LETTER

doi:10.1038/nature13315

A single female-specific piRNA is the primary determiner of sex in the silkworm

Takashi Kiuchi[1], Hikaru Koga[1]*, Munetaka Kawamoto[1]*, Keisuke Shoji[1]*, Hiroki Sakai[2], Yuji Arai[1], Genki Ishihara[1], Shinpei Kawaoka[1], Sumio Sugano[3], Toru Shimada[1], Yutaka Suzuki[3], Masataka G. Suzuki[2] & Susumu Katsuma[1]

The silkworm *Bombyx mori* uses a WZ sex determination system that is analogous to the one found in birds and some reptiles. In this system, males have two Z sex chromosomes, whereas females have Z and W sex chromosomes. The silkworm W chromosome has a dominant role in female determination[1,2], suggesting the existence of a dominant feminizing gene in this chromosome. However, the W chromosome is almost fully occupied by transposable element sequences[3–5], and no functional protein-coding gene has been identified so far. Female-enriched PIWI-interacting RNAs (piRNAs) are the only known transcripts that are produced from the sex-determining region of the W chromosome[6], but the function(s) of these piRNAs are unknown. Here we show that a W-chromosome-derived, female-specific piRNA is the feminizing factor of *B. mori*. This piRNA is produced from a piRNA precursor which we named *Fem*. *Fem* sequences were arranged in tandem in the sex-determining region of the W chromosome. Inhibition of *Fem*-derived piRNA-mediated signalling in female embryos led to the production of the male-specific splice variants of *B. mori doublesex* (*Bmdsx*), a gene which acts at the downstream end of the sex differentiation cascade[7,8]. A target gene of *Fem*-derived piRNA was identified on the Z chromosome of *B. mori*. This gene, which we named *Masc*, encoded a CCCH-type zinc finger protein. We show that the silencing of *Masc* messenger RNA by *Fem* piRNA is required for the production of female-specific isoforms of *Bmdsx* in female embryos, and that Masc protein controls both dosage compensation and masculinization in male embryos. Our study characterizes a single small RNA that is responsible for primary sex determination in the WZ sex determination system.

In *Bombyx mori*, sex determination is probably established at an early stage of embryogenesis. We prepared the sexed RNA from individual silkworm embryos genotyped by three W chromosome-specific randomly amplified polymorphic DNA (RAPD) markers[4] (Extended Data Fig. 1a), and examined the splicing pattern of a *doublesex* orthologue of *B. mori* (*Bmdsx*). *Bmdsx* produces female- and male-specific RNAs by sex-specific alternative splicing[9] that have essential roles in silkworm sexual development[7,8]. Female-specific splice variants of *Bmdsx* were the default transcripts during an early stage of development. Whereas the male-specific splice variants clearly appeared in male embryos, only faint bands were observed in females, from 21–24 h post-oviposition (hpo) (Fig. 1a). This indicated that the feminizing signal is transmitted from the W chromosome before 21 hpo. Thus, we performed deep sequencing of RNAs (RNA-seq) isolated from male and female embryos at 15, 18, 21 and 24 hpo, and identified differentially expressed transcripts between male and female embryos.

One contig, comp73859_c0, was consistently identified in female embryos at all of the developmental times tested (Extended Data Fig. 1b, c). This sequence was amplified by PCR only when female genomic DNA or complementary DNA was used as a template (Fig. 1b). The sequence of comp73859_c0 did not show significant identity with any sequence in the draft male silkworm genome sequence[10], suggesting that this contig is localized on and transcribed from the W chromosome. In addition, this sequence was amplified from genomic DNA isolated from female wild silkmoth *Bombyx mandarina*, but not from males (Fig. 1b). Reverse transcription followed by quantitative PCR (RT–qPCR) showed that the expression level peaked at 18–21 hpo, and then gradually declined during embryogenesis (Fig. 1c). This transcript was also detected in the ovary and other somatic tissues (Extended Data Fig. 2a, b, d). Long PCR demonstrated that there are multiple copies of this sequence on the W chromosome (Extended Data Fig. 2c). The sequences were occasionally arranged in tandem and some were probably expressed as long transcriptional units. Northern blot analysis revealed that transcripts of approximately 0.8 and 1.4 kilobase are major units, and antisense transcripts were not detected (Extended Data Fig. 2d). The copy number of this contig per haploid genome was estimated at more than 30 in the genome of *B. mori*.

This contig did not show homology with any known sequence, nor did it seem to encode a functional protein. Instead, this transcript seemed to be a piRNA precursor. Mapping of embryonic or ovarian piRNAs[6,11,12] onto this transcript and northern blotting revealed a 29-nucleotide-long piRNA-producing region (Fig. 1d and Extended Data Fig. 3a). This piRNA was poorly transmitted from the mother moth, accumulated from 15 hpo,

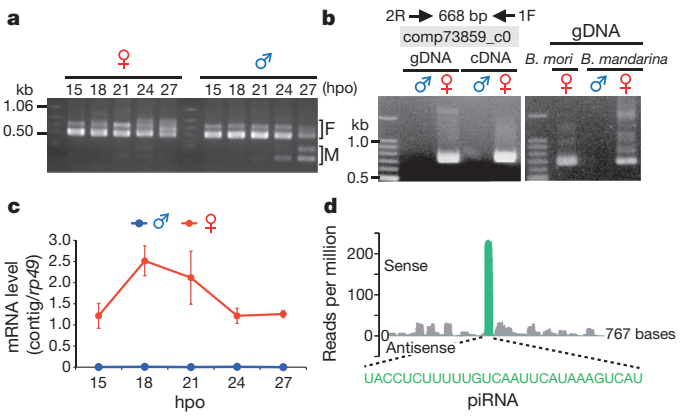


Figure 1 | Characterization of a female-specific piRNA precursor in early silkworm embryos. **a**, Splicing patterns of *Bmdsx* in early embryos. The F and M indicate female- and male-type splicing of *Bmdsx*, respectively. Similar results were obtained in three independent experiments. **b**, Detection of a W chromosome-derived transcript. Genomic DNA (gDNA) and cDNA were prepared from female and male embryos of *B. mori* at 24 hpo (left panel) or adult *B. mandarina* (right panel). **c**, Expression profile of the female-specific contig in early embryos. Data shown are means ± s.d. of three embryos. **d**, Mapping of embryonic piRNAs (24 hpo) onto the comp73859_c0. The relative location, abundance and sequence of the 29 base-long piRNA (shown in green) are indicated.

