1 **CRISPR/Cas9-based heritable targeted mutagenesis in *Thermobia domestica*: a genetic**

2 **tool in an apterygote development model of wing evolution**

3

4 Takahiro Ohde[a, b, c*], Yusuke Takehana[d], Takahiro Shiotsuki[e] and Teruyuki Niimi[a, b]

5

6 [a] Division of Evolutionary Developmental Biology, National Institute for Basic Biology, 38

7 Nishigonaka Myodaiji, Okazaki 444-8585, Japan.

8 [b] Department of Basic Biology, School of Life Science, SOKENDAI (The Graduate

9 University for Advanced Studies), 38 Nishigonaka Myodaiji, Okazaki 444-8585, Japan.

10 [c] Department of Applied Biosciences, Graduate School of Agriculture, Kyoto University,

11 Kitashirakawa Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan.

12 [d] Department of Animal Bio-Science, Faculty of Bio-Science, Nagahama Institute of Bio-

13 Science and Technology, 1266 Tamura, Nagahama, Shiga 526-0829, Japan.

14 [e] Department of Life Science and Technology, Graduate School of Life and Environmental

15 Science, Shimane University, 1060 Nishikawatsu-cho, Matsue, Shimane 690-8504, Japan.

16

17 *Correspondence to TO at ohde.takahiro.4n@kyoto-u.ac.jp

Abstract

Despite previous developmental studies on basally branching wingless insects and

crustaceans, the evolutionary origin of insect wings remains controversial. Knowledge

regarding genetic regulation of tissues hypothesized to have given rise to wings would help to

elucidate how ancestral development changed to allow the evolution of true wings. However,

genetic tools available for basally branching wingless species are limited. The firebrat

Thermobia domestica is an apterygote species, phylogenetically related to winged insects. *T.*

domestica presents a suitable morphology to investigate the origin of wings, as it forms the

tergal paranotum, from which wings are hypothesized to have originated. Here we report the

first successful CRISPR/Cas9-based germline genome editing in *T. domestica.* We provide a

technological platform to understand the development of tissues hypothesized to have given

rise to wings in an insect with a pre-wing evolution body plan.

1. Introduction

A neontological approach is essential for understanding the evolution of insect wings because

the transitional fossil record from wingless to winged insects is sparse. Developmental

comparisons between modern species regarding the pre- and post-evolution status could

elucidate the key developmental changes for trait evolution. For example, developmental

comparisons between the digital skeletons of fins and limbs have previously clarified the

relationship between bones in the pre- and post-evolution status: tetrapods' digits originated

through transition of distal cellular fates in ancestral fins (Nakamura *et al.*, 2016).

The case of insect wings is more challenging since the anatomical origin remains unclear.

Two anatomically distinct portions, the tergum and pleuron, have long been proposed as

possible origins of insect wings, and several developmental studies evaluating the expression

of wing gene orthologs in wingless arthropods have provided new insights into these classical

hypotheses (reviewed in Clark-Hachtel and Tomoyasu, 2016). However, the question remains

unanswered, as results from those studies support both the tergal and pleural origin

hypotheses. Averof and Cohen (1997) showed that Apterous and Pdm, which are two

functional proteins involved in the development of wings in *Drosophila melanogaster*, are

expressed in the limb branch (epipod) in crustaceans (Averof and Cohen, 1997). As the

1   common ancestry between wings and crustacean limb branches had been suggested

2   (Kukalová-Peck, 1983), the study provided the first genetic evidence supporting the pleural

3   origin hypothesis. On the other hand, Niwa *et al*. (2010) examined the expression pattern of

4   the wing gene orthologs *wingless*, *apterous,* and *vestigial* in the jumping bristletail

5   *Pedetontus unimaculatus* and showed that those genes are expressed in both the developing

6   tergal margin and pleural branches. This dual wing gene expression led them to propose that

7   both the tergal and pleural tissues contribute to wing evolution, hence formulating the dual

8   origin hypothesis (Niwa *et al*., 2010). To further clarify how each structure contributed to the

9   first wing development, it is important to delve further into the genetic regulation operating in

10   these structures that possibly originate wings. This approach may shed light on the

11   development of wingless ancestors and the subsequent developmental changes that led to

12   evolution of the first wing. However, we currently lack an apterygote insect model that

13   allows for detailed genetic analysis.

14   The firebrat *Thermobia domestica* belongs to the order Zygentoma, which comprises an

15   only apterygote group within ectognatha, in which the first wing evolved (Misof *et al*., 2014).

16   Because of the phylogenetic proximity with pterygote insects, the order Zygentoma retains

17   key traits that were present in the wingless ancestor. Although *T. domestica* forms no wings,

4

1   it forms a flat sheet structure, known as the paranotum, at the dorsolateral margin of thoracic

2   segments (**Fig. 1**). The paranotum is considered a tissue hypothesized to have given rise to

3   wings because of its location at the tergal lateral margin where pterygote insects form wings

4   (Snodgrass, 1935).

5   The present study aimed to establish the CRISPR/Cas9 (Clustered regularly interspaced

6   short palindromic repeats/CRISPR-associated protein 9) system in *T. domestica,* as the first

7   genome editing tool in an apterygote insect. In CRISPR/Cas9-based genome editing, Cas

8   nuclease introduces double-strand DNA breaks according to the guide RNA nucleotide

9   sequence, and brings about subsequent endogenous DNA repair processes, such as non-

10   homologous end joining and homologous recombination, resulting in targeted gene disruption

11   (insertion/deletion) or targeted gene integration at the cleaved site (Doudna and Charpentier,

12   2014; Hsu *et al*., 2014). The establishment of CRISPR/Cas9 germline genome editing in *T.*

13   *domestica* could potentially expand the available approaches, such as gene functional

14   analysis and transcriptional modulation (Doudna and Charpentier, 2014; Hsu *et al*., 2014),

15   hence accelerating the understanding of apterygote insect development.

16   For the proof-of-principle experiment of this study, we targeted *white*, an evolutionarily

17   conserved gene, encoding the ATP-binding cassette (ABC) membrane transporter. In *D.*

1   *melanogaster,* three paralogous half-type ABC transporters, namely White, Brown, and

2   Scarlet, function to transport pigment precursors into the eye pigment cells (Ewart and

3   Howells, 1998). The White/Brown heterodimer transports a precursor of the red pigment

4   pteridine, whereas the White/Scarlet heterodimer transports a precursor of the brown pigment

5   ommochrome (Ewart and Howells, 1998). Dysfunction of *white* orthologs results in a clear

6   loss of eye-color in non-dipteran insect orders, such as Lepidoptera and Coleoptera (Abraham

7   *et al*., 2000; Grubbs *et al*., 2015). In addition, a white ortholog has been found in a

8   crustacean, which is a closely related group of hexapoda, suggesting that the requirement of

9   White as a core subunit in ABC transporters is evolutionarily conserved from some

10   crustacean lineages to hexapoda (Sturm *et al*., 2009). In this study, we successfully

11   introduced heritable mutations in *white* locus of *T. domestica* by using CRISPR/Cas9 system,

12   which results in white-eyed firebrats. The establishment of CRISPR/Cas9-mediated genome

13   engineering in the apterygote insect open new research opportunities on the evolution of

14   insects, such as on the insect wing origin.

