npg
Heredity (2012) 108, 594–601
© 2012 Macmillan Publishers Limited All rights reserved 0018-067X/12
www.nature.com/hdy

Heredity

ORIGINAL ARTICLE

Genetic basis of stage-specific melanism: a putative role for a cysteine sulfinic acid decarboxylase in insect pigmentation

SV Saenko[1,3], MA Jerónimo[2] and P Beldade[1,2]

Melanism, the overall darkening of the body, is a widespread form of animal adaptation to particular environments, and includes bookcase examples of evolution by natural selection, such as industrial melanism in the peppered moth. The major components of the melanin biosynthesis pathway have been characterized in model insects, but little is known about the genetic basis of life-stage specific melanism such as cases described in some lepidopteran species. Here, we investigate two melanic mutations of *Bicyclus anynana* butterflies, called *Chocolate* and *melanine*, that exclusively affect pigmentation of the larval and adult stages, respectively. Our analysis of Mendelian segregation patterns reveals that the larval and adult melanic phenotypes are due to alleles at different, independently segregating loci. Our linkage mapping analysis excludes the pigmentation candidate gene *black* as the *melanine* locus, and implicates a gene encoding a putative pyridoxal phosphate-dependant cysteine sulfinic acid decarboxylase as the *Chocolate* locus. We show variation in coding sequence and in expression levels for this candidate larval melanism locus. This is the first study that suggests a biological function for this gene in insects. Our findings open up exciting opportunities to study the role of this locus in the evolution of adaptive variation in pigmentation, and the uncoupling of regulation of pigment biosynthesis across developmental stages with different ecologies and pressures on body coloration.
Heredity (2012) **108,** 594–601; doi:10.1038/hdy.2011.127; published online 11 January 2012

Keywords: evo-devo; pigmentation; lepidoptera; melanogenesis; PLP-dependent enzyme; black

INTRODUCTION

One of the most exciting questions in biology is how the diversity of living forms has evolved. Animal coloration is a visually compelling example of such diversification. It is an adaptive trait that plays a role not only in visual communication (for example, interactions with predators and mates), but also in thermoregulation, UV protection and pathogen resistance (True, 2003; Caro, 2005). Variation in coloration is common in natural populations and recently has become a foremost model in studies of the interactions between evolutionary and developmental processes that shape phenotypic variation and diversification (Protas and Patel, 2008; Mills and Patterson, 2009; Wittkopp and Beldade, 2009).

In the past decades, evo-devo studies of animal coloration have advanced our understanding of the genetic mechanisms underlying adaptive changes in overall body pigmentation, and in the color patterns formed by the spatial organization of pigmented cells. Multiple signaling pathways, transcription factors and enzymes have been associated with the evolution of color patterns in vertebrates (Mills *et al*., 2007; Manceau *et al*., 2011) and arthropods (Wittkopp *et al*., 2002; Gompel *et al*., 2005). Some of those have also been implicated in changes in overall body coloration, including dramatic reduction of pigmentation in animals inhabiting caves (Gross *et al*., 2009; Protas *et al*., 2011) or light background environments (Rosenblum *et al*., 2010), as well as dramatic darkening of the body, or melanism, which is addressed here.

Naturally occurring melanic phenotypes are widespread in arthropods and vertebrates as an adaptation for life in particular environments. The genetic basis of melanism has been investigated in different systems where black coloration renders animals cryptic in relatively dark backgrounds; for example, in pocket mice living on lava flows (Nachman *et al*., 2003), in night-hunting cats (Eizirik *et al*., 2003) and in the peppered moth example of industrial melanism (Van't Hof and Saccheri, 2010; Van't Hof *et al*., 2011). The biosynthesis of melanin and other tyrosine-derived pigments, the relative levels of which determine the color of the vertebrate skin and hair and of the insect cuticle, has been studied extensively in model organisms (Protas and Patel, 2008). A number of melanogenesis genes have been associated with larval and adult melanic phenotypes in the insect genetic models *Drosophila melanogaster* and *Bombyx mori* (Wittkopp *et al*., 2002; True *et al*., 2005; Zhan *et al*., 2010). These genes encode enzymes involved in melanin biosynthesis and are obvious candidates for harboring allelic variation contributing to naturally occurring variation in body pigmentation.

Among insects, Lepidoptera (the order of butterflies and moths) show spectacular adaptive variation in color of larval and pupal integument (Hazel, 2002; Jones *et al*., 2007; Noor *et al*., 2008), and in adult wing patterns (Beldade and Brakefield, 2002; Saenko *et al*., 2011). Caterpillars and adults typically exploit different habitats, with specific ecological challenges and selective pressures on their pigmentation. In some cases, stage-specific ecological pressures can lead to

[1]Institute of Biology, Leiden University, Leiden, The Netherlands and [2]Instituto Gulbenkian de Ciência, Oeiras, Portugal
[3]Current address: Department of Genetics & Evolution, University of Geneva, Geneva, Switzerland.
Correspondence: Dr P Beldade, Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, Oeiras P-2780-156, Portugal.
E-mail: pbeldade@igc.gulbenkian.pt or Dr SV Saenko, Department of Genetics & Evolution, University of Geneva, Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland.
E-mail: siouzanna.saenko@unige.ch
Received 20 August 2011; revised 25 October 2011; accepted 17 November 2011; published online 11 January 2012

variation in pigmentation even between larval instars (for example, white-brown early instars of *Papilio xuthus* butterflies mimic bird droppings, whereas the last instar has a green camouflage color; Futahashi and Fujiwara, 2005), or between different wing surfaces of adult butterflies (for example, dorsal and ventral color patterns in *Bicyclus anynana* butterflies involved in mate choice or in predator avoidance, respectively, Oliver *et al*., 2009). The diversified pigmentation phenotypes of different life-stages of lepidopteran species provide a good model for studies of the development and evolution of stage-specific color and color patterns.

The availability of mutants with altered pigmentation or color patterns is especially useful for dissecting the genetic basis of variation in coloration (for example, Weatherbee *et al*., 1999; Koch *et al*., 2000; Futahashi *et al*., 2010; Saenko *et al*., 2010). Here, we focus on two spontaneous melanic mutations of the nymphalid butterfly *B. anynana*, a lab model for studies of wing pattern evolution and development (Beldade and Brakefield, 2002; Beldade *et al*., 2005; Conceição *et al*., 2011) and of developmental plasticity (Beldade *et al*., 2011). Both mutant alleles lead to overall darkening of the body, but while one (named *Chocolate*) affects exclusively the larval stages, the other (named *melanine*) affects only adult pigmentation (Figure 1). We describe the Mendelian segregation of both alleles and show that the larval and adult melanism are not due to genetic variation at the same locus. Furthermore, we provide a detailed genetic investigation of the larval melanism locus, with characterization of sequence and

