



HOKKAIDO UNIVERSITY

Title	LIM-homeodomain transcription factor Awh is a key component activating all three fibroin genes, fibH, fibL and fhx, in the silk gland of the silkworm, Bombyx mori
Author(s)	Kimoto, Mai; Tsubota, Takuya; Uchino, Keiro et al.
Citation	Insect biochemistry and molecular biology, 56, 29-35 https://doi.org/10.1016/j.ibmb.2014.11.003
Issue Date	2015-01
Doc URL	https://hdl.handle.net/2115/58056
Rights	(C) 2014 Elsevier Ltd. All rights reserved.
Type	journal article
File Information	IBMB_56_p29-.pdf



LIM-homeodomain transcription factor Awh is a key component activating
all three fibroin genes, *fibH*, *fibL* and *fhx*, in the silk gland of the silkworm,
Bombyx mori

Mai Kimoto^a, Takuya Tsubota^b, Keiro Uchino^b, Hideki Sezutsu^b and
Shigeharu Takiya^{a,c,1}

^aGraduate School of Life Science, Hokkaido University; ^bTransgenic
Silkworm Research Unit, National Institute of Agrobiological Sciences,
Owashi 1-2, Tsukuba, Ibaraki 305-8634; ^cDivision of Biological Sciences and
Center for Genome Dynamics, Faculty of Science, Hokkaido University,
North 10, West 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan

¹To whom correspondence should be addressed.

Division of Biological Sciences and Center for Genome Dynamics, Faculty of
Science, Hokkaido University, North 10, West 8, Kita-ku, Sapporo, Hokkaido
060-0810, Japan

E-mail: takiya@sci.hokudai.ac.jp

Phone: +81-11-706-3590, Fax: +81-11-706-3588

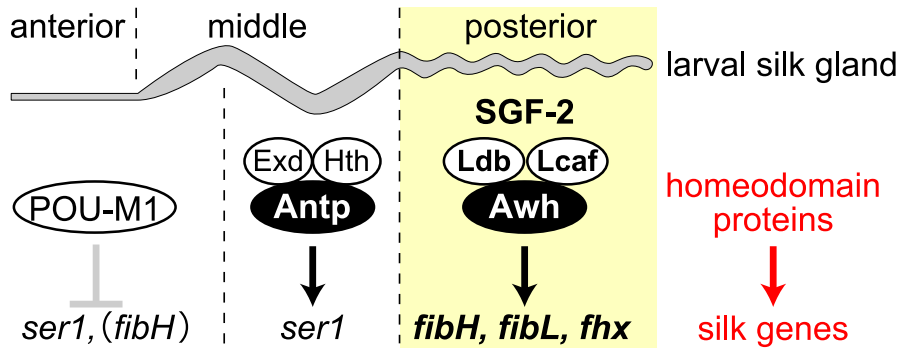
Keywords: *Bombyx mori*; silk gland; *fibroin*; Awh; transgenic

The authors declare no conflict of interest.

Abstract

In the silkworm *Bombyx mori*, three fibroin genes, *fibroin-heavy-chain* (*fibH*), *fibroin-light-chain* (*fibL*) and *fibrohexamerin* (*fhx*), are coexpressed only in the posterior silk gland (PSG) cells, while the sericin genes encoding silk glue proteins are expressed in the middle silk gland (MSG) cells. Silk gland factor-2 (SGF-2) is a PSG-specific activator complex of *fibH*, composed of a LIM-homeodomain protein, Awh, and its cofactors, Ldb and Lcaf. We investigated whether SGF-2 can activate other fibroin genes using transgenic silkworms. The genes for Ldb and Lcaf were expressed ubiquitously in various tissues, while the gene for Awh was expressed strictly specific in PSG of the wild type silkworms. Misexpression of *Awh* in transgenic silkworms induced ectopic expression of *fibL* and *fhx* as well as *fibH* in MSG. Coincidentally with the induction of *fibL* and *fhx* by Awh, binding of SGF-2 to the promoter of *fibL* and *fhx* was detected *in vitro*, and SGF-2 binds directly to the *fhx* core promoter. Ectopic expression of the fibroin genes was observed at high levels in the middle part of MSG. Moreover, *fibL* and *fhx* were induced in the anterior silk gland (ASG) of the transgenic silkworms, but *fibH* was not. These results indicate that Awh is a key activator of all three fibroin genes, and the activity is probably regulated in conjunction with additional factors.

44 **Graphical abstract**



45

46

Highlights

- *Awh* encoding LIM-homeodomain protein, which constitutes SGF-2, is expressed specifically in the posterior silk gland (PSG).
- Genes for other components of SGF-2, Ldb and Lcaf, are expressed in most tissues.
- Misexpression of *Awh* induces expression of the *fibH*, *fibL* and *fhx* genes in the middle silk gland (MSG).
- SGF-2 can activate *fhx* through binding to its core promoter.
- Ectopic expression of the fibroin genes was detected only in the silk gland with sublocal specificity in *Awh* transgenic silkworms.

Abbreviations: ASG, anterior silk gland; Awh, Arrowhead; EMSA, electrophoretic mobility shift assay; FMBP-1, fibroin-modulator-binding protein-1; Lcaf, LIM-homeodomain and Ldb complex-associated factor; Ldb, LIM domain -binding protein; MSG, middle silk gland; PSG, posterior silk gland; SGF-2, silk gland factor-2; *ser1*, *sericin-1*

1. Introduction

The silkworm *B. mori* spins a small amount of silk threads during all larval instars except for each molting stage, and produces massive amount of silk proteins to make a cocoon at the end of the last instar. The silk fibroin synthesized by *B. mori* consists of a fibroin-heavy-chain (FibH), fibroin-light-chain (FibL) and fibrohexamerin (Fhx) proteins, with a 6:6:1 ratio (Inoue *et al.*, 2000). FibH and FibL are linked by a disulfide bond and this linkage is essential for secretion of these proteins from silk gland cells. Fhx associates with the FibH-FibL complex through non-covalent interactions and is thought to play a role in maintaining the integrity of the fiber. The core silk fibers are coated by sericins, a family of glue proteins that bind the fibroin fibers together. These silk proteins are synthesized in the silk gland with strict regional specificity regulated at the transcriptional level. The sericin genes (*ser1*, *2* and *3*) are expressed in the middle silk gland (MSG), which is further divided into three subparts along the anteroposterior axis based on expression patterns of the sericin genes (Couple *et al.*, 1987; Takasu *et al.*, 2010). Three fibroin genes (*fibH*, *fibL* and *fhx*) are expressed in a similar expression pattern specifically in the posterior silk gland (PSG), suggesting that their expression is coordinately regulated, though these genes are located on different chromosomes (Bello *et al.*, 1994; Inoue *et al.*, 2005).

The promoter of *fibH* has been analyzed in detail. Cell-free transcription assays using silk gland extracts revealed that a 200 base-pair upstream region of the promoter is essential for its transcriptional activity (Tsuda and Suzuki, 1983). Sequence -30 to +6 acts as the core promoter and the proximal region from -94 to -32 is occupied by SGF-1/BmFkh, which is proposed to play important roles in silk gland development and silk gland-specific expression of the silk genes (Tsuda and Suzuki, 1981; Takiya *et al.*, 1990; Hui *et al.*, 1990; Julien *et al.*, 2002; Mach *et al.*, 1995; Takiya *et al.*, 1997). Recently, it was reported that Bmsage, a basic helix-loop-helix transcription factor expressed specifically in the silk gland, stimulates expression of *fibH* through interaction with Fkh, and is probably responsible for the massive production of the silk proteins at the last instar (Zhao *et al.* 2014). The distal part from -238 to -73 of *fibH* promoter contains a cluster of homeodomain-binding sites and is required for PSG-specific transcriptional enhancement (Suzuki *et al.*, 1986; Hui *et al.*, 1990). These homeodomain binding sites are bound by SGF-2, SGF-3/POU-M1 and FMBP-1, though FMBP-1 is not a homeodomain protein (Takiya *et al.*, 1997, 2005; Hui *et al.*, 1990; Fukuta *et al.*, 1993).