15

2. Materials and Methods

2.1. Insect rearing

Firebrats were placed in plastic containers and stored inside a dark incubator at 37 ± 1°C with no photoperiodic conditioning. A water bottle without a lid was fitted in the container to maintain approximately 55%–70% humidity. The insects were fed a mixture of the artificial fish foods Tetra Goldfish and Tetra Goldfish Gold Growth (Spectrum Brands, WI, U.S.A.).

2.2 Identification of translated nucleotide sequences of the ABC transporter gene homologs in T. domestica

BLAST search against our unpublished de novo transcriptome of *T. domestica* was performed to find ABC transporter gene homologs. We constructed *T. domestica* cDNA library from total RNA extracted from few terminal abdominal segments of both sexes, and sequenced on HiSeq2000 (Illumina, CA, U.S.A.). Short reads were assembled with Trinity software (Grabherr *et al*., 2011).

2.3. Molecular phylogenetic analysis

The *translated nucleotide sequences of the ABC transporter gene homologs* from *T.*

1   *domestica* were aligned with the following public amino acid sequences using the MUSCLE

2   software with the default configuration on EMBL-EBI server (version 3.8.31; Edgar, 2004):

3   White from *Drosophila melanogaster* (AAF45826.1), *Aedes aegypti* (XP_011492980),

4   *Anopheles gambiae* (XP_310530), *Tribolium castaneum* (AAL56571.1), *Bombyx mori*

5   (NP_001037034.1), *Athalia rosae* (BAD54748.1), and *Pediculus humanus corporis*

6   (XP_002425032.1); Scarlet from *D. melanogaster* (AAF49455.1), *Anopheles gambiae*

7   (XP_321812.4), *T. castaneum* (AJD07060.1), *B. mori* (BAK08373.1), *Apis mellifera*

8   (XP_001122240.1), and *Pediculus humanus corporis* (XP_002423908.1); Brown from *D.*

9   *melanogaster* (AAF47020.3), *A. gambiae* (XP_308215.4), *T. castaneum* (AJD07061.1), and

10   *A. mellifera* (XP_395665.5); and ABC transporter expressed in the trachea (Atet) from *D.*

11   *melanogaster* (AAF45826.1) and *T. castaneum* (AAL56571.1). Atet was included to the

12   analysis as a representative non-eye-color ABC transporter since it is grouped into ABCG

13   subfamily with three eye-color ABC transporters, but expresses in non-eye tissues such as

14   trachea (Kuwana et al., 1996; Dermauw and Van Leeuwen, 2014). Poorly aligned regions

15   were trimmed with Gblocks (version 0.91b; Castresana, 2000), and 327 residues of amino

16   acid from each protein was used for the molecular phylogenetic analysis (sequences used for

17   phylogenetic analysis are provided in Appendix A).

1   A maximum likelihood-based phylogenetic tree was generated using the Garli software

2   (version 2.01; Zwickl, 2006) with 1000 bootstrap replicates. Prottest (version 3.4.2) was used

3   for selecting the LG model of amino acid evolution.

4   The iTOL was used to draw the phylogenetic tree (version 4.0.3; Letunic and Bork, 2016).

5   Nucleotide sequences of *Td22089*, *Td23978*, *Td30354,* and *Td32270* were deposited to

6   DDBJ/ENA/NCBI with accession numbers LC368081, LC368082, LC368083, and

7   LC368084, respectively.

8

9   *2.4. Embryo staging*

10   *T. domestica* embryos were collected from tissue paper at four hours after placed in a

11   container, and incubated until reaching to each time point. Embryos were collected and

12   incubated at 37 ± 1°C. We made a hole on a collected embryo in 4% paraformaldehyde, and

13   fixed overnight at 4°C. After fixation in paraformaldehyde, chorion and vitelline membrane

14   were manually removed with forceps, and then embryos were post-fixed in Carnoy's solution

15   overnight at 4°C. We washed fixed embryos in 0.75% NaCl three times, and stained with

16   SYTOX Green (S7020; Thermo Fischer Scientific, MA, USA) for at least one hour at room

17   temperature.

1

2   *2.5. Single guide RNA (sgRNA) synthesis*

3   The gene-specific sgRNAs were transcribed *in vitro* from synthesized DNA templates and

4   column-purified using a GeneArt Precision gRNA synthesis kit (Thermo Fischer Scientific,

5   MA, USA) according to the manufacturer's instructions. The DNA oligomers for each

6   sgRNA (sgRNA#1: TAATACGACTCACTATAGTAAGTGTTGTGGGAC and

7   TTCTAGCTCTAAAACATCGGTCCCACAACACTTA; sgRNA#2

8   TAATACGACTCACTATAGGGATTCCTGGAAGGC and

9   TTCTAGCTCTAAAACTTCAGCCTTCCAGGAATCC) were annealed and used as

10   templates for sgRNA synthesis. Synthesized sgRNA was stored at −80°C until use.

11

12   *2.6. sgRNA/Cas9 protein microinjection*

13   Eggs laid within eight hours at 37 ± 1°C were collected from layered tissue paper, and

14   aligned on grooves in a 2% agarose gel, which held the eggs during microinjection. We used

15   a plastic mold to make grooves on an agarose gel. The mold design is according to Porazinski

16   *et al*. (2010). A solution containing 1 µg/µL *Streptococcus pyogenes* Cas9 nuclease with three

17   nuclear localization signals (Integrated DNA Technologies, IA, USA) and 100 ng/µL sgRNA

1   was injected into the eggs with Femtotips II (Eppendorf, Hamburg, Germany) using the

2   Femtojet microinjector (Eppendorf). An sgRNA targeting a site was injected at a time.

3   Injected eggs were incubated on a 2% agarose gel in an incubator at $37 \pm 1°C$.

4

5   *2.7. Genotyping*

6   Embryos and hatched nymphs were genotyped using genomic DNA extracted from whole

7   egg, and cerci and a caudal filament, respectively. The tissue was digested with 0.25 mg/mL

8   proteinase K solution in Tris-EDTA buffer for more than three hours at 55°C. The supernatant

9   was stored at −20°C and used as a PCR template. A heteroduplex mobility assay was

10   performed by running the PCR-amplified DNA fragments in MCE-202 MultiNA system

11   (Shimadzu corporation, Kyoto, Japan; Chen *et al*., 2012; Ansai *et al*., 2014). In this method,

12   heteroduplexes are detected as bands with low mobility by running PCR products directly on

13   MultiNA system without denaturing and reannealing (Ansai *et al*., 2014). Some PCR

14   products that showed heteroduplex bands were sub-cloned into pCR4-TOPO (Thermo

15   Fischer Scientific) and sequenced.

16   For high-resolution melt analysis performed on the G1 generation, the PCR product was

17   mixed with SYTO9 dye (Thermo Fischer Scientific), incubated for two minutes at 95°C and

1   five minutes at 25°C, and analyzed in a LightScanner system (Idaho Technologies Inc., UT,

2   U.S.A; Dahlem *et al.*, 2012). The PCR products that showed a differential melting pattern

3   were re-amplified from genomic DNA, and sequenced after treatment with ExoSAP-IT™

4   PCR Product Cleanup Reagent (Thermo Fischer Scientific).

5   PCR was performed using the primer pair#1: 5′-GCTGTTGAACGCTCTGACTTT-3′ and 5′-

6   TGCTGTACGTAAGCGGACAA-3′, and pair#2 5′-CAACATACTTTGATCGGGATTC-3′

7   and 5′-TCTGATGCAAAAGACAATCG-3′ for checking effects of sgRNA#1 and sgRNA#2,

8   respectively.