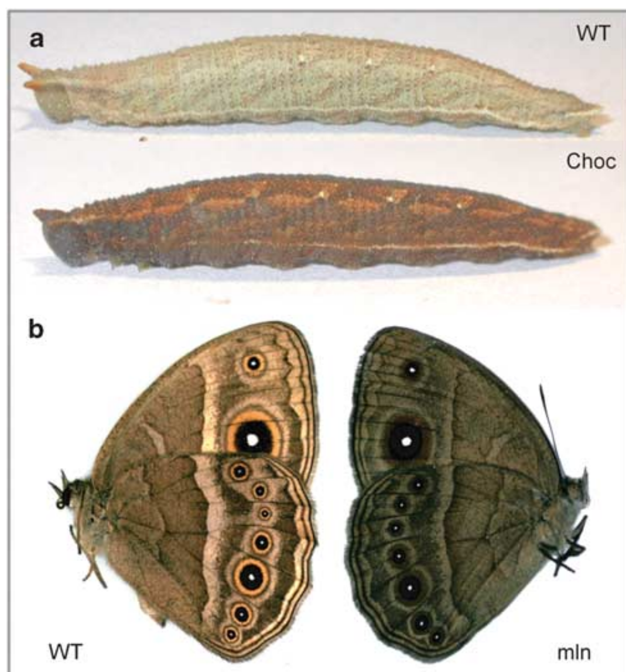


Figure 1 Stage-specific melanic mutants of *Bicyclus anynana*. (**a**) Lateral view of the mid-last instar "WT" and Choc larvae. Typical color of WT head capsule and body integument is light brown, whereas that of the Choc mutants is dark reddish brown (pupal and adult coloration are not affected). The effect of the *Choc* allele is visible but much weaker in earlier larval instars; the dark coloration in these disappears almost completely by the end of each instar, just before each molt. (**b**) Ventral view of WT and melanine (mln) adult females. Typical adult coloration is brown with light bands and concentric rings of white, black and yellow running in the anterior-posterior direction along the wings. Homozygous adults for the *mln* allele are overall dark, almost black (larval and pupal coloration are not affected).

expression differences in pigmentation genes. This study suggests a role in body pigmentation for a gene whose orthologs in other insects have not yet been associated with any biological process.

MATERIALS AND METHODS

Experimental animals and candidate genes

The Chocolate (Choc) and melanine (mln) stocks of *B. anynana* (for clarity, we refer to phenotypes in regular font and to genotypes in italics) were each originally set-up from single individuals isolated from a 'wild-type' (WT) lab population (Brakefield *et al*., 2009), and since maintained with selection in favor of the mutant phenotype. To preserve the health of mutant stocks facing potential effects of inbreeding depression or disease, these can be occasionally outcrossed with healthy WT individuals. All animals in this study were reared at 27 °C as described before (Brakefield *et al*., 2009).

The complete list of *B. anynana* candidate genes analyzed in this study (different subsets in the linkage and expression analyses described below) include those encoding key enzymes in melanogenesis (*pale*, *Ddc*, *black*, *tan*, *yellow*, *ebony* and *purple*) and one suggested by previous linkage mapping work (*CSAD*, details below). The *B. anynana* orthologs of these genes were derived from the species' EST collection available on GenBank (Beldade *et al*., 2006, 2009), as explained below for each of the different cases.

Cloning and phylogenetic analysis of *Ba_black* and *Ba_CSAD*

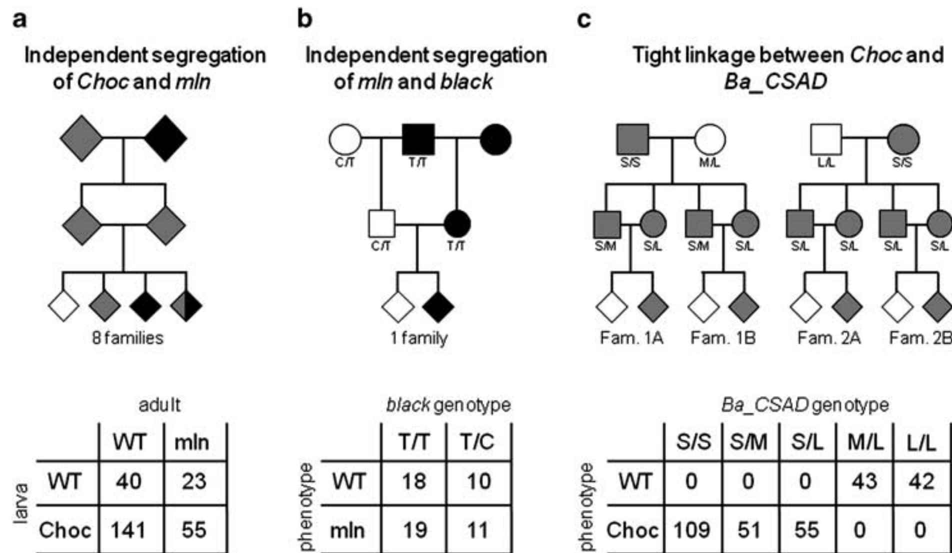
Candidate genes *black* (encoding an enzyme involved in melanin synthesis) and *CSAD* (encoding cysteine sulfinic acid decarboxylase) were chosen based on function in pigmentation in other insects or position in the sequenced genome of the reference lepidopteran *B. mori* (see below), respectively. *B. anynana* ESTs corresponding to these candidate genes were identified by BLAST-screening the NCBI EST database with the predicted protein sequences of *D. melanogaster black* (FlyBase; http://flybase.org/) and *B. mori BGIBMGA010122* (Silkworm Genome Database; http://silkworm.genomics.org.cn/). The full coding sequences of *Ba_black* and *Ba_CSAD* were then obtained with regular PCR or rapid amplification of cDNA ends (RACE), respectively.

Two *B. anynana* ESTs with the highest similarity to *D. melanogaster* gene *black* (FlyBase: CG7811), GE718264 (tblastn e-value=1e-59, 59% identities) and GE718255 (tblastn e-value=5e-59, 64% identities), correspond to the 5′ and 3′ regions of *Ba_black*, respectively. The full coding sequence of this gene was obtained by PCR with primers 5′-ACGTTGCACGCTATTCAGTG-3′ and 5′-CTGCCATAAACGCCAGAAG-3′ on cDNA prepared from total RNA (extracted from whole cuticle of last instar larvae using Trizol, Invitrogen, Paisley, UK) with Reverse Transcription System (Promega, Leiden, The Netherlands). *B. anynana* EST GE676169, that with the highest similarity (tblastn e-value=2e-54, 68% identities) to *B. mori* predicted protein BGIBMGA010122 (Silkworm Genome Database), corresponds to the 3′ region of *Ba_CSAD*. Its full coding sequence was obtained by RACE PCR with primers 5′-CGCTACCATTGCTTGAGATCGCAGTG-3′ and 5′-AACTTCCACCCCAACAAGCATCGACA-3′ on cDNA prepared from total RNA (extracted as above) with SMART RACE Amplification Kit (Clontech, Saint-Germain-en-Laye, France). The nucleotide sequences of *Ba_CSAD* and *Ba_black* complete coding regions were deposited to GenBank under accession numbers JN003848 and JN003850, respectively.

Orthologs of both genes from other insects were identified using BLASTp in NCBI (that is, top hits for different species with e-value below e-150) and aligned with Multiple Sequence Comparison by Log-Expectation (MUSCLE) tool (http://www.ebi.ac.uk/Tools/). A phylogenetic tree was generated using the neighbour-joining method with the bootstrap analysis in the MEGA5 program (http://www.megasoftware.net/).