The SGF-2 complex is detected specifically in PSG (Hui *et al.* 1990), and mainly consists of a LIM-homeodomain transcription factor Arrowhead (Awh), LIM domain-binding protein (Ldb) and Lcaf belonging to a sequence-specific single-stranded DNA-binding protein (SSDP) family (Ohno

et al., 2013). Ldb proteins are known as cofactor for LIM-homeodomain proteins (Matthews and Visvader, 2003), and SSDP proteins are presumed to play a role in the stabilization of Ldb (Chen *et al.*, 2002; Xu *et al.*, 2007; Wang *et al.*, 2010). Furthermore, Lcaf would mediate long-range enhancer-promoter communication on chromatin by its self-oligomerization activity (Ohno *et al.*, 2013).

Enforced expression of *Awh* in transgenic silkworms induces ectopic expression of *fibH* in MSG, but, it has not been tested whether misexpression of *Awh* activates two other fibroin genes *fibL* and *fhx*, and whether is sufficient for activation of *fibH* in other tissues. In this study, we show that *Awh* is expressed in PSG specifically in wild type silkworms, while other components of SGF-2 are expressed ubiquitously in most tissues. Overexpression of *Awh* using transgenic silkworms resulted in ectopic expression of the *fibH*, *fibL* and *fhx* genes in the silk gland with subregional specificities. However, tissues other than the silk gland remained unaffected.

2. Materials and Methods

2.1. Animals

This study used w-c as a wild-type strain of *B. mori*. The transgenic strains used in this study are also maintained in w-c genetic background. Larvae were reared at 25°C under light–dark (LD) exposure of 16:8 hours on an artificial diet obtained from Nippon Nosan-Kogyo (Yokohama, Japan) and staged as described previously (Kiguchi and Agui, 1981; Suzuki *et al.*, 1990). Tissue extracts were prepared using the Kinshu x Showa strain purchased from Ueda Sanshu (Ueda, Japan).

2.2. Transgenic silkworms and heat-shock treatment

Awh transgenic silkworms were prepared as described in Ohno *et al.* (2013). The ORF was amplified by PCR, and inserted into the *BlnI* site of pBacUASMCS vector (Sakudoh *et al.*, 2007) with [3xP3-AmCyan] screening marker. The plasmid was injected into *w1-pnd* embryos. The established UAS-*Awh* strains were crossed with *hsp70*-GAL4 strain (Uchino *et al.*, 2006). Larvae were placed in 100-ml flasks plugged with cotton. Flasks were then submersed in a water bath and heated at 42°C for 2 hours (Uhlirova *et al.*,

2002; Kimoto *et al.*, 2014). Following the heat shock, larvae were returned to food.

2.3. RNA extraction and RT-PCR

Total RNA was extracted using Illustra RNAspin MINI RNA Isolation kit (GE Healthcare), and cDNA was synthesized using PrimeScript RT-PCR kit (Takara) as described previously (Kimoto *et al.*, 2012). PCR amplification was performed using ExTaqHS polymerase (Takara) with primers and appropriate cycles listed in Table S1. PCR products were analyzed on 1% agarose gels.

2.4. Cloning of the *Lcaf* gene

For RT-PCR, the region from 61 to 669 of *Lcaf* (Genbank accession number: AB687556) was amplified with the primer sets indicated in Table S1. The sequence obtained differed partly from the previously reported *Lcaf* transcript (Ohno *et al.*, 2013), and we registered the novel sequence in Genbank (AB910676). The sequence from 520 to 623 of AB687556 was replaced by the sequence from 460 to 515 of AB910676. This difference might be attributed to the distinct silkworm strains, Kinshu x Showa and w-c.

2.5. Electrophoretic mobility shift assay (EMSA)

EMSA was carried out as described previously (Takiya *et al.* 1997, 2005; Ohno *et al.* 2013) with several modifications. Tissue extracts were prepared from entire MSG on day 2 of the fifth instar and PSGs on the fourth molting stage or day 2 of the fifth instar. Double-stranded oligonucleotides (50 nucleotide long) listed in Table S2 were labeled with Klenow polymerase and used as probes. The probes were incubated in each tissue extract (10 µg protein) on ice for 20 min., and analyzed by electrophoresis on 5% polyacrylamide gels in 50 mM Tris borate (pH8.3) and 1 mM EDTA buffer at 4°C. Glycerol was omitted from the gels. Anti-Awh antiserum was prepared by the immunization of mice with the C-terminal region (208 to 274) of recombinant Awh protein.

2.6. Whole mount *in situ* hybridization

Digoxigenin (DIG)-labeled RNA probes for *fibH* (nucleotide sequence 62425-62479 and 63451-63830, AF226688) were synthesized using a DIG RNA Labeling kit (Roche). Silk glands were isolated from wild-type (+/+) or heat-treated transgenic silkworms (hs-Gal4/UAS-Awh of day 2 fifth instars) and *in situ* hybridization was performed as described previously (Kimoto *et al.*, 2014). Coloring reactions were performed using Nitroblue

192 tetrazolium/5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) tablet
193 (Roche). Tissues were then cleared with a graded series of ethanol.
194

3. Results

3.1. Genes for *Ldb* and *Lcaf* are expressed ubiquitously, but *Awh* is expressed specifically in PSG

Ohno *et al.* (2013) showed PSG-specific expression of *Awh* in the silk gland, but the expression in other tissues has not been examined. To explore gene expression patterns of the components of SGF-2 in *B. mori*, we performed RT-PCR with total RNA from each part of the silk gland or other tissues of day 2 fifth instar larvae. Of these genes, *Ldb* and *Lcaf* were expressed in all tissues examined, in addition to their expression throughout the silk gland, whereas *Awh* was expressed strictly in PSG alone (Fig. 1). PSG-specific expression of all three fibroin genes, *fibH*, *fibL* and *fhx*, were confirmed as reported previously (Bello *et al.*, 1994).

3.2. *Awh* induces ectopic expression of the fibroin genes in MSG

The results that *Ldb* and *Lcaf* were expressed in most tissues and PSG-specific expression of *Awh* suggest that the localization of *Awh* restricts the region where a functional SGF-2 complex is formed, leading PSG-specific expression of *fibH*. We showed previously that misexpression of *Awh* induces ectopic expression of *fibH* in MSG at the fourth instar using transgenic

silkworms that allow heat-inducible expression of *Awh*. In this study, the transgenic silkworms were treated at 42°C for two hours on day 1 of the fifth instar, and gene expression was analyzed 1 day after the heat treatment. Expression of *fibH* was observed in MSG at the fifth instar as same as the fourth instar (Fig. 2A). We further investigated time course of the induction of *Awh* and *fibH* after heat treatment. *Gal4* was expressed both immediately and transiently after heat treatment, followed by *Awh* induction in MSG within 6 hours (Fig. 2B). The ectopic expression of *fibH* was then observed after another 6 hours and lasted at least 2 days. Induction of two other PSG-specific fibroin genes, *fibL* and *fhx*, was observed in MSG with almost the same time course in the transgenic silkworms (Fig. 2B). The data suggests that *Awh* acts as a common activator of these fibroin genes.