9

3. Results and Discussion

3.1. Identifying a white ortholog in T. domestica

To identify a *white* ortholog in *T. domestica*, we used the basic local alignment search tool in

our unpublished *T. domestica* transcriptome (total number of transcripts are 237,304 with

1,625 of contig N50) for finding sequences highly similar to the eye-color ABC transporters

white, *brown*, and *scarlet* of other species. We found four amino acid sequences of candidate

ABC transporter orthologs–Td22089, Td23978, Td30354, and Td32270–which were then

aligned with public ABC transporter sequences (as listed in the Materials and Methods) from

other insect species to perform a phylogenetic analysis. We first performed phylogenetic

analysis with all four ABC transporter genes identified from *T. domestica* transcriptome, and

obtained low bootstrap values of 349 and 127 for Brown and Scarlet subfamily clusters,

respectively (Supplementary figure in Appendix A). Since Td22089 has a shorter amino acid

sequence due to the lack of C-terminal region, we omitted Td22089 and reanalyzed the

phylogeny. It results in the clear separation between Brown and Scarlet clusters (**Fig. 2A**).

Td32270 and Td23978 were grouped with the White and Scarlet subfamilies, respectively

(**Fig. 2A**), whereas Td30354 was not grouped with any of the eye-color ABC transporters.

The cluster of White subfamily was supported by a high bootstrap value (971) after 1,000

1   replicates (**Fig. 2A**). We then used the *Td32270* sequence as a *T. domestica white* ortholog

2   (*Td-w*) for the subsequent analysis. An InterProScan search against Td-W revealed that it

3   contains an ATP-binding domain and six transmembrane helices like other ABC transporters

4   (**Fig. 2B**; Dermauw and Van Leeuwen, 2014). To increase the chance of disrupting protein

5   function by introducing insertions or deletions (indels), we designed sgRNAs targeting sites

6   within the ATP-binding domain (**Fig. 2B**).

7

8   *3.2. Establishment of CRISPR/Cas9-mediated gene knockout in* T. domestica

9   More injected materials would be successfully delivered into nucleus before cell membrane

10   barrier are formed (i.e. before the cellular blastoderm stage) as shown in other insects. To

11   determine the timing of injection to *T. domestica* eggs, we briefly examined developmental

12   stage of embryos until 24 hours after egg laying (hAEL) (**Fig. 3**). At 20–24 hAEL most of

13   nuclei arrive at cell surface where cellularization takes place, whereas nuclei divide deep

14   interior of the egg at 4–8 hAEL (**Fig. 3**). Superficial nuclei start appearing at 8–12 hAEL

15   (**Fig. 3**). We thus injected mixture of sgRNA and Cas9 into eggs collected within eight hours,

16   which appear to be prior to cellularization.

17   To test whether the injection of sgRNA and Cas9 protein caused indels, we used a microchip

1   electrophoresis system to perform a heteroduplex mobility assay on the PCR products that

2   include the targeted region. We observed multiple heteroduplex bands suggestive of indels in

3   injected embryos (**Fig. 4A**). Such heteroduplex bands were detected for samples treated with

4   either of two sgRNAs (#1 and #2), suggesting two independent sgRNAs work similarly. We

5   then analyzed the *Td-w* nucleotide sequences of the targeted site of the sgRNA#1-injected

6   eggs, and found various indels (**Fig. 4B**). We performed another injection of sgRNA#1 and

7   observed its effect on the phenotype; 32.4% of the injected embryos hatched (**Table 1**).

8   Consistent with the genotypes, 32.6% of the hatched nymphs presented a loss of black

9   coloration in some ommatidia (Arrowhead in **Fig. 5A, B, Table 1**). *T. domestica* nymphs

10   have a pale red color in some body regions, including the lateral to posterior margin of the

11   thoracic segments (Arrows in **Fig. 5A**). In addition to the eye coloration defect, 80.4% of the

12   hatched nymphs presented a lack of body pigmentation in at least a portion of those body

13   regions (Arrow in **Fig. 5B, Table 1**). Taken together, these results indicate that the loss of

14   coloration in the eye and body was caused by indels at the *Td-w* locus, because of

15   CRISPR/Cas9 in at least some somatic cells.

16   To investigate whether CRISPR/Cas9 induced indels also in germ cells, we individually

17   crossed each of the 24 injected adults showing body color defects with a wildtype individual,

1   and successfully obtained offspring from 23 of the 24 pairs (95.8%) (**Table 2**). To screen for

2   possible knockouts within the G1 generation, we performed a high-resolution melt assay for

3   at least two individual nymphs from each line. Thirteen pairs presented different melting

4   patterns from the wildtype (data not shown). We then checked the G1 genotypes of those 13

5   pairs and found indels in 69.2% of the checked pairs (**Table 2**). We further obtained the next

6   generation (G2) from the *Td-w* knockout allele heterozygotes, and found homozygotes which

7   present complete loss of compound eye coloration (**Fig. 5C–E'**). The homozygotes presented

8   a loss of coloration of the dorsal head capsule (Asterisks in **Fig. 5E and E'**) and of the

9   segment margins of the thorax (Arrows in **Fig. 5F, F'**) and abdomen (Arrows in **Fig. 5G, G'**).

10   Loss of coloration on other body parts such as antennae, maxillary palps, legs, abdominal

11   styli, caudal filament, and cerci are displayed in adults of *Td-w* knockouts (Fig. 6). These

12   observation confirmed that *Td-w* plays a role in the pigmentation of both the eye and body.

13   The ABC transporter has been implicated in the evolutionarily conserved function of the eye

14   color because the crustacean species *Daphnia pulex* encodes *white* and *scarlet* (Sturn *et al.*,

15   2009). However, this functional evidence has been restricted to holometabolous insects.

16   Although it is beyond the scope of this study, the *white* ortholog function observed in *T.*

17   *domestica* supports the proposition that the role of ABC transporters in the eye and body

color is evolutionarily conserved among insects.

The rate of CRISPR/Cas9-mediated gene knockout in somatic cells (80.4%) observed in *T. domestica* is comparable with the somatic transformation efficiency previously observed in other species. For example, the early reports of CRISPR/Cas9-mediated knockout efficiency in *D. melanogaster*, *T. castaneum*, and *B. mori* were approximately 88%, 80%, and 95.6%, respectively (Bassett *et al*. 2013; Wang *et al*., 2013; Gilles *et al*. 2015). Similarly, the germline transmission efficiency in *T. domestica* (39.1%) is comparable with that reported for other species, such as *D. melanogaster*, (~58%), *T. castaneum* (~100%), and *B. mori* (35.6%; Bassett *et al*. 2013; Wang *et al*., 2013; Gilles *et al*. 2015).

In conclusion, our results demonstrate that the CRISPR/Cas9 system provides efficient heritable genome editing in the apterygote insect *T. domestica*. In addition to the RNAi-mediated gene knockdown (Ohde *et al*., 2009), germline genome editing using CRISPR/Cas9 in *T. domestica* provides a platform technology that opens up a new avenue toward understanding ancestral development in wingless insects and evolution of the wing development program.

Acknowledgements

5

6   References

7   Abraham, E. G., Sezutsu, H, Kanda, T, Sugasaki, T, Shimada, T and Tamura, T. 2000.

8   Identification and characterisation of a silkworm ABC transporter gene homologous to

9   *Drosophila white*. Mol. Gen. Genet. 264, 11–19.

