to be located at this locus: *Black* (p^B; Whole body surface is black color, especially ventral epidermis is uniformly black), *Ventral striped* (p^G; Stripes faintly light colored but ventral stripes alone deep-black), *Whitish striped* (p^{Sw}; Eye spots and ventral stripe are clearly detected in heterozygote p^{Sw}/p, but back of abdominal segments is light-black, thorax is nearly white. Crescents and star spots are yellowish. Homozygote is similar to p^S), *Sable* (p^{Sa}; Sable pigmented striped marking, eye spot fairly deep-black, distinctly outlined. Crescents and star spots are extremely light colored. Light dark ventral stripe is present) and *Sable-2* ($p^{Sa\text{-}2}$; Like p^{Sa}. Black pigments are much entirely in epidermis, crescents are dark). The dominance relationship

between the major alleles is: $p^B > p^S > p^M > +^p > p$ (Tazima, 1964).

From the historical point of view, the *p* locus studies have made a significant contribution to genetics. The best known example is the world's first "re-discovery" of Mendelian inheritance in invertebrates (Toyama, 1906). In the paper titled "*Studies on the hybridology of insects. I. On some silkworm crosses, with special reference to Mendel's law of heredity*" published in 1906, Toyama demonstrated that larval markings of p^S, cocoon colors, and other characteristics were inherited in accordance with Mendel's law. Another example is the finding that the crossing over occurs only in the male (Tanaka, 1913; Sturtevant, 1915). Tanaka has carried out back-cross tests in both sexes using linkage relationship between p^S locus and *Yellow* (*Y*) locus (cocoon colors) (Tanaka, 1913), or p^M locus and *Y* locus (Tanaka, 1914). The sex-limited strains, in which *p* locus ($+^p$, p^M, p^B or p^{Sa}) is translocated to the female W chromosome by X-ray irradiation, are useful in sericulture as they make it easy to separate the sexes (Tazima, 1941, 1943a, 1944, 1955; Sasaki, 1955).

In this study, I identified *apontic-like* (*apt-like*) as the gene responsible for these *p* alleles by positional cloning and functional analyses. I discussed its function and possible roles in pigmentation pattern diversity.

Materials and Methods

Silkworm strains.

The p^S mutant strains c05 and f34, $+^p$ strains g01and p20, p^M strains u10 and T02, p^B strain A10, and *L* strain g01 were from the Institute of Genetic Resources of Kyushu University. The $+^p$ strain p50T and *p* strains N4 and C108 were gifts of Dr. T. Shimada (University of Tokyo). The mottled striped strain p^{Sm}872, which had been induced by X-ray irradiation, was a gift of Dr. O. Ninaki (Tokyo University of Agriculture and Technology). Silkworms were reared on mulberry leaves or artificial diets at 25 °C under long-day conditions (16 h light: 8 h dark). Staging of fourth instar larvae was based on the spiracle index, which reflects the characteristic sequence of new spiracle formation (Kiguchi & Agui, 1981).

Positional cloning of the p^S locus.

For positional cloning of the p^S locus, 12 F1 heterozygous males were obtained from a single-pair cross between a *p* strain N4 male and a p^S strain c05 female and each was backcrossed with an N4 female. A total of 1,755 Backcrossed F1 (BC1: p^S phenotype, 934 individuals; *p* phenotype, 821 individuals) fifth instar larvae were used for analysis. Genomic DNA was extracted from original parental moths, F1 moths and each BC1 of fifth instar larvae using DNAzol solution (Invitrogen).

The details of extraction of genomic DNA and genotyping are as follows.

(i) *Extraction of genomic DNA from the original parents and F1 moths.*

First, the flight muscles were dissected from the moths and homogenized in 1.5-ml Eppendorf tubes containing 500 µl of DNA extraction buffer [150 mM NaCl, 10 mM Tris-HCl (pH 8.0), 10 mM ethylenediaminetetraacetic acid (EDTA), 0.1% sodium dodecyl sulfate (SDS), and 100 µg ml^{-1} of Proteinase K (MERCK)]. The samples were incubated overnight at 50 °C. After centrifugation (at 9,100 × *g* for 5 min at 4 °C), the supernatant was transferred to a new 1.5-ml tube; equivalent volume of phenol, Tris-saturated (pH 8.5), was added. The tubes were gently shaken overnight using a rotator. After centrifugation (at 4,550 × *g* for 15 min at 4 °C), the supernatant was transferred to a new 1.5-ml tube. An equivalent volume of phenol–chloroform (neutral phenol:chloroform, 1:1) was added, and then the tubes were gently shaken for 20 min. After centrifugation (at 21,880 × *g* for 15 min at 4 °C), the supernatant was transferred to a new 1.5-ml tube, and it was mixed with an equivalent volume of isopropyl alcohol. The tubes were left for 10 min at room temperature and then centrifuged (at 7,280 × *g* for 5 min at room temperature). The supernatant was decanted and each sample was washed with 300 µl of 70% ethanol. Finally, the residue was left to dry and then resuspended in 500 µl of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

(ii) *Extraction of genomic DNA from BC1 larvae.*

Genomic DNA from each BC1 (fifth instar) larvae was extracted. A hole in the first abdominal leg was made with a sterile needle, and a couple of drops of hemolymph (over 50 µl) were collected in a 1.5-ml Eppendorf tube. After

centrifugation (at 17,800 × *g* for 5 s at room temperature), some supernatant was removed to adjust the volume to 50 µl, and 500 µl of DNAzol (Sigma) was added to each sample; the samples were subsequenly mixed. After centrifugation (at 10,000 × *g* for 10 min at 4 °C), the supernatant was transferred to a new 1.5-ml tube. Subsequent analyses were performed using an automated DNA isolation system PI-1000 (KURABO). To concentrate genomic DNA, 250 µl of ethanol was added to each sample, and tubes were centrifuged at 5,000 × *g* for 5 min at 4 °C. After supernatant removal, the samples were mixed with 75% ethanol and then centrifuged at 5,000 × *g* for 5 min at 4 °C (this procedure was repeated twice). Following this, 75% ethanol was completely removed using a pipette. To dissolve the DNA precipitate, 50 µl of 8 mM NaOH was added to each sample followed by 5 µl of 0.1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, pH 8.0). The samples were centrifuged at 21,880 × *g* for 10 min at 4 °C, and the supernatant was transferred to new 1.5-ml tubes. The DNA samples were stored at 4 °C.

Genomic DNA of fourth instar larvae was extracted from their epidermal tissues. Each epidermis was placed in a 1.5-ml tube and 600 µl of DNAzol was added. After homogenization, the tubes were centrifuged at 17,800 × *g* for 2 min at room temperature. The supernatant (500 µl) was transferred to new 1.5-ml tubes. Subsequent analyses followed the procedure similar to that of the fifth instar larvae, as described above.

(iii) *Genotyping.*

For genetic analysis, 17 SNP-PCR markers were constructed. The primers used for linkage analysis are listed in Table 2. The homozygous (A) or heterozygous (H) state of each BC1 individual was determined (Table 3).

Genotyping was performed by direct PCR sequencing. The PCR reactions were performed using PrimeSTAR GXL DNA Polymerase (Takara, Japan) under the following conditions: 95 °C for 5 min; 40 cycles at 95 °C for 15 s, 58 °C for 15 s, and 68 °C for 1 min; and a final extension at 68 °C for 10 min. The purified genomic DNAs described above (i and ii) were used as PCR templates. Using gel electrophoresis, I confirmed that each PCR fragment migrated as a single band. Each PCR-amplified DNA was cleaned using 1.25 units of Exonuclease I (Takara, Japan) and 0.125 units of Shrimp Alkaline Phosphatase (Roche) at 37 °C for 30 min (enzymatic reaction) and at 80 °C for 20 min (heat inactivation). One-third of each reaction solution was used as a sequencing template. PCR analyses, using the primers chr2-23F and chr2-23R, and 02-027F and 02-027R (Table 2), were conducted at the National Institute of Agrobiological Sciences (NIAS). After cycle-sequencing reaction using 384-well plates, each PCR-amplified DNA was automatically purified using CleanSEQ (Beckman Coulter) and then sequenced in a 3730xl DNA Sequencer (Applied Biosystems). Sequencing data were analyzed using a trial version of Sequencher 4.8 (Gene Codes Corp.).

RT-PCR (Reverse Transcription-PCR) and RACE (Rapid Amplification of cDNA Ends).

Total epidermal RNA was isolated from the silkworm p^S mutant strains c05 and f34, $+^p$ strains p50T, g01 and p20, *p* strain N4, p^M mutant strain T02, and p^B mutant strain A10 using TRI reagent (Sigma), treated with 0.2 U of DNase I (Takara) at 37 °C for 15 min, purified by phenol–chloroform extraction and then reverse transcribed with a random primer (N6) using a First-Strand cDNA Synthesis Kit (GE Healthcare), according to manufacturers' protocol. 5′ and 3′ RACE was performed using SMART RACE cDNA Amplification Kit (Clontech), according to the manufacturer's protocol. PCR products were subcloned into a pGEM-T Easy Vector System (Promega). Nucleotide sequences were determined using an ABI3130xl genetic analyzer (Applied Biosystems). Primers used for PCR are listed in Table 4. The gene for *ribosomal protein L3* (*rpL3*) of *B. mori* was used as an internal control to normalize for equal sample loading.

Phylogenetic analysis.

To determine whether *3234/5* (*apt-like*) orthologs existed in species other than *B. mori*, I performed phylogenetic analysis using genes included in the Blastp hits (E value < 10^{-4}) on the NCBI server (http://www.ncbi.nlm.nih.gov/BLAST/), Manduca base (http://www.agripestbase.org/manduca/), Heliconius Genome Project (http://www.butterflygenome.org/), KONAGAbase (http://dbm.dna.affrc.go.jp/px/), BeetleBase (http://beetlebase.org/), AphidBase (http://www.aphidbase.com/) and RNA-seq data for *P. xuthus*. Apt-like and Apt were used as query for each Blastp search. Codon alignment was generated from

a multiple sequence alignment of predicted amino acid sequences by MUSCLE (Edgar, 2004) and PAL2NAL (Suyama *et al*., 2006) was used to construct nucleotide alignments. The phylogenetic tree was constructed using the maximum likelihood method with the MEGA6 program (GTR + G + I model) (Tamura *et al*., 2013). The confidence levels for various phylogenetic lineages were assessed by bootstrap analysis (1,000 replicates).

The following sequences were used to create the diagram (Fig. 7): Diptera *apt*; *D. melanogaster* (AAB82746.1), *A. gambiae* (XP_308214.4), *A. aegypti* (XP_001652318.1), *C. quinquefasciatus* (XP_001862892.1). Hymenoptera *apt*; *A. mellifera* (XP_006559097.1), *M. rotundata* (XP_003708630.1), *C. floridanus* (EFN69861.1), *H. saltator* (EFN76284.1), *A. echinatior* (EGI67737.1). Lepidoptera *apt-like*; *D. plexippus* (EHJ68239.1), *M. sexta* (Msex004196-PA), *H. melpomene* (HMEL009522-PA), *P. xylostella* (PXGS_V2_024627).

Comparative genomics using mVISTA plots.

The around $+^{p}$: 33-kb genomic regions of *B. mori* (chr 2: 1319300-1325400), *M. sexta* (scaffold00032), *H. melpomen*e (HE670913_scf7180001249883) and *D. plexippus* (DPSCF300283) were extracted with KAIKOBASE (http://sgp.dna.affrc.go.jp/KAIKObase/), Manduca Base (http://agripestbase.org/manduca), Heliconius Genome Database (http://www.butterflygenome.org/) and MonarchBase (http://monarchbase.umassmed.edu/), respectively. Genomic regions were

aligned using the program LAGAN (Brudno *et al*., 2003) and alignments were visualised using the program mVISTA (http://genome.lbl.gov/vista/mvista/submit.shtml).

Quantitative RT-PCR.

To determine *apt-like* mRNA expression, I used a Real-Time PCR System with power SYBR green PCR master mix using a StepOne System (Applied Biosystems). Specific primer sets used for target genes are listed in Table 4. The relative expression of a target gene was determined from a standard curve. Expression levels for each sample were determined with no fewer than three biological replicates. I set the normalized *apt-like* expression level of the far left sample as 1.0 for comparisons with other columns. Statistical comparisons were made by Student's *t*-test or paired Student's *t*-test.

Plasmid construction for *piggyBac*-based transgene expression.

I constructed recombinant plasmids from a *piggyBac* target vector (*pPIGA3GFP::A3DsRed2*) containing an EGFP and a DsRed2 marker, each driven by a *Bmactin A3* promoter, as described below.

(i) *Insert preparation.*

First, the full-length open reading frame (ORF) DNA of *apt-like* was amplified using the PCR primers Bm3234_Exon2-F2 and Bm3234_Exon7-R1 (listed in Table 4) and iProof high-fidelity DNA polymerase (Bio-Rad, Hercules, CA). The

conditions were: 98 °C for 30 s; 32 cycles at 98 °C for 10 s, 60 °C for 30 s, and 72 °C for 30 s; and a final extension at 72 °C for 5 min. cDNAs from the p^S strain c05, *p* strain N4, p^M strain u10, and p^B strain A10 were used as the PCR templates. The enzyme protein was removed from PCR reaction solution using the phenol/chloroform/isoamyl alcohol (PCI) mixture. The PCR product was purified by ethanol precipitation. The pellet was rinsed with 70% ethanol and resuspended in 7 μl of ultrapure water (Merck Millipore). The dATP addition to the 3′ end of the PCR product was carried out using Ex-Taq (Takara) at 70 °C for 30 min. Each PCR product was subcloned into the pGEM-T Easy vector (Promega). Then, colony PCR was conducted using the PCR primers piggyBac_Bm3234-F1 and piggyBac_Bm3234-R1 (Table 4) and iProof high-fidelity polymerase (Bio-Rad, Hercules, CA) under the following conditions: 98 °C for 30 s; 22 cycles at 98 °C for 10 s, 60 °C for 30 s, and 72 °C for 30 s; and a final extension at 72 °C for 5 min. The pGEM-T Easy Vector (Promega) from each PCR reaction solution was removed using *DpnI* (Toyobo) at 37 °C for 1 h. The PCR product was purified using a GenElute PCR Clean-Up Kit (Sigma).

(ii) *Donor plasmid preparation.*

DNA vector was based on pPIGA3DsRed2 (Ando & Fujiwara, 2013). The vector was linearized with *BamHI* (Takara) and *NotI* (Toyobo) at 37 °C for 2 h. The enzyme protein was removed using the PCI mixture. After gel electrophoresis, the target fragment was purified using a QIAquick Gel Extraction Kit (Qiagen).

(iii) *Ligation reaction.*

In-Fusion enzyme (Clontech) was used to ligate the two DNA fragments, as mentioned above (i and ii). The ligation product was transformed into *Escherichia coli* strain DH5α. Recombinant plasmids were sequenced using an ABI3130xl Genetic Analyzer (Applied Biosystems) to verify nucleotide sequences.

Transgene expression.

piggyBac transposase-based electroporation was performed using a previously described method (Ando & Fujiwara, 2013). The target plasmid, which contained both the gene of interest and *EGFP* marker driven by *BmActin A3* promoter, and the helper plasmid *pHA3PIG*, which included *piggyBac* transposase driven by *BmActin* A3 promoter, were mixed at a final concentration of 1 μg μl^{-1} each, and 0.5–0.75 μl of this mixture was injected into the haemolymph of second instar larvae of *p* strain N4 using a glass needle connected to a microinjector (FemtoJet, Eppendorf). Shortly after this injection, electrical stimulation was applied (5 square pulses of 20–25 V, 280 ms width) via two separated small drops of phosphate-buffered saline (PBS) on the larva, which were used as electrodes to avoid serious damage.

Confocal microscopy analysis.

Epidermal tissues expressing EGFP were dissected from *apt-like* transgenic fifth instar larvae of *p* allele. The dissected tissues were rinsed at once with PBS and

fixed with 4% paraformaldehyde for 2 h at room temperature. Each sample was washed at three times with PBS and incubated in 10 μg ml^{-1} of 4', 6-Diamidino-2-phenylindole, dihydrochloride (DAPI, Wako) for 5 min and washed with PBS. Then, the samples were mounted on slides with Gel Mount. Images were collected using a LASER confocal microscope (Fluo View FV1000, Olympus, Japan).

siRNA for gene knockdown.

siRNAs for *apt-like* were designed that targeted 5′-AACTATAAGAAGAAATGCAAAAT-3′ and 5′-GAGCATACATCCCTTAGACGTCA-3′. siRNAs for *BGIBMGA003237* and *BGIBMGA003260* were designed to target 5′-TACAATAAAGCAGCCGAAAATGT-3′ or 5′-TTGCATACTCATCTAAACAGACA-3′, respectively. siRNAs for *Distal-less* (*Dll*) was designed to target 5′-TAGATCATTAGGTTATCCTTTCC-3′ (FASMAC Co., Japan). Universal negative control siRNA (Nippongene, Tokyo, Japan) was used as a negative control. siRNA (250 μM; 0.5–0.75 μl) was injected from the lateral side of the abdominal fifth segment into haemolymph at the fourth instar larval stage just after molting. Immediately after injection, PBS droplets were placed nearby and an appropriate voltage was applied (20 V for the p^S, p^M and p^B alleles; 18 V for the $+^p$ allele). Phenotypic effects were observed for fifth instar larvae. To quantify the mRNA levels of target genes by quantitative

RT-PCR, the “left-half” (into which siRNAs were introduced) and the “right-half” of the epidermis were separately dissected from the abdominal third segment (p^S, p^M and p^B alleles) or the abdominal second segment ($+^p$ allele) of fourth instar larvae (C2 period).

Scanning electron microscopy analysis.

Cuticular nanostructure of p^S and *p* strains was observed by scanning electron microscopy using a Hitachi TM-1000 scanning electron microanalyzer (Hitachi High-Technologies Corp., Tokyo, Japan).

TALEN-mediated mosaic analysis.

A TALEN was designed using TALEN Targeter 2.0 (https://TALE-nt.cac.cornell.edu/). TALEN expression vectors (*pBlue-TAL*) (Takasu *et al*., 2013) were supplied by Dr. T. Daimon (NIAS). Both *apt-like*-TALEN Left and *apt-like*-TALEN Right did not have homologous sequences in the genome. The designed TALEN sequences were confirmed by sequencing analysis.

(i) *In vitro synthesis of TALEN mRNAs.*

TALEN expression vectors were linearized using *XbaI* at 37 °C for 3 h (enzymatic reaction) and 65 °C for 30 min (heat inactivation). The linearized vectors were used as templates in mRNA synthesis using an mMESSAGE mMACHINE T7 Ultra Kit (Ambion). The TALEN mRNAs were precipitated with lithium chloride (LiCl) and washed three times with 70% ethanol. The

TALEN mRNAs were dissolved in 50 µl of 0.5 mM phosphate buffer (pH 7.0).

(ii) *Microinjection of the TALEN mRNAs.*

The capped and poly (A)-tailed mRNAs encoding TALEN pairs were dissolved in 0.5 mM phosphate buffer (pH 7.0) to a final concentration of 200 ng μl^{-1} for each mRNA. The procedure was performed with the technical assistance of Dr. Y. Kondo (University of Tokyo). An aliquot of 3–5 nl of RNA solution was injected through the chorion into silkworm embryos at the syncytial preblastoderm stage (4–8 h after oviposition) (Takasu *et al*., 2010). The small hole in the chorion caused by the microinjection was sealed with cyanoacrylic glue (Aron Alpha, Konishi Co, Osaka), and the embryos were incubated at 25 °C in a humidified atmosphere. TALEN mRNAs were microinjected into 64 embryos of p^S strain. Of the microinjected embryos, 11 larvae hatched and two larvae had a mosaic appearance having white patches dispersed in a black background. To determine whether mutations occurred at the specified locus and to determine the types of mutations that were possibly introduced by TALEN, I extracted genomic DNA from each mosaic mutant. The regions surrounding the target sites were amplified by PCR (using the primers Bm3234_Exon6-F1 and Bm3234_Exon6-R1, Table 4), subcloned into a bacterial plasmid, and then checked for target sequence modifications. Sixteen clones were examined for each of the two mosaic mutants. The frequencies of detected mutant clones were 62.5% (10 per 16 clones) and 18.8% (3 per 16 clones).