3.3 SGF-2 binds to the promoter of *fibL* and *fhx*

To investigate whether SGF-2 can bind to the promoter region of *fibL* and *fhx*, we first compared nucleotide sequence of *fibH*, *fibL* and *fhx* promoters. Although paired binding elements existing in the E and C sites of *fibH* promoter was not observed around the promoter of *fibL* and *fhx*, half elements like the -1620 region in the *fibH* upstream enhancer were found at the -230 and -50 regions of *fibL* promoter as well as the +100 region of *fhx* intron (Fig. 3).

Binding of SGF-2 to these elements was examined in EMSAs (Fig. 4A, B, C). A specific band was detected with the fibH-200 probe corresponding to the E site using PSG extracts on day 2 of the fifth instar, but not on PSG extracts of the fourth molting stage or MSG extracts on day 2 of the fifth instar (Fig 4A). The band was supershifted by the addition of anti-Awh antiserum into EMSA reactions, and was competed with fibL-50 and fhx+100 oligonucleotides as well as fibH-200 and fibH-120 oligonucleotides, suggesting that the fibL-50 and fhx+100 regions contain SGF-2-binding elements. Using these oligonucleotides as probes, EMSAs were performed, and SGF-2-like complexes were detected, though the background signals were often high under the conditions used here, and mobility of the band with the fibL-50 probe was delayed (Fig. 4B).

3.4 The *fhx* core promoter is a binding site for SGF-2

Horard *et al.* (1997) showed that PSGF site (-45 to -39) is a key element for PSG-specific expression of *fhx*, and combination of SGFB site (-74 to -56) with the PSGF site is sufficient for the PSG-specific expression. The SGFB site contains binding elements for Fkh and FMBP-1 (Fig. 3), but the sequences of the SGFB and PSGF sites do not coincide with the typical sequence (CAATTA) for SGF-2 binding. However, we found that the nucleotide sequence of the PSGF site and TATA box (TATAATTA) of *fhx*

(designated fhx-30 region) is similar to the paired SGF-2-binding element in the E site of *fibH* (Fig. 4D), and investigated whether SGF-2 can bind to the fhx-30 region. As shown in Fig. 4A, fhx-30 oligonucleotide competed the binding of SGF-2 to the fibH-200 probe, and gave shifted band as a probe in EMSA at similar position with the fibL-50 probe (Fig. 4B and C). On the other hand, fhx-30M oligonucleotide containing the canonical TATA box sequence (TATAAAAA) did not compete the SGF-2 binding as a competitor nor did it give a shifted band as a probe. The band with the fhx-30 probe was competed with fibH-200 and fibL-50 oligonucleotides as well as the fhx-30 oligonucleotide itself but not with the fhx-30M oligonucleotide (Fig. 4C). Furthermore, the band was supershifted with anti-Awh antiserum, but not with anti-POU-M1 antiserum as a control. These results indicate that the fhx-30 region is a target site of SGF-2 binding.

3.5 Subregional localization of ectopic expression of the fibroin genes in the silk gland

MSG is further divided into three distinct parts both anatomically and physiologically, referred to as the anterior, middle and posterior parts of MSG (MSG-A, MSG-M and MSG-P). We analyzed expression levels of *fibH* induced by the misexpression of *Awh* in each region of MSG. With RT-PCR, expression of *fibH* was detected clearly in MSG-M, but the levels in MSG-A

and MSG-P were extremely low in spite of the expression of *Awh* in all parts of MSG (Fig. 5A). To confirm the expression pattern of *fibH* in the silk gland of *Awh* transgenic silkworms, whole-mount *in situ* hybridization of *fibH* was performed using silk glands of heat-treated fifth instar larvae. In addition to the normal expression of *fibH* in PSG, ectopic accumulation of *fibH* mRNA was detected in MSG-M, as in the RT-PCR findings (Fig. 5B). Similarly, expression levels of the *fibL* and *fhx* genes were higher in MSG-M than in MSG-A and MSG-P of the transgenic silkworms.

We further investigated whether ectopic expression of the fibroin genes is induced in the anterior silk gland (ASG) or other tissues. Although *Awh* overexpression occurred throughout the body, expression of the fibroin genes was restricted to the silk gland (Fig. 5A). In ASG, where no silk genes are expressed in the wild-type silkworms, *fibL* and *fhx* were induced by the *Awh* misexpression, whereas the expression of *fibH* was not detected.

4. Discussion

SGF-2 was identified as a complex that binds to the E (-214/-180) and C (-133/-96) sites in the distal upstream element of *fibH* promoter and is detected specifically in PSG extracts (Hui *et al.*, 1990). Ohno *et al.* (2013) revealed that SGF-2 contains Awh, Ldb, and Lcaf proteins, and that misexpression of *Awh* induces ectopic expression of *fibH* in MSG. Since *Awh* is expressed only in PSG in the wild type silkworms, in contrast to the ubiquitous expression of *Ldb* and *Lcaf*, Awh could be a key factor in the PSG-specific expression of *fibH*.

Two other fibroin genes, *fibL* and *fhx*, were also induced in MSG by the misexpression of *Awh*, indicating that all three fibroin genes share the same transcriptional regulation mechanism for their PSG-specific expression (Figs. 2 and 5). SGF-2-binding sites were found in the immediate 5' flanking region or intron near the promoter of *fibL* and *fhx* (Fig. 3). These SGF-2-binding sites contain a half element (CAATTA) of the E or C site in *fibH* promoter, and SGF-2 can bind these half elements (Fig. 3). A key element around at -1620 in the far upstream enhancer of *fibH* also contains the half element of the E and C sites (Shimizu *et al.*, 2007).

Promoter activity of the fibroin genes was analyzed using transgenic *Drosophila* carrying the promoter of *fibH* (-1500/+66) or *fhx* (-1451/+76) fused to the *beta-galactosidase* gene, with expression of both transgenes being

observed in the salivary gland (Bello *et al.*, 1994; Nony *et al.*, 1995). It is notable that both promoters of *fibH* and *fhx* were activated in a restricted territory of the salivary gland, located between the glue-protein-secreting cells and the imarginal cells, which resemble the expression domain of the *Drosophila Awh* gene (Curtiss and Heiling, 1995, 1997). These observations support our results that the transcription of *fibH* and *fhx* is under the control of Awh.

Horard *et al.* (1997) showed that the combination of the SGFB site and PSGF site is sufficient to drive PSG-specific expression of *fhx* introduced into the silk gland with the use of a particle gun. The SGFB site contains binding-elements for Fkh and FMBP-1 (Fig. 3), and the PSGF site has been demonstrated to be a part of SGF-2-binding element in this study. The *fibL*-50 probe used covers TATA box of *fibL*. This may be a reason as to why the *fib*-50L probe gave a retarded band that was similar with the *fhx*-30 probe in EMSA. Moreover, the TATA box (TATATAAATA) of *fibL* contains suitable binding-sequences for Fkh and FMBP-1 (Fig. 3). The core promoters themselves of *fibL* and *fhx* might regulate their tissue-specific expression.