[1]Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. [2]Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan. [3]Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.
*These authors contributed equally to this work.

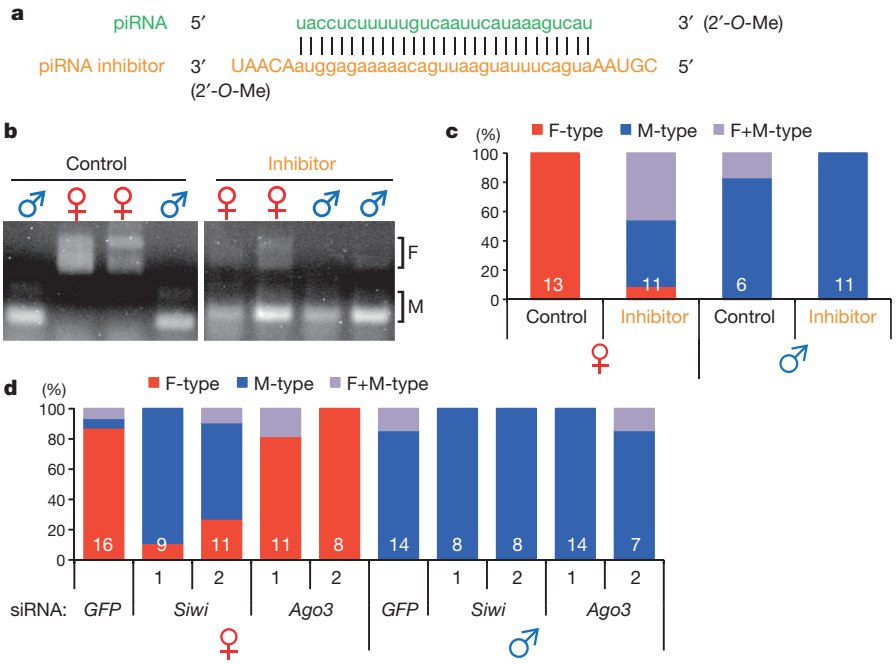


Figure 2 | Female-specific piRNA is a primary sex determinant of *B. mori*. **a**, Structure of the female-specific piRNA and its inhibitor. **b**, **c**, Effect of the RNA inhibitor on the *Bmdsx* splicing. The splicing patterns were examined at 72 h post-injection. Representative splicing patterns are shown in **b** and the data are summarized in **c**. The number indicates the sample size. F-type, female-type; M-type, male-type; F+M-type, both variants are mixed. **d**, Splicing of *Bmdsx* in embryos that were injected with *Siwi* or *BmAgo3* siRNAs. Two types of siRNAs for each target were used. The splicing patterns were examined at 72 h post-injection. The numbers in the columns indicate sample size.

and increased rapidly between 18 and 21 hpo (Extended Data Fig. 3b, c). By screening piRNA libraries that were generated from three *B. mori* strains that each possess a unique W chromosome structure[6], we found that this piRNA was produced from the sex-determining region of W chromosome (Extended Data Fig. 3d).

The silkworm KG strain, which possesses a mutation(s) in the W chromosome, shows various degrees of female masculinization features[13]. Expression of the contig-derived piRNA in the masculinized ovary was markedly lower than in wild type (Extended Data Fig. 4a), indicating that a certain amount of this piRNA might be required for complete feminization of adult moths. To investigate the role of this piRNA, we used a unique RNA-based inhibitor that seemed to function in BmN4 cells (Fig. 2a and Extended Data Fig. 4b) and investigated the effect of the inhibitor on *Bmdsx* splicing in early embryos. The *Bmdsx* splicing was markedly altered to produce the male-type isoform when the inhibitor was injected into female embryos, whereas the pattern was not affected in male embryos (Fig. 2b–c). The splicing pattern, however, was not altered in newly hatched larvae, presumably because this inhibitor does not possess the long-term inhibitory activity (Extended Data Fig. 4c). These results demonstrated that the targeted piRNA is required for the female-type splicing of *Bmdsx*. Thus, we named the precursor of this piRNA, *Feminizer* (*Fem*).

We performed RNA interference (RNAi) experiments that targeted two core components of the silkworm piRNA biogenesis pathway called *Siwi* and *BmAgo3* (Extended Data Fig. 4d)[14,15]. Small interfering RNA (siRNA)-mediated knockdown of *Siwi* in female embryos commonly led to the production of the male-type *Bmdsx* transcripts, whereas little effect was observed in male embryos (Fig. 2d and Extended Data Fig. 4e). *BmAgo3* RNAi did not affect *Bmdsx* splicing in either female or male embryos (Fig. 2d and Extended Data Fig. 4e). These results indicated that *Siwi* expression is crucial for the female-type splicing of *Bmdsx* in female embryos. Our hypothesis as to why *BmAgo3* knockdown did not affect *Bmdsx* splicing in early embryos is discussed later.

We identified only one genomic locus where the *Fem* piRNA sequence was extensively complementary (Fig. 3a). This locus was present within the ninth exon of an uncharacterized gene located on the Z chromosome (Fig. 3a and Extended Data Fig. 5a). We named this gene *Masculinizer* (*Masc*). *Masc* potentially encoded a novel CCCH-tandem zinc finger protein (Extended Data Fig. 5b). Phylogenetic analysis suggested that this protein forms a novel lepidopteran-specific protein family (Extended Data Fig. 5c). PIWI–piRNA complexes are known to cleave their complementary target sequences across from positions 10 and 11 of the guide piRNA[16,17]. By a modified 5′ rapid amplification of cDNA ends (modified RACE) method[18], we found that all of the cloned 5′ ends of the