Injection of siRNA into eggs.

The preparation of eggs and injection of siRNA were performed according to the method described by Yamaguchi *et al*. (2011). Eggs of the p^M and *p* strain (both were non-diapause strains) were collected within 5 h of oviposition. The newly laid eggs were washed using tap water and left in water for 2–3 min. The floating eggs were transferred to glass slides and aligned in the same direction under a dissection microscope. The eggs were left to dry for about 1 h at room temperature and bonded using an adhesive. The injection system comprised an injector (Eppendorf, FemtoJet), manipulator (SURUGA SEIKI, M401), stepping motor controller (SURUGA SEIKI, M331), and a dissection microscope (Nikon, SMZ1500). The two siRNAs that targeted *apt-like*, as described above, were mixed together and used at a final concentration of 50 μM. First, a small hole was made on the ventral side of the egg using a tungsten needle, then a glass capillary was inserted, and 1–5 nl of the siRNA solution was injected. The eggs were incubated in a Petri dish in a moist plastic box at 25 °C until hatching. The procedure was performed with the technical assistance of Ms. K. Chagi and Ms. K. Arita (University of Tokyo).

Results

Linkage analyses of *B. mori p* alleles.

Previously reported linkage analyses mapped 15 *p* alleles in *B. mori* to a single chromosomal locus (Linkage Group 2, 3.0 cM) (Banno *et al*., 2005) (Table 1). To identify the genomic region responsible for these *p* alleles, I carried out a positional cloning in p^S, collaborated with National Institute of Agrobiological Sciences (NIAS). I performed low-resolution mapping for 45 back-crossed F1 (BC1) progeny between *p* and p^S individuals using standard single nucleotide polymorphism (SNP) markers (A–F, a and k). I found that the 1.1 Mb genomic region between markers "a" and "k" of Linkage Group2 was responsible for the p^S phenotype (Fig. 2a). Next, I performed screening using the remaining 1,710 BC1 progeny in which the recombination had occurred between the markers "a" and "k." I collected 107 individuals from this progeny and subjected them to a more detailed mapping. Using the newly designed SNP markers (Table 2; b–j), I genotyped these 107 individuals and found that the 233-kb genomic region between the markers "e" and "h" was responsible for the p^S phenotype (Fig. 2a).

This region included eight predicted genes (*BGIBMGA003234*, *BGIBMGA003235*, *BGIBMGA003236*, *BGIBMGA003237*, *BGIBMGA003238*, *BGIBMGA003258*, *BGIBMGA003259* and *BGIBMGA003260*) (Fig. 2a). Among these genes, cDNA for *BGIBMGA003234* was found in four different expressed sequence tag (EST) databases for larval tissues (corpora allata, prothoracic gland,

wing disk and maxilla) with the $+^{p}$ phenotype (CYBERGATE), which suggested that this gene was actively transcribed. In addition, the cDNA for *BGIBMGA003234* included sequences for *BGIBMGA003235*. This indicated that *BGIBMGA003234* and *BGIBMGA003235* were actually the same gene. Thus, I subsequently designated this gene as *3234/5*.

Using a similar strategy, Okamoto (2005) mapped the genomic region responsible for the *p* alleles using BC1 progeny of *p* and $+^{p}$ individuals, collaborated with NIAS. They performed fine-mapping for 2,893 BC1 progeny and found a 33-kb genomic region responsible for the $+^{p}$ phenotype, which included only the first exon of the gene *3234/5* (Fig. 2b). These results indicate that the *3234/5* gene is the most probable candidate for causing the various phenotypes of the *p* alleles because *3234/5* gene was included not only $+^{p}$: 33-kb but also p^{S}: 233-kb.

Structural features and phylogenetic analysis of the *3234/5* gene.

To determine the structural features of the *3234/5* gene, I used rapid amplification of cDNA ends (RACE) and reverse transcription-polymerase chain reaction (RT-PCR) for the full-length cDNA structure of the *3234/5* gene from individuals with three different alleles: *p*, $+^{p}$ and p^{S}. All these genes comprised eight exons, which included a 951-bp open reading frame (ORF), encoding 317 amino acids (aa) (Fig. 3). Of exons of *3234/5* gene, only exon 1b encoding the 5′-untranslated region was included in $+^{p}$: 33-kb. When comparing ORFs among the

3234/5 genes for these three alleles, p^S *3234/5* included one nucleotide change compared with those of other two alleles, *p* and $+^p$, which altered Met at the 279-aa site to Thr, while all other amino acid sequences were the same among these three alleles (Fig. 3–Fig. 5). I further obtained full-length cDNA sequences of the *3234/5* gene from p^M and p^B alleles using RT-PCR. When comparing ORFs between $+^p$ and p^M, $+^p$ and p^B *3234/5* genes, p^M *3234/5* gene included one nucleotide change which altered Ala at the 189 aa site to Val, and p^B *3234/5* gene had one amino acid deletion at 100 aa site compared with that of $+^p$ (Fig. 3–Fig. 5). All other amino acid sequences of p^M or p^B were identical to $+^p$ *3234/5* gene.

To estimate the function of the gene *3234/5*, I performed a BLAST search using its full-length sequence and found a transcription factor Apontic (Apt) which had a sequence homologous to that of the *3234/5* gene. *apt*, previously identified in *Drosophila melanogaster*, is known to be involved in various morphological events, including modifications of the *Hox* gene, tracheal systems and others (Gellon *et al*., 1997; Eulenberg & Schuh, 1997; Su *et al*., 1999; Lie & Macdonald, 1999; Yoon *et al*., 2011; Liu *et al*., 2014). The Apt protein included a Myb/SANT DNA-binding motif and a leucine zipper protein-protein interaction motif, both of which were also retained in the *3234/5* gene (Fig. 3, Fig. 5). Thus, I subsequently designated *3234/5 apt-like*.

In the lepidopteran species *Manduca sexta*, *Plutella xylostella*, *Danaus plexippus*, *H. melpomene* and *P. xuthus*, an orthologous sequence of silkworm *apt-like* was found. When comparing the full-length ORFs for several

lepidopteran species (the p^M strain of *B. mori*, *B. mandarina*, *M. sexta*, *P. xuthus*, *P. machaon*, *P. aristrochiae*, *H. melpomene* and *D. plexippus*), I found that most regions, except for the leucine zipper-like motif, were conserved among Apt-like of *B. mori* and these other lepidopteran species (Fig. 6). Interestingly, the amino acid sequence of *B. mandarina* was completely consistent with that of the p^M strain (Fig. 6). A phylogenetic tree constructed based on the nucleotide sequences for *apt*, *apt-like* from several insect species (E value < 10^{-4}) revealed that all *apt-like*s in Lepidoptera comprised a monophyletic group, which suggested that this gene was highly conserved among lepidopteran species (Fig. 7). In addition, orthologous sequences of *Drosophila apt* were observed among dipteran and hymenopteran species (Fig. 7). The phylogenetic tree suggested that *apt-likes* in Lepidoptera were closely related to *apts* in Diptera and Hymenoptera (bootstrap value of 96 %; 1,000 replications). The *apt-like*-related sequence CG32813 (NP_569909.2) in *D. melanogaster*, the *apt*-related sequence LOC101742283 (XP_004933604.1) and LOC101737608 (XP_004933178.1) in *B. mori* were also found, although these genes did not encode the leucine zipper-like motif (Fig. 7). The *apt*-related sequence LOC101742283 and LOC101737608 in *B. mori* were located at Linkage Group 28 and Linkage Group 1, respectively. I further compared the genomic structures surrounding *apt-like* in *B. mori*, *D. plexippus* and *H. melpomene*, and found that the gene sets in this region were conserved among these three distantly related species (Fig. 8). This suggested that *apt-like* in Lepidoptera may have evolved from the same ancestral sequence. The

accession numbers for *apt-like* from p^S, $+^p$ and *p* strains were registered as AB860412, AB860413 and AB860414, respectively.

Furthermore, I compared the $+^p$: 33-kb genomic region among several lepidopteran species, *B. mori* (Surrounding $+^p$: 33-kb as the reference sequence), *M. sexta*, *Papilio polytes*, *H. melpomene* and *D. plexippus* by mVISTA, but found no clear homologous regions among them (Fig. 9).

Pigmentation-associated expression of *apt-like*.

To determine the cause of phenotypic differences associated with these five alleles (*p*, $+^p$, p^M, p^S and p^B), I examined the *apt-like* mRNA expression profiles in the larval epidermis by quantitative RT-PCR. During the fourth instar stage, I prepared cDNA from the larval epidermis of the second abdominal segment, for which pigmentation occurs in phenotypes $+^p$ (crescent markings), p^M (branch-like markings), p^S (broad band) and p^B (uniformly black) but not in *p*. The genotypes of all alleles excepted *p* allele were *p* heterozygote. Interestingly, *apt-like* mRNAs were highly expressed in the epidermis of p^S, p^M and p^B but not that of *p*, both at the feeding stage (C2) and molting stage (E1), the period just before occurring cuticle pigmentation (Fig. 10a). Furthermore, in the second abdominal segment, *apt-like* mRNA expression was higher in the epidermis of the $+^p$ phenotype compared with that in the epidermis of the *p* phenotype at the feeding stage but not at the molting stage (Fig. 10a). Notably, *apt-like* mRNA expression was not induced in the third abdominal segment of $+^p$, which had no

pigmentation patterns (Fig. 10b), whereas *apt-like* mRNA expression in the third abdominal segment of p^S, which had broad band pigmentation as with the second abdominal segment, was at the same level as the second abdominal segment (Fig. 10b). In addition, in the second abdominal segment of $+^p$ and p^S, *apt-like* mRNAs in the marking region were highly expressed compared with non-marking region (Fig. 10c). These results clearly show that ectopic expression of *apt-like* is associated with larval marking pattern formation.

Functional analyses of *apt-like* for larval pigmentation.

To establish whether the candidate gene *apt-like* for the *p* locus caused larval pigmentation, I used a recent developed technique to ectopically express this gene *in vivo* (Ando & Fujiwara, 2013). This method enabled the DNA constructs to become integrated into the epidermal genomes using electroporation and DNA transposon, which resulted in long-term expression of the target gene throughout the larval period. I injected a plasmid that contained both *apt-like* isolated from the p^S strain and enhanced *GFP* (*EGFP*) genes (*A3-*p^S*apt-like/A3-EGFP*), along with a *piggyBac* helper plasmid, into the second instar larva of *p*, which had no pigmentation (Fig. 11a). These *apt-like*-transformed cell lineages could be traced by *EGFP* fluorescent signals. I used an *actin A3* promoter to constitutively express *apt-like*.

When both plasmids were simultaneously introduced into the left half of a larva, *EGFP*-positive cells could be observed in the subsequent larval third,

fourth and fifth instars (Fig. 11b; left sides of fifth instars, EGFP, *A3-p^{S}apt-like/A3-EGFP*). Additional melanin pigmentation was observed in the corresponding area of EGFP positive cells (Fig. 11b; Bright field, *A3-p^{S}apt-like/A3-EGFP*). However, this pigmentation pattern was not observed if negative control *DsRed2* was expressed instead of *apt-like* (Fig. 12). This suggests that pigmentation on the *p* larval surface is truly induced by transgenic *apt-like* and not because of unanticipated technical problems. Furthermore, I used quantitative RT-PCR to confirm that *apt-like* expression introduced in the epidermal cells by the above method was actually up-regulated in the *EGFP*-positive regions (Fig. 13).

As described above, I noted that p^{S}-Apt-like, p^{M}-Apt-like and p^{B}-Apt-like had one amino acid change compared with *p* (or $+^{p}$), respectively (p^{S}, Met to Thr at the 279-aa site; p^{M}, Ala to Val at the 189-aa site; p^{B}, one amino acid deletion at the 100 aa site) (Fig. 3, Fig. 5). Thus, I tested whether these alteration affected the phenotypic differences among p^{S}, p^{M}, p^{B} and *p* (or $+^{p}$) by electroporation-mediated somatic transgenesis as described above. I constructed additional three plasmids (*A3-p apt-like/A3-EGFP*, *A3-p^{M} apt-like/A3-EGFP* and *A3-p^{B} apt-like/A3-EGFP*) that contained *apt-like* from the *p*, p^{M} or p^{B} strain rather than *apt-like* from the p^{S} strain used in the construct described above, which was injected along with the *piggyBac* helper plasmid into the second instar larva of *p* and introduced into epidermal cells by electroporation. Similar to the previous results for *p^{S}-apt-like*, *p-apt-like*, *p^{M} apt-like* and *p^{B} apt-like* ectopic expressions in larval

epidermis could also induce pigmentation in the *EGFP*-positive regions, respectively (Fig. 11b; Bright field, *A3-p apt-like/A3-EGFP*, *A3-p*^{M} *apt-like/A3-EGFP* and *A3-p*^{B} *apt-like/A3-EGFP*). These indicate that the amino acid changes observed among the five alleles of the *p* locus are not involved in the protein function of Apt-like, at least with respect to larval pigmentation.

In addition, I assessed the detailed pigmentation patterns caused by *apt-like* ectopic expression using confocal microscopy (Fig. 11c). Ectopic pigmentation (Fig. 11c; Bright field) was generated in the same cells that had *EGFP* signals (Fig. 11c; EGFP), indicated that *apt-like* regulated the downstream gene activation network during pigmentation processes in a cell-autonomous manner.

To confirm the reproducibility of above results, I performed larval RNAi of *apt-like*. I injected short interfering RNA (siRNA) that targeted *apt-like* into epidermis using *in vivo* electroporation because in lepidopteran larvae, only siRNA injection is not sufficient to induce RNAi at *in vivo* level. When siRNA was introduced into the left side epidermis of *p*^{S} or +^{p} at the fourth instar larval stage, melanin pigmentation was blocked for both alleles at the fifth instar stage (Fig. 14a; right side). However, there was no effect on pigmentation using negative-control siRNA (Fig.14a; left side). In addition, inhibition of black pigmentation was observed after introducing *apt-like* siRNA into *p*^{M} or *p*^{B} larvae (Fig. 14a; right side). I also confirmed whether *apt-like* mRNA expression levels decreased by RNAi. All RNAi-treated epidermis showed significantly reduction

of *apt-like* expression compared with those of non-treated epidermis (Fig. 14b; right side), whereas all negative-controls did not show significant reduction of *apt-like* expression (Fig. 14b; left side). These results indicate that a single gene, *apt-like*, regulates entirely different markings associated with *p* alleles, and that its expression is necessary for actual pigmentation to occur with all *p* alleles (at least five) that I tested.

To determine whether *apt-like* of other lepidopteran species operated on larval pigmentation, I cloned homologous sequences from three swallowtail butterflies, *Papilio xuthus*, *Papilio machaon* and *Pachliopta aristolochiae,* which are distantly related to *B. mori*. When these *apt-like* sequences from butterflies were expressed by *in vivo* electroporation in the larval epidermis of a *p* silkworm, melanin pigmentation was observed in regions where it had been introduced (Fig. 15). These results imply that the putative function of *apt-like* in melanin pigmentation may be conserved among Lepidoptera. The accession numbers for *apt-like* homologs from *P. xuthus*, *P. machaon* and *P. aristolochiae* were registered as LC017746, LC017747 and LC017748, respectively.

Six predicted genes in p^S are not involved in pigmentation.

Linkage analysis for the p^S locus suggested that, in addition to *apt-like*, six genes (*BGIBMGA003236*, *BGIBMGA003237*, *BGIBMGA003238*, *BGIBMGA003258*, *BGIBMGA003259* and *BGIBMGA003260*) were predicted to be within the 233-kb region (Fig. 2a, Fig. 16a). To determine the possible involvement of these

genes in larval pigmentation, I compared the expressions of these genes in the larval epidermis of p^S and *p* strains at the feeding (C2) and molting (E1) stages (Fig. 16b). *BGIBMGA003237* and *BGIBMGA003260* showed higher expressions in the p^S epidermis during the feeding stage, whereas the other four genes did not show statistically significant differences in expression between the two alleles at any stage. Of these six genes, *BGIBMGA003237* and *BGIBMGA003260* were predicted to encode for transcription factors. I knocked down *BGIBMGA003237* and *BGIBMGA003260* expression by siRNA with electroporation, which resulted in no effect on pigmentation (Fig. 16c, d). These results suggest that the six predicted genes are not involved in p^S pigmentation and that *apt-like* is the only promising candidate for p^S pigmentation.

Isoform expression and functional domains of *apt-like*.

I detected several *apt-like* isoforms in EST databases and by RT-PCR and RACE analyses (Fig. 17a). To determine the isoform expression for *p*, $+^p$ and p^S alleles, I used RT-PCR to amplify the region between exons two and eight of *apt-like* cDNA prepared from the second abdominal segment of fourth instar larvae. Only a single PCR product was detected in all samples from *p*, $+^p$ and p^S larvae; however, its expression appeared to be more abundant in the p^S strain (Fig. 17b). This indicates that the full-length transcript of *apt-like* is the most abundantly expressed and induced melanin pigmentation in the p^S strain. The function of remaining isoforms is unknown, but at least, they may not have potential to

induce black pigmentation because they lack most of Myb/SANT motif or Leucine zipper-like motif (Fig. 17a).

Tanaka (1970) has reported that the cuticle nanostructure of the p^S strain is strikingly different from that of the *p* strain. To confirm this observation, I examined the cuticle surface from animals with these alleles using electron microscopy. Tubercles and fold structures were found only in the p^S cuticle, as described in the study of Tanaka (1970) (Fig. 18). This ultrastructural difference may be caused by the mRNA expression of full-length *apt-like* or *apt-like* isoforms.

To characterize the functional domain of *apt-like* exons, I constructed a series of deletion mutants from the C-terminus of full-length cDNA (Fig. 19). Using electroporation-mediated somatic transgenesis, I co-introduced each plasmid and the *piggyBac* helper plasmid into second instar larvae and found that exons 3–6 were essential for pigmentation; however, the construct that lacked exon 7 exhibited only weak melanin formation, which suggested that the leucine zipper-like motif was essential for melanin formation in epidermal cells.

To confirm this possibility, I tested TALEN-mediated targeted gene knockout in a p^S homozygote. To disrupt *apt-like*, a pair of TALENs that targeted *apt-like* exon 6, which was just before the leucine zipper-like motif, was designed (Fig. 20a; *apt-like*-TALEN Left and *apt-like*-TALEN Right). Among TALEN microinjected embryos, two larvae showed white patches among the black stripes of p^S (Fig. 20b). In these white patched regions, the essential function of *apt-like*

may have been lost in epidermal cells, which could not induce black pigmentation. To determine what kinds of mutations were introduced into the *apt-like* gene by TALEN, I extracted genomic DNA from two mosaic larvae and amplified the region surrounding the target site by PCR. The sequences of PCR clones showed that two different mutants (Δ6 + I11, Δ16) potentially caused the frameshift of *apt-like*, which broke the coding region of the leucine zipper-like motif (Fig. 20c). These results suggest that the leucine zipper-like motif is involved in the essential function of *apt-like* and support the results described above for knockdown experiments with siRNA.

Embryonic and tissue expression of *apt-like*.