Mapping families and linkage analysis

To determine whether mutations responsible for Choc and mln phenotypes occurred in the same gene, reciprocal crosses between one Choc and one mln homozygous individuals were set-up: two crosses between a Choc male and a mln female and two crosses between a mln male and a Choc female (Figure 2a). From each of the four F1 families, two males were crossed to their sisters to obtain eight F2 families with a total of 259 individuals (details in Supplementary Table S1). All these were phenotyped for larval (Choc vs WT; Figure 1a)



larva \ adult	WT	mln
WT	40	23
Choc	141	55

phenotype \ *black* genotype	T/T	T/C
WT	18	10
mln	19	11

phenotype \ *Ba_CSAD* genotype	S/S	S/M	S/L	M/L	L/L
WT	0	0	0	43	42
Choc	109	51	55	0	0

Figure 2 Analysis of linkage of melanism loci. Top: schematic representations of the experimental crosses used to test for linkage between loci responsible for larval Choc and adult mln phenotypes (**a**), between *mln* and candidate gene *Ba_black* (**b**), and between *Choc* and candidate gene *Ba_CSAD* (**c**). Different phenotypes are represented by color: gray for Choc melanic larvae (with WT-colour adults), black for mln melanic adults (with WT-colour larvae), and white for WT larval and adult coloration. Males are shown as squares, females as circles and diamonds represent offspring of both sexes. Bottom: tables showing the numbers of offspring in each phenotypic/genotypic class (details about the mapping families in Supplementary Tables S1 and S2). DNA sequence polymorphisms used to genotype the individuals in (**b**) and (**c**) are explained in the text.

and adult (mln vs WT; Figure 1b) pigmentation. The numbers of progeny in each phenotypic class were pooled over the eight families, and totals were tested against the numbers expected under the hypothesis that the loci responsible for the larval and the adult melanic phenotypes are not the same (that is, that the *Choc* and *mln* alleles segregate independently) using the χ^2 goodness-of-fit test.

For the analysis of co-segregation between the mln phenotype and variation at candidate gene *Ba_black* (explained in the Results and Discussion section), a single backcross panel consisting of 30 mln and 28 WT individuals was obtained by crossing one heterozygote male (offspring of a cross between a mln and a WT butterfly) to one mln female (Figure 2b). All butterflies were phenotyped and genomic DNA was extracted from thoraces using the DNeasy Tissue Kit (Qiagen, Hilden, Germany), following manufacturer's recommendations. A 260-bp fragment of the coding sequence of candidate gene *Ba_black* was amplified and sequenced in the grandparents with primers 5′-ACGTTG CACGCTATTCAGTG-3′ and 5′-AATGGCGTATCCGTTAGCAA-3′. A single-nucleotide polymorphism (SNP), C/T in position 119 of the amplified fragment (corresponding to position 381 in the coding sequence of JN003850), is located within the recognition site for the endonuclease *Acc*I, which cuts the fragments only in the presence of C. Genotyping assays based on this polymorphism were carried out in the parents and 58 F2 offspring of the mapping family as follows: the fragments of *Ba_black* were amplified with the same primers as above, digested overnight with 5U *Acc*I (New England Biolabs, Ipswich, USA) following the manufacturer's recommendations, and scored on agarose gels.

For the analysis of co-segregation between the Choc phenotype and variation at candidate gene *Ba_CSAD* (explained in the Results and Discussion section), each of two Choc individuals were crossed to a WT butterfly from an unrelated laboratory population, and an F2 mapping panel consisting of 264 Choc and 85 WT individuals was established (details in Figure 2c and Supplementary Table S2). Genomic DNA for individuals in this panel was obtained as described above. Fragments of the candidate gene *Ba_CSAD* were amplified in the grandparents with primers 5′-ATTGACGCGTTCAAACTGTG-3′ and 5′-ACAAGACCTGGAATTTCCCTA-3′, cloned and sequenced to confirm PCR specificity. We identified a length polymorphism in the last intron of the gene, discriminating between the grand-parental alleles: 1086-bp allele in the Choc individuals, and a 1189-bp plus a 1305-bp alleles in the WT individuals. A genotyping assay based on this length polymorphism was carried out in the eight parents and all 349 F2 offspring of the four families by amplifying fragments of *Ba_CSAD* with the same primers as above and by scoring allele lengths on agarose gels.

Sequence and expression variation in candidate genes

For the analysis of coding sequence variation in *Ba_CSAD*, primers 5′-GAACCACACGTTCGAATTTCA-3′ and 5′-ACAAGACCTGGAATTTCCC-TA-3′ were used to amplify a 1592-bp fragment (base pairs 46–1637 in JN003848, Figure 3a) including the gene's open reading frame from cDNA (prepared from larval cuticle, as above) of four WT and four Choc offspring of a cross between a *Choc*/+ male and a WT female. The products were cloned into the pCRII-TOPO vector (Invitrogen), and five clones per individual were Sanger-sequenced with M13 primers in both directions. All sequences were aligned in BioEdit to identify SNPs and indels.

For the analysis of expression variation, we performed semi-quantitative PCR on cDNA representing transcripts in different stages of larval integument development in individuals from the WT and the pure-breeding Choc stocks. Using pure-breeding stocks, rather than melanic and non-melanic siblings of a single cross, allows us to be certain of individual genotypes even at stages when phenotypes are not obvious. Six individuals were dissected for each of four developmental stages, characterized by the following criteria: (1) late 4th instar (4–5 days after the third molt), (2) pre-molt (next-instar head capsule is visible through the cuticle, which becomes lighter), (3) 1st day of the 5th instar (18–24 h after the 4th molt), (4) 2nd day of the 5th instar (42–48 h after the 4th molt). The integument was dissected from segments 1–5 of the abdomen, and muscles and fat tissue attached to the epidermis were removed. Total RNA was extracted as described above, and 200 ng of RNA of each individual was used to prepare cDNA with Reverse Transcription System (Promega) and oligo(dT) primers, following manufacturer's instructions. Expression levels of *Ba_CSAD*, *Ba_black* and six genes with known role in melanin biosynthesis were examined with semi-quantitative PCR using primers that amplify fragments, which span introns, to avoid noise due to potential residual contamination of cDNA with genomic DNA (see Supplementary Table S3 for GenBank accession numbers and primer sequences). During protocol optimization, negative controls were performed on water, and, to establish the optimal number of cycles, each primer pair was tested with 25, 30 and 35 PCR cycles on a mix of cDNA prepared from all WT individuals. The amplicons were also sequenced to confirm gene identity. The housekeeping gene *FK506 binding protein 2* (*FK506*) was used as an internal control (Pijpe *et al*., 2011). PCRs were performed on