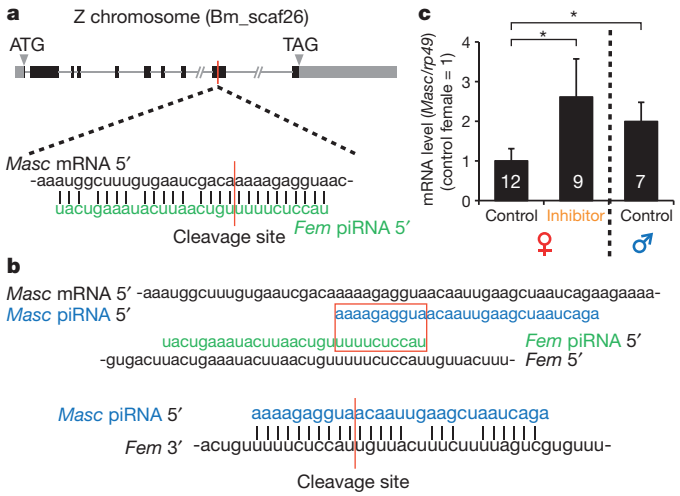


Figure 3 | *Masc* mRNA is the target of *Fem* piRNA. **a**, Genomic structure of the *Masc* gene on the Z chromosome of *B. mori*. The putative cleavage site by the *Fem* piRNA–Siwi complex is shown by the red line. **b**, Identification of *Masc* piRNA. The ping-pong signature (within the red box) and putative cleavage site (red line) of *Fem* by the *Masc* piRNA–BmAgo3 complex are shown. **c**, *Masc* expression in the piRNA inhibitor-injected embryos at 18 hpo. The numbers in the columns indicate sample size. Data shown are means + s.d. Data were subjected to Kruskal–Wallis analysis with post hoc Dunn's test. $*P < 0.05$.

Masc-derived RNA fragments from early embryos mapped precisely to the predicted *Fem* piRNA cleavage site (Extended Data Fig. 6a), indicating that *Masc* mRNA is the target of *Fem* piRNA.

The piRNA biogenesis occurs through a ping-pong mechanism that involves two different PIWI proteins. A 10-nucleotide overlap between sense and antisense piRNAs, called a ping-pong signature, is the hallmark of the cleavage reaction catalysed by PIWI proteins[16,17]. We found piRNAs that have a perfect 10-nucleotide overlap with *Fem* piRNA. The most abundant of these piRNAs were those that perfectly matched to the *Masc* coding region (Fig. 3b and Extended Data Fig. 6b, c), indicating that the *Masc* mRNA-derived piRNA (*Masc* piRNA) is a ping-pong partner of *Fem* piRNA. The *Masc* piRNA was extensively complementary to *Fem* (Fig. 3b), indicating that the PIWI–*Masc* piRNA complex will reliably slice *Fem* RNA. *Fem* piRNA preferentially bound to Siwi, whereas *Masc* piRNA preferentially bound to BmAgo3 (Extended Data Fig. 6d). Thus, our findings indicate a ping-pong amplification model for *Fem* and *Masc* piRNAs (Extended Data Fig. 6e). This model is experimentally supported by the introduction of the inhibitor for *Fem* piRNA (Fig. 3c) or siRNAs for *Siwi* (Extended Data Fig. 7a) into female embryos showing enhanced *Masc* levels. Unlike *Fem* piRNA, a moderate amount of *Masc* piRNA was maternally transmitted (Extended Data Fig. 6f, g). Together with the fact that BmAgo3 is also maternally transmitted[12], this suggests that a moderate amount of the *Masc* piRNA–BmAgo3 complex exists even in newly laid eggs. The presence of this complex helps to explain why *BmAgo3* RNAi in female embryos had little effect on *Bmdsx* splicing (Fig. 2d). Embryonic RNAi for *BmAgo3* did not alter the *Bmdsx* splicing, but enhanced *Masc* expression in newly hatched female larvae (Extended Data Fig. 7b–d), supporting the role of the *Masc* piRNA–BmAgo3 complex in sex determination. Higher levels of *Masc* piRNA were detected from 21–27 hpo (Extended Data Fig. 6f, g); this increase correlated with a massive accumulation of *Fem* piRNA (Extended Data Fig. 3b, c).

In male embryos, *Masc* expression rapidly increased, then rapidly decreased between 15 and 18, and 18 and 21 hpo, respectively (Fig. 4a). In contrast, *Masc* expression in female embryos gradually declined from 15 hpo, and remained at a low level compared with that found in males (Fig. 4a). These data indicate that *Fem* piRNA-mediated cleavage of *Masc* mRNA results in low-level accumulation of *Masc* mRNA in female embryos. Injection of *Masc* siRNA into male embryos reduced *Masc* expression to levels that were found in control female embryos at 18 hpo (Fig. 4b), and resulted in the production of female-type variants of *Bmdsx* throughout the embryonic stage (Fig. 4c and Extended Data Fig. 8a, b). Female embryos injected with *Masc* siRNA hatched normally, whereas male embryos did not (Fig. 4d), indicating that inhibition of the *Masc* pathway at the embryonic stage results in male-specific lethality. This probably mimics the way that an arthropod pathogen *Wolbachia* induces a male-killing phenotype in lepidopteran insects[19]. The *Fem* piRNA-resistant *Masc* (*Masc-R*) mRNA was more accumulated than the wild-type *Masc* mRNA in *Masc* cDNA-transfected BmN4 cells, whereas *Masc* piRNA was poorly detected in *Masc-R* cDNA-transfected cells (Extended Data Fig. 9a–c). The *Masc-R* mRNA was not cleaved, which changed the *Bmdsx* splicing pattern in BmN4 cells to the male-type completely, and induced a growth inhibition (Extended Data Fig. 9d–f), indicating that the *Fem* piRNA-mediated cleavage of *Masc* mRNA is essential for silkworm feminization.