When siRNA that targeted *apt-like* was injected into an embryo with the p^M or *p* allele, it caused embryonic lethality in both cases (Fig. 21). Most embryos injected with *apt-like*-siRNA exhibited abnormal morphogenesis, such as a defective dorsal closure (Fig. 21a) or a twisted abdomen phenotype (Fig. 21b). This result suggests that *apt-like* is essential in the embryonic development of silkworms. To identify putative embryonic function of this gene, I examined temporal mRNA expression patterns in p^S and *p* embryos (non-diapausing eggs) using RT-PCR. During the embryonic development, I prepared cDNA from day 1–day 9 after oviposition. In non-diapausing eggs, the embryogenesis is completed in 10 days at 25 °C (Takami, 1969). In both alleles, *apt-like* mRNA expressed during early embryogenesis (day 1–day3), organogenesis (day 4–day7),

and late embryogenesis (day 8–day9) (Fig. 22). Embryos between day 8 and day 9 developed body pigmentation (Takami, 1969). These results suggest that *apt-like* is involved not only in the induction of larval pigmentation but also in organogenesis during embryonic development. Furthermore, I found that *apt-like* mRNA expression in p^S embryos was almost higher than that of *p* embryos (Fig. 22). This result implicates that the transcriptional activity of p^S *apt-like* is high not only in larval epidermis (Fig. 10) but also in embryogenesis (Fig. 22).

I also examined *apt-like* full-length mRNA expression in several tissues in the fourth instar larvae with $+^p$ allele (feeding period; stage C2). RT-PCR analysis showed transcriptional activity of *apt-like* in all nine tissues examined (the head, silk gland, fat body, trachea, Malpighian tubules, ovary, testis, mid-gut, and epidermis) (Fig. 23). This result suggests that *apt-like* has different biological roles in each larval tissue.

Regulatory gene network between *apt-like* and *Wnt1* signaling.

Although more than 100 Mendelian mutations have been described in the silkworm larvae, the phenotype of several of them can be dramatically affected by the *p* locus. For example, in the strain with mutant *multi lunar* (*L*), a dominant trait mapped to a single locus, twin black spots appear on the dorsal side of several neighboring segments. Recently, Yamaguchi *et al*. (2013) have identified *Wnt1* as the gene responsible for this pattern. Twin black spots appear only when the *L* individuals are crossed with $+^p$ allele carriers (Fig. 24a). These

characteristic markings do not appear in *L* and *p* allele crosses (Fig. 24b). These observations suggest an interaction between *Wnt1* and *apt-like*. Here I attempted to determine the genetic pathways among *apt-like, Wnt1* and *Distal-less* (*Dll*), using *in vivo* electroporation as described above. *Dll* is known to respond directly to morphogen *Wnt1* during leg and wing development in *Drosophila* (Neumann & Cohen, 1997; Zecca *et al*., 1996), and eye-spot formation on wings of Nymphalidae butterflies (Carroll *et al*., 1994). When *Wnt1* was ectopically introduced into the larval epidermis of the $+^{p}$ strain, increased expressions of *Dll* and *apt-like* were observed and melanin pigmentation occurred (Fig. 25a, b; left side). However, *Wnt1* ectopic expression in *p* epidermis could induce the expression of *Dll* but not that of *apt-like* or melanin pigmentation (Fig. 25c, d; left side). In contrast, *apt-like* ectopic expression in *p* larvae caused melanin pigmentation, as shown in Fig. 11, and induced the expression of *Dll* but not that of *Wnt1* (Fig. 25e, f; left side). I further investigated whether *Dll* expression in the epidermis causes black pigmentation. When siRNA that targeted *Dll* was introduced into the epidermis of $+^{p}$ and p^{B} as described above, no effects were observed in the both larval pigmentation, which suggested that *Dll* was not involved in black pigmentation (Fig. 26). These results demonstrate that *apt-like* expression is essential for melanin pigmentation caused by *Wnt1* in the $+^{p}$ strain. To confirm the reproducibility of the results, I performed a detailed time-course analysis of *apt-like* mRNA expression in the second abdominal segment of the fourth instar *L* (*L*; $+^{p}$) larvae. As previously reported by Yamaguchi *et al*. (2013),

I confirmed that the amount of *Wnt1* mRNA increased at the end of the feeding period, quickly decreased at the onset of molting (Fig. 27a). Furthermore, I found that similar to the *Wnt1* mRNA, *apt-like* mRNA was strongly expressed at the end of the feeding period; its expression quickly decreased at the onset of molting (Fig. 27b). This result shows that the changes in the expression of *apt-like* mRNA and *Wnt1* mRNA coincide during this stage of larval development.

Possible models for a gene network among these three genes are shown in Fig. 25g and h. In $+^{p}$ epidermis, *Wnt1* induces *apt-like* and *Dll* expression, which further induces black pigmentation (Fig. 25g). In contrast, in *p* epidermis, *Wnt1* can induce the expression of *Dll*, but not that of *apt-like* and black pigmentation (Fig. 25h). In both cases, *Dll* cannot induce melanin pigmentation.

Discussion

In this study, I focused on differences among five *p* alleles (*p*, $+^{p}$, p^{M}, p^{S} and p^{B}) and showed by fine mapping, transgene expression analysis and siRNA knockdown that *apt-like* was responsible for these alleles. The above results suggest that *apt-like* up-regulation is crucial for these alleles, and its regulatory changes may result in dramatically different larval color patterns. Based on classical linkage analyses for *B. mori*, more than 10 pigmentation mutants were mapped exactly at the same *p* locus (Table 1) using thousands of larvae for each analysis. From the genetic-historical view, the *p* locus has been shown to be the most remarkable one among hundreds of silkworm mutants (Banno *et al*., 2005). From the results shown here, I anticipate that all of the multiple phenotypes associated with *p* alleles will be found to be controlled by a single gene, *apt-like.*

Biological function of *apt* in *D. melanogaster*.

Apt has been characterized in *Drosophila* in several studies, which reported different aspects of the biological functions of this protein. Apt was shown to be a modifier of some *Hox* gene functions during gnathal development (Gellon *et al*., 1997) and involved in tracheal cell movement (Eulenberg & Schuh, 1997), heart morphology (Su *et al*., 1999), eye development (Liu *et al*., 2014) and derepression of *oskar* translation (Lie & Macdonald, 1999). Three of these reports indicated that *apt* may encode for a transcription factor, which may act as

a nuclear co-factor for *Hox* genes (Eulenberg & Schuh, 1997; Gellon *et al*., 1997; Liu *et al*., 2014). The Apt protein has glutamine-rich residues and a bZIP motif that includes a potential leucine zipper motif, which implies that this protein homodimerizes or forms a heterodimer with another protein. In addition, Apt includes a putative Myb/SANT DNA-binding domain, which comprises three helix–turn–helix types of alpha-helices. Liu *et al*. (2014) reported that Apt regulates the G1-to-S phase transition by directly inducing *cyclin E* during eye development. Montell's group discovered that Apt was involved in the expression of an miRNA that regulated JAK/STAT morphogen signaling (Yoon *et al*., 2011). Alternatively, Apt was reported to be a cytoplasmic RNA-binding protein involved in the translational control of *oskar* mRNA (Lie & Macdonald, 1999); however, no identifiable RNA-binding motif has been identified for this protein. No reports have indicated that Apt of *Drosophila* is involved in pigmentation or the melanin synthesis pathway.

Putative biological function of *apt-like* in Lepidoptera.

In my current study, several Apt-like proteins that retained the Myb/SANT motif were found among distantly related lepidopteran species (Fig. 6, Fig. 7). Monophyletic *apt-like* homologs in Lepidoptera were observed at the corresponding chromosomal location of silkworm Linkage Group 2 (Fig. 8) and showed highly conserved structures not only in the Myb/SANT motif but also in other coding regions (Fig. 6). During embryogenesis of the silkworm, *apt-like*

expressed in the organogenesis (Fig. 22). Injecting siRNA for *apt-like* into a *Bombyx* embryo caused its death before hatching (Fig. 21), similar to the embryonic lethality of a *Drosophila apt* mutant (Eulenberg & Schuh, 1997). In addition, homozygote of p^B, which shows the most darkness phenotype in *p* alleles, results in death before hatching (Kawaguchi, 1933). Furthermore, the EST database (CYBERGATE) and RT-PCR analysis showed that silkworm *apt-like* was expressed in various tissues, including the wing disk, head, silk gland, fat body, trachea, Malpighian tubles, ovary, testis and mid-gut (Fig. 23). These observations imply that lepidopteran Apt-like has some essential functions for various morphological events other than pigmentation formation on the larval cuticle similar to *Drosophila* Apt. The possible function of Apt-like for larval pigmentation may be conserved, at least among Lepidoptera (Fig. 15).

***apt-like* induces black pigmentation in a cell-autonomous manner.**

The important finding of this study is that only one putative transcription factor, Apt-like, is responsible for a single genetic locus, *p*, which possibly generates phenotypic diversity for pigment patterns. In addition, *piggyBac*-based somatic transgenesis by electroporation clearly showed that *apt-like* caused melanin synthesis in a cell-autonomous manner (Fig. 11). Furthermore, TALEN-mediated mosaic analysis showed that somatic *apt-like* knockout caused white patches among p^S markings in a cell-autonomous manner (Fig. 20). The transplantation of larval integuments between p^S and *p* suggests that transplanted integuments of

p^S form melanin pigments in the donor's epidermis (Nagashima, 1958). Fujiwara *et al*. demonstrated that somatic loss of a chromosomal fragment, including the p^S locus, in an epidermal cell-lineage of a genetic mosaic p^{Sm}, mottled striped strain (Fig. 1), produced white patches (variegated pigmentation) among the dorsal black stripes (Fujiwara & Maekawa, 1994; Fujiwara *et al*., 2000). The random, impartial loss of this p^S chromosomal fragment occurred at any developmental stage, which caused various-sized white spots, even at the single cell level. These observations supported cell-autonomous functions of *apt-like*, as shown in this study.

***apt-lik*e regulates multiple melanin-related genes in a module manner.**

In the swallowtail butterfly *P. xuthus*, expressions of seven melanin-related genes, *yellow*, *tan*, *ebony*, *DDC*, *TH*, *Laccase 2* and *GTPCHI*, are strongly correlated with larval eye-spot markings (Futahashi & Fujiwara, 2005, 2006, 2007, 2008; Futahashi *et al*. 2008, 2010). Most of these expressions coincide with species-specific cuticule markings of other *Papilio* species, *P. polytes* and *P. machaon* (Shirataki *et al*., 2010). These observations in butterflies indicate that multiple melanin-related genes located in a separate chromosomal locus act as a module to cause various interspecies and intraspecies pigmentation patterns. In contrast, in *Drosophila*, alterations in the spatial expression patterns of a small set of melanin synthesis genes, such as *yellow*, *tan* and *ebony*, causes species-specific variations in color patterns; the expression control of *tan* and *yellow* is involved

in the divergence of abdominal pigmentation between *D. santomea* and *D. yakuba* (Jeong *et al*., 2008; Rebeiz *et al*., 2009). In addition, these few melanin-related genes have been reported to be involved in pigmentation diversity in several *Drosophila* species (Wittkopp *et al*., 2003; Wittkopp *et al*., 2009).

It is known that the pigments caused by the p^S and $+^p$ alleles consist of melanin (Aruga *et al*., 1965). In addition, Futahashi *et al*. (2008) showed that *yellow* and *ebony* mRNA expressions were induced at molting stages (E1 and E2, respectively), and were more abundant in the p^S phenotype compared with that in the $+^p$ phenotype. Although the Apt-like-regulated gene pathway remains unclear, the abundant expression of *yellow* and *ebony* is considered to be involved in the dorsal black stripes of the silkworm p^S strain (Futahashi *et al*., 2008). Okamoto (2005) compared the mRNA expressions of genes involved in the melanin synthesis pathway and found that the mRNA expressions for *TH*, *DDC* and *Laccase 2* were induced at higher levels during molting stages with the p^S allele than with the $+^p$ allele, whereas *Laccase 1* mRNA expression was at similar levels in both phenotypes. These results suggest that *apt-like* induces downstream melanin synthesis pathway, primarily through *yellow* and *ebony* gene expression.

It is noteworthy that the *yellow* expression coincides with the black regions and is not detected in the white striped region (Futahashi *et al*., 2008). The weak induction of *TH*, *DDC* and *Laccase 2* in the p^S phenotype after molting stages (Okamoto, 2005) support the above observations and suggest that these melanin synthesis genes act as a module that is co-regulated by Apt-like.

I found one amino acid change among each five *p* alleles; however, these alterations were not involved in the various marking patterns (Fig. 11). *apt-like* expression was specifically observed in the black marking regions of p^M, p^S, p^B and the crescent markings of $+^p$; however, it was severely repressed in *p* (Fig. 10). In addition, *apt-like* expression in the pigmentation regions was observed during the feeding stages, which preceded the gene expression of the melanin-synthesis module during the molting phase. These expression profiles of *apt-like* suggest that the level of gene expression regulates directly or indirectly this melanin synthesis module in a region-specific and stage-specific manner. Notably, larval markings should be repeatedly redrawn at each larval molt, which is different from adult pigmentation that occurs only once during metamorphosis. Recent works showed that the molting hormone ecdysteroid possibly regulated larval pigmentation through the function of *Wnt1* in lepidopteran spot markings (Yamaguchi *et al*., 2013) or *E75* in eye spots of the swallowtail butterfly (Futahashi *et al*., 2012). I speculate that in contrast to adult pigmentation, module-type control of gene expression of the melanin-synthesis genes is necessary for repeated pigmentation in the larval epidermis.

Pleiotropic effect of *apt-like* in the larval integument.

Morphologically, the cuticle of various *p* allele carriers shows ultrastructural differences that correlate with pigmentation. As described in the study of Tanaka (1970), I found, using scanning electron microscopy, that a number of tubercles

and fold structures emerged on the p^S cuticles, whereas such ultrastructures were rarely observed in *p* cuticle (Fig. 18). Okamoto (2005) has shown that *LCP17*, encoding cuticular protein RR-1 motif 2, and *drop dead-like* gene, encoding acyltransferase (involved in the development of chitin-based cuticle), cause specific markings in p^S but not in *p* epidermis. These results suggest that *apt-like* regulates the expression of cuticle-related genes in parallel with the induction of pigmentation.

Ashino (1940) and Aruga (1943) have reported that the *Knobbed* mutant, which forms characteristic knobs (protuberances) at specific dorsal regions of several segments, presents different phenotypes depending on the *p* allele. Ashino (1940) has suggested that the knob formation is because of excessive epidermal cell proliferation during the molting stage. When *Knobbed* is combined with *p*, the knobs protrude visibly. When *Knobbed* is combined with p^S, the knobs are reduced markedly (Shimura, 2013). It is possible that the accumulation of melanin granules suppresses knob development (Aruga, 1943), or that the increased expression of *apt-like* affects *Knobbed* gene expression.

Besides melanin pigments, the formation of urate granules in the larval epidermis is also important for larval body coloration in Lepidoptera; the granules make the larval integument white or opaque (Tamura & Sakate, 1983). The epidermal cells of the silkworm contain large quantities of uric acids (Yoshitake & Aruga, 1952). Some physiological studies have suggested that the uric acid content in the larval epidermis and hemolymph is different in p^S and *p*

larvae (Yoshitake & Aruga, 1952; Hu *et al*., 2013). The p^S larvae have very little urate content in the epidermis, in contrast to the urate content in that of the *p* larvae. Interestingly, the uric acid content in the hemolymph is in inverse relation to its content in the integument. The above results suggest that the uric acid uptake into the epidermis is affected by *apt-like*. In several lepidopteran larvae, the epidermis of white integument contains much more urate than that of black integument (Timmerman & Berenbaum, 1999; Ninomiya *et al.*, 2006; Ninomiya & Hayakawa, 2007). This implies an inverse relationship between the levels of urate and melanin in the larval epidermis. Hu *et al.* (2013) have found that melanin and urate prevent ultraviolet damage to the larval integument in the silkworm. If *apt-like* contributed to black pigmentation in lepidopteran larvae in nature, it may be involved in this antagonistic relationship.

A putative molecular mechanism for producing $+^p$ phenotype.

How the region-specific expression of *apt-like* occurs in the crescent markings of $+^p$ is a question that remains to be answered. Previous study found a specialized cuticular structure designated “bulge” in the crescent marking area, in which *Wnt1* was ectopically expressed and that *Wnt1* expression in the larval epidermis had the potential to cause twin spot markings (Yamaguchi *et al*., 2013). In this study, using *in vivo* electroporation, I observed that ectopically expressed *Wnt1* induced *apt-like* expression and melanin pigmentation in $+^p$ epidermis but not in *p* epidermis (Fig. 25). These results suggest that the *Wnt1* signaling pathway may

be involved in the formation of the crescent markings in $+^{p}$.

It is noteworthy that *apt-like* ectopic expression could cause melanin pigmentation even in *p* epidermis, which suggests that the phenotypic differences between *p* and $+^{p}$ may be generated by their differential responses to the *Wnt1* pathway. I speculate that *apt-like* may be downstream of the *Wnt1* pathway in the $+^{p}$ larva (Fig. 25g), but not in the *p* larva (Fig. 25h) probably because the *cis*-regulatory elements of *apt-like* is mutated and does not respond to the *Wnt1* pathway.

Okamoto (2005) compared the sequences in most regions of $+^{p}$: 33-kb between *p* and $+^{p}$ strains and observed numerous indels and sequence variations. Thus, some of these variations may be involved in the *cis*-regulatory changes responsible for binding of *Wnt1* signaling factors. In addition, I compared sequences of the $+^{p}$: 33-kb genomic region among several lepidopteran species by mVISTA, but found no clear homologous regions among them (Fig. 9). Detailed, long-range sequence comparisons of each *p* allele should clarify the regulatory mechanisms for producing phenotypic diversity by only a single novel gene such as *apt-like*.

The molecular basis for generating multiple *p* phenotypes.

Understanding the molecular basis underlying the diverse larval coloration patterns associated with the multiple *p* alleles will be the future goal of our studies. What kind of genetic changes within *apt-like* contribute to various

coloration patterns? The evolution of diversity of the insect color patterns is generally associated with changes in the spatial expression of genes governing the development (Carroll *et al*., 2005). In the recent studies of *Drosophila* species, the divergence of particular morphological traits has been linked to changes in specific enhancers (*cis*-regulatory elements) of individual loci (Gompel *et al*., 2005; Prud'homme *et al*., 2006; Jeong *et al*., 2006, 2008; Williams *et al*., 2008). The *apt-like* may have also undergone some functional changes within its non-coding region, causing the phenotypic diversity of the *p* alleles in the silkworm. Genomic instability has often been observed in the *p* locus. For instance, p^{G} and p^{Sa} have been obtained from the p^{S} allele by X-ray irradiation (Tazima, 1938; Shimodaira, 1947). $p^{Sa\text{-}2}$ and p^{St} have also been induced by X-ray irradiation (Tazima, 1943b; Takasaki, 1947), and p^{L} has been derived from the p^{S} allele using high-temperature treatment (Tazima, 1938). High sensitivity to these mutagens suggests that the genomic instability contributed to the emergence of the functional *cis*-regulatory elements in the *p* locus. To examine the genetic differences causing phenotypic variation in the *p* allele carriers, we need to perform sequence comparisons of the regions surrounding *apt-like* in the genomes with different *p* alleles. The functional changes should be identified using reporter assays and/or genome editing (TALEN and/or CRISPR-Cas9 System).

The results of this study showed that *Wnt1* signaling may be involved in the induction of *apt-like* expression in $+^{p}$. Because *Wnt1* is regulated by the

ecdysteroid hormone (Yamaguchi *et al*., 2013), *apt-like* may be indirectly regulated by that hormone. According to Kiguchi (1972), it is likely that *apt-like* is also regulated by the juvenile hormone (JH). Deep melanization in $+^{p}$ markings caused by deligation or allatectomy in the larvae suggests that *apt-like* expression levels is JH-dependent. The larval markings characteristic for the *p* alleles only emerge in the larval stages; to induce black pigmentation, *apt-like* may need some interactions with both ecdysone and JH signaling.