The ectopic expression of *fibH* was restricted only to MSG cells, particularly in the middle part, suggesting existence of additional regulators that reduce the occurrence of inappropriate expression of *fibH* in other tissues (Fig. 5). Fkh is known to serve in the regulation of *fibH* and *fhx* as well as *ser1* (Mach *et al.* 1995; Takiya *et al.* 1997; Julien *et al.* 2002), and a

silk gland-specific factor, sage, activates *fibH* expression through interaction with Fkh (Zhao *et al.* 2014). Another factor capable of restricting *fibH* gene expression is a POU-homeodomain protein POU-M1 expressed highly in ASG and MSG-A (Kokubo *et al.* 1997; Matsunami *et al.* 1998; Kimoto *et al.* 2012). Our previous analysis showed that POU-M1 represses *ser1* expression through competition with the binding sites of an activator complex MIC for *ser1* (Kimoto *et al.*, 2012). Since a typical and high-affinity POU-M1-binding sequence exists within the C site of *fibH* promoter (Fig. 6, Hui *et al.*, 1990), it is possible that POU-M1 competes the binding of SGF-2 to the C site and regulates *fibH* negatively. On the other hand, no typical POU-M1-binding sequence is found near the SGF-2-binding sites around promoter of *fibL* and *fhx*, which were expressed ectopically in ASG by the misexpression of *Awk* (Figs. 3 and 5).

Expression levels of the fibroin genes were generally low in MSG-P under the conditions of *Awk* overexpression. The homeobox gene *invected* (*in*), a homolog of *engrailed* (*en*) (Simmonds *et al.*, 1995; Gustavson *et al.*, 1996), was expressed highly in the MSG-P (Kimoto *et al.*, 2014). The engrailed-family proteins are known to recruit the co-repressor groucho and repress target genes through the use of chromatin modifications (Tolkunova *et al.*, 1998; Peel *et al.*, 2006). These homeodomain proteins might be involved in regulation of the fibroin genes. Inhibition of mRNA accumulation of the fibroin genes by a territory-specific digestion in MSG-P cells or

region-specific translational regulation of Awh mRNA is also possible in the region-specific induction of the fibroin genes in the transgenic silkworms, though such a mechanism is not yet known.

We have demonstrated that LIM-homeodomain protein Awh activates all fibroin genes in the silk gland. Expression of *ser1* is activated by the Hox protein Antp, and repressed by the POU-homeodomain protein POU-M1, (Kimoto *et al.*, 2012, 2014; Takiya *et al.*, 2011). Thus, homeodomain-containing transcription factors appear to play roles in region-specific expressions of the silk genes in the larval silk gland in conjunction with other non-homeodomain transcription factors expressed in the silk gland such as Fkh, sage and FMBP-1.

375 **Acknowledgments**

376

377 We are grateful to Ms. A. Kosaka and Mr. S. Ishikawa for their technical
378 assistance. We are also grateful to Mr. K. Nakamura, Mr. K. Hashimoto and
379 Mr. T. Misawa for rearing the silkworms. This work was partly supported by
380 the Ministry of Agriculture, Forestry, and Fisheries (MAFF) of Japan (to
381 H.S.) and a Grant-in-aid for JSPS fellows (24•5323 to M.K.).

382

383

References

- Bello, B., Horard, B., Couble, P., 1994. The selective expression of silk-protein-encoding genes in *Bombyx mori* silk gland. *Bull. Inst. Pasteur* **92**, 81-100.
- Chen, L., Segal, D., Hukriede, N.A., Podtelejnikov, A.V., Bayarsaihan, D., Kennison, J.A., Ogryzko, V.V., Dawid, I.B., Westphal, H., 2002. Ssdp proteins interact with the LIM-domain-binding protein Ldb1 to regulate development. *Proc. Nat. Acad. Sci.* **99**, 14320-14325.
- Couble, P., Chevillard, M., Moine, A., Revel-Chapuis, P., Jean-Claude, P., 1985. Structural organization of the P₂₅ gene of *Bombyx mori* and comparative analysis of its 5' flanking DNA with that of the fibroin gene. *Nucl. Acids Res.* **13**, 1801-1814.
- Couble, P., Michaille, J.J., Garel, A., Couble, M.,L., Prudhomme, J.,C., 1987. Developmental switches of sericin mRNA splicing in individual cells of *Bombyx mori* silk gland. *Dev. Biol.* **124**, 431-440.
- Curtiss, J., Heilig, J.S., 1995. Establishment of *Drosophila* imaginal precursor cells is controlled by the *Arrowhead* gene. *Development* **121**,

406 3819-3828.

407

408 Curtiss, J., Heilig, J.S., 1997. *Arrowhead* encodes a LIM homeodomain
409 protein that distinguishes subsets of *Drosophila* imaginal cells. *Dev. Biol.*
410 **190**, 129-141.

411

412 Fukuta, M., Matsuno, K., Hui, C.-C., Nagata, T., Takiya, S., Xu, P.-X., Ueno,
413 K., Suzuki, Y., 1993. Molecular cloning of a POU domain-containing factor
414 involved in the regulation of the *Bombyx sericin-1* gene. *J. Biol. Chem.* **268**,
415 19471-19475.

416

417 Gustavson, E., Goldsborough, A.S., Ali, Z., Kornberg, T.B., 1996. The
418 *Drosophila engrailed* and *invected* genes: partners in regulation, expression
419 and function. *Genetics* **142**, 893-906.

420

421 Horard, B., Julien, E., Nony, P., Garel, A., Couble, P., 1997. Differential
422 binding of the Bombyx silk gland-specific factor SGFB to its target DNA
423 sequence drives posterior-cell-restricted expression. *Mol. Cell Biol.* **17**,
424 1572-1579.

425

426 Hui, C.-C., Matsuno, K., Suzuki, Y., 1990. Fibroin gene promoter contains a
427 cluster of homeodomain binding sites that interact with three silk gland

428 factors. *J. Mol. Biol.* **213**, 651-670.

429

430 Inoue, S., Kanda, T., Imamura, M., Wuan, G-C., Kojima, K., Tanaka, H.,
431 Tomita, M., Hino, R., Yoshizato, K., Mizuno, S., Tamura, T., 2005. A fibroin
432 secretion-deficient silkworm mutant, *Nd-S^D*, provides an efficient system for
433 producing recombinant proteins. *Insect Biochem. Mol. Biol.* **35**, 51-59.

434

435 Inoue, S., Tanaka, K., Arisaka, Fu., Kimura, S., Ohtomo, K., Mizuno, S., 2000.
436 Silk fibroin of *Bombyx mori* is secreted, assembling a high molecular mass
437 elementary unit consisting of H-chain, L-chain, and P25 with a 6:6:1 molar
438 ratio. *J. Biol. Chem.* **275**, 40517-40528.

439

440 Julien, E., Bordeaux, M.-C., Garel, A., Couble, P., 2002. Fork head
441 alternative binding drives stage-specific gene expression in the silk gland of
442 *Bombyx mori*. *Insect Biochem. Mol. Biol.* **32**, 377-387.

443

444 Kiguchi, K., Agui, N., 1981. Ecdysteroid levels and developmental events
445 during larval moulting in the silkworm, *Bombyx mori*. *Insect Physiol.* **27**,
446 805-812.