10   Ansai, S., Inohaya, K., Yoshiura, Y., Schartl, M., Uemura, N. *et al*. 2014. Design , evaluation ,

11   and screening methods for efficient targeted mutagenesis with transcription activator-like

12   effector nucleases in medaka. 98–107.

13   Averof, M. and Cohen, S. M. 1997. Evolutionary origin of insect wings from ancestral gills.

14   Nature 385, 627–630.

15   Bassett, A. R., Tibbit, C., Ponting, C. P. and Liu, J.-L. 2013. Highly efficient targeted

16   mutagenesis of *Drosophila* with the CRISPR/Cas9 system. Cell Rep. 4, 220–228.

17   Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in

1   phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

2   Chen, J., Zhang, X., Wang, T., Li, Z., Guan, G. and Hong, Y. 2012. Efficient detection,

3   quantification and enrichment of subtle allelic alterations. DNA Res. 19, 423–433.

4   Clark-Hachtel, C. M. and Tomoyasu, Y. 2016. Exploring the origin of insect wings from an

5   evo-devo perspective. Curr. Opin. Insect Sci. 13, 77–85.

6   Dahlem, T. J., Hoshijima, K., Jurynec, M., Gunther, D., Starker, C. G. *et al*. 2012. Simple

7   methods for generating and detecting locus-specific mutations induced with TALENs in the

8   zebrafish genome. PLoS Genet. 8, e1002861.

9   Doudna, J. A. and Charpentier, E. 2014. The new frontier of genome engineering with

10   CRISPR-Cas9. Science 346, 1258096.

11   Dermauw, W. and Van Leeuwen, T. 2014. The ABC gene family in arthropods: Comparative

12   genomics and role ininsecticide transport and resistance. Insect Biochem. Mol. Biol. 45, 89–

13   110.

14   Edgar R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high

15   throughput. Nucleic Acids Research. 32, 1792–1797.

16   Ewart, G. D. and Howells, A. J. 1998. ABC transporters: biochemical, cellular, and molecular

17   aspects. Methods Enzymol. 292, 213–224.

1 Gilles, A. F., Schinko, J. B. and Averof, M. 2015. Efficient CRISPR-mediated gene targeting

2 and transgene replacement in the beetle *Tribolium castaneum*. Development 142, 2832–2839.

3 Grabherr M. G., Haas B. J., Yassour M., Levin J. Z., Thompson D. *et al*. 2011. Full-length

4 transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. 15,

5 644–652.

6 Grubbs, N., Haas, S., Beeman, R. W. and Lorenzen, M. D. 2015. The ABCs of eye color in

7 *Tribolium castaneum*: Orthologs of the *Drosophila white*, *scarlet*, and *brown* genes. Genetics

8 199, 749–759.

9 Hsu, P. D., Lander, E. S. and Zhang, F. 2014. Development and applications of CRISPR-Cas9

10 for genome engineering. Cell 157, 1262–1278.

11 Kukalová-Peck, J. 1983. Origin of the insect wing and wing articulation from the arthropodan

12 leg. Can. J. Zool. 61, 1618–1669.

13 Kuwana, H., Shimizu-Nishikawa, K., Iwahana, H. and Yamamoto, D. 1996. Molecular

14 cloning and characterization of the ABC transporter expressed in trachea (*ATET*) gene from

15 *Drosophila melanogaster*. Biochim. Biophys. Acta 1309, 47–52.

16 Letunic, I. and Bork, P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display

17 and annotation of phylogenetic and other trees. Nucleic Acids Research, 44, W242–W245.

1   Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A. *et al.* 2014. Phylogenomics

2   resolves the timing and pattern of insect evolution. Science 346, 763–767.

3   Nakamura, T., Gehrke, A. R., Lemberg, J., Szymaszek, J. and Shubin, N. H. 2016. Digits and

4   fin rays share common developmental histories. Nature 537, 1–16.

5   Niwa, N., Akimoto-Kato, A., Niimi, T., Tojo, K., Machida, R. and Hayashi, S. 2010.

6   Evolutionary origin of the insect wing via integration of two developmental modules. Evol.

7   Dev. 12, 168–176.

8   Ohde, T., Masumoto, M., Yaginuma, T. and Niimi, T. 2009. Embryonic RNAi analysis in the

9   firebrat , *Thermobia domestica* : *Distal-less* is required to form caudal filament. J. Insect

10   Biotechnol. Sericol. 105, 99–105.

11   Porazinski, S. R., Wang, H. and Furutani-seiki, M. 2010. Microinjection of medaka embryos

12   for use as a model genetic organism. J. Vis. Exp. 46, doi:10.3791/1937.

13   Snodgrass, R. E. 1935. Principles of insect morphology. McGraw Hill, New York.

14   Sturm, A., Cunningham, P. and Dean, M. 2009. The ABC transporter gene family of *Daphnia*

15   *pulex*. BMC Genomics 10, 1–18.

16   Wang, Y., Li, Z., Xu, J., Zeng, B., Ling, L. *et al.* 2013. The CRISPR/Cas System mediates

17   efficient genome engineering in *Bombyx mori*. Cell Res. 23, 1414–1416.

1   Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large

2   biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The

3   University of Texas at Austin.

4

Figure Captions

Fig. 1 The paranotum in *T. domestica*

A, Lateral view of the female adult *T. domestica*. **B,** A transverse section of the *T. domestica*

thorax stained with hematoxylin and eosin Y. The arrow indicates the paranotum.

1 **Fig. 2 Identification of a *white* ortholog in *T. domestica***

2 **A,** Phylogenetic tree of eye-color ABC transporters. Bootstrap values after 1000 replicates

3 are shown for each node. **B,** Schematic diagram of the Td32270 domain architecture. The

4 arrow indicates the site where sgRNA is designed.

5 Ae, *Aedes aegypti*; Ag, *Anopheles gambiae*; Am, *Apis mellifera*; Ar, *Athalia rosae*; Atet, ABC

6 transporter expressed in trachea; Bm, *Bombyx mori*; Dm, *Drosophila melanogaster*; Ph,

7 *Pediculus humanus corporis*; Tc, *Tribolium castaneum*; Td, *Thermobia domestica*.

8

1　**Fig. 3 Eggs in eight hours after laying are in early cleavage stage**

2　*T. domestica* embryos stained with SYTOX Green. Labels indicate hours after egg laying. A

3　number in parenthesis is the number of samples. A representative image of an embryo for

4　each time window is displayed. Scale bar is 0.5 mm.

5

1   **Fig. 4 Microinjection of sgRNA and Cas9 protein induces indels in *T. domestica***

2   **A,** Heteroduplex shift assay for *Td-w* sgRNA/Cas9 protein injected embryos (G0). Asterisks

3   indicate samples showing multiple heteroduplex bands. The lane 'C' represents no injection

4   control. **B,** Nucleotide sequences around the *Td-w* CRISPR/Cas9 target site in *Td-w*

5   sgRNA#1/Cas9 protein injected embryos (G0). The target sequence of our *Td-w* sgRNA#1 is

6   underlined. Protospacer adjacent motif is represented as a blue box. The red dash and

7   character indicate detected mutations. The number of deleted (−) or inserted (+) nucleotide(s)

8   and the number of sample(s) that showed each nucleotide sequence (N = 19) are indicated at

9   the right side of the sequences.