a

```
Ba_CSAD         MELNGHTFLDKVLGIVKEETDHETALIQFKHPKELEEILRENLDISKKLNDQDLEKAVRK 60
BGIBMGA010122   MDLS---FLDRVLQIVKDERKKDVPLVRFKHPEELEAIL--DLDIGQEVNDDDLERCVRQ 55
                *:*.   ***:** ***:* .::..*::****:*** **  :***.:::**:***:.**:

Ba_CSAD         VLQYSVKTNKPTFRNQLYGGTDPYGLAGAWIAEAFNTSQYTFEVAPVFTLIELKVLDHIL 120
BGIBMGA010122   VLQYSVKTDKATFKNQLYGCTDPYGLAGAWIAEAFNTSQYTFEVAPVFTLIELKVLNHIL 115
                ********:*.**:***** ************************************:***

Ba_CSAD         KLFGISDGDGIFSPGGSISMLYALVAAASKAFPEVKRIGMGGLPEMVIVTSEDSHYSILK 180
BGIBMGA010122   NLFGIPNGDGIFSPGGSVSMLYALVAARFKAFPEVKRKGMRNLPEMVIFTSEDSHYSIRK 175
                :****.:**********:*********  ******** ** .******.********* *

Ba_CSAD         AAHWLGFGIENVKVIKTNDHGQMSATDLENTLQQELALGRTPLMVNATAGTTVLGAIDDL 240
BGIBMGA010122   AAHWLGFGTESLRPIKANEHGQMIVSELEKAIEREKKENRYPVLVNATAGTTVLGAIDDL 235
                ******** *.:: **:*:**** .::**:::::*   .* *::****************

Ba_CSAD         ENVAAICEKYGVWMHVDACWGGSLILSKKYRSKLKGINRASSISWNPHKMMGVPLQCSVF 300
BGIBMGA010122   NDIANVCKKHGVWMHVDACWGGSLMLSPKYRSKLLGVKRADSISWNPHKMAGAPLQCSVF 295
                :::* :*:*:**************:** ****** *::**.********* *.*******

Ba_CSAD         LLREKGVLHEANSAAAQYLFQQDKFYDVSYDTGDKSIQCGRKIDAFKLWMIWKARGDIGL 360
BGIBMGA010122   LIKERGLLHEANSAAAQYLFQQDKFYDVSYDTGDKSVQCGRKIDAFKLWMMWKSRGDIGL 355
                *::*:*:*****************************:*************:**:******

Ba_CSAD         NHLTDHVMEIAEFCIQTVAHRSNFRLVTMHMQCPNVCFWYIPTFMNNKMENEEWWNTTHK 420
BGIBMGA010122   SVLTDKVMSIAQFCLSATSKMEGFRLVSEKLQCPNVCFWYIPKFMRNKDENEEWWDLMHK 415
                . ***:**.**:**:.:.:: ..****: ::***********.**.** ******:  **
                                                                          ↓
Ba_CSAD         ITPKIKEQLTLRSQAMVAYSPLRERKNFFRLAFTFHPVLEKSHVLDILQAIEECGEGITL 480
BGIBMGA010122   ITPKIKEQLTLTSKLMVAYSPLRQHKNFFRLAFTFHPELNENQVIEMLNAVKLSGEAITK 475
                *********** *: ********::************ *::.:*:::*:*:: .**.**

Ba_CSAD         PVL- 483
BGIBMGA010122   EMFP 479
                 ::
```

b

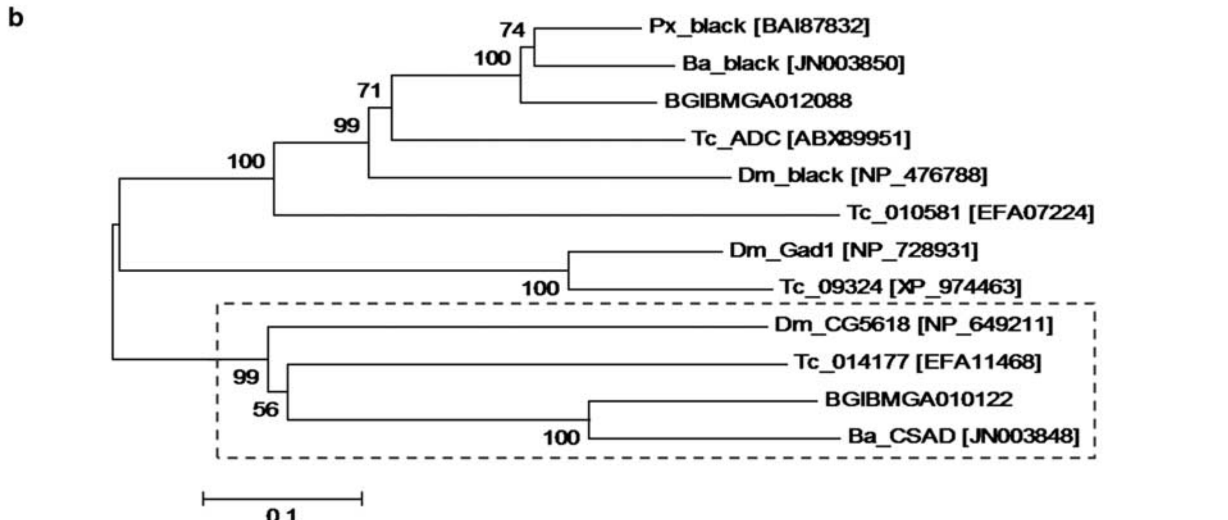


Figure 3 Sequence annotation and phylogenetic analysis of *Ba_CSAD*. (**a**) Alignment of the predicted protein sequences of Ba_CSAD (483 amino acids) and its silkworm ortholog BGIBMGA010122 (479 amino acids). Numbers on the right show amino acid sequence position; sequence identities are marked with (*), conserved substitutions with (:) and semi-conserved substitutions with (.) (*cf.* MUSCLE, see Materials and methods). Dashed line indicates the PLP-dependent DOPA decarboxylase domain, gray boxes indicate pyridoxal 5′-phosphate-binding residues and the catalytic lysine (cf. NCBI Conserved Domains Search; see Materials and methods), and the arrow indicates the threonine, which is substituted by lysine in the mutant allele. (**b**) Neighbor-joining phylogenetic tree of Ba_CSAD and related insect proteins based on their amino acid sequences (with corresponding GenBank accession numbers), with bootstrap values for 1000 replications. Ba_CSAD, *B. mori* BGIBMGA010122, *T. castaneum* 014177 and *D. melanogaster* CG5618 form a separate clade, indicated with the dashed square. The scale bar indicates the evolutionary distance between the groups.