447

448 Kikuchi, Y., Mori, K., Suzuki, S., Yamaguchi, K., Mizuno, S., 1992. Structure
449 of the *Bombyx mori* fibroin light-chain-encoding gene: upstream sequence

450 elements common to the light and heavy chain. *Gene* **110**, 151-158.
 451
 452 Kimoto, M., Kitagawa, T., Kobayashi, I., Nakata, T., Kuroiwa, A., Takiya, S.,
 453 2012. Inhibition of the binding of MSG-intermolt-specific complex, MIC, to
 454 the *sericin-1* gene promoter and *sericin-1* gene expression by POU-M1/SGF-3.
 455 *Dev. Genes Evol.* **222**, 351-359.
 456
 457 Kimoto, M., Tsubota, T., Uchino, K., Sezutsu, H., Takiya, S., 2014. Hox
 458 transcription factor Antp regulates *sericin-1* gene expression in the terminal
 459 differentiated silk gland of *Bombyx mori*. *Dev. Biol.* **386**, 64-71.
 460
 461 Kokubo, H., Xu, P.-X., Xu, X., Matsunami, K., Suzuki, Y. 1997. Spatial and
 462 temporal expression pattern of POU-M1/SGF-3 in *Bombyx mori*
 463 embryogenesis. *Dev. Genes Evol.* **206**, 494-502.
 464
 465 Mach, V., Takiya, S., Ohno, K., Handa, H., Imai, T., Suzuki, Y., 1995. Silk
 466 gland factor-1 involved in the regulation of *Bombyx* sericin-1 gene contains
 467 fork head motif. *J. Biol. Chem.* **270**, 9340-9346.
 468
 469 Matthews J.M., Visvader, J.E., 2003. LIM-domain-binding protein 1: a
 470 multifunctional cofactor that interacts with diverse proteins. *EMBO Rep.* **4**,
 471 1132-1137.

472

473 Matsunami, K., Kokubo, H., Ohno, K., Suzuki, Y. 1998. Expression pattern
474 analysis of SGF-3/POU-M1 in relation to sericin-1 gene expression in the silk
475 gland. *Develop. Growth Differ.* **40**, 591-597.

476

477 Nony, P., Prudhomme, J.C., Couble, P., 1995. Regulation of the p25 gene
478 transcription in the silk gland of *Bombyx*. *Biol. Cell.* **84**, 43-52.

479

480 Ohno, K., Sawada, J., Takiya, S., Kimoto, M., Matsumoto, A., Tsubota, T.,
481 Uchino, K., Hui, C-C., Sezutsu, H., Handa, H., Suzuki, Y., 2013. Silk gland
482 factor-2 involved in fibroin gene transcription, consists of LIM homeodomain,
483 LIM-interacting, and single-stranded DNA-binding proteins. *J. Biol. Chem.*
484 **288**, 31581-31591.

485

486 Peel, A.D., Telford, M.J., Akam, M., 2006. The evolution of hexapod
487 engrailed-family genes: evidence for conservation and concerted evolution.
488 *Proc. R. Soc. B.* **273**, 1733-1742.

489

490 Sakudoh, T., Sezutsu, H., Nakashima, T., Kobayashi, I., Fujimoto, H., Uchino,
491 K., Banno, Y., Iwano, H., Maekawa, H., Tamura, T., Kataoka, H. and
492 Tsuchida, K., 2007. Carotenoid silk coloration is controlled by a
493 carotenoid-binding protein, a product of the Yellow blood gene. *Proc. Natl.*

494 *Acad. Sci. USA* **104**, 8941-8946.

495

496 Shimizu, K., Ogawa, S., Hino, R., Adachi, T., Tomita, M., Yoshizato, K., 2007.

497 Structure and function of 5'-flanking regions of *Bombyx mori* fibroin heavy

498 chain gene: identification of a novel transcription enhancing element with a

499 homeodomain protein-binding motif. *Insect Biochem. Mol. Biol.* **37**, 713-725.

500

501 Simmonds, A.J., Brook, W.J., Cohen, S.M., Bell, J.B., 1995. Distinguishable

502 functions for *engrailed* and *invected* in anterior-posterior patterning in the

503 *Drosophila* wing. *Nature* **376**, 424-427.

504

505 Suzuki, Y., Obara, ., Takiya, S., Hui, C.-C., Matsuno, K., Suzuki, T., Suzuki,

506 E., Ohkubo, M., Tamura, T., 1990. Differential transcription of the fibroin

507 and sericin-1 genes in cell-free extracts. *Develop Growth Differ* **32**, 179-187.

508

509 Suzuki, Y., Tsuda, M., Takiya, S., Hirose, S., Suzuki, E., Kameda, M., Ninaki,

510 O., 1986. Tissue-specific transcription enhancement of the fibroin gene

511 characterized by cell-free systems. *Proc. Natl. Acad. Sci. USA* **83**, 9522-9526.

512

513 Takasu, Y., Hata, T., Uchino, K., Zhang, Q., 2010. Identification of Ser2

514 proteins as major sericin components in the non-cocoon silk of *Bombyx mori*.

515 *Insect Biochem. Mol. Biol.* **40**, 339-344.

516

517 Takiya, S., Inoue, H., Kimoto, M., 2011. Novel enhancer and promoter
518 elements indispensable for the tissue-specific expression of the *sericin-1* gene
519 of the silkworm *Bombyx mori*. *Insect Biochem. Mol. Biol.* **41**, 592-601.

520

521 Takiya, S., Kokubo, H., Suzuki, Y., 1997. Transcriptional regulatory elements
522 in the upstream and intron of the fibroin gene bind three specific factors
523 POU-M1, BmFkh and FMBP-1. *Biochem. J.* **321**, 645-653.

524

525 Takiya, S., Hui, C.-C., Suzuki, Y., 1990. A contribution of the core-promoter
526 and its surrounding regions to the preferential transcription of the fibroin
527 gene in posterior silk gland extracts. *EMBO J.* **9**, 489-496.

528

529 Takiya, S., Ishikawa, T., Ohtsuka, K., Nishita, Y., Suzuki, Y. (2005)
530 Fibroin-modulator-binding protein-1 (FMBP-1) contains a novel
531 DNA-binding protein, repeats of the score and three amino acid peptide
532 (STP), conserved from *Caenorhabditis elegans* to humans. *Nucleic Acids Res.*
533 **33**, 786-795

534

535 Tsuda, M., Suzuki, Y., 1981. Faithful transcription initiation of fibroin gene
536 in a homologous cell-free system reveals an enhancing effect of 5' flanking
537 sequence far upstream. *Cell* **27**, 175-182.

538

539 Tsuda, M., Suzuki, Y., 1983. Transcription modulation *in vitro* of the fibroin
540 gene exerted by a 200-base-pair region upstream from the "TATA" box. *Proc.*
541 *Natl. Acad. Sci. USA* **80**, 7442-7446.

542

543 Tolkunova, E.N., Fujioka, M., Kobayashi, M., Deka, Deepali, Jaynes, J.B.,
544 1998. Two distinct types of repression domain in engrailed: One interacts
545 with the groucho corepressor and is preferentially active on integrated target
546 genes. *Mol. Cell. Biol.* **18**, 2804-2814.

547

548 Tsujimoto, Y., Suzuki, Y., 1979. The DNA sequence of *Bombyx mori* fibroin
549 gene including the 5' flanking, mRNA coding, entire intervening and fibroin
550 protein coding regions. *Cell* **18**, 591-600.

551

552 Uchino, K., Imamura, M., Sezutsu, H., Kobayashi, I., Kojima, K., Kanda, T.,
553 Tamura, T., 2006. Evaluating promoter sequences for trapping an enhancer
554 activity in the silkworm *Bombyx mori*. *J. Insect Biotech. Sericol.* **75**, 89-97.

555

556 Uhlirova, M., Asahina, M., Riddiford, L., Jindra, M., 2002. Heat-inducible
557 transgenic expression in the silkworm *Bombyx mori*. *Dev. Genes Evol.* **212**,
558 145-151.