The pigmentation intensity may be correlated with *apt-like* mRNA expression levels (Fig. 10). The analysis of relative levels of *apt-like* expression showed that for p^{B}, which represents the darkest pigmentation among the *p* alleles, *apt-like* expression was about 5 times higher during the feeding period and more than 20 times higher during the molting period, than for the *p* allele. In contrast, in p^{S} and p^{M} individuals *apt-like* expression was only twice as high as in those with the *p* allele. These results suggest that the pigmentation intensity in the epidermis depends on *apt-like* mRNA expression level. p^{B} *apt-like* has one amino acid deletion in Myb/SANT motif (involved in DNA binding), and such alteration may affect the transcriptional activity of downstream genes.

The silkworm has been domesticated over the past 5,000 years from the wild progenitor *B. mandarina* (Goldsmith *et al*., 2005). *B. mandarina* displays branch mimicking and characteristic eye-spot markings similar to those in the moths with $+^{p}$. Although p^{M} phenotype closely resembles that of *B. mandarina*, there is no direct evidence to show introgression of *p* locus between *B.*

mandarina and *B. mori*. However, recently, Dr. Y. Banno (Kyushu University) (2014, personal letter) has reported some interesting results. Heterozygous F1 females have been obtained from a single-pair cross between a *B. mandarina* (strain Sakado) female and a $+^{p}$ (strain p50) male. Each offspring was backcrossed with a $+^{p}$ (strain p50) female (line breeding, at least 12 times). Moreover, those backcrossed individuals showed p^{M} phenotype; they were established as strain T02 (used as the experimental material in that study). Using molecular markers, Drs. Y. Banno and T. Shimada (University of Tokyo) confirmed that in backcrossed individuals, all chromosomes, except for chromosome 2 (Linkage Group 2), were replaced by those from $+^{p}$. This result indicates that p^{M} is the trait of *B. mandarina*. In the present study, when siRNA targeting *apt-like* was introduced into the p^{M} epidermis of the fourth instar larvae, melanin pigmentation was severely blocked at the fifth instar stage (Fig. 14). Furthermore, the amino acid sequence of p^{M} Apt-like protein (derived from the spontaneous mutant strain u10) was consistent with that of *B. mandarina* (Fig. 6). Altogether, these results suggest that *apt-like* is involved in the development of the *B. mandarina* markings, and it contributed to the branch mimicking.

References

Ando, T. & Fujiwara, H. Electroporation-mediated somatic transgenesis for rapid functional analysis in insects. *Development* **140,** 454–458 (2013).

Aruga, H. Histological studies of larval marking mutants in the silkworm, *Bombyx mori*. *Bull. Seric. Exp. Sta.* **11,** 387–439 (1943).

Aruga, H., Kawase, S., Nagashima, E., Watanabe, H. & Yoshitake, N. *Insect Genetics* (Azumi-Shobo, Tokyo, 1965). (in Japanese)

Ashino, M. Morphological study on the manifestation of knobbed factor in the silkworm, *Bombyx mori*. *J. Seric. Sci. Jpn.* **11,** 209–210 (1940).

Banno, Y. *et al.* A Guide to the Silkworm Mutants—*Gene Name and Gene.* Silkwork Genetics Division, Institute of Genetic Resources, Kyushu University, Fukuoka, Japan (2005).

Berenbaum, M. R. Aposematism and mimicry in caterpillars. *J. Lepid. Soc.* **49,** 386–396 (1995).

Bernardi, G., Pierre, J. & Nguyen, T. H. Le polymorphisme et le mimétisme de Papilio dardanus *Brown. Bul. Soc. Entomo. France* **90,** 106–155 (1985).

Bond, A. B. The evolution of color polymorphism: crypticity, searching images, and apostatic selection. *Annu. Rev. Ecol. Evol. Syst.* **38,** 489–514 (2007).

Brakefield, P. M. *et al.* Development, plasticity and evolution of butterfly eyespot patterns. *Nature* **384,** 236–242 (1996).

Brodbeck, D. *et al.* Molecular and biochemical characterization of the *aaNAT1*

(*Dat*) locus in *Drosophila melanogaster*: differential expression of two gene products. *DNA Cell Biol.* **17,** 621–633 (1998).

Brudno, M. *et al*. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Research* **13,** 721–731 (2003).

Carroll, S. B. *et al*. Pattern formation and eyespot determination in butterfly wings. *Science* **265,** 109–114 (1994).

Carroll, S. B., Grenier, J. K. & Weatherbee, S. D. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design.* (Blackwell, Malden, MA. ed. 2, 2005).

Clark, R. *et al.* Color pattern specification in the Mocker swallowtail *Papilio dardanus*: the transcription factor *invected* is a candidate for the mimicry locus *H*. *Proc. Biol. Sci.* **275,** 1181–1188 (2008).

Cook, L. M. The rise and fall of the *carbonaria* form of the peppered moth. *Q. Rev. Biol.* **78,** 399–417 (2003).

Edgar, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinfom*. **5,** 113 (2004).

Eulenberg, K. G. & Schuh, R. The tracheae defective gene encodes a bZIP protein that controls tracheal cell movement during *Drosophila* embryogenesis. *EMBO J.* **16,** 7156–7165 (1997).

Ford, E. B. The genetics of *Papilio dardanus*, Brown (Lep.). *Trans. R. Entomol. Soc. London* **85,** 435–466 (1936).

Ford, E. B. *Ecological Genetics*. (Methuen, London, 1964).

Futahashi, R. & Fujiwara, H. Melanin-synthesis enzymes coregulate stage-specific larval cuticular markings in the swallowtail butterfly, *Papilio xuthus*. *Dev. Genes Evol.* **215,** 519–529 (2005).

Futahashi, R. & Fujiwara, H. 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 (2006).

Futahashi, R. & Fujiwara, H. 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 (2007).

Futahashi, R. & Fujiwara, H. Juvenile hormone regulates butterfly larval pattern switches. *Science* **319,** 1061 (2008).

Futahashi, R. *et al*. *yellow* and *ebony* are the responsible genes for the larval color mutants of the silkworm *Bombyx mori*. *Genetics* **180,** 1995–2005 (2008).

Futahashi, R., Banno, Y. & Fujiwara, H. Caterpillar color patterns are determined by a two-phase melanin gene prepatterning process: new evidence from *tan* and *laccase 2*. *Evol. Dev.* **12,** 157–167 (2010).

Futahashi, R., Shirataki, H., Narita, T., Mita, K. & Fujiwara, H. Comprehensive microarray-based analysis for stage-specific larval camouflage pattern-associated genes in the swallowtail butterfly, *Papilio xuthus*. *BMC Biol.* **10,**

46 (2012).

Fujiwara, H. & Maekawa, H. Mosaic formation by developmental loss of a chromosomal fragment in a “mottled striped” mosaic strain of the silkworm, *Bombyx mori*. *Roux's Arch. Dev. Biol.* **203,** 389–396 (1994).

Fujiwara, H., Nakazato, Y., Okazaki, S. & Ninaki, O. Stability and telomere structure of chromosomal fragments in two different mosaic strains of the silkworm, *Bombyx mori*. *Zool. Sci.* **17,** 743–750 (2000).

Gellon, G., Harding, K. W., McGinnis, N., Martin, M. M. & McGinnis, W. A genetic screen for modifiers of *Deformed* homeotic function identifies novel genes required for head development. *Development* **124,** 3321–3331 (1997).

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

Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A. & Carroll, S. B. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**, 481–487 (2005).

Greene, E. A diet-induced developmental polymorphism in a caterpillar. *Science* **243,** 643–645 (1989).

Hebert, P. D., Penton, E. H., Burns, J. M., Janzen, D.H. & Hallwachs, W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. USA* **101,** 14812–14817 (2004).

Hu, Y. G., Shen, Y. H., Zhang, Z. & Shi, G. Q. Melanin and urate act to prevent

ultraviolet damage in the integument of the silkworm, *Bombyx mori*. *Arch. Insect Biochem. Physiol*. **83,** 41–55 (2013).

International Silkworm Genome Consortium. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochem. Mol. Biol.* **38,** 1036–1045 (2008).

Janzen, D. H., Hallwachs, W. & Burns, J. M. A tropical horde of counterfeit predator eyes. *Proc. Natl. Acad. Sci. USA* **107,** 11659–11665 (2010).

Jeong, S., Rokas, A. & Carroll, S. B. Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell* **125,** 1387–1399 (2006).

Jeong, S. *et al.* The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* **132,** 783–793 (2008).

Jiggins, C. D., Naisbit, R. E., Coe, R. L. & Mallet, J. Reproductive isolation caused by colour pattern mimicry. *Nature* **411**, 302–305 (2001).

Joron, M. *et al.* A conserved supergene locus controls color pattern diversity in *Heliconius* butterflies. *PLoS Biol.* **4,** e303 (2006).

Joron, M. *et al*. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477,** 203–206 (2011).

Kawaguchi, E. *Jap. J. Genet.* **8,** 97–107 (1933). (in Japanese)

Kiguchi, K. Hormonal control of the coloration of larval body and the pigmentation of larval markings in *Bombyx mori* (1) Endocrine organs affecting the coloration of larval body and the pigmentation of markings. *J.*

Seric. Sci. Jap. **41,** 407–412 (1972).

Kiguchi, K. & Agui, N. Ecdysteroid levels and developmental events during larval molting in the silkworm, *Bombyx mori*. *J. Insect Physiol.* **27,** 805–812 (1981).

Kronforst, M. R. *et al*. Unraveling the thread of nature's tapestry: the genetics of diversity and convergence in animal pigmentation. *Pigment Cell Melanoma Res*. **25,** 411–433 (2012).

Lie, Y. S. & Macdonald, P. M. Apontic binds the translational repressor Bruno and is implicated in regulation of *oskar* mRNA translation. *Development* **126,** 1129–1138 (1999).

Liu, C. *et al*. Repression of tyrosine hydroxylase is responsible for the sex-linked chocolate mutation of the silkworm, *Bombyx mori*. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **107,** 12980–12985 (2010).

Liu, Q. X. *et al*. Evolutionarily conserved transcription factor Apontic controls the G1/S progression by inducing *cyclin E* during eye development. *Proc. Natl. Acad. Sci. USA* **111,** 9497–9502 (2014).

Martin, A. *et al*. Diversification of complex butterfly wing patterns by repeated regulatory evolution of a *Wnt* ligand. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **109,** 12632–12637 (2012).

Nagashima, E. Physiogenetical studies on the transplantation of integument in the silkworm, *Bombyx mori* L. (IV) On the transplantation of larval integument of lemon yellow and *od*-translucent. *J. Seric. Sci. Jap.* **27,** 393-

399 (1958).

Neumann, C. J. & Cohen, S. M. Long-range action of Wingless organizes the dorsal-ventral axis of the *Drosophila* wing. *Development* **124,** 871–880 (1997).

Nijhout, H. F. Polymorphic mimicry in *Papilio dardanus*: mosaic dominance, big effects, and origins. *Evol. Dev.* **5,** 579–592 (2003).

Nijhout, H. F. Molecular and physiological basis of color pattern formation. *Adv. Insect Physiol.* **38,** 219–265 (2010).

Ninomiya, Y. *et al*. Mechanisms of black and white stripe pattern formation in the cuticles of insect larvae. *J. Insect Physiol.* **52,** 638–645 (2006).

Ninomiya, Y. & Hayakawa, Y. Insect cytokine, growth-blocking peptide, is a primary regulator of melanin-synthesis enzymes in armyworm larval cuticle. *FEBS J.* **274,** 1768–1777 (2007).

Ohashi, M., Tsusue, M., Yoshitake, N., Sakate, S. & Kiguchi, K. Epidermal pigments affecting the larval coloration of the silkworm, *Bombyx mor*i. *J. Seric. Sci. Jpn.* **52,** 498–504 (1983).

Okamoto, S. *Genome-wide approaches for identifying melanogenesis and larval epidermis-specific genes in Lepidoptera.* (Master thesis, University of Tokyo, 2005). (in Japanese)

Poulton, E. B. *Papilio dardanus*, the most wonderful butterfly in the world. *J. East Afr. Uganda Nat. Hist. Soc.* **209**, 4–22 (1924).

Prud'homme, B. *et al*. Repeated morphological evolution through *cis*-regulatory

changes in a pleiotropic gene. *Nature* **440,** 1050–1053 (2006).

Rebeiz, M., Pool, J. E., Kassner, V. A., Aquadro, C. F. & Carroll, S. B. Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* **326,** 1663–1667 (2009).

Rebeiz, M. *et al.* Evolution of the *tan* locus contributed to pigment loss in *Drosophila santomea*: a response to Matute et al. *Cell* **139,** 1189–1196 (2009).

Reed, R. D. *et al*. *optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* **333,** 1137–1141 (2011).

Rogers, W. A. *et al*. Recurrent modification of a conserved *cis*-regulatory element underlies fruit fly pigmentation diversity. *PLoS Genet*. **9,** e1003740 (2013).

Saenko, S. V., Jerónimo, M. A. & Beldade, P. Genetic basis of stage-specific melanism: a putative role for a cysteine sulfinic acid decarboxylase in insect pigmentation. *Heredity* **108,** 594–601 (2012).

Sasaki, S. *Jpn. J. Genet.* **30,** 185 (1955). (in Japanese)

Shamim, G., Ranjan, S. K., Pandey, D. M. & Ramani, R. Biochemistry and biosynthesis of insect pigments. *Eur. J. Entomol.* **111,** 149–164 (2014).

Shimodaira, M. *Jap. J. Genet.* **22,** 86–92 (1947). (in Japanese)

Shimura, S. Mechanism of dermal knob formation in the *Knobbed* (*K*) mutant of the silkworm, Bombyx mori: Characterization of the epidermal cells. *Misc. Publ. Natl. Inst. Agrobiol. Sci.* **10,** 1–53 (2013).

Shirataki, H., Futahashi, R. & Fujiwara, H. Species-specific coordinated gene expression and trans-regulation of larval color pattern in three swallowtail butterflies. *Evol. Dev.* **12,** 305–314 (2010).

Sturtevant, A. H. No crossing over in the female of the silkworm moth. *Am. Nat.* **49,** 42–44 (1915).

Su, M. T., Venkatesh, T. V., Wu, X., Golden, K. & Bodmer, R. The pioneer gene, *apontic*, is required for morphogenesis and function of the *Drosophila* heart. *Mech. Dev.* **80,** 125–132 (1999).

Suyama, M., Torrents, D. & Bork, P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34,** W609–W612 (2006).

Takami, T. *General Principles of Silkworm Variety*. (National Silkworm Association, 1969). (in Japanese)

Takasaki, T. *Silkworm Information Service* **2**, 3 (1947). (in Japanese)

Takasu, Y. *et al*. Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection. *Insect Biochem. Mol. Biol.* **40,** 759–765 (2010)

Takasu, Y. *et al*. Efficient TALEN construction for *Bombyx mori* gene targeting. *PLoS ONE* **8,** e073458 (2013).

Tamura, T. & Sakate, S. Relationship between the expression of oily character and uric acid incorporation in the larval integument of various oily mutants of the silkworm, *Bombyx mori*. *Bull. Seric. Exp. Stn.* **28,** 719–740 (1983).

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **30,** 2725–2729 (2013).

Tanaka, Y. A study of mendelian factors in the silkworm, *Bombyx mori*. *Jour. Coll. Agric. Tohoku Imp. University, Sapporo, Japan* **5,** 91–113 (1913).

Tanaka, Y. Further data on the reduplication in silkworms. *Jour. Coll. Agric. Tohoku Imp. University, Sapporo, Japan* **6,** 1–16 (1914).

Tanaka, K. Studies on the structure of the integuments of *Bombyx mori* L. and several other kind of Lepidoptera. I. On the surface structure of the larval integuments of the silkworm. *J. Seric. Sci. Jpn.* **39,** 94–103 (1970).

Tazima, Y. *Jap. J. Genet.* **14,** 117–128, 191–203 (1938). (in Japanese)

Tazima, Y. *J. Seric. Sci. Jpn.* **12,** 184–188 (1941). (in Japanese)

Tazima, Y. *J. Seric. Sci. Jpn.* **14,** 76–89 (1943a). (in Japanese)

Tazima, Y. *Bull. Seric. Exp. Sta. Jap.* **11,** 525–604 (1943b). (in Japanese)

Tazima, Y. *Bull. Imp. Sericult. Expt. Sta. Japan* **12,** 109–181 (1944). (in Japanese)

Tazima, Y. *Rep. Silk Sci. Res. Inst.* **5,** 5–24 (1955). (in Japanese)

Tazima, Y. *The Genetics of the Silkworm* (Logos Press and Prentice Hall, 1964).

Tazima, Y. *The Silkworm: An Important Laboratory Tool* (Kodansha, Tokyo, 1978).

Tazima, Y. *Kaiko no hinsyu-ikusei* (Science house, Japan, 1993). (in Japanese)

The *Heliconius* Genome Consortium. Butterfly genome reveals promiscuous

exchange of mimicry adaptations among species. *Nature* **487,** 94–98 (2012).

Timmermann, S. & Berenbaum, M. R. Uric acid deposition in larval integument of black swallowtails and speculation on its possible functions. *J. Lepidopterists' Soc.* **53**, 104–107 (1999).

Timmermans, M. J. T. N. *et al*. Comparative genomics of the mimicry switch in *Papilio dardanus*. *Proc. R. Soc. B* **281**, 20140465 (2014).

Toyama, K. Studies on the hybridology of insects. I. On some silkworm crosses with special reference to Mendel's law of heredity. *Bull. Coll. Agr. Tokyo Imp. Univ.***7,** 321–334 (1906).

True, J. R. Insect melanism: the molecules matter. *Trends Ecol. Evol.* **18,** 640–647 (2003).

True, J. R. *et al*. *Drosophila tan* encodes a novel hydrolase required in pigmentation and vision. *PLoS Genet.* **1,** e63 (2005).

Turner, J. R. G. Adaptation and evolution in *Heliconius*: a defense of neo-Darwinism. *Annu. Rev. Ecol. Syst.* **12**, 99–121 (1981).

Turner, J. R. G. *The Biology of Butterflies* (eds Vane-Wright, R. I. & Ackery, P. R.) (Academic Press, 1984).

Wei, G. Q., Yu, L., Liu, C. L., Zhu, B. J. & Ding, H. J. Linkage and mapping analyses of the normal marking gene *+P* in the silkworm (*Bombyx mori*) using SSR markers. *Genet. Mol. Res*. **12,** 2351–2359 (2013).

Werner, T., Koshikawa, S., Williams, T. M. & Carroll, S. B. Generation of a novel wing color pattern by the *Wingless* morphogen. *Nature* **464,** 1143–

1148 (2010).

Williams, T. M. *et al*. The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* **134**, 610–623 (2008).

Wittkopp, P. J., True, J. R. & Carroll, S. B. Reciprocal functions of the *Drosophila* Yellow and Ebony proteins in the development and evolution of pigment patterns. *Development* **129,** 1849–1858 (2002).

Wittkopp, P. J., Williams, B. L., Selegue, J. E. & Carroll, S. B. *Drosophila* pigmentation evolution: divergent genotypes underlying convergent phenotypes. *Proc. Natl. Acad. Sci.USA* **100,** 1808–1813 (2003).

Wittkopp, P. J. & Beldade P. Development and evolution of insect pigmentation: Genetic mechanisms and the potential consequences of pleiotropy. *Semin. Cell Dev. Biol.* **20**, 65–71 (2009).

Wittkopp, P. J. *et al.* Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in *Drosophila*. *Science* **326,** 540–544 (2009).

Wright, T. R. F. The genetics of biogenic amine metabolism, sclerotization, and melanization in *Drosophila melanogaster*. *Adv. Genet.* **24**, 127–222 (1987).