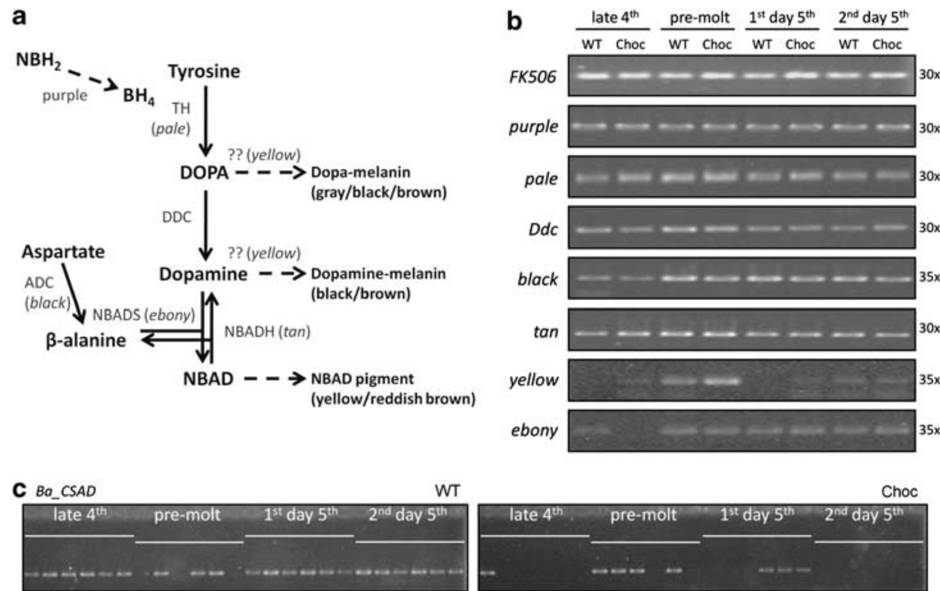


Figure 4 Melanin pathway and semi-quantitative PCR of key genes in WT and Choc larvae. (**a**) A simplified schematic representation of the melanin pathway (adapted from Arakane *et al*., 2009; Zhan *et al*., 2010; Futahashi *et al*., 2010): enzymes are shown in gray with corresponding genes in brackets; arrows indicate reactions catalyzed by each enzyme (dashed arrows correspond to multi-step reactions). TH, tyrosine hydrolase; putative DCE, dopachrome conversion enzyme (note that the exact function of the *yellow* gene is still unclear); DOPA decarboxylase (DDC), dopa decarboxylase; ADC, aspartate decarboxylase; NBADS, N-beta-alanyldopamine synthase; NBADH, N-beta-alanyldopamine hydrolase; NBH_2, dihydroneopterin; BH_4, tetrahydrobiopterin. (**b**) Relative expression levels of the control *FK506* and seven melanin pathway genes in larval integument around the last larval molt (stages indicated on top). PCR reactions (number of cycles on the right) were run on six replicate individuals per stage per phenotype (Supplementary Figure S1) and, for illustration, also on pools of replicates (shown here). (**c**) Semi-quantitative analysis of *Ba_CSAD* expression in WT and Choc larvae, performed at 35 cycles on six individuals for each stage and phenotype (data for 30 and 38 cycles is shown in Supplementary Figure S1). The fact that PCRs worked for both phenotypic groups at some stage assures that lack of amplification is unlikely to be caused by putative sequence differences between populations.

cDNA of each individual and, for illustration in Figure 4b, on pools of 5 μl cDNA of individuals belonging to the same stage and phenotypic class. The PCR conditions were: 95 °C for 3 min, followed by 30 and/or 35 (and also 38 for *Ba_CSAD*) cycles of 94 °C for 20 s, 60 °C for 20 s, 72 °C for 20 s and a final extension at 72 °C for 5 min. Differences in expression were assessed by running 5 μl of PCR products on 1.2% agarose gels.

RESULTS AND DISCUSSION

Mutations in two different loci cause melanism at different developmental stages

The color of the integument of young *B. anynana* larvae is typically green and changes to light brown in the final (fifth) instar; the adults are largely brown with more or less conspicuous wing patterns (Figure 1). Although there is subtle quantitative variation for different aspects of larval and adult coloration (Beldade *et al*., 2002; Allen *et al*., 2008), particular alleles can have very large effects and produce phenotypes well outside the normal distribution range (Beldade *et al*., 2008, Saenko *et al*., 2008; Saenko *et al*., 2010). This is the case for two spontaneous mutations which affect overall pigmentation, each at a specific developmental stage. The recessive *mln* and the dominant *Choc* mutations cause dramatic darkening of adults and larvae, respectively, without having any detectable effects during other developmental stages (Figure 1). Note that the Choc phenotype is virtually indistinguishable from that caused by the *B. anynana dark larvae* allele described elsewhere (Bear *et al*., 2010), but, while the latter is recessive, the *Choc* allele we investigate is dominant (Beldade *et al*., 2009). Because the locus carrying the *dark larvae* allele was not mapped, we cannot establish if the two mutations are at the same locus.

As the *Choc* and *mln* mutations cause overall darkening of the integument, we proceeded to investigate whether they correspond to alleles of the same or different genes. We performed crosses between homozygotes for either of the two mutant alleles and investigated the frequency of different phenotypes in 259 F2 progeny of eight families (Figure 2a and Supplementary Table S1). The total number of progeny in each phenotypic class was not different from that expected if *Choc* and *mln* are alleles of two independently segregating loci (goodness-of-fit chi-square test=5.62, d.f.=3, P=0.87). Moreover, 40 out of 259 individuals had WT coloration at both larval and adult stages, which is only possible if *Choc* and *mln* are not alleles at the same locus and can, therefore, complement.

Analysis of linkage between mutant loci and candidate genes

Several genes have been associated with melanization of the whole body or specific body parts in insects (True, 2003; Wittkopp and Beldade, 2009). In particular, loss-of-function mutations in the gene *black* (encoding Aspartate 1-decarboxylase, involved in production of melanin substrate beta-alanine, Kramer *et al*., 1984) result in overall darkening of adult fruitflies (Phillips *et al*., 2005) and beetles (Arakane *et al*., 2009), a phenotype that resembles that of *B. anynana* mln mutants. Hence, we performed linkage analysis between the *mln* allele, whose location in the *B. anynana* genome is unknown, and a polymorphic marker in *Ba_black*, previously mapped to *B. anynana* linkage group (LG) 11 (Beldade *et al*., 2009), in a single backcross family consisting of 28 WT and 30 mln offspring (Figure 2b). We identified a synonymous C/T polymorphism in the coding sequence of *Ba_black* in the grandparents and genotyped the parents and the 58 offspring (see Materials and methods). In case of tight linkage between

Ba_black and the locus responsible for the mutant phenotype, mln offspring would be expected to carry only the T allele at the *Ba_black* locus (like their mln grandfather), while WT offspring should have the C/T genotype (like their WT grandmother). However, we found that 11 out of 30 mln individuals had C/T and 18 out of 28 WT individuals had T/T genotypes (Figure 2b), which is close to 50% recombination rate between *Ba_black* and the locus carrying the *mln* mutation. This indicates that the two loci are either on different chromosomes or at a large distance on the same chromosome, and that *mln* is not an allele of *Ba_black*.

The *Choc* allele was previously mapped to a 35 cM region on *B. anynana* LG 7 (Beldade *et al*., 2009). This excludes *Ba_black*, mapped to LG 11, and other melanin pathway genes found on other LGs in *B. anynana* (for example, *Ddc* on LG 4, Beldade *et al*., 2009), and in the syntenic reference lepidopteran genome of the silkworm *B. mori* (for example, *ebony* on LG 26, Futahashi *et al*., 2008). The region of the silkworm genome orthologous to the *B. anynana* genomic region where *Choc* maps to was identified as nscaf2986 (Beldade *et al*., 2009 see Supplementary Table S4 for gene content of this *B. mori* genomic region). This region contains the gene *BGIBMGA010122*, which shows high sequence similarity to *D. melanogaster* gene *CG5618* (encoding a predicted cysteine sulfinic acid decarboxylase, CSAD; NCBI tblastn e-value 3e-156, 49% identities), gene *CG14994* (encoding a glutamic acid decarboxylase 1, GAD1; e-value 2e-123, 42% identities) and gene *black* (e-value 1e-147, 45% identities). As Aspartate-1 decarboxylase (the product of *black*) is involved in cuticular melanization in other insects (see above), we wished to investigate whether the *B. anynana* ortholog of *BGIBMGA010122* was responsible for melanization of the larval integument in the Choc mutant.