559

560 Wang, Y., Klumpp, S., Amin, H.M., Liang, H., Li, J., Estrov, Z.,
561 Zweidler-McKay, P., Brandt, S.J., Agulnick, A., Nagarajan, L., 2010. *SSBP2*
562 is an *in vivo* tumor suppressor and regulator of LDB1 stability. *Oncogene* **29**,
563 3044-3053.

564

565 Xu, Z., Meng, X., Cai, Y., Liang, H., Nagarajan, L., Brandt, S.J., 2007.
566 Single-stranded DNA-binding proteins regulate the abundance of LIM
567 domain and LIM domain-binding proteins. *Genes Dev.* **21**, 942-955.

568

569 Zhao, X., Liu, C., Hu W., Zhou H., Zhang, Y., Peng Z., Zhao P., Xia, Q., 2014.
570 Basic helix-loop-helix transcription factor Bmsage is involved in regulation
571 of fibroin-H-chain gene *via* interaction with SGF1 in *Bombyx mori*. PLoS one
572 DOI: 10.1371/journal/pone.0094091.

573

Figure legends

Fig. 1. Gene expression of the components of SGF-2 and three fibroin proteins in the silk gland and various tissues. RT-PCR was performed with total RNA from each part of the silk gland and other tissues of wild-type day 2 fifth instars. *Awh*, *Ldb* and *Lcaf* are SGF-2 components. The ribosomal protein 49 (Rp49) gene was used as a control. MSG-A, -M and -P are the anterior, middle and posterior parts of the middle silk gland (MSG), respectively. CNS: central nervous system.

Fig. 2. Misexpression of *Awh* induces ectopic expression of *fibH* in MSG. (A) RT-PCR was performed with total RNA from the MSG or PSG of non-heated wild-type strains (+/+) on day 2 of the fifth instar or +/-*hs-Gal4* strain carrying *hs-Gal4* transgene alone or two *hs-Gal4/UAS-Awh* strains (*Awh* TG1, 2) 24 hr after heat treatment on day 1 of the fifth instar. (B) Time course of the gene expression in MSG of transgenic silkworms after heat treatment. RT-PCR was performed with total RNA from MSG of *Awh* transgenic silkworms (*hs-Gal4/UAS-Awh*) after heat treatment on day 1 of the fifth instar. The dissection time after heat treatment is shown at the top of the panels. Expression of *fhx* was detected at 12 hours after heat treatment by RT-PCR with 24 cycles in this figure, but further data for *fhx* used in this paper was obtained through RT-PCR with 21 cycles.

596

597 **Fig. 3.** Binding sites of transcription factors on the promoter of *fibH*, *fibL* and
598 *fhx*. Nucleotide sequence around the promoter including the first intron of
599 *fibH*, *fibL* and *fhx* is shown (Tsujimoto and Suzuki, 1979; Kikuchi *et al.*,
600 1992; Couble *et al.*, 1985). Number +1 represents the transcriptional start
601 site. TATA box is marked with a box. Single underlines indicate protein
602 coding regions. Double underlines indicate LIM-homeodomain recognition
603 sequences (*fibH*) and putative LIM-homeodomain binding sequences (*fibL*
604 and *fhx*). Broken lines indicate FMBP-1-binding elements, and arrows
605 indicate Fkh-binding elements. Wavy underline indicates a typical and
606 high-affinity POU-M1 binding sequence. Brackets indicate the E and C sites
607 of *fibH* promoter.

608

609 **Fig. 4.** Binding of SGF-2 to the promoter of *fibL* and *fhx*. (A) EMSA was
610 carried out with *fibH*-200 probe and extracts from entire MSG on day 2 of the
611 fifth instar (V2 MSG), PSG on the fourth molting stage (IVm PSG) or PSG on
612 day 2 of the fifth instar (V2 PSG). For competition assays, 100-fold excess of
613 unlabeled oligonucleotides indicated at the top of panel was added into
614 EMSA reactions. For supershift assay, anti-Awh antiserum (+αAwh) or
615 anti-POU-M1 antiserum (+αPOU) was added. (B) Each double stranded
616 oligonucleotide probe was used for EMSA using IVm PSG extract or V2 PSG
617 extract. Filled circles indicate a shifted band specific to V2 PSG extract. (C)

SGF-2 complex was detected with fhx-30 probe in EMSA. (D) Nucleotide sequences around SGF-2-binding element of fhx-30, fhx-30M and fibH-200 are compared. The TATA box of fhx-30 region is changed to the canonical TATA box sequence in fhx-30M oligonucleotide. xx indicates mutations from the typical SGF-2-binding element.

Fig. 5. Effects of ectopic expression of *Awh* in each part of the silk gland and various tissues. (A) Expressions of *Awh* and *fibroin* genes were analyzed by RT-PCR. Tissues were obtained from *Awh* transgenic silkworms (hs-Gal4/UAS-Antp) 24 hr after heat treatment of day 1 fifth instar. (B) Localization of *fibH* mRNA in the silk gland was analyzed using *in situ* hybridization. Silk glands were collected from day 2 fifth instar larvae of wild type (+/+) or a heat-treated *Awh* transgenic strain (hs-Gal4/UAS-Antp). Arrows point to the boundary between *fibH* positive and -negative cells. Bar: 2 mm.