10

1   **Fig. 5 CRISPR/Cas9-based germline genome editing in *T. domestica***

2   **A, B,** Nymph phenotypes in injected generation (G0). Lateral view of the first instar nymphs

3   after no injection (A) and *Td-w* sgRNA#1/Cas9 protein injection (B). **C–G',** G2 nymph

4   phenotype. Dorsal view of the first instar nymphs of the wild type (C) and *Td-w* knockout

5   allele homozygote (D). The regions in white boxes in C and D are magnified in E–G and E'–

6   G', respectively. The red arrowhead indicates a compound eye. The arrow indicates

7   coloration at the tergal edge. T1, prothorax; A1, first abdominal segment. Scale bar

8   corresponds to 0.25 mm in A–D, and 0.1 mm in E–G'.

9

1   **Fig. 6 *Td-w* mutant homozygotes display loss of coloration in body parts of adults**

2   **A,** Dorsal view of adult heads of the wildtype (top) and *Td-w* mutant homozygote (bottom).

3   Arrowheads indicate eyes. Arrows indicate segments on maxillary palps that are colored in

4   the wildtype. Asterisks indicate antennae. **B, C,** Legs on prothorax, mesothorax and

5   metathorax (left to right) in the wildtype (B) and *Td-w* mutant homozygote (C). Arrows

6   indicate colored regions in the wildtype. **D, E,** Ventral view of caudal regions in the wildtype

7   (D) and *Td-w* mutant homozygote (E). Arrows indicate abdominal styli. Asterisks indicate

8   cerci. Scale bars are 1 mm.

1 **Table 1 Summary of G0 phenotype after *Td-w* sgRNA/Cas9 protein injection**

No. of injected egg	No. (%) of hatched egg	No. (%) of nymph with each phenotype		
		Wildtype	Loss of body pigmentation	White ommatidium
145	46 (32.4)	10 (15.6)	37 (80.4)	15 (32.6)

2

1 **Table 2 Summary of fertility and G1 genotype in *Td-w* knockouts**

Sex of *Td-w* sgRNA/Cas9 protein injected G0	No. crossed to wildtype	No. (%) of fertile pair	No. of pairs in which offspring genotype was checked	No. (%) of line with Indels
Male (M)	6	6 (100)	4	3 (50.0)
Female (F)	18	17 (94.4)	9	6 (35.3)
Total (M+F)	24	23 (95.8)	13	9 (39.1)

2

A Thorax

B

Fig. 1 (color)

A

White subfamily

Td 30354

Tc Atet
Dm Atet

Dm W
Aa W
Ag W
Ph W
Tc W
Bm W-3
Ar W
Td 32270

1000

773
1000
573
671
997
992

0.5 substitution per site

920 614
970 849 836
986

Tc Bw

Dm St

Ag St
Tc St
Bm W-2
Ph St
Am St
Td 23978

Am Bw

Dm Bw
Ag Bw

Scarlet subfamily

Brown subfamily

B

#1 #2

ATP-binding domain   Transmembrane helices

Fig. 2

Fig. 3 (color)