Yamaguchi, J., Mizoguchi, T. & Fujiwara, H. siRNA induce efficient RNAi response in *Bombyx mori* embryos. *PLoS ONE* **6,** e25469 (2011).

Yamaguchi, J. *et al*. Periodic *Wnt1* expression generates twin-spot markings on the larval epidermis. *Nat. Commun.* **4,** 1857 (2013).

Yamamoto, K. *et al.* A BAC-based integrated linkage map of the silkworm

Bombyx mori. *Genome Biol.* **9,** R21 (2008).

Yoshitake, N. & Aruga. H. On the uric acid content in the integument and the blood of several mutants in the silkworm. *J. Seric. Sci. Jpn.* **21,** 7–14 (1952).

Yoon, W. H., Meinhardt, H. & Montell, D. J. miRNA-mediated feedback inhibition of JAK/STAT morphogen signalling establishes a cell fate threshold. *Nat. Cell Biol.* **13,** 1062–1069 (2011).

Yu, H. S. *et al*. Evidence of selection at melanin synthesis pathway loci during silkworm domestication. *Mol. Biol. Evol.* **28,** 1785–1799 (2011).

Zecca, M., Basler, K. & Struhl, G. Direct and long-range action of a wingless morphogen gradient. *Cell* **87,** 833–844 (1996).

Zhan, S. *et al*. Disruption of an N-acetyltransferase gene in the silkworm reveals a novel role in pigmentation. *Development* **137,** 4083–4090 (2010).

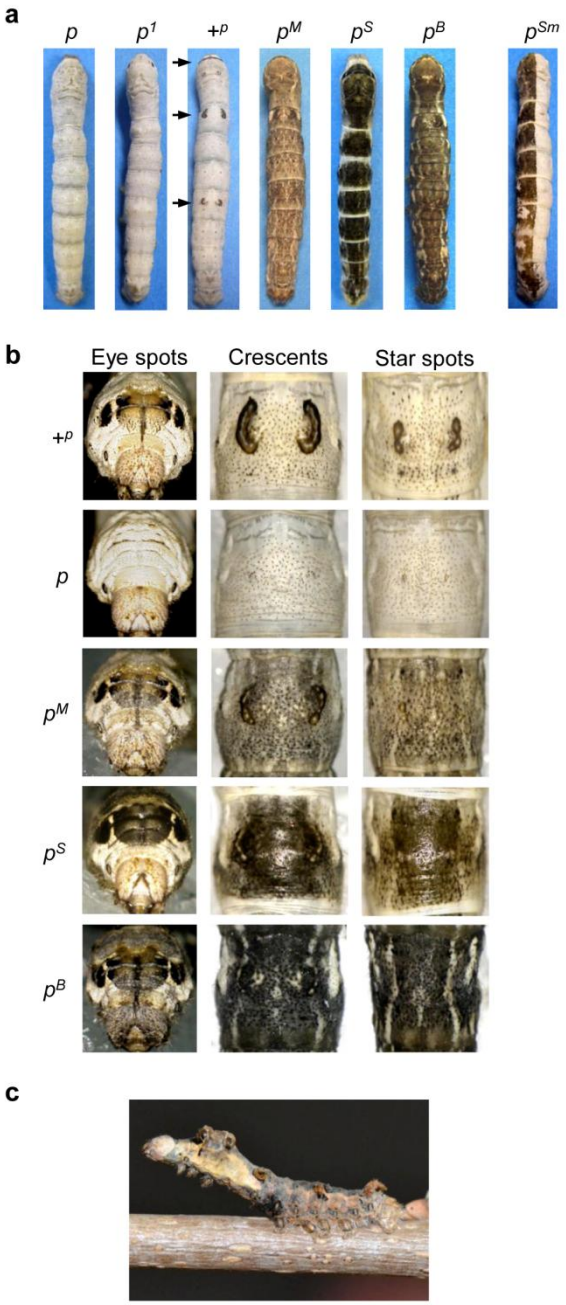


Figure 1 | Larval marking phenotypes of *Bombyx mori* mutants and *B. mandarina*.

(**a**) Multiple phenotypes of fifth instar larvae for *p* locus alleles (except for genetic mosaic strain p^{Sm}) of the silkworm *B. mori*. The p^{Sm} strain was originally derived from the p^{S} allele by X-ray irradiation. Arrows indicate eye spots (upper), crescents (middle) and star spots (bottom). (**b**) Three kinds of spots reflecting the *normal pattern* ($+^{p}$) and the corresponding region of these markings in *plain* (*p*), *Moricaud* (p^{M}), *Striped* (p^{S}) and *Black* (p^{B}). Note that with the *p* allele, disk-like structures of *normal marking* develop but lack black pigmentation. (**c**) Branch mimicking phenotype of the wild silkworm *B. mandarina* with noticeable eye-spot markings on its second thoracic segment.

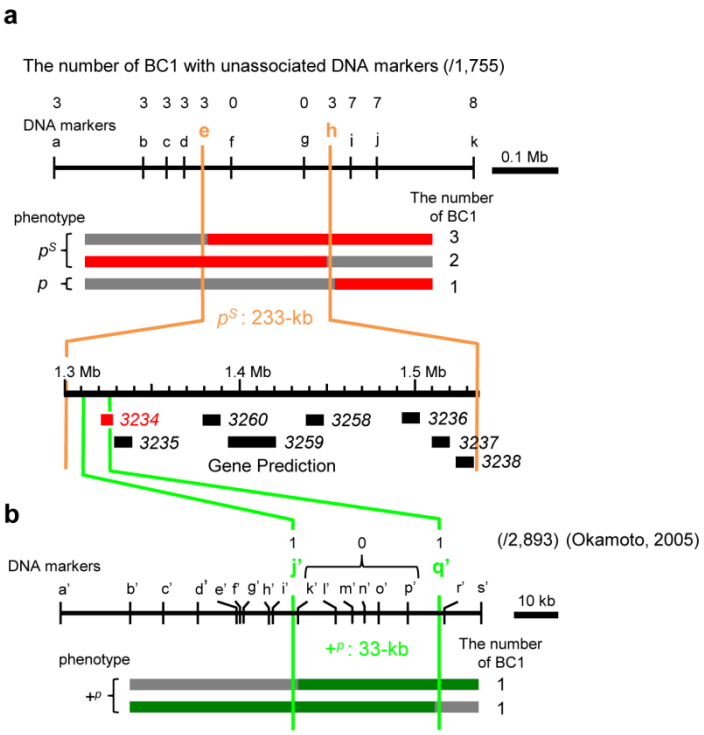


Figure 2 | *p* locus positional cloning.

(**a**) Fine mapping of the p^S region and genotypes of informative BC1 selected from 1,755 individuals. Red and gray bars indicate genotypes of heterozygote (p^S/p) and homozygote (p/p), respectively. The 233-kb region responsible for p^S is denoted by: "p^S: 233-kb". The numbers above SNP markers indicate recombination events. Gene prediction was derived from the *B. mori* china gene model. (**b**) Fine mapping of the $+^p$ region and genotypes of informative BC1 selected from 2,893 individuals (Okamoto, 2005). Green and gray bars indicate genotypes of heterozygote ($+^p/p$) and homozygote (p/p), respectively. The 33-kb region responsible for $+^p$ is denoted by: "$+^p$: 33-kb". The numbers above SNP markers indicate recombination events.

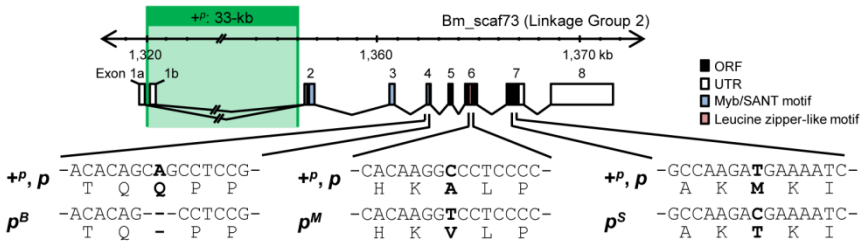


Figure 3 | Structural features of *apt-like* (*3234/5* gene).

Genomic structure of *apt-like* is shown. Open reading frame (black), untranslated region (white), Myb/SANT motif (blue) and leucine zipper-like motif (red).

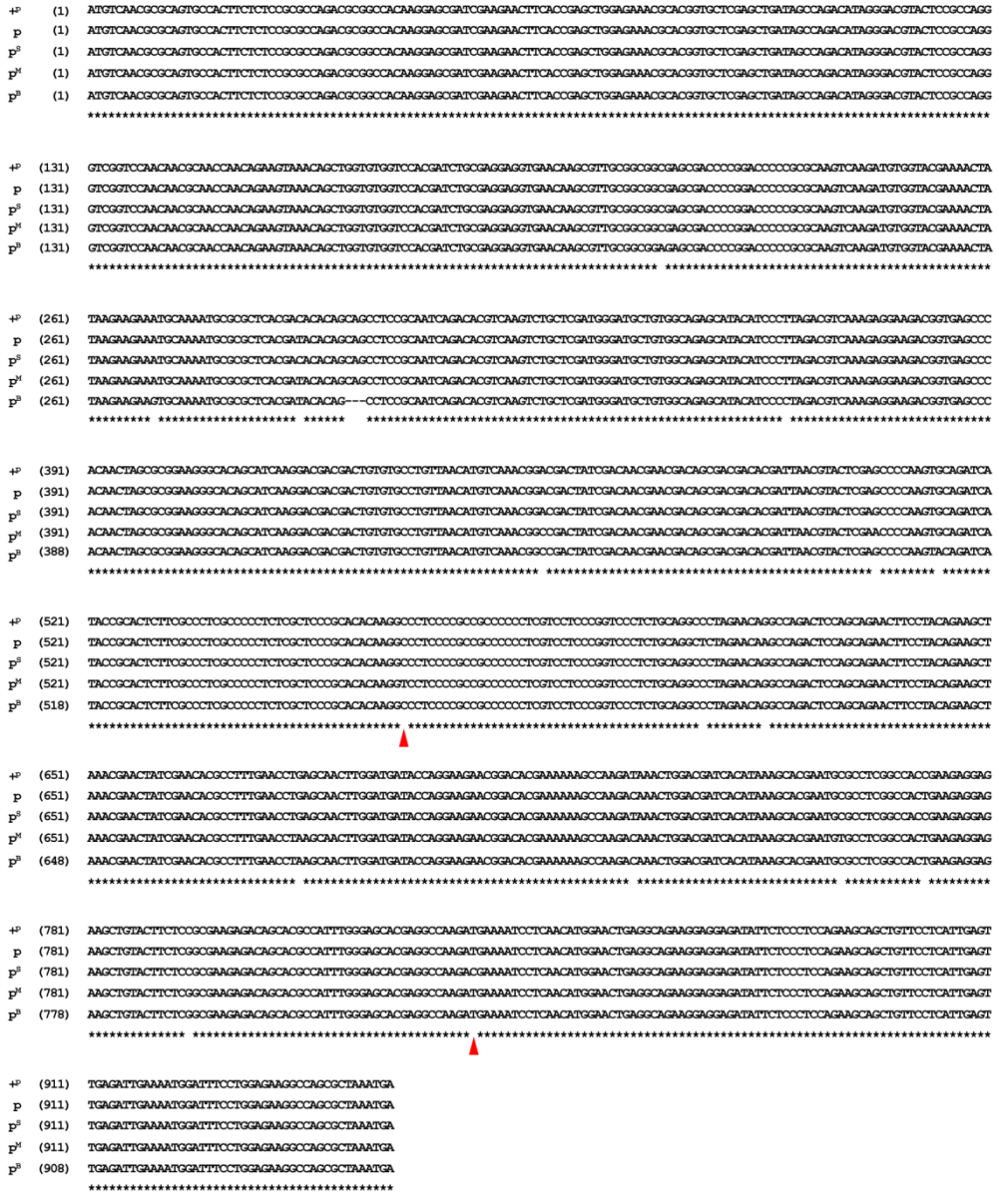


Figure 4 | Nucleotide sequence alignments of *apt-like*.

DNA sequence alignments for $+^P$, p, p^S, p^M and p^B alleles. Asterisks and dashes indicate identical nucleotides and nucleotide deletions, respectively. Red triangles indicate non-synonymous substitutions.

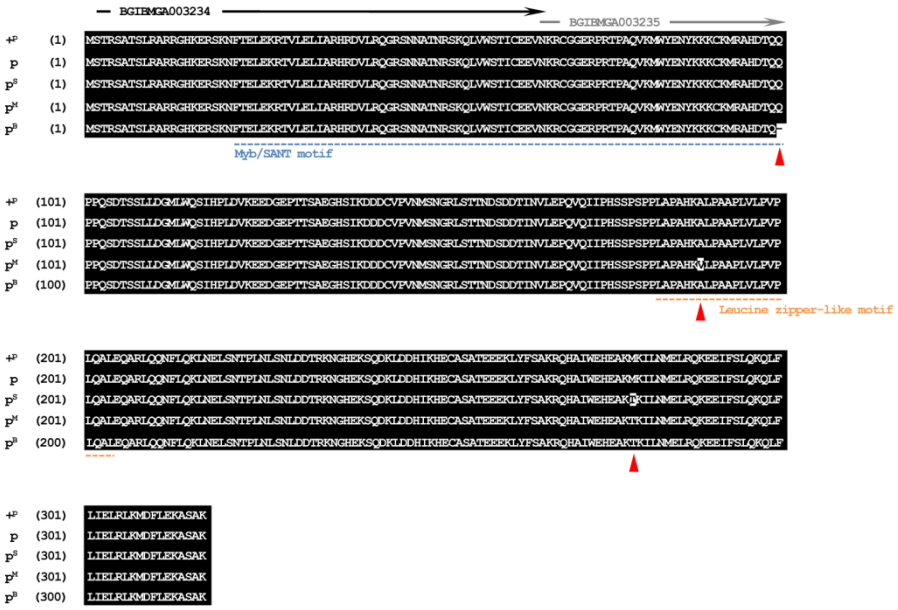


Figure 5 | Amino acid sequence alignments of Apt-like.

Amino acid sequence alignments for $+^{P}$, p, p^{S}, p^{M} and p^{B} alleles. Black shading with white letters indicates identical amino acid residues. Myb/SANT motifs and leucine zipper-like motifs are underlined. Reported BGIBMGA003234 and BGIBMGA003235 sequences are shown by arrows at the top of the alignment. Red triangles indicate amino acid changes among five alleles.

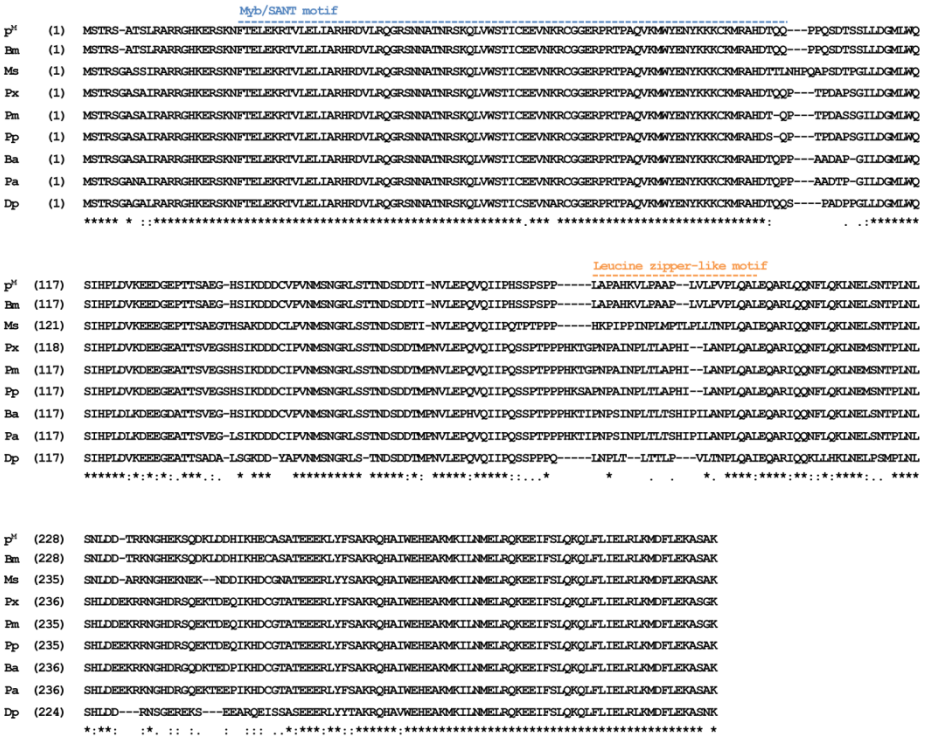


Figure 6 | Amino acid sequence alignments of Apt-like among lepidopteran species.

Amino acid sequences of Apt-like in the p^M strain of *B. mori* (p^M), *B. mandarina* (Bm), *M. sexta* (Ms), *P. xuthus* (Px), *P. machaon* (Pm), *P. polytes* (Pp), *B. alcinous* (Ba), *P. aristolochiae* (Pa) and *D. plexippus* (Dp) were aligned. Identical residues in all species are marked with asterisk below the alignment; weakly conserved residues in all species are marked by a single period; strongly conserved residues in all species are marked by double period. Myb/SANT motifs and leucine zipper-like motifs in *B. mori* are underlined.

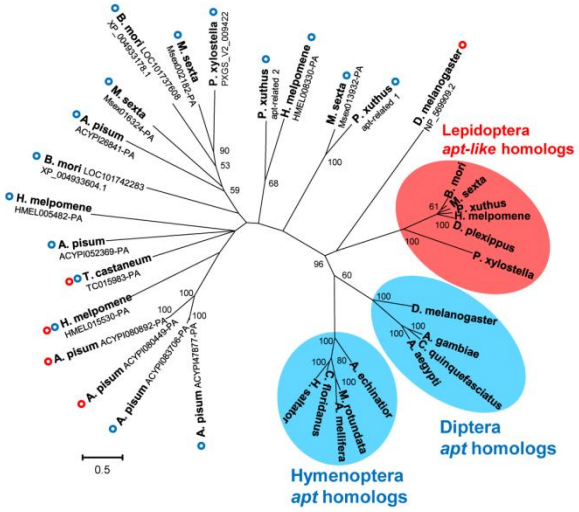


Figure 7 | Phylogenetic analysis of *apt-like* (*3234/5* gene).

Unrooted maximum likelihood tree for *apt-like* and *apt* based on their nucleotide sequences. The numbers at the tree edges represent bootstrap values (1,000 replicates). Only bootstrap value > 50% are shown. Branch lengths are scaled to the number of substitutions per site. Red and blue circles indicate Blastp sequence hits (*E* value < 10^{-4}), obtained with Apt-like or Apt as a query, respectively. These sequences include a Myb/SANT motif, but lack a leucine zipper-like motif.

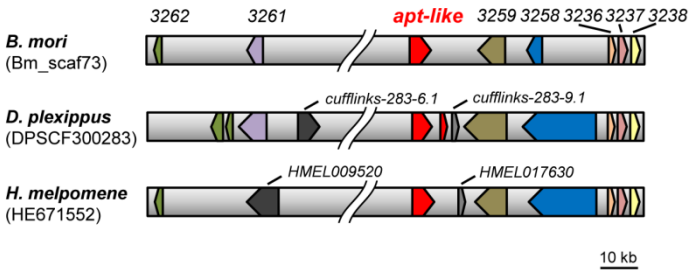


Figure 8 | Gene synteny in the region surrounding *apt-like*.

The genomes of *Bombyx mori* (top), *Danaus plexippus* (middle) and *Heliconius melpomene* (bottom) are shown based on MonarchBase (http://monarchbase.umassmed.edu/). Genes are depicted by colored polygons and transcriptional orientation is indicated at the angled end of each gene. Similar colors indicate homologous sequences among these species. Species-specific genes (*cufflinks-283-6.1* and *HMEL009520*) are indicated by dark gray polygons. Light gray polygons (*cufflinks-283-9.1* and *HMEL017630*) indicate orthologous genes of *B. mori BGIBMGA001704* and *BGIBMGA001706*, respectively.