We identified the *B. anynana* EST with the highest similarity to *BGIBMGA010122* (GE676169, see Materials and methods) and carried out linkage analysis in four F2 families segregating for Choc and WT larval coloration phenotypes (Figure 2c and Supplementary Table S2). An intron length polymorphism in this *B. anynana* gene was identified in the grandparents: the Choc grandparents carried two 1086-bp alleles (small, S), while the WT grandparents carried a 1189-bp (middle, M) and a 1305-bp (large, L) alleles. These were used to genotype the eight F1 parents and 349 F2 offspring in the mapping panel. None of the 85 F2 offspring with WT larval coloration (+/+) had the Choc-specific S allele. On the other hand, 109 Choc larvae (putative *Choc*/*Choc* homozygotes) carried two S alleles, and 155 Choc larvae (putative *Choc*/+ heterozygotes) carried one S allele and either M or L alleles. This perfect co-segregation between the polymorphic marker and larval phenotype indicates that the *Choc* locus is tightly linked (within 0.29 cM, equivalent to fewer than 1 recombination event in 349 meiosis) to EST GE676169 and is likely an allele of the corresponding gene.

The *Choc* locus encodes a putative pyridoxal phosphate (PLP)-dependent cysteine sulfinic acid decarboxylase

We obtained the full-length *B. anynana* ortholog of the *B. mori* gene *BGIBMGA010122* (see Materials and methods) and established that its 1452-bp open reading frame encodes a 483-amino acid protein with a DOPA decarboxylase domain (Figure 3a). This domain belongs to a superfamily of highly conserved domains, the PLP-dependent aspartate aminotransferases. These are characterized by seven conserved pyridoxal 5′-phosphate-binding residues and one catalytic residue lysine (see Figure 3a), which forms the aldimine bond with the coenzyme (John, 1995). Phylogenetic analysis based on the amino acid sequences of the *B. anynana* enzyme and eleven putative homologs from four other insect species (see Materials and methods for selection details) revealed that it is closely related to predicted proteins of *B. mori* (BGIBMGA010122 in Silkworm Genome Database), *Tribolium castaneum* (014177 in BeetleBase; http://beetlebase.org/), and *D. melanogaster* (CG5618 in FlyBase). The function of these proteins, which are clearly separated from insect Black and Gad1 enzymes (Figure 3b), has not been shown experimentally, but *D. melanogaster* CG5618 is predicted (cf. FlyBase) to be a PLP-dependent enzyme with cysteine sulfinic acid decarboxylase activity, or CSAD (enzyme nomenclature code EC 4.1.1.29). This suggests that its *B. anynana* ortholog, hereafter referred to as *Ba_CSAD*, can perform a similar role.

PLP-dependent enzymes catalyse a wide variety of biochemical reactions (for example, decarboxylation or transamination) that almost exclusively involve amino acids (Percudani and Peracchi, 2003; Mozzarelli and Bettati, 2006). Approximately 0.5% of genes in any eukaryotic genome encode PLP-dependent enzymes, and over 100 distinct enzymatic activities are carried by these proteins (Christen and Mehta, 2001). Well-studied enzymes of this group include DOPA decarboxylase, which catalyzes the decarboxylation of tyrosine, and glutamate and aspartate decarboxylases (Richardson *et al*., 2010). CSAD enzymes have been implicated in taurine biosynthesis in mammals (Guion-Rain *et al*., 1975), but, to our knowledge, have not been associated with pigmentation or any other biological process in insects. Our linkage mapping puts the *Choc* allele within 0.29 cM of the putative *Ba_CSAD*. The relationship between genetic and physical maps has not been determined for *B. anynana*, but in the highly syntenic (Beldade *et al*., 2009) reference genome of *B. mori* 0.29 cM is estimated to correspond to 87–174 Kb (Sato *et al*., 2008). Within 174 Kb to the left and to the right of *B. mori BGIBMGA010122* in nscaf2986 there are 12 predicted genes (cf. Silkworm Genome Database), none with described role in melanogenesis (see Supplementary Table S4).

Many cases of larval melanism have been documented in association with thermoregulation and high larval densities (Goulson, 1994; Hazel, 2002; Noor *et al*., 2008), yet little is known about the genetic basis of this phenotype. Our study suggests a novel candidate locus for larval melanism and open up new prospects for further analysis of the role of CSADs in the adaptive evolution of insect coloration.

Regulatory and coding sequence variation of *Ba_CSAD*

To assess whether the causal mutation occurred in the coding or regulatory sequence of *Ba_CSAD*, we looked for differences in protein-coding sequence and in expression levels between Choc and WT individuals. The complete coding sequence of the gene was analyzed in four WT and four Choc sibling offspring of a single cross between a *Choc*/+ male and a +/+ female. No indels causing a frameshift in the open reading frame, or SNPs resulting in a premature stop codon were found in the mutant. Of the five SNPs identified, three are synonymous and two adjacent SNPs (AA in the Choc allele vs CC in the WT at positions 1436+1437 bp in JN003848) cause a threonine to lysine substitution in position 479 of the 483-aa predicted protein, which is outside the functional domain (Figure 3a). However, this substitution might be functionally relevant because amino acid changes outside catalytic sites can still affect enzymatic function, for example, by changing the structure of the protein (Freeman *et al*., 2011).

Expression levels of *Ba_CSAD* and of seven genes with known function in insect pigmentation were examined by semi-quantitative PCR. This analysis targeted four stages around the last larval molt (that is, when cuticular pigmentation takes place) and during the first 2 days of the last larval instar (that is, when the differences between phenotypes are the most obvious), in replicate individuals with either

WT or Choc coloration (Figure 4). The genes *pale*, *Ddc*, *black, tan*, *yellow* and *ebony* encode enzymes of the melanin biosynthesis pathway, whereas the product of *purple* is involved in the biosynthesis of a cofactor for Pale (see Wittkopp *et al*., 2003, Futahashi *et al*., 2010, and references therein). No variation was observed between the phenotypes or stages in the expression levels of the control gene *FK506* and pigmentation genes *purple*, *pale*, *Ddc*, *black,* and *tan*. Expression of *yellow* was relatively high around the molt (consistent with what has been described in other butterflies; Futahashi *et al*., 2010), whereas *ebony* (mapped to LG 26 in the syntenic genome of *B. mori*) was expressed at relatively low levels across all stages, and downregulated in late 4th instar Choc individuals (Supplementary Figure S1), suggesting a regulatory feedback with allelic variation at *Ba_CSAD* (the locus the Choc phenotype maps to).