RNA-seq analyses of *Masc* siRNA-injected embryos revealed that the transcripts differentially expressed in males were mapped predominantly onto the Z chromosome (chromosome 1, 51%), whereas such a bias was not observed in females (Fig. 4e). Most of the Z-chromosome-derived transcripts expressed differentially in males (97%) were expressed higher in *Masc* RNAi embryos (Fig. 4e) and randomly dispersed throughout this chromosome (Extended Data Fig. 10). These results demonstrate that Masc protein globally represses gene expression from the male Z chromosome at the embryonic stage. Taken together, Masc protein controls both dosage compensation and masculinization (Fig. 4f). In *Drosophila*, Sex-lethal, a master switch for sex determination, controls dosage

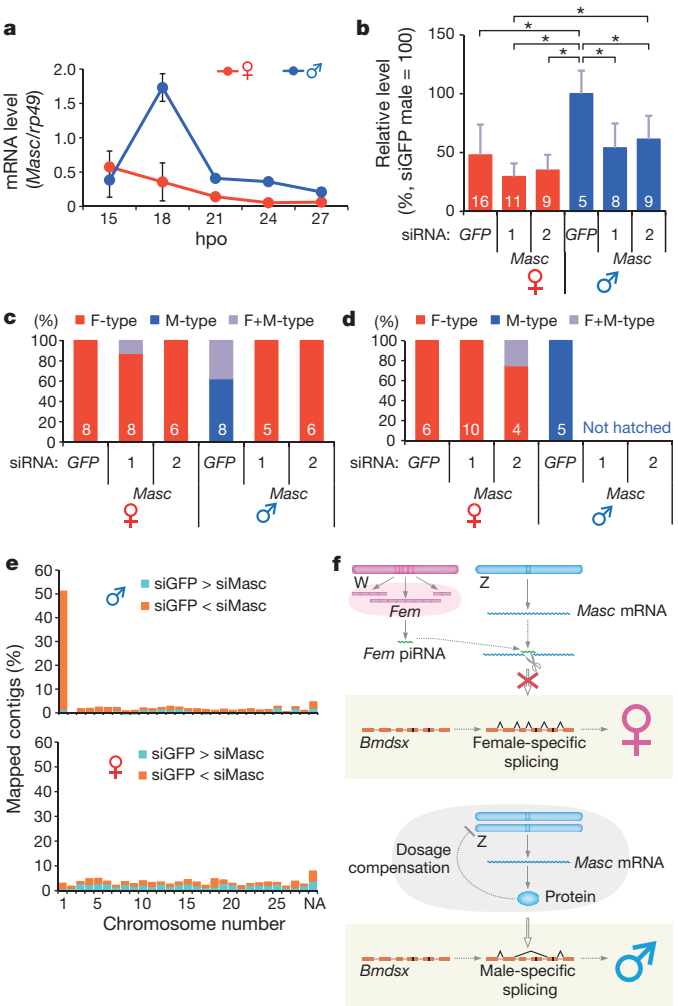


Figure 4 | Masc protein controls both masculinization and dosage compensation in male embryos. **a**, Expression profile of *Masc* in early embryos. Data shown are means ± s.d. of three embryos. **b**, Knockdown of *Masc* mRNA in *B. mori* embryos. The embryos were injected with two types of siRNAs for *Masc*, and *Masc* expression was examined by RT–qPCR at 18 hpo. Data shown are means + s.d. The number indicates the sample size. One-way ANOVA was performed with post hoc Tukey's test. $*P < 0.05$. **c**, **d**, Splicing of *Bmdsx* in *Masc* siRNA-injected embryos. The splicing pattern was determined at 72 h (**c**) and about 240 h (**d**, immediately after hatching) post-injection. The number indicates the sample size. **e**, Differentially expressed transcripts in *Masc* RNAi embryos. siGFP, *GFP* siRNA-injected embryos; siMasc, *Masc* siRNA-injected embryos; NA, not assigned. **f**, A proposed model for the sex determination pathway in *B. mori*.

compensation by inhibiting translation of *male-specific lethal 2* (*msl-2*)[20]. Loss of *msl-2* causes male lethality, owing to the failure of hypertranscription from the male X chromosome. A failure of dosage compensation is probably involved in male-specific lethality of *Masc* mRNA-depleted male embryos.

We unravelled a question that has perplexed insect geneticists for more than eight decades. Our study answers the question of how the W chromosome determines the femaleness of the silkworm *B. mori*. The silkworm feminizer *Fem* is the precursor of a 29-nucleotide-long small RNA. To our knowledge, this is the first example of the identification of a primary sex-determining factor in Lepidoptera, and the first experimental evidence showing a piRNA-mediated sex determination mechanism. Our findings also suggest that *Masc* levels may be involved in sex determination in lepidopteran species that are monosomic (Z0) in females and ZZ in males[21]. We are now experimentally surveying this hypothesis using moth species with a Z0/ZZ sex chromosome constitution.

METHODS SUMMARY

Sex-specific splicing of *Bmdsx*, piRNA mapping, qRT–PCR of piRNA, and transfection experiments using BmN4 cells were performed as described previously[6,12,13,14]. Embryonic RNAi was performed by injecting embryos with two different siRNAs for each gene investigated.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 2 October 2013; accepted 8 April 2014.
Published online 14 May 2014.