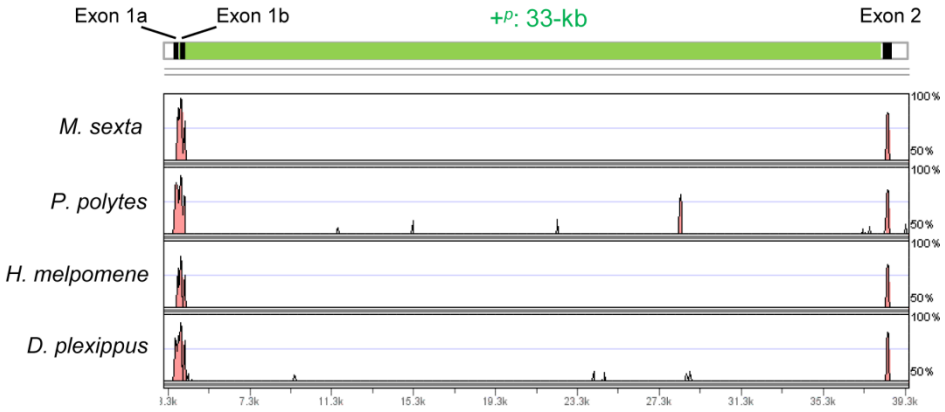


Figure 9 | +P: 33-kb LAGAN alignments visualized with mVISTA.

Sequences surrounding +P:33-kb were used as a reference, but clear sequence similarities were not found. Red peaks indicate conserved nucleotide sequences. Crossbar indicates the genomic structure of the *apt-like* locus (black and green indicate exons and +P: 33-kb, respectively).

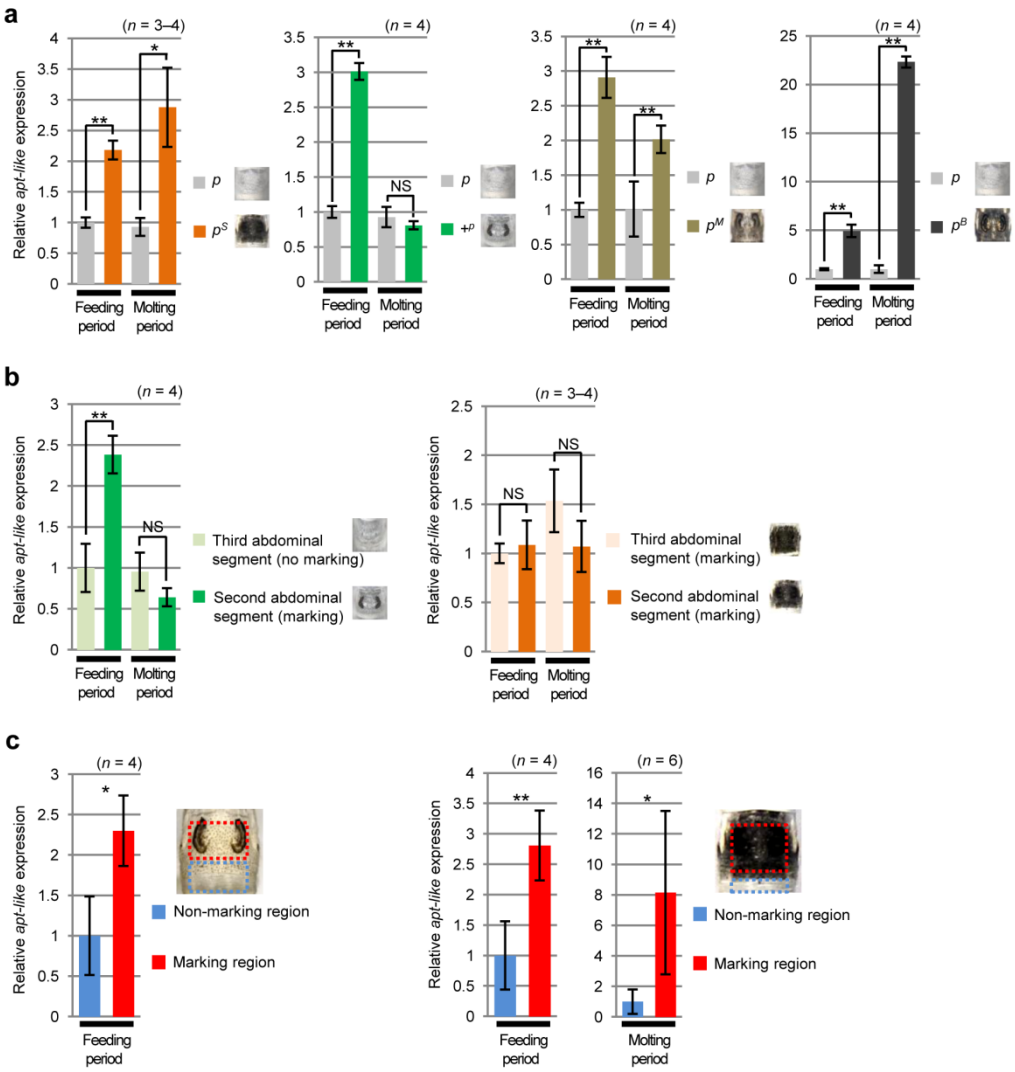


Figure 10 | Pigmentation associated with *apt-like* expression.

(**a**) Relative expression profiles of *apt-like* among p^S, $+^p$, p^M and p^B alleles as determined by quantitative RT-PCR analysis and plotted as means ± s.d. (n = 3–4). cDNAs were prepared from the larval epidermis of the second abdominal segment during the fourth instar stage. Feeding period (stage C2) and molting period (stage E1) were determined based on the spiracle index. *$P < 0.05$, **$P < 0.01$, Student's *t*-test. (NS, not significant). (**b**) Relative expression profiles of *apt-like* in $+^p$ (left) and p^S (right) alleles as determined by quantitative RT-PCR analysis and plotted as means ± s.d. (n = 3–4). cDNAs were prepared from the larval epidermis of the second and third abdominal segments during the fourth instar stage. C2 and E1 stages were determined based on the spiracle index. **$P < 0.01$, Student's *t*-test (NS, not significant). (**c**) Differential expression of *apt-like* in the abdominal second segment with the $+^p$ (left) and p^S (right) alleles. In both $+^p$ and p^S, non-marking region (blue dashed box) and marking region (red dashed box) were separately dissected from the same individual and the *apt-like* mRNA levels were determined by quantitative RT-PCR (n = 4–6). *$P < 0.05$, **$P < 0.01$, paired Student's *t*-test.

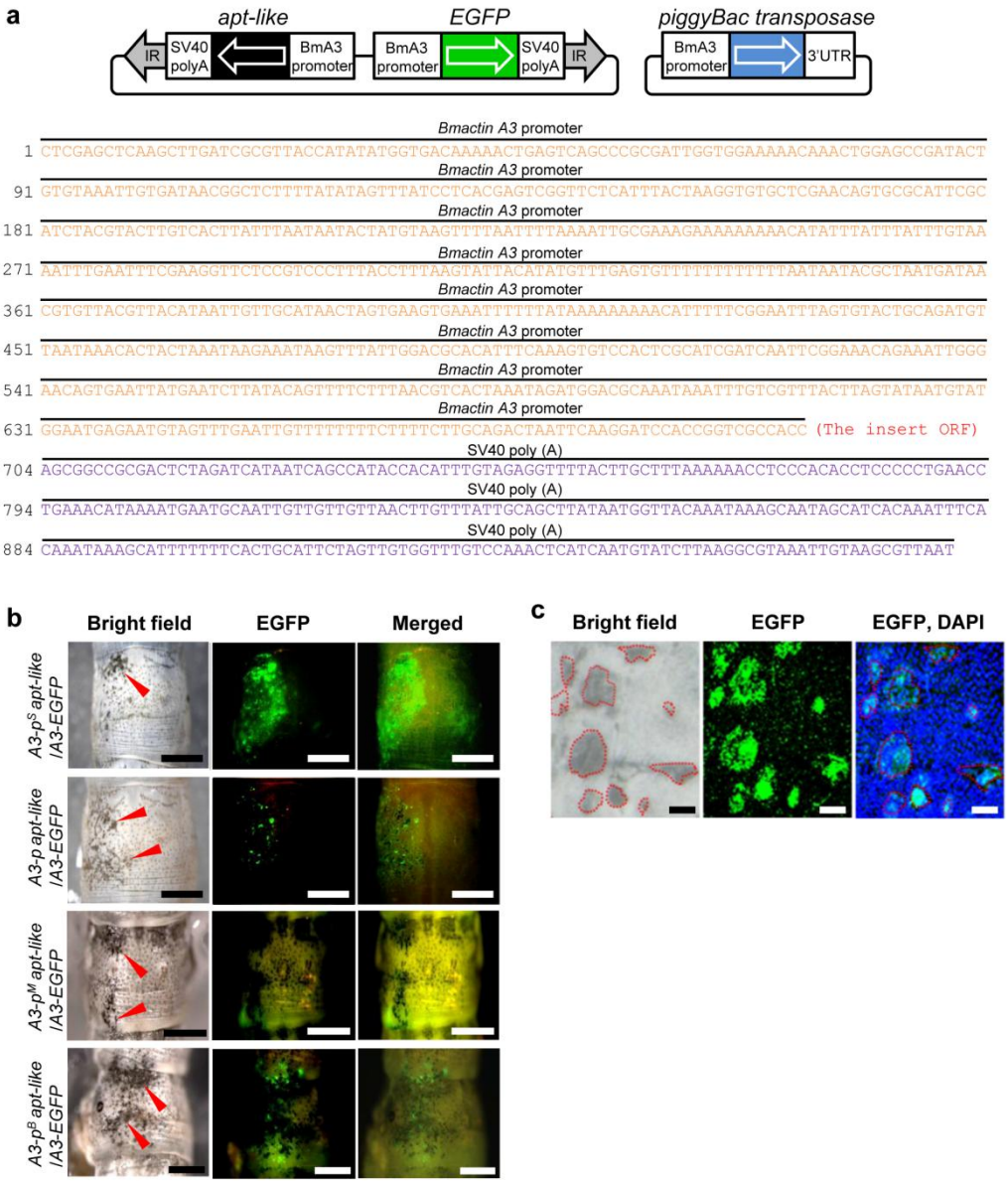


Figure 11 | Transgenic ectopic expression of *apt-like* induces black pigmentation on larval cuticles.

(**a**) Schematic representation of the *piggyBac* construct. Colored boxes show inserted genes. Detailed sequences from *Bmactin A3* promoter to SV40 poly (A) are shown below. (**b**) Transgenic expression in the *p* alleles by electroporation at the fifth instar stage. Upper and lower panels show the results of transgenic expression of *apt-like* isolated from p^S, *p*, p^M *and* p^B strains, respectively. Red arrowheads indicate ectopic black pigmentation. Scale bars, 0.5 cm. (**c**) Enlarged areas of ectopic pigmentation (middle panel shows EGFP expression). Epidermal cell nuclei were stained with DAPI (right panel) and pigmented areas (left panel) are enclosed by red lines. Scale bars, 100 μm.

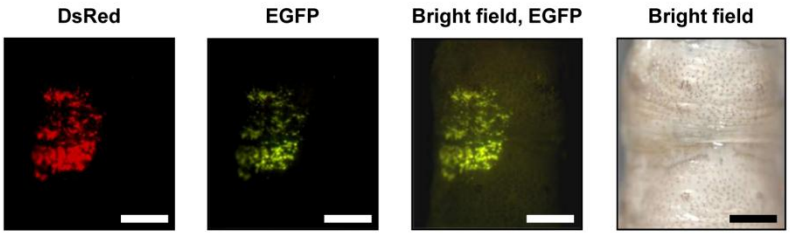


Figure 12 | Negative control experiments for *in vivo* electroporation.

The *DsRed2* gene with the *Bmactin A3* promoter induced no pigmentation. Scale bars, 0.5 cm.

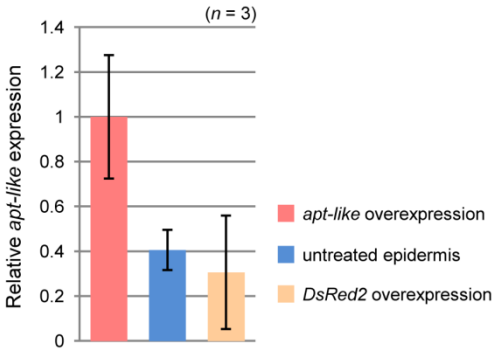


Figure 13 | Quantitative RT-PCR confirmation of *apt-like* overexpression.

Relative *apt-like* expression levels in the *p* epidermis with *apt-like* overexpression (red bar), untreated (blue bar) and *DsRed2* overexpression (yellow bar) as determined by quantitative RT-PCR analysis and plotted as mean ± s.d. ($n = 3$).

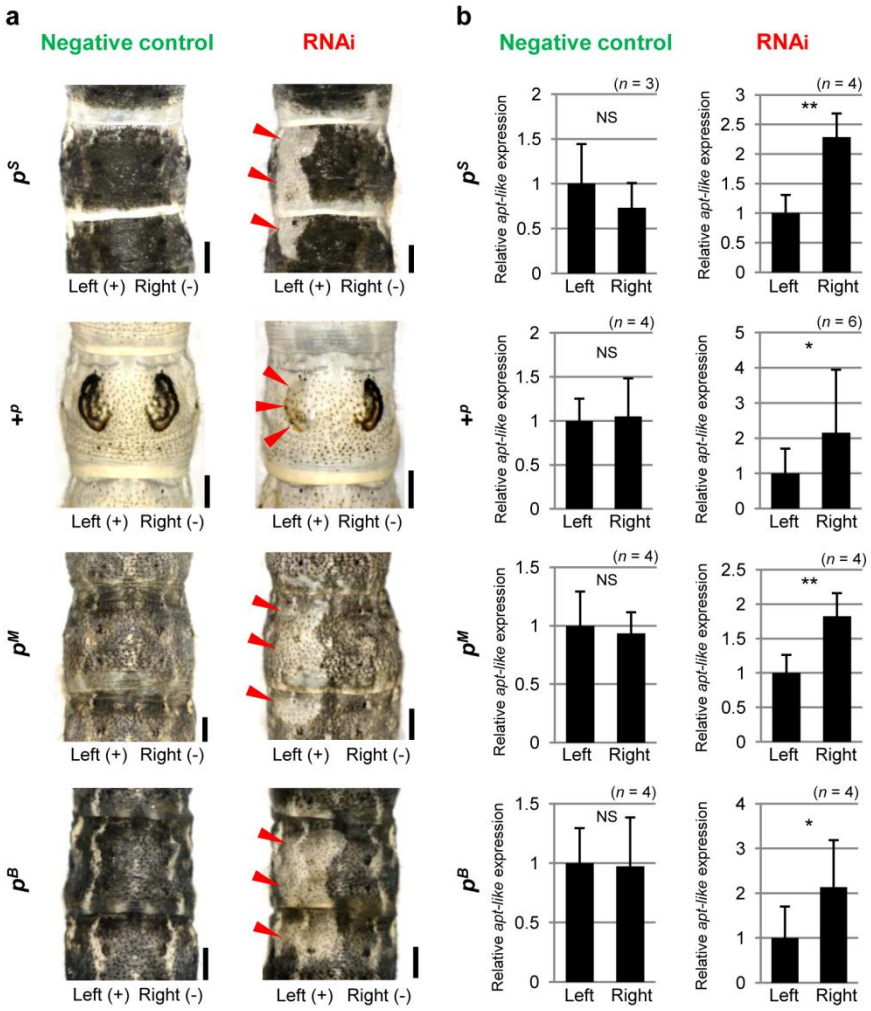


Figure 14 | Effects of *apt-like* siRNA injection on pigmentation formation for different *p* alleles.

(**a**) RNAi effects for the p^S, $+^p$, p^M and p^B alleles. siRNA was introduced by microinjection followed by electroporation. "+" and "-" indicate the electrical current direction during electroporation. For all four different *p* alleles, reduced pigmentation were observed only on the "+" sides by siRNA that targeted *apt-like* (right panel), whereas no effects were observed on pigmentation by negative control siRNA (left panel). Scale bars, 1 mm. (**b**) Relative *apt-like* expression levels in a negative control (left row) and with RNAi (right row) as determined by quantitative RT-PCR analysis and plotted as means ± s.d. Numbers of samples are shown at the upper right in each graph. $^*P < 0.05$, $^{**}P < 0.01$, paired Student's *t*-test. (NS, not significant).

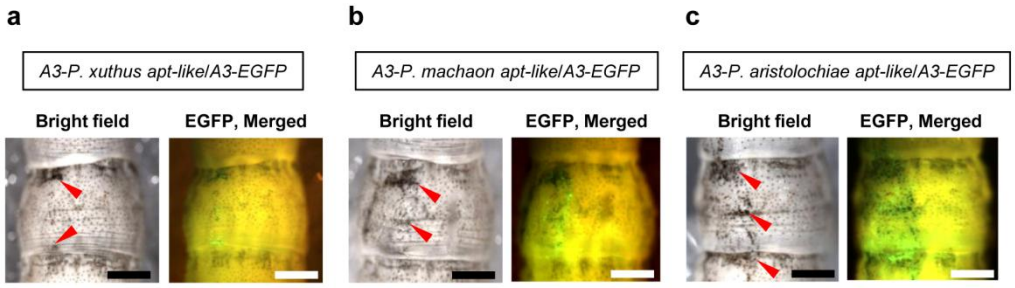


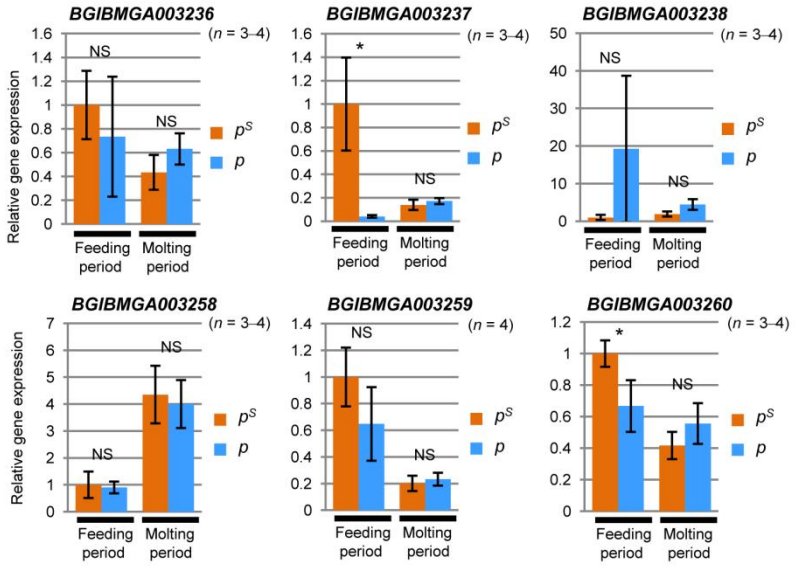
Figure 15 | Functional analyses of Papilionidae species *apt-like* homologs using electroporation mediated somatic transgenesis.

(**a**–**c**) Transgenic expression in the *p* allele by electroporation at the fourth instar stage. Left, middle and right panels show the results of transgenic expression of *apt-like* homologs isolated from *P. xuthus* (**a**), *P. machaon* (**b**) and *P. aristolochiae* (**c**), respectively. Red arrowheads indicate ectopic black pigmentation. Scale bars, 1 mm.