The candidate gene *Ba_CSAD* was expressed in 21 out of 24 WT individuals, and in only 8 out of 24 Choc larvae examined (Figure 4c and Supplementary Figure S1). The contrast between the WT and mutant phenotypes was particularly obvious during the late 4th instar (*Ba_CSAD* expressed in all six WT and in just one out of six Choc individuals), and on the second day of the 5th instar (transcripts of *Ba_CSAD* detected in all six WT, but in none of the Choc larvae). The correlation between larval pigmentation phenotype and *Ba_CSAD* expression is not perfect, which could be explained by: (1) discrepancies in 'physiological' and 'chronological' stages (differences in physiological state among larvae within same group could result from our staging method, based on a combination of morphological markers and chronological time), (2) differences between males and females (larval sex was not determined here), and (3) cryptic variation in gene expression (cf. Reed *et al*., 2007). Taking together our mapping data, that place the locus responsible for the Choc phenotype within 0.29 cM of *Ba_CSAD*, and these differences in expression of the same gene between WT and Choc larvae, suggests that a mutation in *Ba_CSAD* is causing the darkening of the larval cuticle in the *B. anynana* Choc individuals. With all data taken together, it seems unlikely that the expression difference we see for *Ba_CSAD* between Choc and WT is not related to the difference in pigmentation, or that it is controlled by some trans-acting factor diverged between the two lab populations.

Concluding remarks

Analysis of laboratory mutants has proven to be useful in evo-devo studies of animal pigmentation (Protas and Patel, 2008). We investigate two melanic mutations in the butterfly model, *B. anynana*, and provide several lines of evidence suggesting that allelic variation at the *Ba_CSAD* locus could be responsible for the larval color phenotype in the Choc mutant. First, the *Ba_CSAD* and *Choc* loci are mapped to within 0.29 cM of each other in a region that does not contain any other pigmentation candidate genes (cf. the analysis of the syntenic, fully sequenced genome of the silkworm; see Supplementary Table S4). Second, we identified differences between WT and mutant individuals, both in the coding sequence (an amino acid substitution outside the catalytic domain) and in expression levels (downregulation around the last larval molt) of *Ba_CSAD*. Additional studies are required to analyze the function (for example, substrate specificity) of this enzyme, to characterize the putative effects of the amino acid changes we detected, and to identify the nucleotide changes responsible for expression differences.

We also show that the larval and adult melanic phenotypes are due to alleles at two different, independently segregating loci, and that the latter is not due to a mutation in the candidate gene *Ba_black* whose orthologs have been implicated in similar phenotypes in other insects. Mutant alleles such as *Choc* and *mln* provide the opportunity to explore the genetic basis of melanism, and to investigate whether the loci identified in lab mutants are important in micro- and macro-evolutionary processes in nature (Haag and True, 2001). Furthermore, mutations with such a stage-restricted effect on pigmentation indicate that larval and adult coloration are, to some extent, regulated independently, and offer a chance to explore experimentally this type of developmental uncoupling.

DATA ARCHIVING

Sequence data have been submitted to GenBank: accession numbers JN003848–JN003850. Phenotype and genotype data from the linkage mapping experiments have been deposited in the Dryad repository: doi:10.5061/dryad.1925530c.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the students Ana Marcelino (for cloning of *ebony* and genotyping of *Ba_black*), Eline Reijmerink (for crosses between Choc and mln mutants) and Paul Brakefield for establishing the mutant stocks and for discussions about them. This work was supported by grants to PB from the Portuguese Foundation for Science and Technology (FCT, PTDC/BIA-BDE/099808/2008) and the Dutch Science Organization (NWO, VIDI 864.08.010).

Allen CE, Beldade P, Zwaan BJ, Brakefield PM (2008). Differences in the selection response of serially repeated color pattern characters: standing variation, development, and evolution. *BMC Evol Biol* **8**: 94.
Arakane Y, Lomakin J, Beeman RW, Muthukrishnan S, Gehrke SH, Kanost MR *et al.* (2009). Molecular and functional analyses of amino acid decarboxylases involved in cuticle tanning in *Tribolium castaneum*. *J Biol Chem* **284**: 16584–16594.
Bear A, Simons A, Westerman E, Monteiro A (2010). The genetic, morphological, and physiological characterization of a dark larval cuticle mutation in the butterfly, *Bicyclus anynana*. *PLoS One* **5**: e11563.
Beldade P, Brakefield PM (2002). The genetics and evo-devo of butterfly wing patterns. *Nat Rev Genet* **3**: 442–452.
Beldade P, Brakefield PM, Long AD (2005). Generating phenotypic variation: prospects from "evo-devo" research on *Bicyclus anynana* wing patterns. *Evol Dev* **7**: 101–107.
Beldade P, French V, Brakefield PM (2008). Developmental and genetic mechanisms for evolutionary diversification of serial repeats: eyespot size in *Bicyclus anynana* butterflies. *J Exp Zool B Mol Dev Evol* **310**: 191–201.
Beldade P, Koops K, Brakefield PM (2002). Developmental constraints versus flexibility in morphological evolution. *Nature* **416**: 844–847.
Beldade P, Mateus ARA, Keller RA (2011). Evolution and molecular mechanisms of adaptive developmental plasticity. *Mol Ecol* **20**: 1347–1363.
Beldade P, Rudd S, Gruber JD, Long AD (2006). A wing expressed sequence tag resource for *Bicyclus anynana* butterflies, an evo-devo model. *BMC Genomics* **7**: 130.
Beldade P, Saenko SV, Pul N, Long AD (2009). A gene-based linkage map for *Bicyclus anynana* butterflies allows for a comprehensive analysis of synteny with the lepidopteran reference genome. *PLoS Genet* **5**: e1000366.
Brakefield P, Beldade P, Zwaan BJ (2009). The African butterfly *Bicyclus anynana*: a model for evolutionary genetics and evolutionary developmental biology. In: Behringer RR, Johnson AD, Krumlauf RE (eds). *Emerging Model Organisms: A Laboratory Manual*. Cold Spring Harbor Laboratory Press **Vol.1**, Chapter 12.
Caro T (2005). The adaptive significance of coloration in mammals. *BioScience* **55**: 125–136.
Christen P, Mehta PK (2001). From cofactor to enzymes. The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes. *Chem Rec* **1**: 436–447.
Conceição IC, Long AD, Gruber JD, Beldade P (2011). Genomic sequence around butterfly wing development genes: annotation and comparative analysis. *PLoS One* **6**: e23778.
Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ (2003). Molecular genetics and evolution of melanism in the cat family. *Curr Biol* **13**: 448–453.
Freeman AM, Mole BM, Silversmith RE, Bourret RB (2011). Action at a distance: amino acid substitutions that affect binding of the phosphorylated CheY response regulator and catalysis of dephosphorylation can be far from the CheZ phosphatase active site. *J Bacteriol* **193**: 4709–4718.
Futahashi R, Banno Y, Fujiwara H (2010). Caterpillar color patterns are determined by a two-phase melanin gene prepatterning process: new evidence from tan and laccase2. *Evol Dev* **12**: 157–167.