1. Hasimoto, H. The role of the W-chromosome in the sex determination of *Bombyx mori* [in Japanese]. *Jpn. J. Genet.* **8,** 245–247 (1933).
2. Tajima, Y. Studies on chromosome aberrations in the silkworm. II. Translocation involving second and W-chromosomes [in Japanese]. *Bull. Seric. Exp. Stn.* **12,** 109–181 (1944).
3. Abe, H. *et al.* Identification of novel random amplified polymorphic DNAs (RAPDs) on the W chromosome of the domesticated silkworm, *Bombyx mori*, and the wild silkworm, *B. mandarina*, and their retrotransposable element-related nucleotide sequences. *Genes Genet. Syst.* **73,** 243–254 (1998).
4. Abe, H. *et al.* Partial deletions of the W chromosome due to reciprocal translocation in the silkworm *Bombyx mori*. *Insect Mol. Biol.* **14,** 339–352 (2005).
5. Abe, H., Mita, K., Yasukochi, Y., Ohshiki, T. & Shimada, T. Retrotransposable elements on the W chromosome of the silkworm, *Bombyx mori*. *Cytogenet. Genome Res.* **110,** 144–151 (2005).
6. Kawaoka, S. *et al.* The silkworm W chromosome is a source of female-enriched piRNAs. *RNA* **17,** 2144–2151 (2011).
7. Suzuki, M. G., Funaguma, S., Kanda, T., Tamura, T. & Shimada, T. Analysis of the biological functions of a *doublesex* homologue in *Bombyx mori*. *Dev. Genes Evol.* **213,** 345–354 (2003).
8. Suzuki, M. G., Funaguma, S., Kanda, T., Tamura, T. & Shimada, T. Role of the male BmDSX protein in the sexual differentiation of *Bombyx mori*. *Evol. Dev.* **7,** 58–68 (2005).
9. Ohbayashi, F., Suzuki, M. G., Mita, K., Okano, K. & Shimada, T. A homologue of the *Drosophila doublesex* gene is transcribed into sex-specific mRNA isoforms in the silkworm, *Bombyx mori*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **128,** 145–158 (2001).
10. The International Silkworm Genome Consortium. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochem. Mol. Biol.* **38,** 1036–1045 (2008).
11. Kawaoka, S. *et al. Bombyx* small RNAs: genomic defense system against transposons in the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **38,** 1058–1065 (2008).
12. Kawaoka, S. *et al.* Zygotic amplification of secondary piRNAs during silkworm embryogenesis. *RNA* **17,** 1401–1407 (2011).
13. Hara, K. *et al.* Altered expression of testis-specific genes, piRNAs, and transposons in the silkworm ovary masculinized by a W chromosome mutation. *BMC Genomics* **13,** 119 (2012).
14. Kawaoka, S. *et al.* The *Bombyx* ovary-derived cell line endogenously expresses PIWI/PIWI-interacting RNA complexes. *RNA* **15,** 1258–1264 (2009).
15. Kawaoka, S., Minami, K., Katsuma, S., Mita, K. & Shimada, T. Developmentally synchronized expression of two *Bombyx mori* Piwi subfamily genes, *SIWI* and *BmAGO3* in germ-line cells. *Biochem. Biophys. Res. Commun.* **367,** 755–760 (2008).
16. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128,** 1089–1103 (2007).
17. Gunawardane, L. S. *et al.* A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* **315,** 1587–1590 (2007).
18. Watanabe, T. *et al.* Role of piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse *Rasgrf1* locus. *Science* **332,** 848–852 (2011).
19. Sugimoto, T. N. & Ishikawa, Y. A male-killing *Wolbachia* carries a feminizing factor and is associated with degradation of the sex-determining system of its host. *Biol. Lett.* **8,** 412–415 (2012).
20. Penalva, L. O. & Sánchez, L. RNA binding protein sex-lethal (Sxl) and control of *Drosophila* sex determination and dosage compensation. *Microbiol. Mol. Biol. Rev.* **67,** 343–359 (2003).
21. Traut, W., Sahara, K. & Marec, F. Sex chromosomes and sex determination in Lepidoptera. *Sex Dev.* **1,** 332–346 (2007).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. G. Kamita for critical reading of the manuscript; Y. Tomari for critical reading of the manuscript and technical suggestions. This work was supported by the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry to Su.K. and Grants-in-Aid for Scientific Research on Innovative Areas (Nos. 22115502 and 22128004) to Su.K. and T.S.

Author Contributions Su.K., T.K. and M.G.S. conceived and designed the experiments. T.K., H.K., K.S., H.S., G.I., Y.A., Sh.K., M.G.S. and Su.K. performed molecular biological experiments. M.K. and K.S. performed most of the bioinformatic analyses. S.S. and Y.S. performed deep sequencing and data analysis. T.S. provided essential reagents and expertise. All of the authors discussed the data and helped manuscript preparation. Su.K. wrote the manuscript with intellectual input from all authors. Su.K. supervised the project.

Author Information The nucleotide sequences of *Fem* and *Masc* have been deposited in the DDBJ/EMBL/GenBank data bank under the accession numbers AB840787 and AB840788. Deep sequencing data obtained in this study are available under the accession numbers DRA001104 and DRA001338 (DDBJ), respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Su.K. (katsuma@ss.ab.a.u-tokyo.ac.jp).

METHODS

Insects and cell lines. Larval *B. mori* (p50T, N4, and F1 hybrid Kinshu × Showa) and *B. mandarina* were reared as described previously[6]. The silkworm ovary-derived BmN4 cells were grown at 27 °C in TC-100 or IPL-41 medium supplemented with 10% fetal bovine serum[14].

Molecular sexing. Total RNA and genomic DNA were prepared simultaneously from a single embryo using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. We previously reported that the polar-body-derived W chromosome fragment can be detected at the early stage of embryogenesis[22]. To perform accurate molecular sexing of each embryo, we used three sets of W chromosome primers for PCR (Supplementary Table 1) or performed RT–qPCR for *Fem*.

RNA-seq. Libraries for RNA sequencing were generated from 15, 18, 21, 24 hpo of molecularly sexed embryos using the TruSeq RNA Sample Preparation kit (Illumina) and were analysed using the Illumina HiSeq 2000 platform with 101-bp paired-end reads (normal embryo samples, 8 data set) or HiSeq 2500 platform with 100-bp paired-end reads (RNAi embryo samples, 4 data set) according to the manufacturer's protocol[23].

Quantifications of *Fem* copy number. We estimated *Fem* copy number per haploid genome by quantitative PCR as reported previously[24]. Genomic DNA was extracted from larval tissues using standard procedures. *Siwi* was used as a single copy control gene on the autosome. qPCR analyses were performed using a KAPA SYBR FAST qPCR kit (Kapa Biosystems) and specific primers listed in Supplementary Table 1.

RT–PCR. Total RNA was prepared using TRIzol reagent (Invitrogen) according to the manufacturer's protocol and subjected to reverse transcription using avian myeloblastosis virus (AMV) reverse transcriptase with an oligo-dT primer (TaKaRa). PCR was carried out with KOD FX-neo DNA polymerase (TOYOBO). Sex-specific splicing of *Bmdsx* was examined by PCR with primers listed in Supplementary Table 1[25]. RT–qPCR analyses were performed using a KAPA SYBR FAST qPCR kit (Kapa Biosystems) and specific primers listed in Supplementary Table 1. RT–qPCR of piRNAs was performed as described previously[6]. In brief, small RNA fractions were enriched with the aid of a mirVana miRNA isolation kit (Ambion) and reverse transcribed using a miScript Reverse Transcription Kit (QIAGEN). qPCR was performed using a miScript PCR System (QIAGEN). The qPCR products were verified by cloning and DNA sequencing. *let-7*, one of the well-known silkworm microRNAs, was used as a control. The primers used in this experiment are described in Supplementary Table 1.