A *Td-w* sgRNA#1 + Cas9 protein

C * * * * * *

Td-w sgRNA#2 + Cas9 protein

C * * * * *
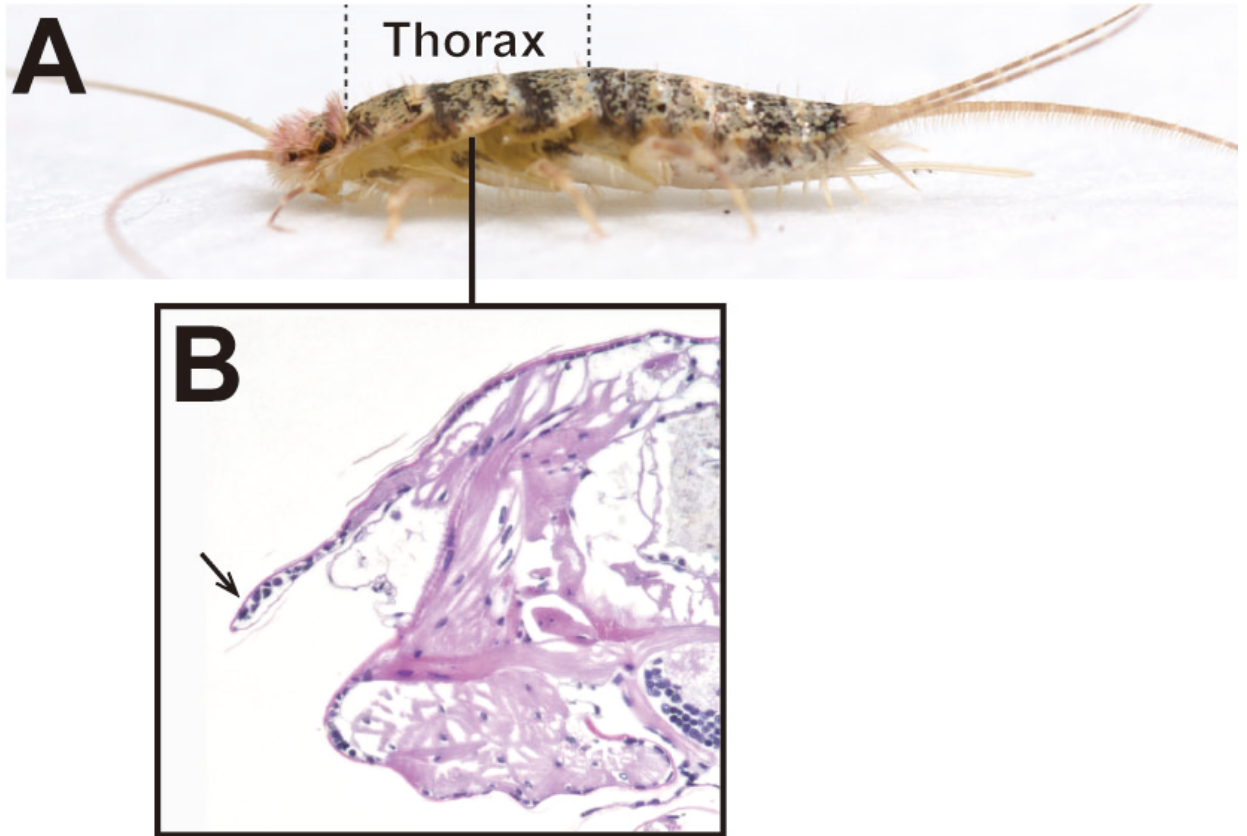