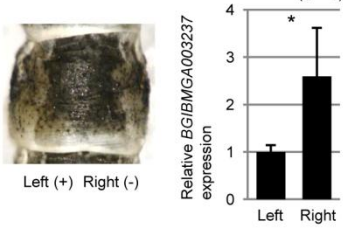
a

Gene	Annotation	Accession number	E-value
BGIBMGA003236	Troponin C-akin-1 protein (*Anopheles darlingi*)	ETN65107.1	$4e^{-30}$
BGIBMGA003237	G patch domain containing 4 (*Anopheles darlingi*)	ETN58705.1	$8e^{-34}$
BGIBMGA003238	NADH dehydrogenase 1 alpha subcomplex subunit 12 (*Anopheles darlingi*)	ETN57796.1	$2e^{-66}$
BGIBMGA003258	pyruvate dehydrogenase kinase (*Bombyx mori*)	NP_001108115.1	0.0
BGIBMGA003259	methionine-tRNA synthetase (*Aedes aegypti*)	XP_001649418.1	0.0
BGIBMGA003260	Zinc finger domain containing protein (*Haemonchus contortus*)	CDJ88226.1	$4e^{-13}$

b



c



d

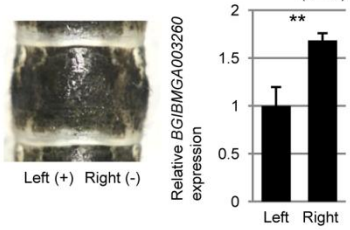


Figure 16 | Characterization of six candidate genes in p^S: 233-kb.

(**a**) List of six candidate genes in p^S: 233-kb and match probabilities resulting from a BLAST X search. (**b**) Relative expression profiles of six candidate genes between p^S and p alleles as determined by quantitative RT-PCR analysis and plotted as means ± s.d. (n = 3–4). cDNAs were prepared from the larval epidermis of the second abdominal segment during the fourth instar stage. Feeding period (stage C2) and molting period (stage E1) were determined based on the spiracle index. *$P < 0.05$, Student's t-test. (NS, not significant) (**c**, **d**) Knockdown of *BGIBMGA003237* and *BGIBMGA003260* in p^S epidermis by siRNA. No effects were observed in pigmentation using siRNA that targeted *BGIBMGA003237* (**c**) and *BGIBMGA003260* (**d**). Reductions in the mRNA levels of *BGIGMGA3237* (**c**) and *BGIBMGA003260* (**d**) were estimated by quantitative RT-PCR analysis and plotted as means ± s.d. (**c**, n = 4; **d**, n = 3), *$P < 0.05$, **$P < 0.01$, one-tailed paired t-test.

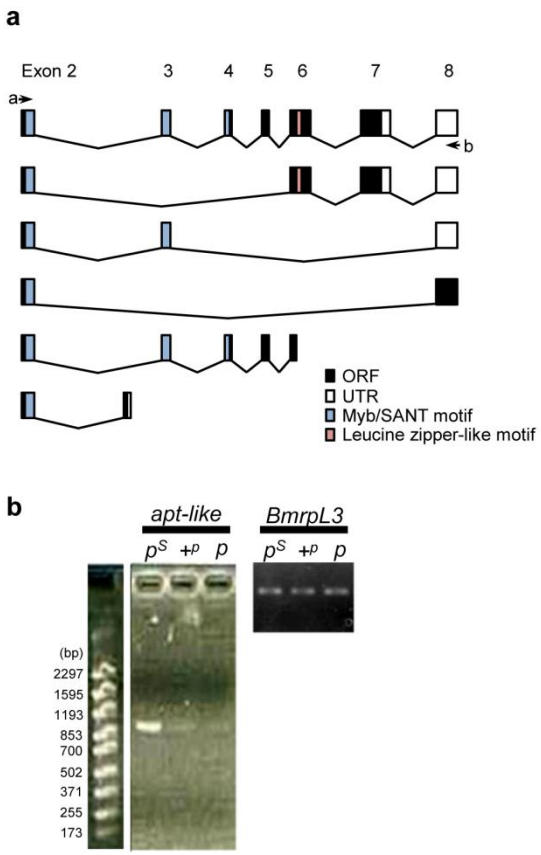


Figure 17 | *apt-like* isoform expression.

(**a**) Schematic structures of *apt-like* mRNA isoforms identified from EST databases, RT-PCR and RACE analyses. “a” and “b” indicate a primer set used for RT-PCR. Open reading frame (black), untranslated region (white), Myb/SANT motif (blue) and leucine zipper-like motif (red). (**b**) *apt-like* expression determined by RT-PCR. Larval epidermis of the second abdominal segment during the fourth instar (E1 period) was used as a template. The *Bombyx mori ribosomal protein L3* (*BmrpL3*) gene expressed constitutively in these cells was used as a control.

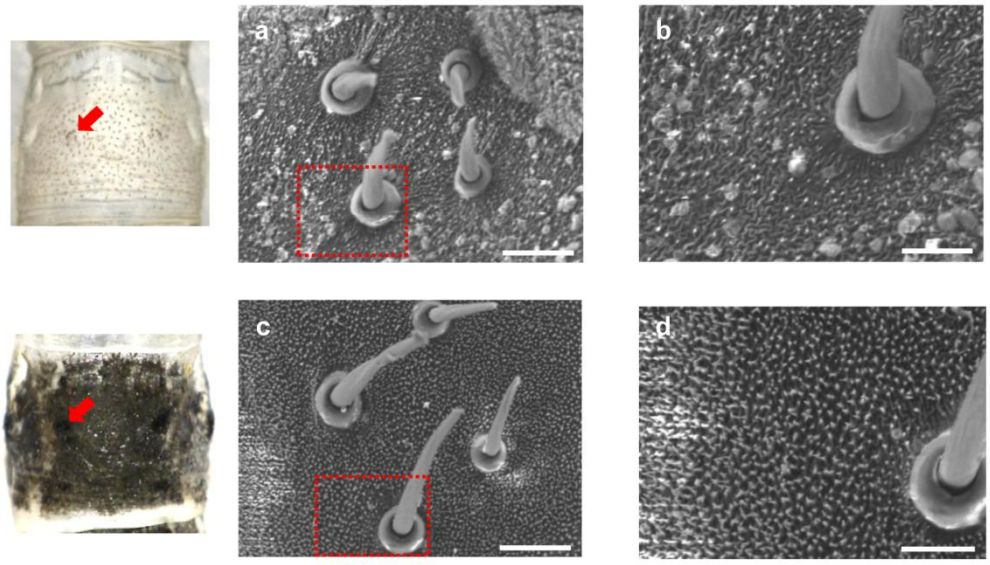


Figure 18 | The ultrastructural differences of the epicuticle between *p* and p^S allele.
Cuticular nanostructures (scanning electron microscopy image) in the fifth instar abodominal third segment in *p* (**a**, **b**) and p^S (**c**, **d**) are shown. Characteristic four bristles (indicated as red arrow) were used as a landmark for the morphological observation. Magnified view of red dashed box (**a**, **c**) is shown to the right (**b**, **d**). Scale bars, 50 μm (**a**, **c**), 25 μm (**b**, **d**).

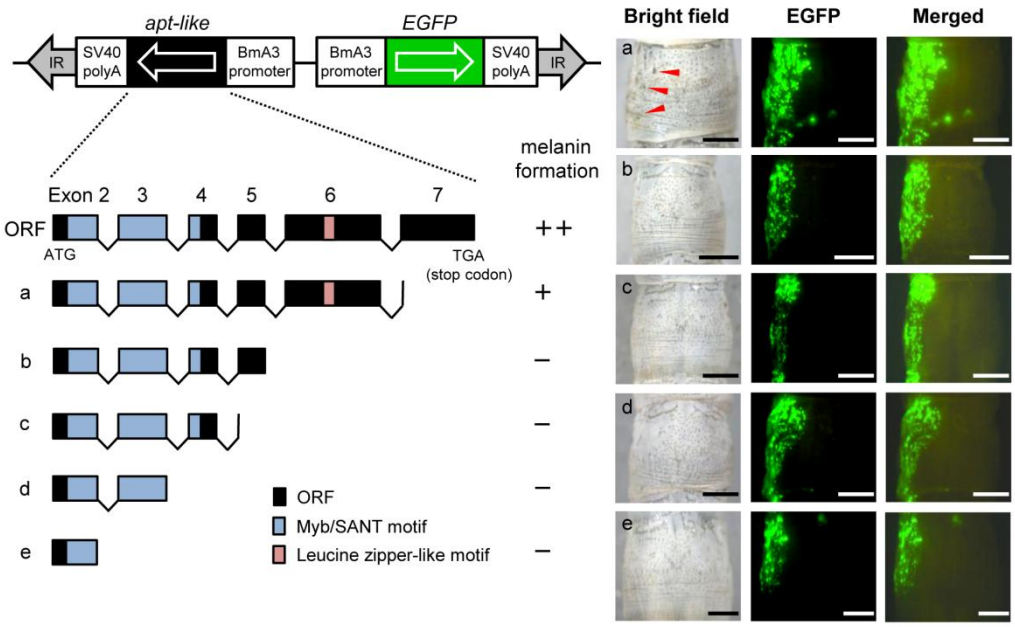


Figure 19 | Characterization of the *apt-like* functional domain.

Schematic of deletion *apt-like* constructs ("a" to "e") and transgenic expression in *p* alleles by *in vivo* electroporation at the fifth instar stage. Dark black pigmentation (++), light black pigmentation (+) and no pigmentation (−). Red arrowheads indicate light black pigmentation. Scale bars, 0.5 cm.

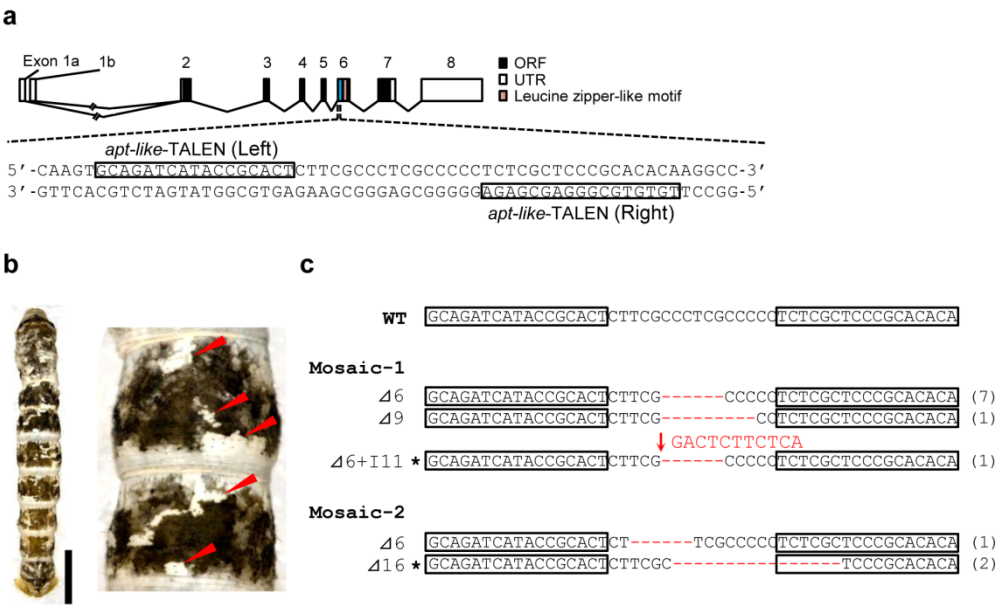


Figure 20 | Disruption of *apt-like* using designed transcriptional activator-like effecter nucleases (TALEN).
(**a**) Genomic structure of *apt-like* is shown. Open reading frame (black), untranslated region (white) and leucine zipper-like motif (red). The blue box of exon 6 is expanded to show the gene sequence that includes the 17 bp TALEN binding sites (shown in boxes). (**b**) Larval mosaic phenotype in G_0 fifth instar larva of p^S targeted for *apt-like*. Red arrowheads indicate remarkable white patches. Photos indicate dorsal overview and magnified view (second and third abdominal segment) in a p^S mosaic larva, respectively. Scale bar, 1 cm. (**c**) Sequences of induced mutations in *apt-like* for two mosaic mutants. The wild-type (WT) sequence is shown above the mutant sequences that contain deletions (indicated by red dashes) and insertions (shown by a red arrow and letters). Boxed sequences indicate TALEN binding sites. Asterisks indicate frameshift mutations. The number of clones having identical mutations is indicated to the right of each sequence.

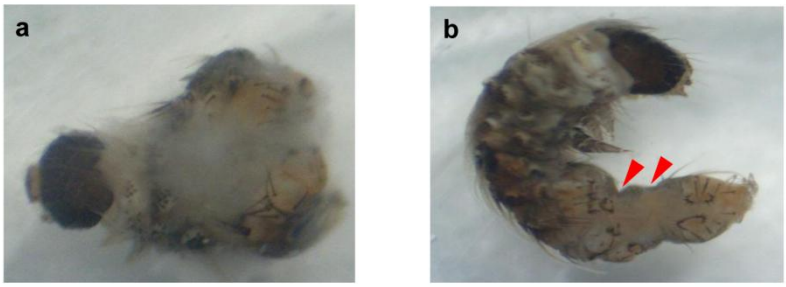


Figure 21 | Phenotypes of embryonic RNAi targeting *apt-like*.

(**a**) A dorsal view of the *apt-like* siRNA injected embryo. Dorsal closure was defective. (**b**) A lateroventral view of the *apt-like* siRNA injected embryo. Red arrowheads denote a twisted abdomen phenotype.

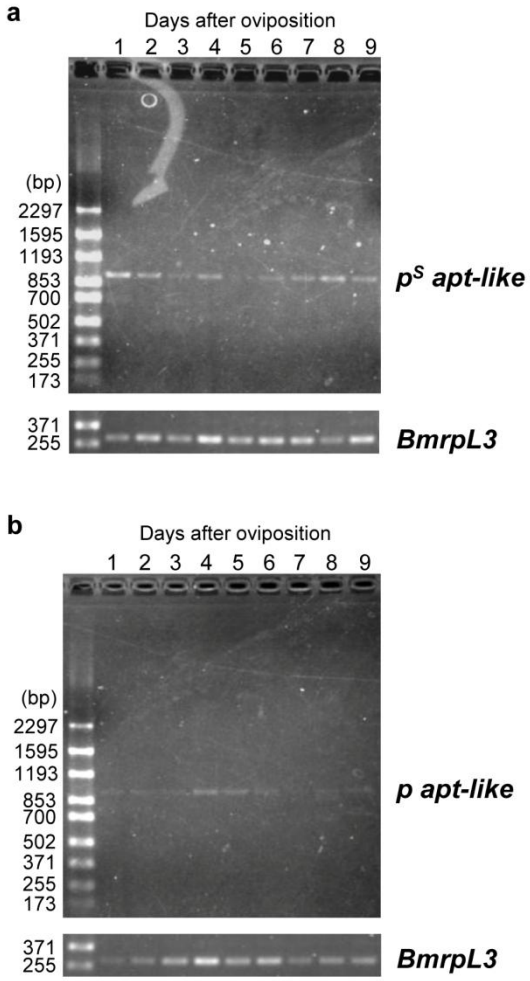


Figure 22 | Temporal embronic *apt-like* expression in non-diapause egg (p^S and *p* strain).
(**a**, **b**) Temporal embryonic expression of *apt-like* in p^S (**a**) and *p* strain (**b**). Each number of lanes indicates days after ovipositon. Amplification cycles were 32 for *apt-like* and 25 for the internal control gene *BmrpL3*.

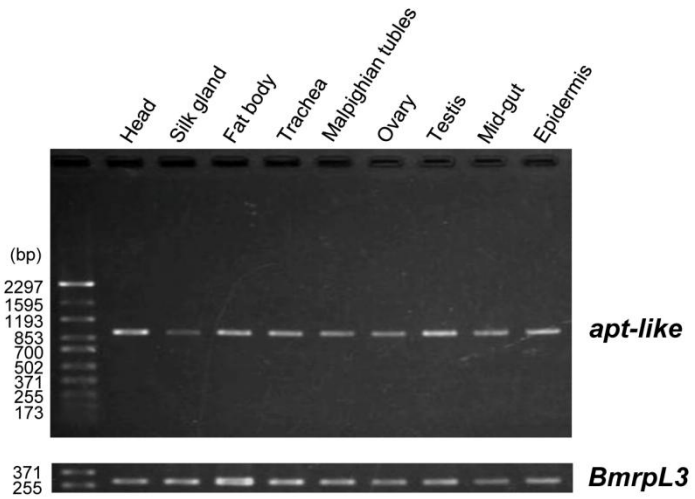


Figure 23 | Tissue expression profile of *apt-like* in the $+^{p}$ allele.

Expression profile of the full-length *apt-like* mRNA in various tissues of $+^{p}$ allele. RT-PCR templates were prepared for head, silk gland, fat body, trachea, malpighian tubules, ovary, testis, mid-gut and epidermis (abdominal second segment). Amplification cycles were 35 for *apt-like* and 25 for the internal control gene *BmrpL3*.

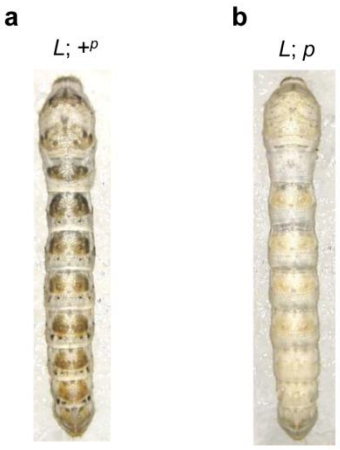


Figure 24 | Larval marking phenotypes of *Bombyx mori* mutant *L*.

(**a**, **b**) Phenotypes of fifth instar larvae for the mutant *L* of the silkworm *B. mori*. Twin black spots appear only when the *L* individuals are crossed with $+^{p}$ allele carriers (**a**). These characteristic markings do not appear in *L* and *p* allele crosses (**b**).

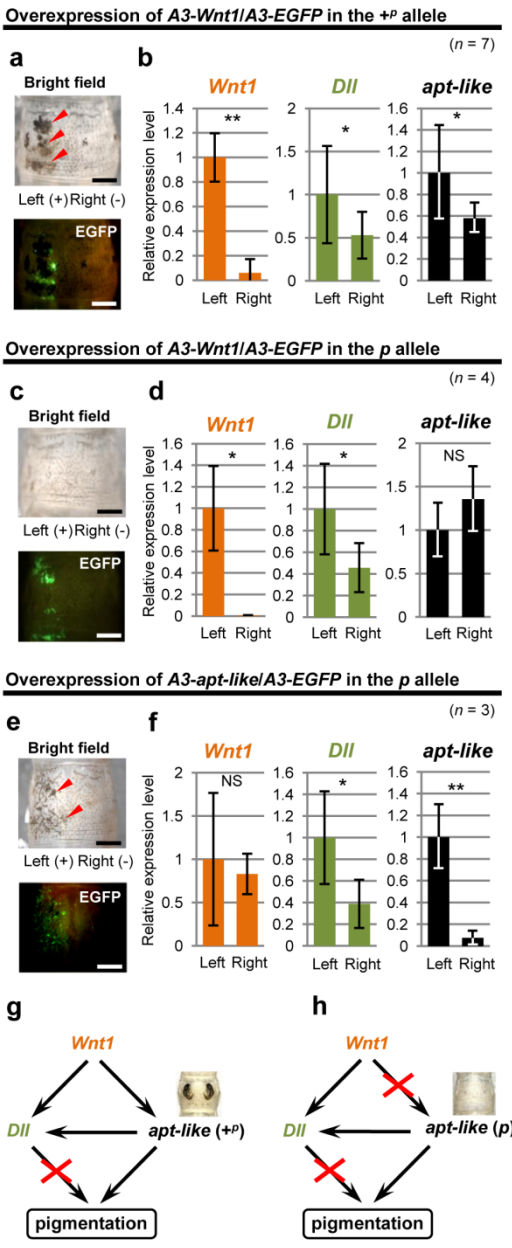


Figure 25 | Putative gene regulatory network among *Wnt1*, *Dll* and *apt-like*.