Embryonic RNAi. The short interfering RNA (siRNA) sequences listed in Supplementary Table 1 were designed based on the ORF sequences of the target genes and enhanced green fluorescent protein (*GFP*, control). Two different siRNAs were designed for each gene (that is, Siwi-1 and Siwi-2). Double-stranded siRNAs were purchased from FASMAC Corp (Japan), dissolved in annealing buffer (100 mM potassium acetate, 2 mM magnesium acetate, 30 mM HEPES-KOH; pH 7.4), and stored at −80 °C for later use. The *B. mori* N4 eggs used for siRNA injection were prepared as described previously[26]. Injection was performed according to the method described previously[27] using a microinjector (IM 300 Microinjector, Narishige Japan). One to 5 nl of each siRNA solution (50 μM for *Siwi*, 100 μM for *BmAgo3* and *Masc* (18 hpo and 72 h post-injection), and 500 μM for *Masc* (144, 216 and about 240 h post-injection)) was injected into each egg within 4–8 h after oviposition. The injected embryos were incubated at 25 °C in a humidified Petri dish. At 72 h post-injection, the expression level of the target gene was quantified by RT–qPCR, and samples whose target mRNA level (*Siwi* and *BmAgo3*) was knocked down by at least 80% was used for further analysis. *Masc* expression levels in embryos that were injected with siRNA were analysed at 18 hpo. Randomization and blinding were not applied to determine how embryo samples were allocated to experimental groups, because it is not possible to visually distinguish female and male embryos of silkworm N4 strain. The expression levels of *rp49* were used to normalize transcript levels. Primers used for RT–qPCR are listed in Supplementary Table 1.

Injection of the piRNA inhibitor into embryos. We designed a unique RNA-based inhibitor by modification of a previously described strategy[28] (Fig. 2a). We first tested the efficacy of our inhibitor using BmN4 cell line, a silkworm-ovary-derived, W-chromosome-harbouring cell line. BmN4 cells express the corresponding piRNA precursors (Extended Data Fig. 2b, d) as well as female-type *Bmdsx* transcripts (Extended Data Fig. 4b), and possess a complete piRNA pathway[14]. The male-type splice variant of *Bmdsx* was enhanced in BmN4 cells when transfected with the inhibitor (Extended Data Fig. 4b), indicating that our RNA inhibitor functioned to inhibit the piRNA-mediated signalling cascade.

One to five nl of a 1 mM RNA solution (anti-*Fem* piRNA or anti-*GFP* piRNA, Supplementary Table 1) was injected into the *B. mori* N4 strain eggs within 4–8 h after oviposition as described above. *Masc* expression levels in embryos that were injected with the inhibitor of *Fem* piRNA were analysed at 18 hpo.

RNA transfection in BmN4 cells. BmN4 cells (2.5×10^5 cells per 60-mm diameter dish) were transfected with single-stranded RNAs (250 pmol per dish, Supplementary Table 1) using X-tremeGENE HP (Roche)[29]. Following incubation for 12 h, the culture medium was removed and fresh medium was added. Cells were collected at 48 h after transfection, and total RNA was isolated. For transfection experiments using BmN4 cells, at least three independent experiments were performed.

Transient expression of *Masc* mRNA in BmN4 cells. The *Fem* piRNA-resistant *Masc* (*Masc-R*) cDNA was constructed by using PrimeSTAR Mutagenesis Basal Kit (TaKaRa). Five nucleotide mutations that do not result in amino acid substitutions for the Masc protein were introduced (Extended Data Fig. 9a). *Masc* or *Masc-R* cDNA was cloned into the pIZ/V5-His vector (Invitrogen). BmN4 cells (2.5×10^5 cells per 35-mm diameter dish) were transfected with plasmid DNAs (0.5 μg) using FuGENE HD (Promega)[29]. Cells were collected at 72 h after transfection. mRNA was prepared using Micro-FastTrack 2.0 Kit (Invitrogen) and subjected to RT–qPCR. *Masc* mRNA level was normalized to that of *rp49*. *Masc* piRNA was also quantified by RT–qPCR as described above.

Generation of BmN4 cells stably expressing Masc proteins. BmN4 cells stably expressing empty vector (pIZ/V5-His), *Masc* or *Masc-R* were generated as described previously[14]. Three days after transfection, zeocin (final concentration, 500 μg ml^{-1}) was added to the medium. Six days after drug selection, the splicing patterns of *Bmdsx* were examined by RT–PCR.

Northern blot analysis. Total RNA was separated by electrophoresis, transferred to a nylon membrane, and probed with strand-specific oligonucleotide probes as described previously[30] with some modifications. Small RNA fractions for piRNA detection were prepared from early embryos whose diapause was artificially terminated. The probe sequences are listed in Supplementary Table 1.

Modified RACE. The *Masc* mRNA-derived RNA fragments were determined by a modified RACE procedure as described previously[18]. To detect the cleaved fragments from exogenously introduced *Masc*, we used the primers designed on the pIZ/V5-His vector (Extended Data Fig. 9d, Supplementary Table 1).

RNA-seq analysis. *De novo* assembly of RNA-seq data from 8 data sets (15, 18, 21, 24 hpo of each sex, 303,483,056 reads in total) was performed using Trinity[31], and 221,677 contigs (170,255 kinds of transcripts) were produced. Transcript abundance in each contig was quantified by RSEM[32]. Differentially expressed transcripts (adjusted P value < 0.05) between female and male embryos were identified by the R/Bioconductor package, DESeq[33]. Contigs with more than 10 transcripts per million at any data set were selected and 157 contigs were used for further analysis. *Fem* contig was the only transcript showing significantly statistical scores between female and male at all time points examined (adjusted P values were 1.89×10^{-6} at 15 hpo, 1.42×10^{-28} at 18 hpo, 1.92×10^{-7} at 21 hpo, and 3.24×10^{-83} at 24 hpo). The R-code for this analysis is available as Supplementary Information.

Analysis of RNA-seq data from *Masc* RNAi experiments (*GFP* and *Masc* RNAi embryos of each sex, 72 h post-injection, 4 data sets) was performed as described above. We selected 585 and 608 differentially expressed transcripts (P value < 0.05, *GFP* siRNA versus *Masc* siRNA-1) in male and female, respectively. The chromosome on which each transcript is localized was identified by mapping the contigs to the silkworm genome scaffolds.