Futahashi R, Fujiwara H (2005). Melanin-synthesis enzymes coregulate stage-specific larval cuticular markings in the swallowtail butterfly, *Papilio xuthus*. *Dev Genes Evol* **215**: 519–529.
Futahashi R, Sato J, Meng Y, Okamoto S, Daimon T, Yamamoto K *et al*. (2008). *yellow* and *ebony* are the responsible genes for the larval color mutants of the silkworm *Bombyx mori*. *Genetics* **180**: 1995–2005.
Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB (2005). Change caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**: 481–487.
Goulson D (1994). Determination of larval melanization in the moth, *Mamestra brassicae*, and the role of melanin in thermoregulation. *Heredity* **73**: 471–479.
Gross JB, Borowsky R, Tabin CJ (2009). A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genet* **5**: e1000326.
Guion-Rain MC, Portemer C, Chatagner F (1975). Rat liver cysteine sulfinate decarboxylase: purification, new appraisal of the molecular weight and determination of catalytic properties. *Biochim Biophys Acta* **384**: 265–276.
Haag ES, True JR (2001). From mutants to mechanisms? Assessing the candidate gene paradigm in evolutionary biology. *Evolution* **55**: 1077–1084.
Hazel WN (2002). The environmental and genetic control of seasonal polyphenism in larval color and its adaptive significance in a swallowtail butterfly. *Evolution* **56**: 342–348.
John RA (1995). Pyridoxal phosphate-dependent enzymes. *Biochim Biophys Acta* **1248**: 81–96.
Jones M, Rakes L, Yochum M, Dunn G, Wurster S, Kinney K *et al*. (2007). The proximate control of pupal color in swallowtail butterflies: implications for the evolution of environmentally cued pupal color in butterflies (Lepidoptera: Papilionidae). *J Insect Physiol* **53**: 40–46.
Koch PB, Behnecke B, Weigmann-Lenz M, Ffrench-Constant RH (2000). Insect pigmentation: activities of beta-alanyldopamine synthase in wing color patterns of wild-type and melanic mutant swallowtail butterfly *Papilio glaucus*. *Pigment Cell Res* (Suppl)**8**: 54–58.
Kramer KJ, Morgan TD, Hopkins TL, Roseland CR, Aso Y, Beeman RW *et al*. (1984). Cateholamines and β-alanine in the red flour beetle, *Tribolium castaneum* - roles in cuticle sclerotization and melanization. *Insect Biochem* **14**: 293–298.
Manceau M, Domingues VS, Mallarino R, Hoekstra HE (2011). The developmental role of Agouti in color pattern evolution. *Science* **331**: 1062–1065.
Mills MG, Nuckels RJ, Parichy DM (2007). Deconstructing evolution of adult phenotypes: genetic analyses of kit reveal homology and evolutionary novelty during adult pigment pattern development of *Danio* fishes. *Development* **134**: 1081–1090.
Mills MG, Patterson LB (2009). Not just black and white: pigment pattern development and evolution in vertebrates. *Semin Cell Dev Biol* **20**: 72–81.
Mozzarelli A, Bettati S (2006). Exploring the pyridoxal 5′-phosphate-dependent enzymes. *Chem Rec* **6**: 275–287.
Nachman MW, Hoekstra HE, D'Agostino SL (2003). The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci USA* **100**: 5268–5273.
Noor MA, Parnell RS, Grant BS (2008). A reversible color polyphenism in American peppered moth (*Biston betularia cognataria*) caterpillars. *PLoS One* **3**: e3142.
Oliver JC, Robertson KA, Monteiro A (2009). Accommodating natural and sexual selection in butterfly wing pattern evolution. *Proc Biol Sci* **276**: 2369–2375.
Percudani R, Peracchi A (2003). A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO reports* **4**: 850–854.
Phillips AM, Smart R, Strauss R, Brembs B, Kelly LE (2005). The *Drosophila* black enigma: the molecular and behavioural characterization of the *black1* mutant allele. *Gene* **351**: 131–142.
Pijpe J, Pul N, van Duijn S, Brakefield PM, Zwaan BJ (2011). Changed gene expression for candidate ageing genes in long-lived *Bicyclus anynana* butterflies. *Exp Gerontol* **46**: 426–434.
Protas ME, Patel NH (2008). Evolution of coloration patterns. *Annu Rev Cell Dev Biol* **24**: 425–446.
Protas ME, Trontelj P, Patel NH (2011). Genetic basis of eye and pigment loss in the cave crustacean, *Asellus aquaticus*. *Proc Natl Acad Sci USA* **108**: 5702–5707.
Reed RD, Chen PH, Frederik Nijhout H (2007). Cryptic variation in butterfly eyespot development: the importance of sample size in gene expression studies. *Evol Dev* **9**: 2–9.
Richardson G, Ding H, Rocheleau T, Mayhew G, Reddy E, Han Q *et al*. (2010). An examination of aspartate decarboxylase and glutamate decarboxylase activity in mosquitoes. *Mol Biol Rep* **37**: 3199–3205.
Rosenblum EB, Römpler H, Schöneberg T, Hoekstra HE (2010). Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc Natl Acad Sci USA* **107**: 2113–2117.
Saenko SV, Brakefield PM, Beldade P (2010). Single locus affects embryonic segment polarity and multiple aspects of an adult evolutionary novelty. *BMC Biol* **8**: 111.
Saenko SV, French V, Brakefield PM, Beldade P (2008). Conserved developmental processes and the formation of evolutionary novelties: examples from butterfly wings. *Philos Trans R Soc Lond B Biol Sci* **363**: 1549–1555.
Saenko SV, Marialva MS, Beldade P (2011). Involvement of the conserved *Hox* gene *Antennapedia* in the development and evolution of a novel trait. *EvoDevo* **2**: 9.
Sato K, Matsunaga TM, Futahashi R, Kojima T, Mita K, Banno Y *et al*. (2008). Positional cloning of a *Bombyx* wingless locus *flugellos* (*fl*) reveals a crucial role for fringe that is specific for wing morphogenesis. *Genetics* **179**: 875–885.
True JR (2003). Insect melanism: the molecules matter. *Trends Ecol Evol* **18**: 640–647.
True JR, Yeh SD, Hovemann BT, Kemme T, Meinertzhagen IA, Edwards TN *et al*. (2005). *Drosophila tan* encodes a novel hydrolase required in pigmentation and vision. *PLoS Genet* **1**: e63.
van't Hof AE, Edmonds N, Dalíková M, Marec F, Saccheri IJ (2011). Industrial melanism in british peppered moths has a singular and recent mutational origin. *Science* **332**: 958–960.
van't Hof AE, Saccheri IJ (2010). Industrial melanism in the peppered moth is not associated with genetic variation in canonical melanisation gene candidates. *PLoS One* **5**: e10889.
Weatherbee SD, Nijhout HF, Grunert LW, Halder G, Galant R, Selegue J *et al*. (1999). Ultrabithorax function in butterfly wings and the evolution of insect wing patterns. *Curr Biol* **9**: 109–115.
Wittkopp PJ, Beldade P (2009). Development and evolution of insect pigmentation: genetic mechanisms and the potential consequences of pleiotropy. *Semin Cell Dev Biol* **20**: 65–71.
Wittkopp PJ, Carroll SB, Kopp A (2003). Evolution in black and white: genetic control of pigment patterns in *Drosophila*. *Trends Genet* **19**: 495–504.
Wittkopp PJ, Vaccaro K, Carroll SB (2002). Evolution of *yellow* gene regulation and pigmentation in *Drosophila*. *Curr Biol* **12**: 1547–1556.
Zhan S, Guo Q, Li M, Li M, Li J, Miao X *et al*. (2010). Disruption of an N-acetyltransferase gene in the silkworm reveals a novel role in pigmentation. *Development* **137**: 4083–4090.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)