Fig. 1

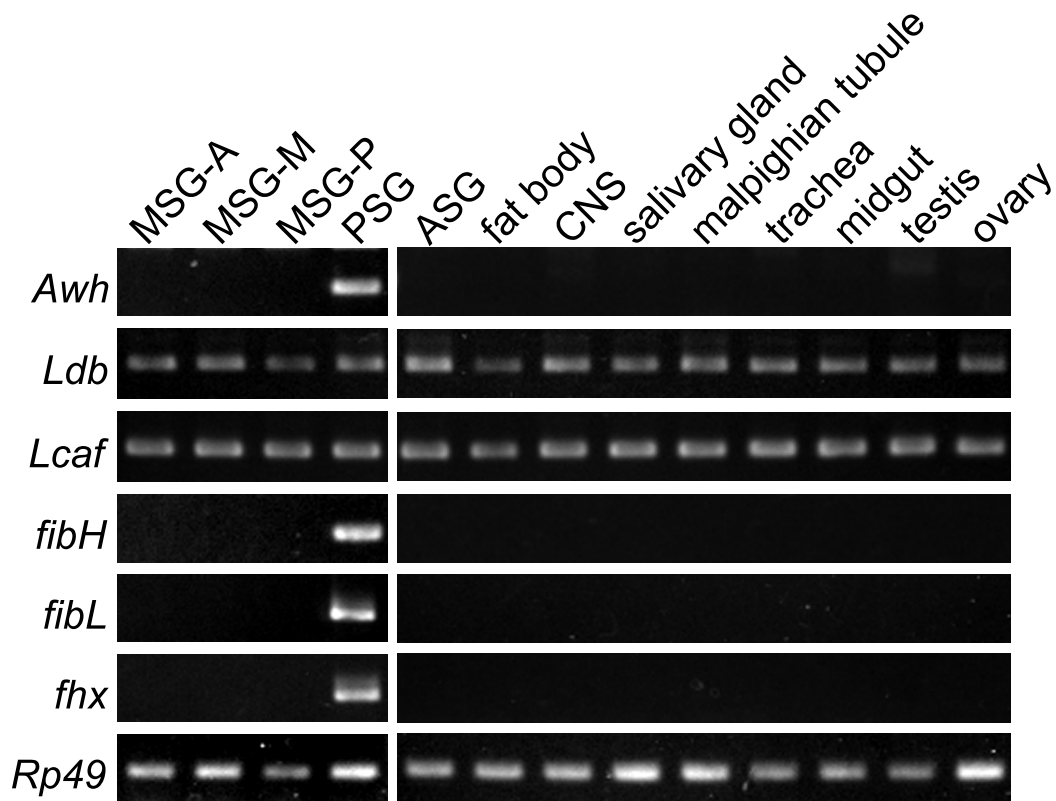


Fig. 2

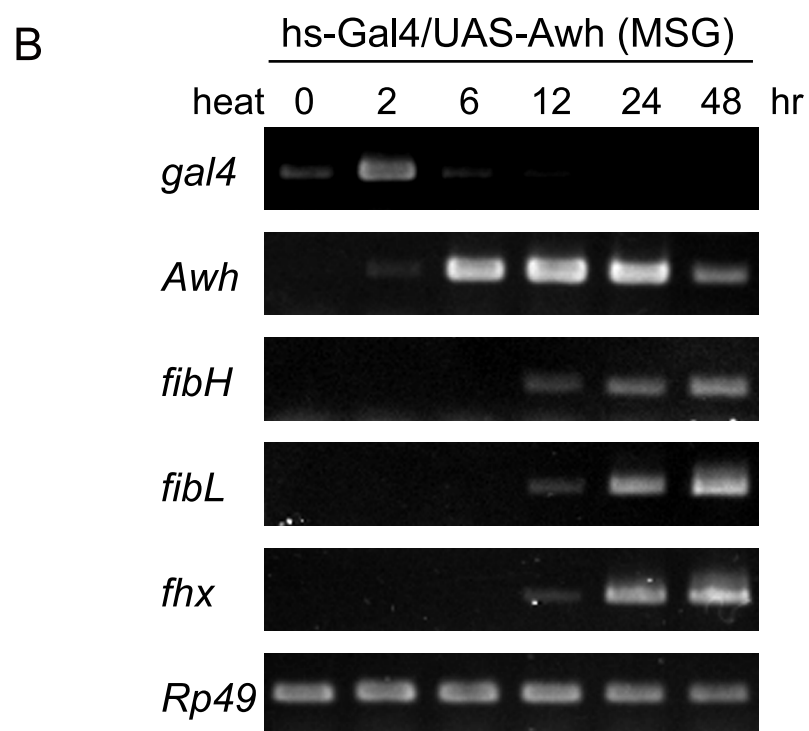
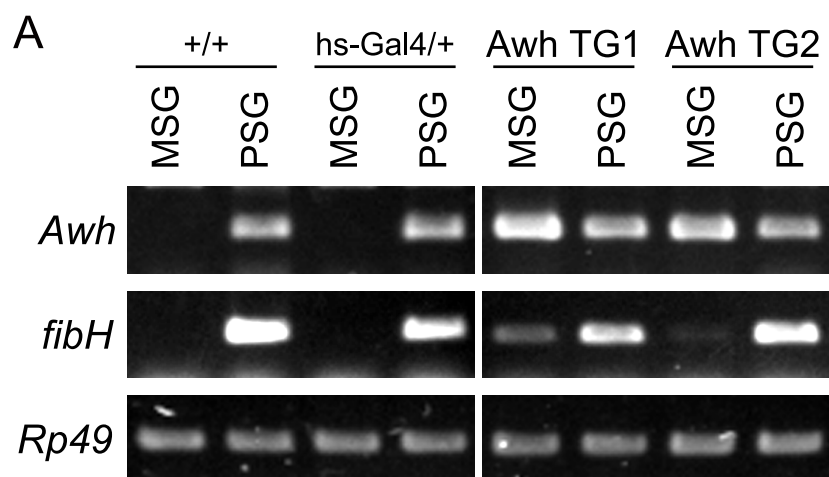


Fig. 3

fibH

-240 AAAATTCTTTAAAATATTTAAAAGTAA GAACAATAAGAT CAATTAAAT CATAATTAATCAC E site
-180 ATTGTTTCATGATCACAATTTAATTTACTTCATACGTTGTATTGTTAT GTTAAATAAAAAG
-120 ATTAATTTCTAT GTAATTGTATCTG TACAATACAATGTGTAGATGTTTATTCTATCGAAA C site
-60 GTAATAACGTCAAAACTCGAAAATTTTCAG TATA AAAAGGTTCAACTTTTTCAAATCAGC
+1 ATCAGTTCGGTTCCAACCTCTCAAGAT GAGAGTCAAAACCTTTGTGATCTTGTGCTGCGCT

fibL

-240 GGAT CAATTAGATCGCTTTGTTTCGAACAACACTTAGTTTAACTAGAGGCGTACACCTCA
-180 AGAAATCATCTTCATTAGAACTAAACCTTAAAATCGCAATAATAAAGCATAGTCAATTT
-120 TAACTGAAATGCAAAGTCTTTTGAACGTTAGATGCTGTCAGCGTTCGTTGGTACAGTTGT
-60 TTGATATTTATTTT AATTGTCTTTTTATAT TATA AATAGTGGAACATTAATCACGGAATCC
+1 TGTATAGTATATACCGATTGGTCACATAACAGACCACTAAAAT GAAGCCTATATTTTTGG

fhx

-120 CGCGGCGCAACAATAAGAACTTAATTCGTGCAATTGTTCCACGACGCTATTTATTTAAC
-60 GTTATTCGTTGTGAGGAACAATACTTTG TATA ATTAATGTTGATCAGTGCCTAACGACGC
+1 AGTTGTTTATTATTCTGCGCAACAT GTGCGCGGTGTCTACCTGTAGCCGCTGTGGCAGT
+61 TTTGGCTTCTGCAGGTAAGTCCGATGTTTCTTGATTAGGT CAATTATAACTACACGTAAC
+121 GCTTTGTGATTAATCGGGAAGACTATTGATTAATTATATTCCGCAACAGGGGGTATATTA

Fig. 4

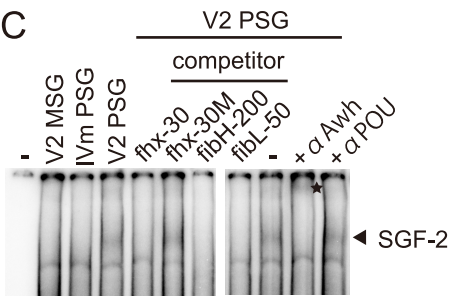
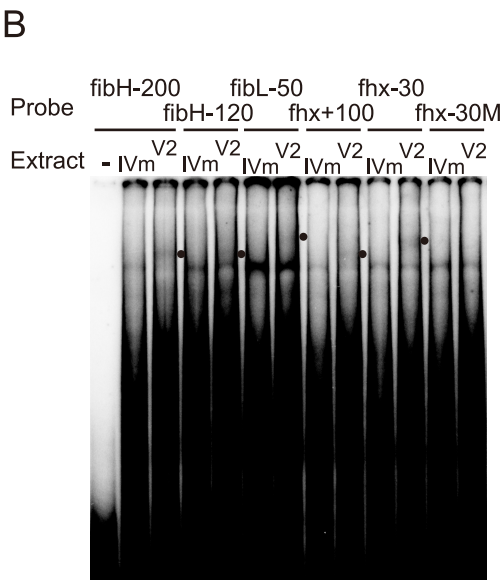
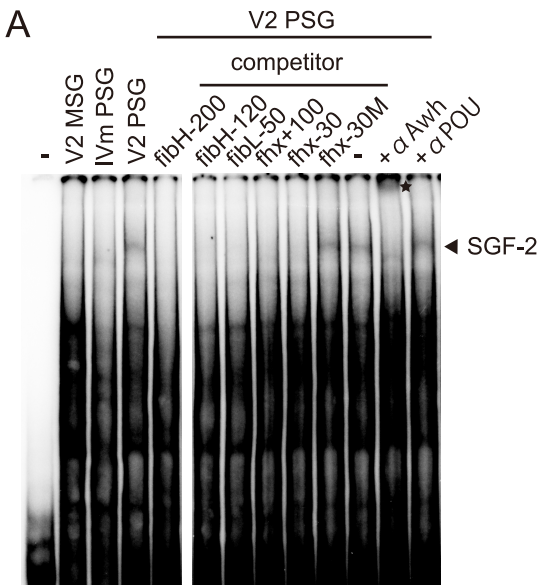