Raw RNA-seq data from control and *Masc* RNAi embryos were also mapped to the silkworm genome scaffolds by Bowtie[34] without mismatches. The coverage at each nucleotide position was estimated by coverageBed (included in BEDtools). The total mapped reads in each RNA-seq library to the scaffolds were used for normalization. The average coverage across each 1-kb window was determined and compared between the two RNA-seq libraries. The genome regions where the average coverage was more than 10 in either library were selected, grouped into three categories (siGFP/siMasc > 2, $0.5 \leq$ siGFP/siMasc ≤ 2, and siGFP/siMasc < 0.5) and visualized as Extended Data Fig. 10.

piRNA mapping. piRNA mapping was performed allowing two mismatches by Bowtie as described previously[6]. The total mapped reads in each piRNA library[6,11–14] to *B. mori* repetitive sequences (121 annotated transposons and 1,690 ReAS clones) were used for normalization.

To determine the genomic locus from which *Fem* piRNA is produced, we used piRNA libraries prepared from three *B. mori* strains that each possess a unique W chromosome structure[6] (Extended Data Fig. 3d). The sex-limited yellow (LY) strain[35] has a W chromosome that is approximately 90% shorter than the W chromosome of wild-type *B. mori*. This extensively truncated W chromosome, however, retains the ability to determine femaleness[35], indicating that this W fragment contains the putative sex-determining region. Of 12 RAPD markers identified in the normal W chromosomes, the LY W chromosome contained only one (W-Rikishi). The sex-determining region can be defined as the region where the W-Rikishi marker exists[35]. The DfZ-DfW strain ('without *Fem*', WF) on the other hand has a truncated W chromosome (approximately 75% shorter than the wild-type chromosome) that is attached to a Z chromosome[36]. This W chromosome fragment is not sufficient for determining femaleness, and indicates that it does not contain the

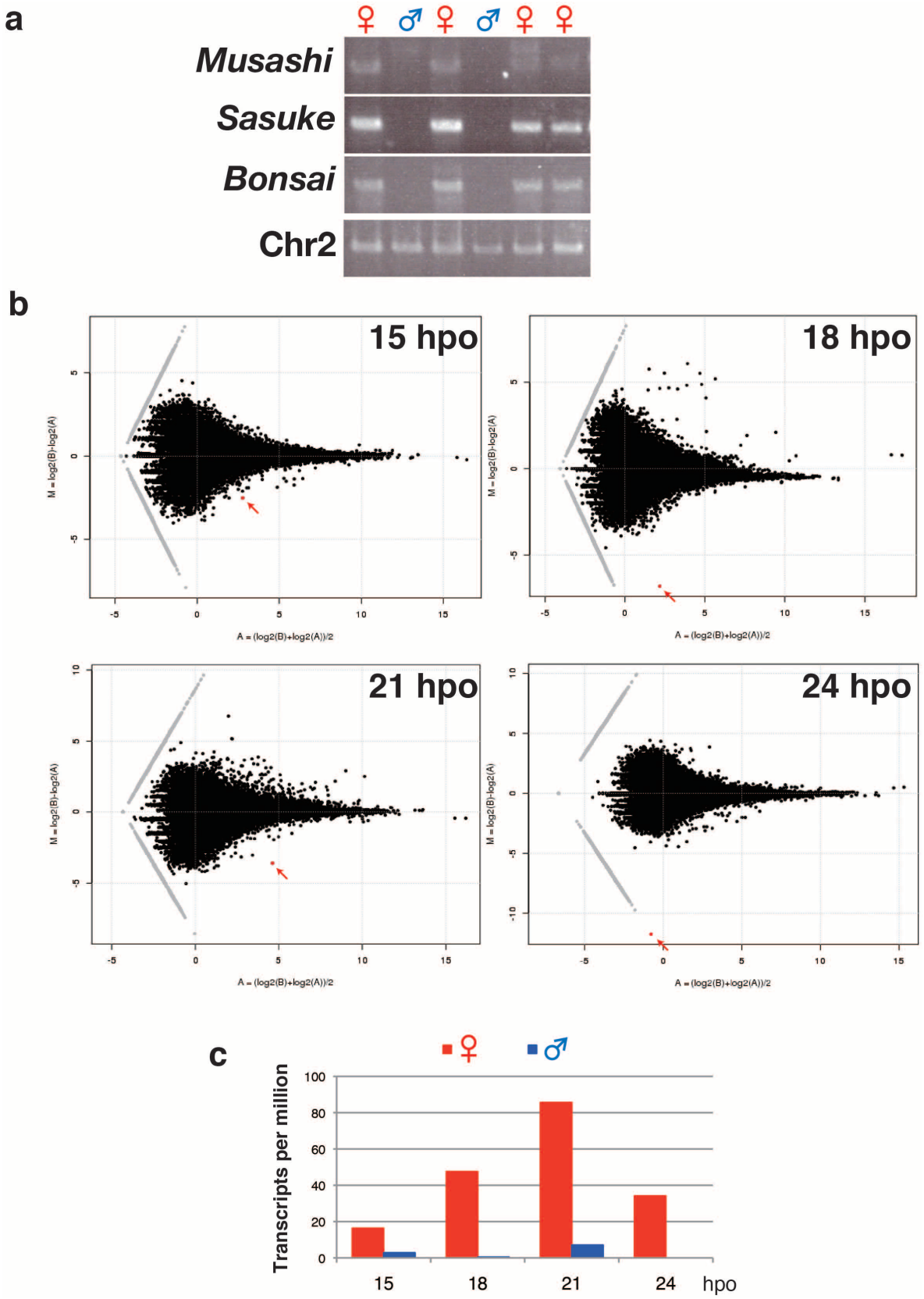
sex-determining region[36]. The Mandarina W (MW) strain of *B. mori* has a W chromosome that originates from *B. mandarina*[6]. When *B. mandarina* is crossed with *B. mori*, fertile hybrids are produced, indicating that the W chromosome of *B. mandarina* can determine the femaleness of *B. mori*, and implying that both species use the same sex-determination system. Examining abundance of *Fem* piRNA in each piRNA library[6,11,12,14] revealed that *Fem* piRNA was expressed in the ovaries of wild-type *B. mori* but not in the testes (Extended Data Fig. 3d). In ovaries from the LY and MW strains, this piRNA was expressed at 9% and 26%, respectively, of the level found in wild-type *B. mori*. Testes from the WF strain expressed an extremely low level (0.4% of the wild-type) of this piRNA even though the Z chromosome of this strain retains one-fourth of the W chromosome (Extended Data Fig. 3d). These results indicated that the sex-determining region of W chromosome produces *Fem* piRNA.