(**a**, **c**) *Wnt1* transgenic expression by electroporation in the $+^p$ allele (**a**) and *p* allele (**c**) for a fifth instar larva, respectively. Red arrowheads indicate ectopic pigmentation. "+" and "–" indicate the electrical current direction during electroporation. Ectopic pigmentation was observed with the $+^p$ allele (**a**), but not with the *p* allele (**c**). Scale bars, 0.5 cm. (**b**, **d**) Relative expression levels of *Wnt1*, *Dll* and *apt-like* with the $+^p$ allele (**b**) and *p* allele (**d**), respectively. Error bars indicate s.d. (b, $n = 7$; d, $n = 4$). $^*P < 0.05$, $^{**}P < 0.01$, one-tailed paired *t*-test. (NS, not significant). cDNAs were prepared separately from the left half (including *Wnt1*-positive cells) and from the right half (without *Wnt1*-positive cells as a control) of the epidermis of the abdominal third segment during the fourth instar larval stage (stage C2). (**e**) *apt-like* transgenic expression in the *p* allele by electroporation for a fifth instar larva. Scale bars, 0.5 cm. (**f**) Relative expression levels of *Wnt1*, *Dll* and *apt-like* with the *p* allele. cDNAs were prepared separately from the left half (including *apt-like*-positive cells) and from the right half (without *apt-like*-positive cells as a control) of the epidermis of the abdominal third segment during the fourth instar larval stage (stage C2). Error bars indicate s.d. ($n = 3$). $^*P < 0.05$, $^{**}P < 0.01$, one-tailed paired *t*-test. (NS, not significant). (**g**, **h**) Schematic models of the gene regulatory network among *Wnt1*, *Dll* and *apt-like*. (**g**) With the $+^p$ allele, *Wnt1* has the potential to induce both *Dll* and *apt-like* expressions. *apt-like* also has the potential to induce *Dll* expression. Black pigmentation is induced by *apt-like*, but not by *Dll*. (**h**) With the *p* allele, *apt-like* originally has the potential to induce black pigmentation. However, because *apt-like* does not have the potential to respond to *Wnt1*, and consequently melanin pigmentation is not induced.

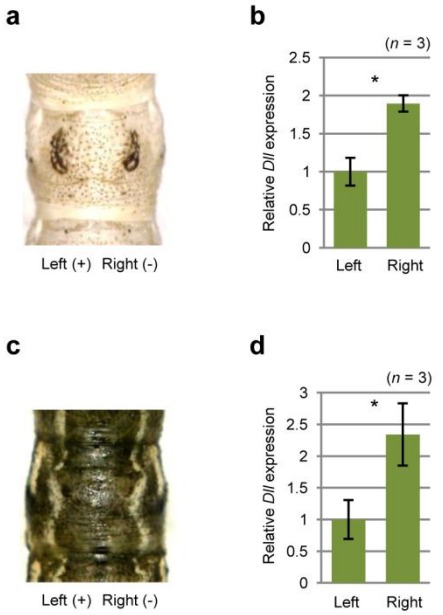


Figure 26 | Knockdown of *Dll* in the $+^{p}$ and p^{B} epidermis by siRNA.

(**a**, **c**) No effects were observed in black pigmentation using siRNA targeted *Dll* in both the $+^{p}$ (**a**) and p^{B} (**c**) alleles. siRNA was introduced by microinjection followed by electroporation. "+" and "-" indicate the electrical current direction during electroporation. (**b**, **d**) Reductions in the mRNA levels of *Dll* in the $+^{p}$ (**b**) and p^{B} epidermis (**d**) were estimated by quantitative RT-PCR analysis and plotted as means ± s.d. ($n = 3$). *$P < 0.05$, one-tailed paired *t*-test.

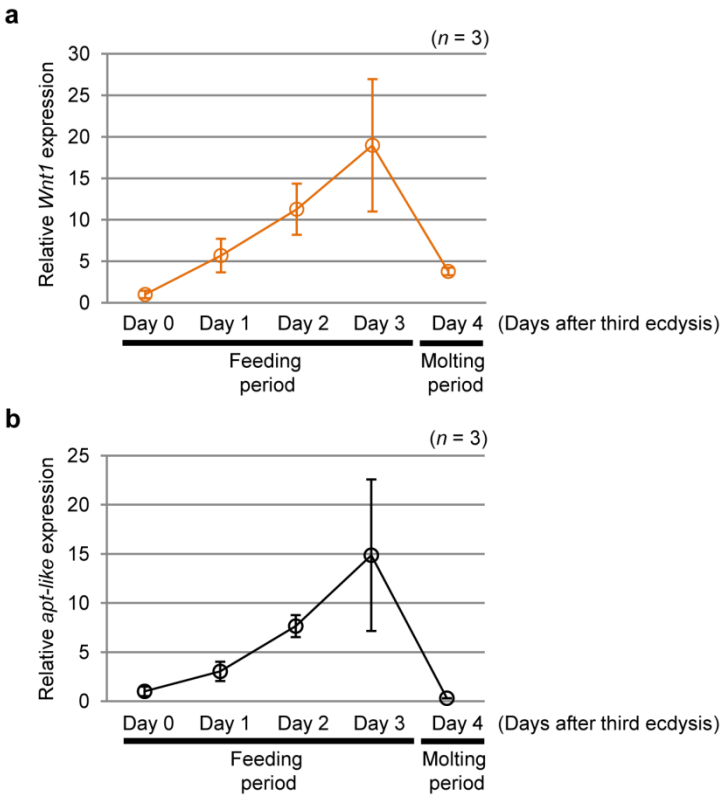


Figure 27 | Temporal expression pattern of *Wnt1* and *apt-like* in the mutant *L*.

(**a**, **b**) Temporal expression profiles of *Wnt1* (**a**) and *apt-like* (**b**) in the mutant *L* (*L*; $+^{p}$) as determined by quantitative RT-PCR analysis and plotted as means ± s.d. ($n = 3$). cDNAs were kindly provided by Dr. J. Yamaguchi (University of Tokyo). These samples were prepared from the larval epidermis of the second abdominal segment during the fourth instar stage. Note that temporal expression pattern of *apt-like* (**b**) is consistent with that of *Wnt1* (**a**).

Table 1 | Allelic information for the *p* locus from "Tajima, Y. *The Genetics of the Silkworm* (Logos Press and Prentice Hall, 1964)" with some modifications.

Gene symbol	Allele name	Location*	References**
p	*plain*	2-3.0	Toyama (1906)
p^{l}	*lightest normal*	2-3.0	Tanaka (1916)
p^{2}	*light normal*	2-3.0	Tanaka (1916)
$+^{p}$ (p^{3})	*Normal marking*	2-3.0	Toyama (1906)
p^{4}	*dark normal*	2-3.0	Tanaka (1916)
p^{B}	*Black*	2-3.0	Kawaguchi (1933)
p^{D}	*Dorsal spot*	2-3.0	Shimodaira (1947)
p^{G}	*Ventral striped*	2-3.0	Takasaki (1947)
p^{L}	*Light crescent*	2-3.0	Chikushi (1938)
p^{M}	*Moricaud*	2-3.0	Tanaka (1913)
p^{S}	*Striped*	2-3.0	Toyama (1909)
p^{Sa}	*Sable*	2-3.0	Tazima (1938)
$p^{Sa\text{-}2}$	*Sable-2*	2-3.0	Tazima (1943)
p^{St}	*Pale striped*	2-3.0	Takasaki (1947)
p^{Sw}	*Whitish striped*	2-3.0	Tanaka (1951)

* Although the *p* locus was originally mapped at position 0.0 cM on chromosome 2 (Linkage Group 2), Banno *et al*. added modifications to this and defined the *p* locus that mapped at 3.0 cM based on additional linkage analysis [Banno *et al*., Summary of The Japanese Society of Sericultural Science Conference, **67,** 93 (1997) (in Japanese).; Silkworm Base, http://www.shigen.nig.ac.jp/silkwormbase/index.jsp; Banno *et al*., (2005)].

Allelic relationships of genes listed in Table 1 were confirmed by allelic tests using authorised genes for the *p* locus, whenever a gene that was expected to be located at the *p* locus was discovered. For example, for p^{B}, the p^{S} gene was used as an authorised gene for the *p* locus and had shown no crossing-over between p^{B} (Kawaguchi, 1933).

** References in Table 1 are listed in the Appendix 1.

Table 2 | Primers for the SNP markers used for constructing the p^S linkage map.

Number	Primer Name	Sequence (5' → 3')	Primer Name	Sequence (5' → 3')
A	02-004F	TGACCCATTATTAATTACGTGAAGA	02-004R	CATGGGTAGAATCGTGTTCATT
B	chr2-21F	ATCAAGGAGCAGCCGGACAAGT	chr2-21R	CATCAGGGTACCGGACTGCAT
C	02-010F	TACGATTCGGAAAATGAGCTG	02-010R	CATGACCATCTGCTCGGAC
D	02-051F	AGAAAATACTGTTACTCGCTAGG	02-051R	CAAACCTAAAATGCAAATGG
E	chr2-7F	AGCTCACGGTATTCATCAGCGGA	chr2-7R	GTTCGAGGTGTCGTTGTGGAC
F	chr2-11F	TTCAGTCTCTGCCAGTGATCTTC	chr2-11R	GTGGCGAAACCCGAAGCCAT
a	chr2-23F	ATTTGGCCTTCATCTCGAGCGA	chr2-23R	CTTTACTGCTACCAGGTGCTTGGT
b	chr2-4F	TCAGTAGCTTGGTGCGTTCCAACTTG	chr2-4R	TGCGGCGATGAGATATGCGT
c	chr2-44F2	AGCTGCGAATGTGAGCCTTAAGCT	chr2-44R2	CTGCTTCTTCGTTTGGGTTTCGAT
d	chr2-47F	CACACTATGCCATTTGCTGGTTTG	chr2-47R	GGCTAGTCAGGGACCTTCGAT
e	chr2-H5F	TCGTTAAAGCACGAATGACTCACGACA	chr2-H5R	AGTGTAGCCTTCGTTATGTATATGGCC
f	chr2-44F2	AGCTGCGAATGTGAGCCTTAAGCT	chr2-44R2	CTGCTTCTTCGTTTGGGTTTCGAT
g	chr2-T24F	AGCGTGTTGTAATCACATAACCGAC	chr2-T24R	CGGACATATTTGCAAGGTGACATATC
h	chr2-T45F	CCAAGCTGACCCGTACCTGCA	chr2-T45R	CTGGTGATATCCCTCTGATACCTCA
i	chr2-T30F	GCGTCTGTCCGAGAGGTTCG	chr2-T30R	TGAGAATCCGGGCCCAGGC
j	chr2-T31F	CAGTAGCTACTCAATGCCAAGACG	chr2-T31R	CGATAACACGGTGACCTTGCTAC
k	02-027F	CCATTAATAACAGCGCTACC	02-027R	TATTTACTGGCTTTGGAACC

Table 3 | p^S linkage analysis of BC1 segregants.

Phenotype	BC1 number	SNPs marker															
		A	B	a	b	c	d	e	f	g	h	i	j	k	C	D	E	F
p^S	pS-100922-1-62	H	H	H	H	H	H	H	H	H	A	A	A	A	A	A	A	A
p^S	pS-100922-1-99	A	A	A	A	A	A	A	H	H	H	H	H	H	H	H	H	H
p^S	pS-100922-1-447	A	A	A	A	A	A	A	H	H	H	H	H	H	H	H	H	H
p^S	pS-100922-1-1123	H	H	H	H	H	H	H	H	H	A	A	A	A	A	A	A	A
p^S	pS-100922-1-1259	A	A	A	A	A	A	A	H	H	H	H	H	H	H	H	H	H
p	p-100922-1-1032	A	A	A	A	A	A	A	A	A	H	H	H	H	H	H	H	H

Table 4 | Primers used for RACE, RT-PCR, quantitative RT-PCR and plasmid construction.

Target gene	Primer name	Sequence (5' → 3')
Primers for 5'- and 3'-RACE		
apt-like	Bm3234-5RACE-R1	GTCTGGCTATCAGCTCGAGCACCGTGCG
	Bm3234-5RACE_nest-R1	GCGGAGAGAAGTGGCACTG
	Bm3234-3RACE-F1	TGCCACTTCTCTCCGCGCCAGACGCGGC
	Bm3234-3RACE_nest-F1	GTCAAGTCTGCTCGATGGGA
Primers for RT-PCR		
apt-like	Bm3234_Exon2-F2	ATGTCAACGCGCAGTGCCACTTC
	Bm3234_Exon7-R1	TCATTTAGCGCTGGCCTTCTCCA
	chr2-87R	TCAAACCTTCCACGGAGATCAGCG
	Bm3234_Exon6-F1	GACTATCGACAACGAACGACAGC
	Bm3234_Exon6-R1	CGTTGTCGATAGTCGTCCGTTTG
	Bm3234_Exon6-R2	CCAAGTTGCTCAGGTTCAAAGGC
rpL3	rpL3-F1	AGCACCCCGTCATGGGTCTA
	rpL3-R1	TGCGTCCAAGCTCATCCTGC
Primers for quantitative RT-PCR		
apt-like	q3234_E2F3	TAGCCAGACATAGGGACGTACTCC
	q3234_E3R2	TTGACTTGCGCGGGGGTC
Wnt1	Bm_Wnt1-F1_qPCR	TTCGTTGATACCGGAGAACGG
	Bm_Wnt1-R1_qPCR	CCAGCAAGTCTTCACGGTGC
Dll	Bm_Dll-F1_qPCR	CTCACGAGGCATCAGGGTTTC
	Bm_Dll-R1_qPCR	TTGGGCGGACTTGCACACTG
BGIBMGA003236	Bm3236-F1_qPCR	AAGTATCCGTTGATCATAGCAACA
	Bm3236-R1_qPCR	GATGTAACAATTAGGGTGATCTTCC
BGIBMGA003237	Bm3237-F1_qPCR	GAACATTGGTGGAGTGTGCTC
	Bm3237-R1_qPCR	TGTTCTCCTCGTGACTATCACACTC
BGIBMGA003238	Bm3238-F3_qPCR	CGATGTAAAGGACGGTGTTCTAG
	Bm3238-R2_qPCR	CAAGTAATACTTGTCCGAGTACTCG
BGIBMGA003258	Bm3258-F1_qPCR	GAACCAGATCCTCCGACGC
	Bm3258-R1_qPCR	GTATGGAGTTCTCCGTTCCGG
BGIBMGA003259	Bm3259-F1_qPCR	GACAAAATCTCAGCGGCCGAC
	Bm3259-R1_qPCR	CAGATATCTGCCCGAATGGTCC
BGIBMGA003260	Bm3260-F1_qPCR	GACGACGACACCCAAAGCCTA
	Bm3260-R1_qPCR	GTATTAACCGGGGGTTCGTTTTC
rpL3	qrpL3-F1	ATCAAGGGTTGCTGCATGGGACCT
	qrpL3-R1	GGTTGATCTTTTCTAGTGCAGCCCTC
Primers for plasmid construction		
apt-like	Bm3234_Exon2-F2	ATGTCAACGCGCAGTGCCACTTC
	Bm3234_Exon7-R1	GCATTCGTGCTTTATGTGATCGTC
	piggyBac_Bm3234-F1	ACTAATTCAAGGATCCATGTCAACGCGCAGTGCCACTTC
	piggyBac_Bm3234-R1	TCTAGAGTCGCGGCCGCTCATTTAGCGCTGGCCTTCTCCAG
	piggyBac_Bm3234-Exon6-R2	TCTAGAGTCGCGGCCGCTCAGCTTTTTTCGTGTCCGTTCTTCC
	piggyBac_Bm3234-Exon5-R2	TCTAGAGTCGCGGCCGCTCAGTTAACAGGCACACAGTCGTC
	piggyBac_Bm3234-Exon4-R2	TCTAGAGTCGCGGCCGCTCACTCTTTGACGTCTAAGGGATGT
	piggyBac_Bm3234-Exon3-R2	TCTAGAGTCGCGGCCGCTCACTGTGTGTCGTGAGCGCG
	piggyBac_Bm3234-Exon2-R1	TCTAGAGTCGCGGCCGCTCACTGTTTACTTCTGTTGGTTGCG
apt-like homolog (*P. xuthus*)	Px3234-F1	ATGTCTACGAGGAGCGGTGCGA
	Px3234-R1	TTTCCCGCTCGCTTTTTCCAGAAAG
	piggyBac_Px_apt-like-F1	ACTAATTCAAGGATCCATGTCTACGAGGAGCGGTGC
	piggyBac_Px_apt-like-R1	TCTAGAGTCGCGGCCGCCTATTTCCCGCTCGCTTTTTCC
apt-like homolog (*P. machaon*)	Pm_apt-like-F1	ATGTCTACGAGGAGCGGTGC
	Pm_apt-like-R1	CTATTTCCCGCTCGCTTTTTCC
	piggyBac_Pm_apt-like-F1	ACTAATTCAAGGATCCATGTCTACGAGGAGCGGTGC
	piggyBac_Pm_apt-like-R1	TCTAGAGTCGCGGCCGCCTATTTCCCGCTCGCTTTTTCC
apt-like homolog (*P. aristolochiae*)	Pa_apt-like-F1	ATGTCGACGCGGAGTGGC
	Pa_apt-like-R1	CTATTTGGCGCTCGCTTTCTC
	piggyBac_Pa_apt-like-F1	ACTAATTCAAGGATCCATGTCGACGCGGAGTGGC
	piggyBac_Pa_apt-like-R1	TCTAGAGTCGCGGCCGCCTATTTGGCGCTCGCTTTCTC

Appendix 1: References in Table 1

Banno, Y. *et al. Summary of The Japanese Society of Sericultural Science Conference,* **67,** 93 (1997).

Banno, Y. *et al.* A Guide to the Silkworm Mutants—*Gene Name and Gene.* Silkwork Genetics Division, Institute of Genetic Resources, Kyushu University, Fukuoka, Japan (2005).

Chikushi, H. *J. Seric. Sci. Jap.* **9,** 144–146 (1938).

Kawaguchi, E. *Jap. J. Genet.* **8,** 97–107 (1933).

Shimodaira, M. *Jap. J. Genet.* **22,** 86–92 (1947).

Takasaki, T. *Silkworm Information Service* **2**, 3 (1947).

Tanaka, Y. *J. Coll. Agr. Tohoku Imp. Univ.* **5,** 115–148 (1913).

Tanaka, Y. *J. Coll. Agr. Tohoku Imp. Univ.* **5,** 115–148 (1916).

Tanaka, Y. *J. Coll. Agr. Tohoku Imp. Univ.* **7,** 123–255 (1916).

Tanaka, Y. *Gene Analysis of the Silkworm.* (Gihodo, Tokyo. 1951).

Tazima, Y. *Jap. J. Genet.* **14,** 117–128, 191–203 (1938).

Tazima, Y. *Bull. Seric. Exp. Sta. Jap.* **11,** 525–604 (1943).

Tazima, Y. *The Genetics of the Silkworm* (Logos Press and Prentice Hall, 1964).

Toyama, K. *Bull. Coll. Agr. Tokyo Imp.Univ.* **7,** 259–393 (1906).

Toyama, K. *Sangyo Shimpo* **196,** 4–7 (1909).

Appendix 2: Accession codes

Full-length coding sequences of *apt-like* from p^S, $+^p$ and *p* strains, and *apt-like* homologs from *P. xuthus*, *P. machaon* and *P. aristolochiae* have been deposited in GenBank/EMBL/DDBJ nucleotide core database under the accession codes AB860412, AB860413, AB860414, LC017746, LC017747 and LC017748, respectively.