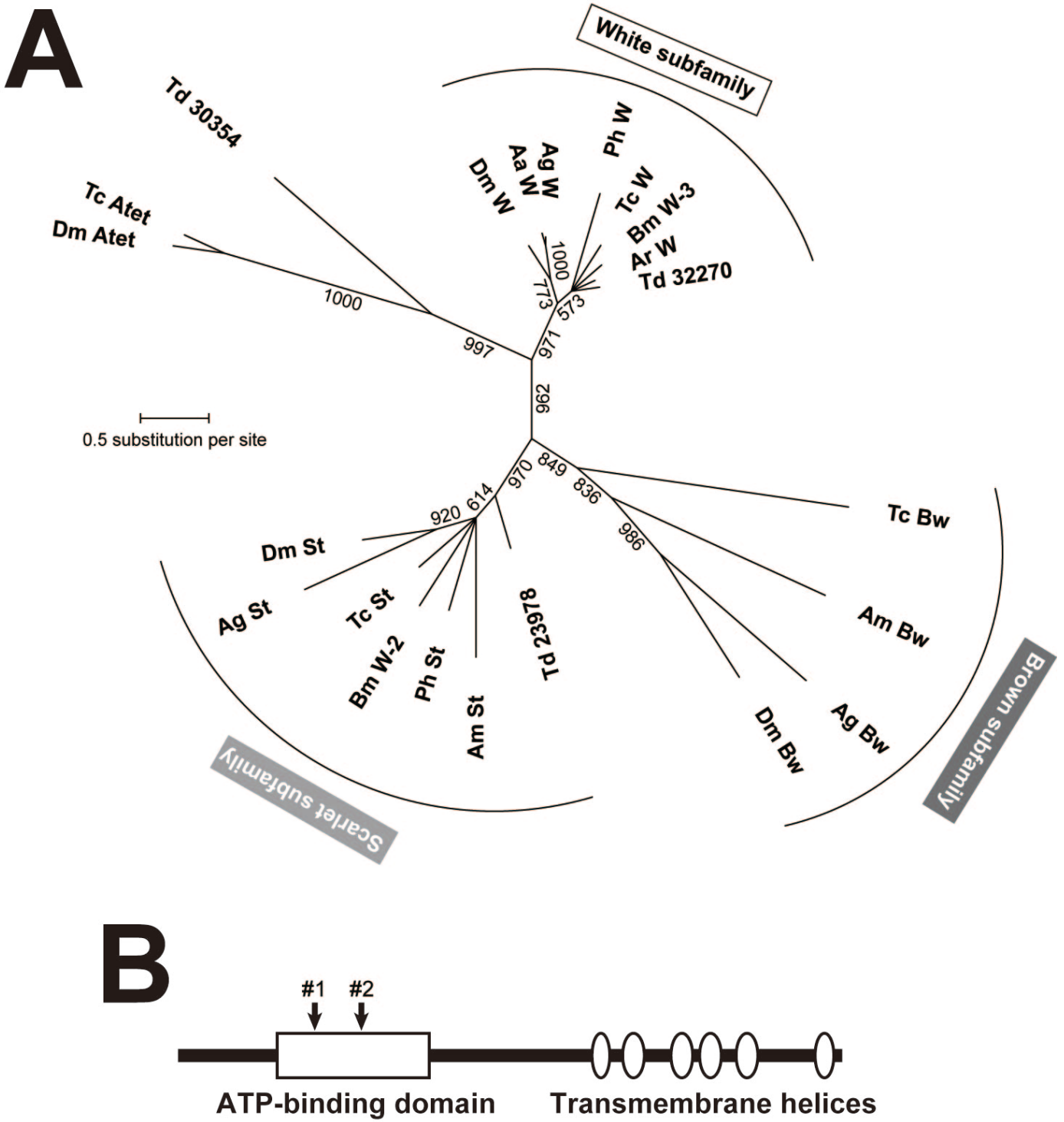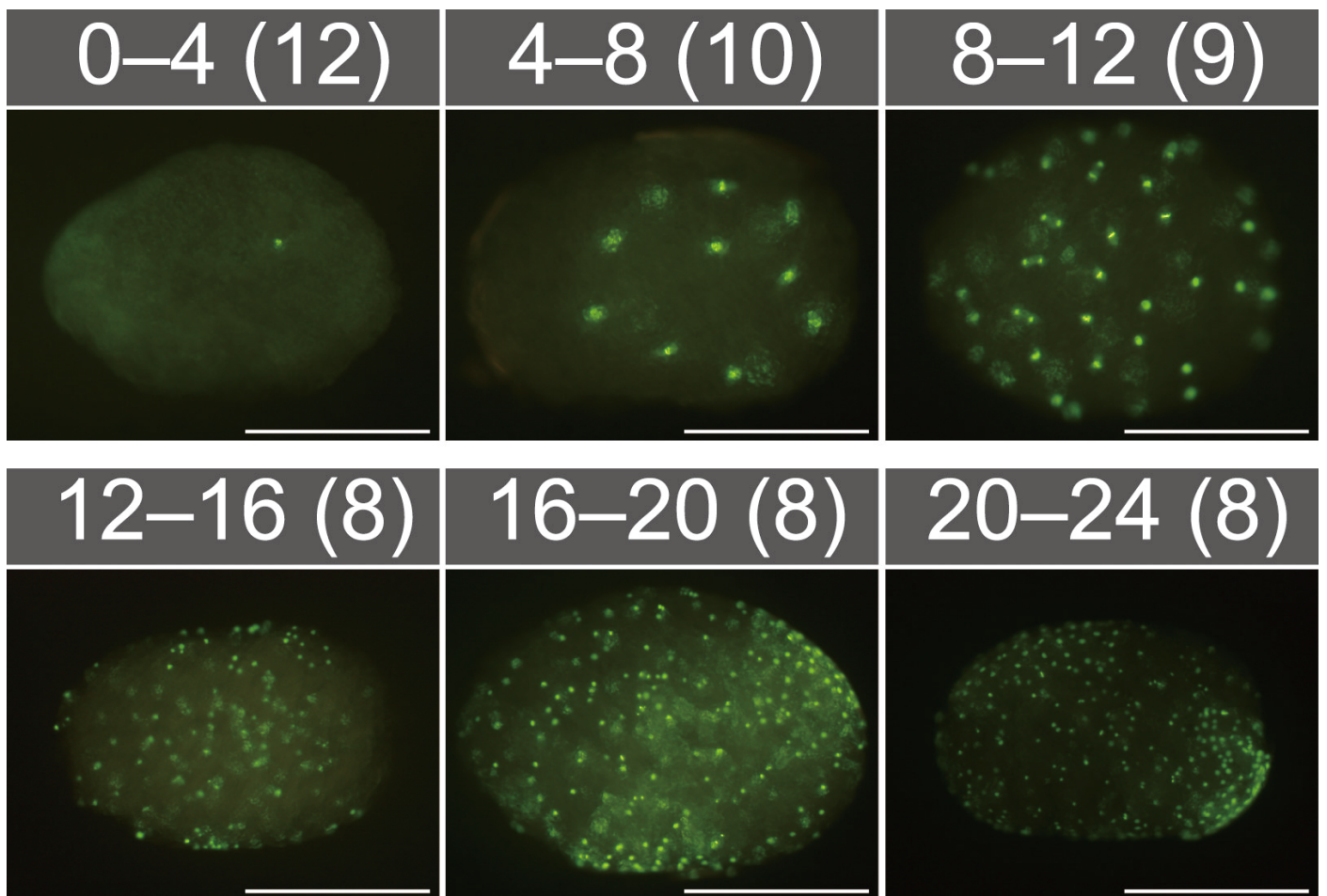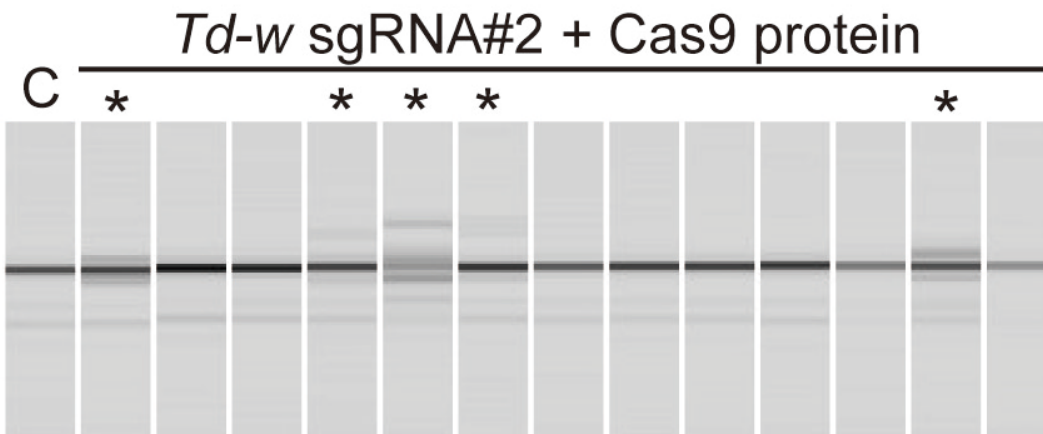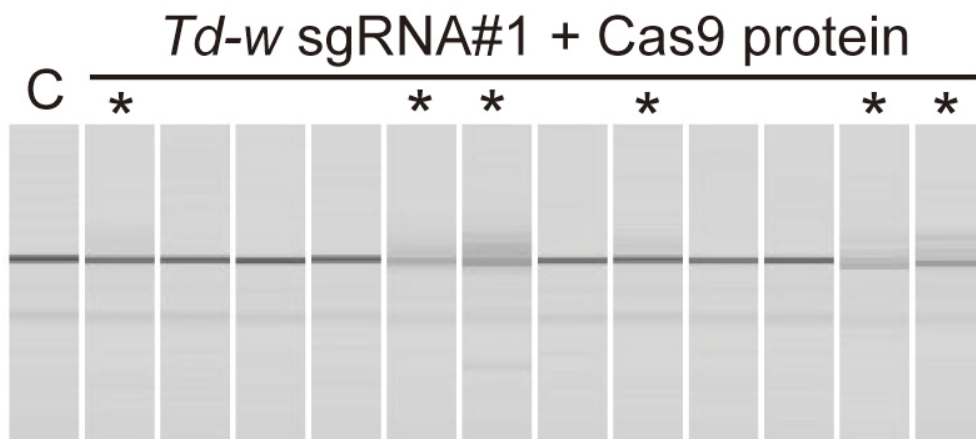
B