Fig. 5

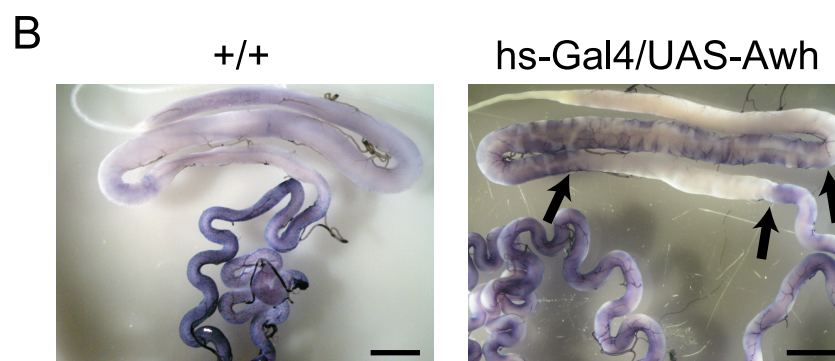
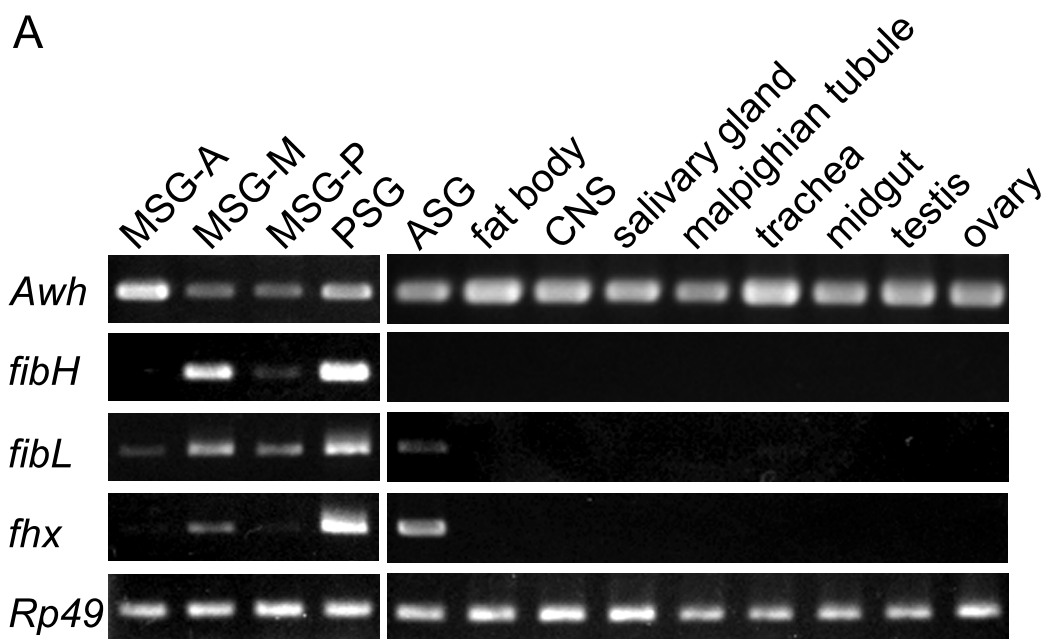


Table S1. Primers and PCR cycles for each gene.

Gene		Primers	Accession no.	Cycles
<i>fibH</i>	F	5'-TCCAACCTCTCAAGATGAG-3'	AF226688	21
	R	5'-TATCCTTGATGAGTGCTG-3'		
<i>fibL</i>	F	5'-ACCATCGGTGACCATCAATC-3'	AF541967	21
	R	5'-TTAGACGTGAACCTGGCTG-3'		
<i>fhx</i>	F	5'-TGGCAGTTTTGGCTTCTGCAG-3'	X04226	21 ^{*2}
	R	5'-TCAGTGCTGGCAACCGAAGTC-3'		
<i>Awh</i>	F	5'-ATGAAGACGGAGCACCGCA-3'	AB687553	35
	R	5'-TTGTGATAGCCCTCCGAGTCAC-3'		
<i>Ldb</i>	F	5'-AAGAGGTTACAGCAGAGGAC-3'	AB687554	33
	R	5'-TTTCTGGAAGAGCGTTGTC-3'		
<i>Lcaf</i>	F	5'-ATGTATGCCAAGGGCAAGG-3'	AB687556	35
	R	5'-TGGCGGTCCATTGAAATCATTC-3'		
<i>Rp49</i>	F	5'-CATACAAGATGGCTATAAGACC-3'	AY769302	21
	R	5'-GTGAGCTGCTGGGCTCTTTC-3'		

*1 F: Forward primer, R: Reverse primer

*2 Expression of *fhx* was analyzed by RT-PCR with 24 cycles in the Fig. 3, but other data for *fhx* in this paper were obtained by RT-PCR with 21 cycles.

Table S2. Probes and competitors used in EMSA.

Probe		Nucleotide sequence*2
fibH-200	T *1	5'-aattCAATAAGATCAATTAAATCATAATTAATCACATTGTTTCATGATCAC-3'
	B	5'-aattGTGATCATGAACAATGTGATTAATTATGATTTAATTGATCTTATTG-3'
fibH-120	T	5'-aattATGTAAATAAAAAAGATTAAATTTCTATGTAATTGTATCTGTACAAT-3'
	B	5'-aattATTGTACAGATACAATTACATAGAAATTAATCTTTTTATTTAACAT-3'
fibL-50	T	5'-aattGGTACAGTTGTTTGATATTTATTTTAATTGTCTTTTTTATATATAAA-3'
	B	5'-aattTTTATATATAAAAAAGACAATTAAAATAAATATCAAACAACGTGACC-3'
fhx-30	T	5'-aattCGTTGTGAGGAACAATACTTTGTATAATTAATGTTGATCAGTGCCT-3'
	B	5'-aattAGGCACTGATCAACATTAATTATACAAAGTATTGTTTCCTCACAACG-3'
fhx-30M	T	5'-aattCGTTGTGAGGAACAATACTTTGTATAAAAAATGTTGATCAGTGCCT-3'
	B	5'-aattAGGCACTGATCAACATTTTTTTATACAAAGTATTGTTTCCTCACAACG-3'
fhx+100	T	5'-aattGATGTTTCTTGATTAGGTCAATTATAACTACACGTAACGCTTTGTG-3'
	B	5'-aattCACAAAGCGTTACGTGTAGTTATAATTGACCTAATCAAGAAACATC-3'

*1 T: Top strand, B: Bottom strand

*2 The lower case letters show extra sequences for labeling with Klenow polymerase and [α -³²P]dATP.