Deletions

```
GACAGAGCAGTGAACGGAAAACCAATCGGTCCCACAACACTTACCAGCTT   WT      3/19
GACAGAGCAGTGAACGGAAAACCAATC-GTCCCACAACACTTACCAGCTT   -1      1/19
GACAGAGCAGTGAACGGAAAACCAA----TCCCACAACACTTACCAGCTT   -4      4/19
GACAGAGCAGTGAACGGAAAACCA----GTCCCACAACACTTACCAGCTT   -4      1/19
GACAGAGCAGTGAACGGAAAACCAA----TCCC-CAACACTTACCAGCTT   -5      1/19
GACAGAGCAGTGAACGGAAAACCAA----TCC----ACACTTACCAGCTT   -8      1/19
GACAGAGCAGTGAACGGAAAACCAA------------CACTTACCAGCTT   -12     1/19
GACAGAGCAGTGAACG---ACCCA----------CAACACTTACCAGCTT   -13     1/19
GACAGAGCAGTGAACGGAAAACCAAA--GTCCCACAACACTTACCAGCTT   -3,+1   1/19
```

Insertions

```
GACAGAGCAGTGAACGGAAAACCAA-TCGGTCCCACAACACTTACCAGCT    WT
GACAGAGCAGTGAACGGAAAACCAAAATAGTCCCACAACACTTACCAGCT    -3,+4   1/19

GACAGAGCAGTGAACGGAAAACCAATCGG-TCCCACAACACTTACCAGCT    WT
GACAGAGCAGTGAACGGAAAACCAATCCGTTCCCACAACACTTACCAGCT    +1      1/19

GACAGAGCAGTGAACGGAAAACCAATCG-GTCCCACAACACTTACCAGCT    WT
GACAGAGCAGTGAACGGAAAACCAAACCAGTCCCACAACACTTACCAGCT    -3,+4   1/19

GACAGAGCAGTGAACGGAAAACCA----ATCGGTCCCACAACACTTACCA    WT
GACAGAGCAGTGAACGGAAAACCAGTGAACCGGTCCCACAACACTTACCA    -2,+6   1/19

GACAGAGCAGTGAACGGAAAACCAATC-------GGTCCCACAACACTTA    WT
GACAGAGCAGTGAACGGAAAACCAATCCCACACCAGTCCCACAACACTTA    -2,+8   1/19
```
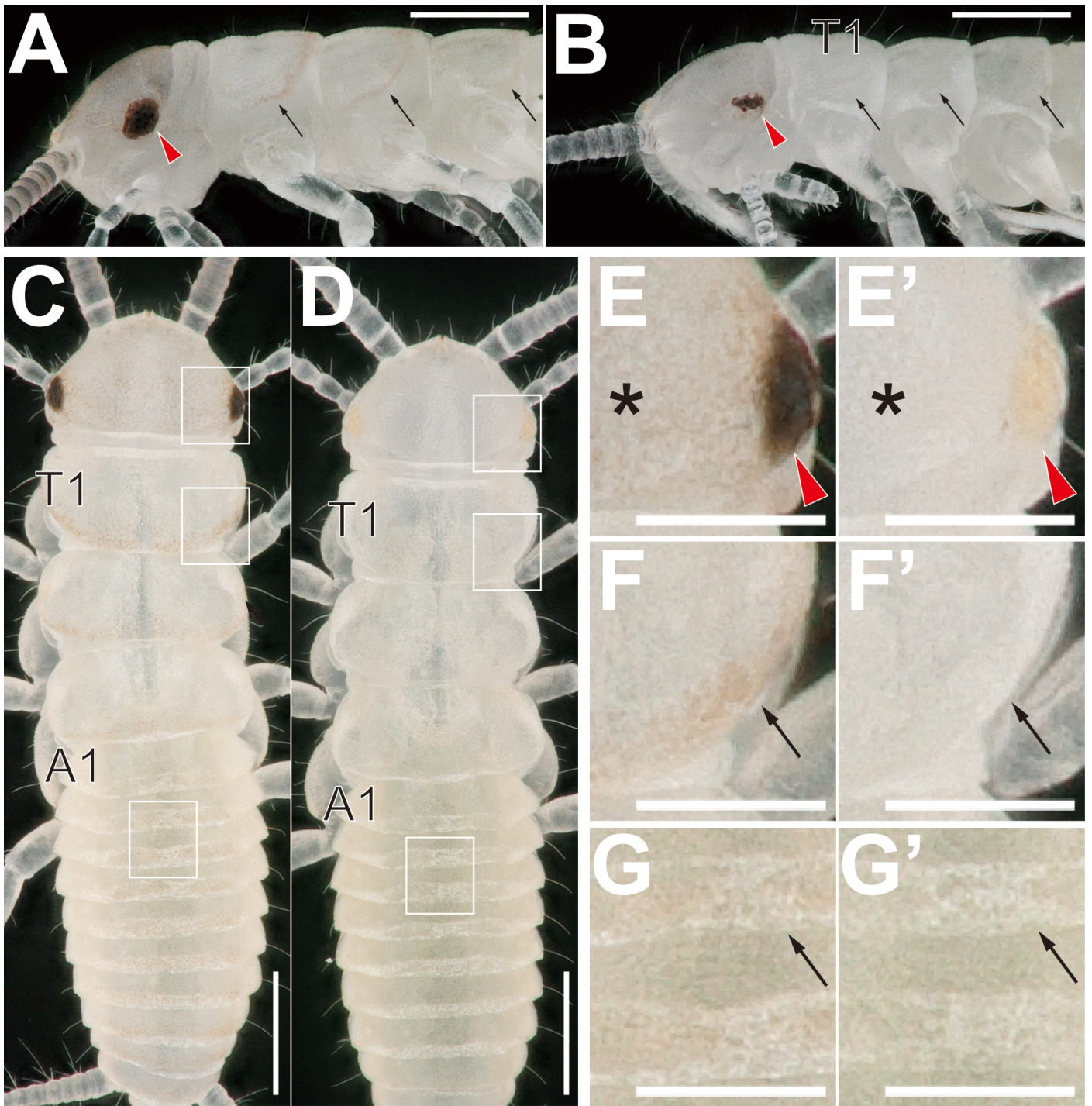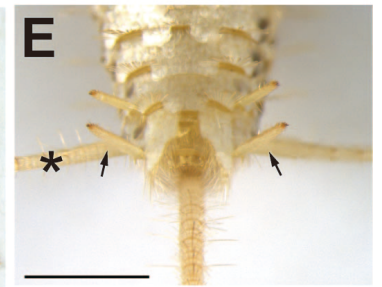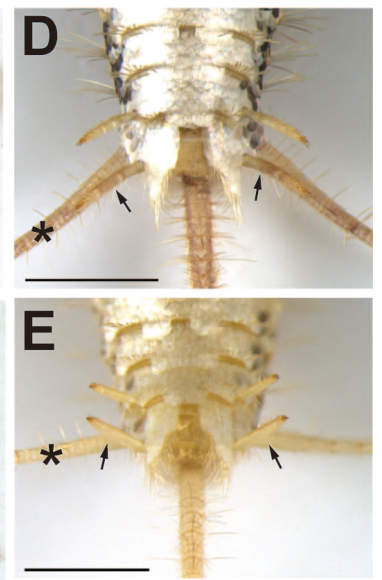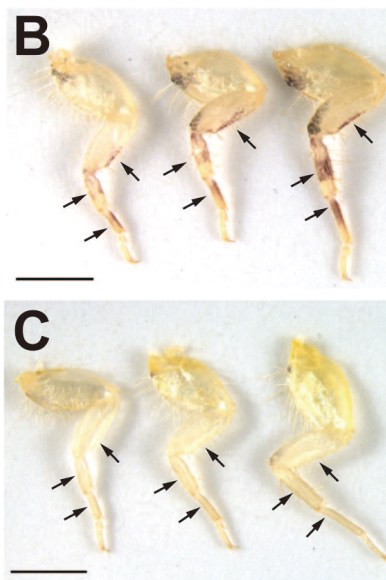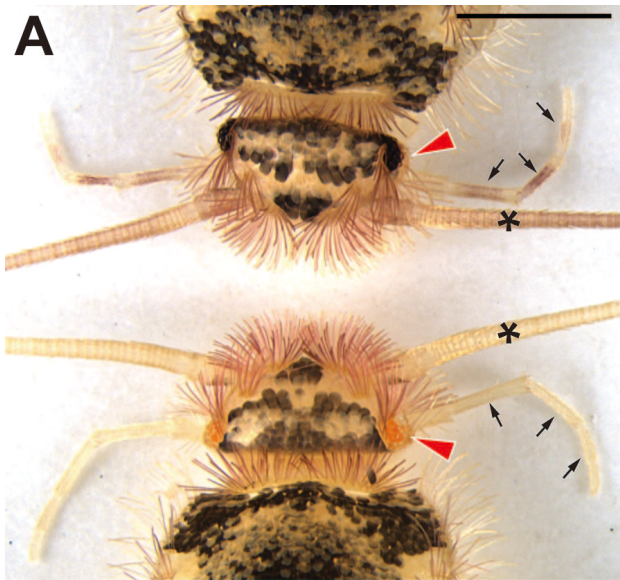
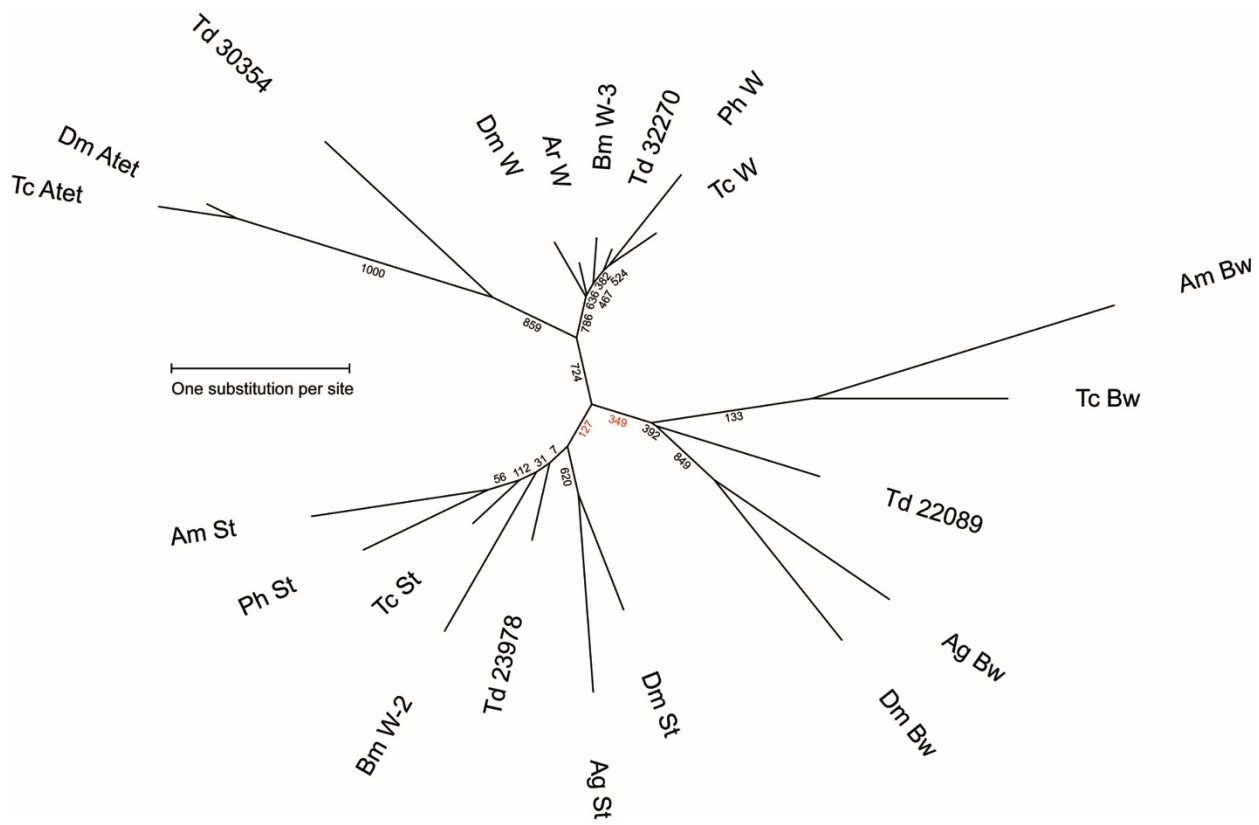Fig. 4 (color)

Fig. 5 (color)

1 **Supplementary figure**

2 Phylogenetic tree of eye-color ABC transporters with Td22089. Bootstrap values after 1000

3 replicates are shown for each node. Low bootstrap values for Brown and Scarlet subfamilies are

4 shown in red characters.

1

2

3

4

5