Target search for *Fem* piRNA. Base pairing of 11 or 12 nucleotides (nucleotides 2–12 or 2–13) at the 5′ end of a target sequence of the piRNA is required for efficient target cleavage by the mouse Piwi protein homologue Miwi[37]. To identify a potential target of *Fem* piRNA, we searched for genomic sequences of *B. mori* that were completely identical to nucleotides 2–12 of the 5′-end *Fem* piRNA. From this search we identified three candidate loci, among which *Masc* showed the lowest *E* value of 0.008, whereas the other two candidate loci showed *E* values that were >0.1. Bioinformatic analysis using the silkworm transcriptome and genome databases revealed that these two loci were not located within a predicted protein-coding gene or transcriptional unit. The *Masc* locus was thus predicted as the primary target of *Fem* piRNAs.

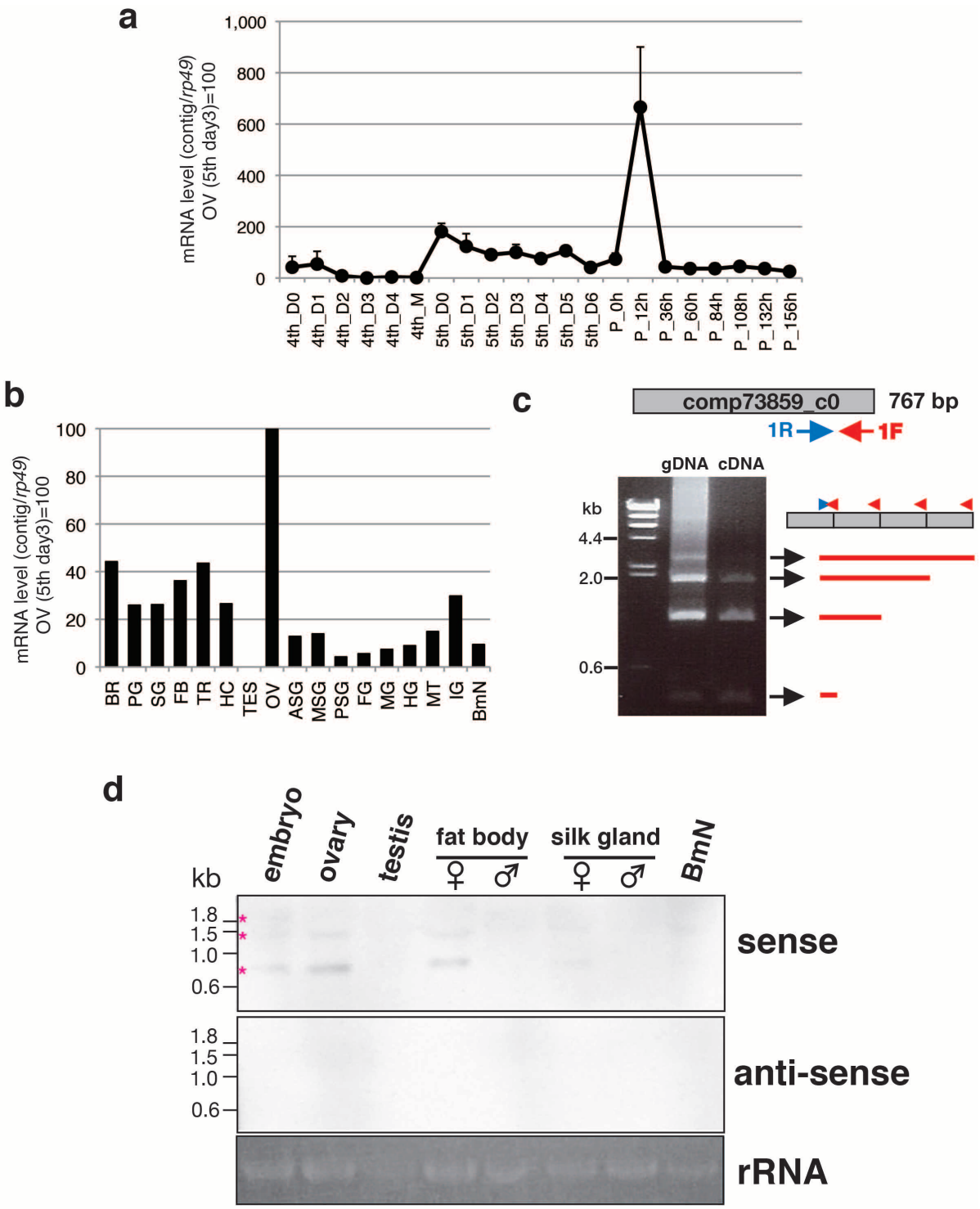
Phylogenetic analysis. The amino acid sequences of proteins in the NCBI database that showed significant homology (*E* value of $< 1 \times 10^{-9}$) to residues 51–122 of Masc were identified using the BLAST program. A neighbour-joining tree was constructed using 39 sequences and the reliability of the tree was tested by bootstrap analysis with 1,000 replications.

Statistical analysis. The sample size in each experiment was adjusted depending on the initial experimental results. Data distribution and normality were assessed by Prism 5 software (Graphpad). The data for *Fem* piRNA inhibitor (Fig. 3c) and *Siwi* RNAi (Extended Data Fig. 7a) experiments were subjected to Kruskal–Wallis analysis with post hoc Dunn's test. For *Masc* RNAi (Fig. 4b) experiment, one-way analyses of variance (ANOVA) was performed with post hoc Tukey's test. The data for *BmAgo3* RNAi (Extended Data Fig. 7b, d) experiments were subjected to Mann–Whitney test.

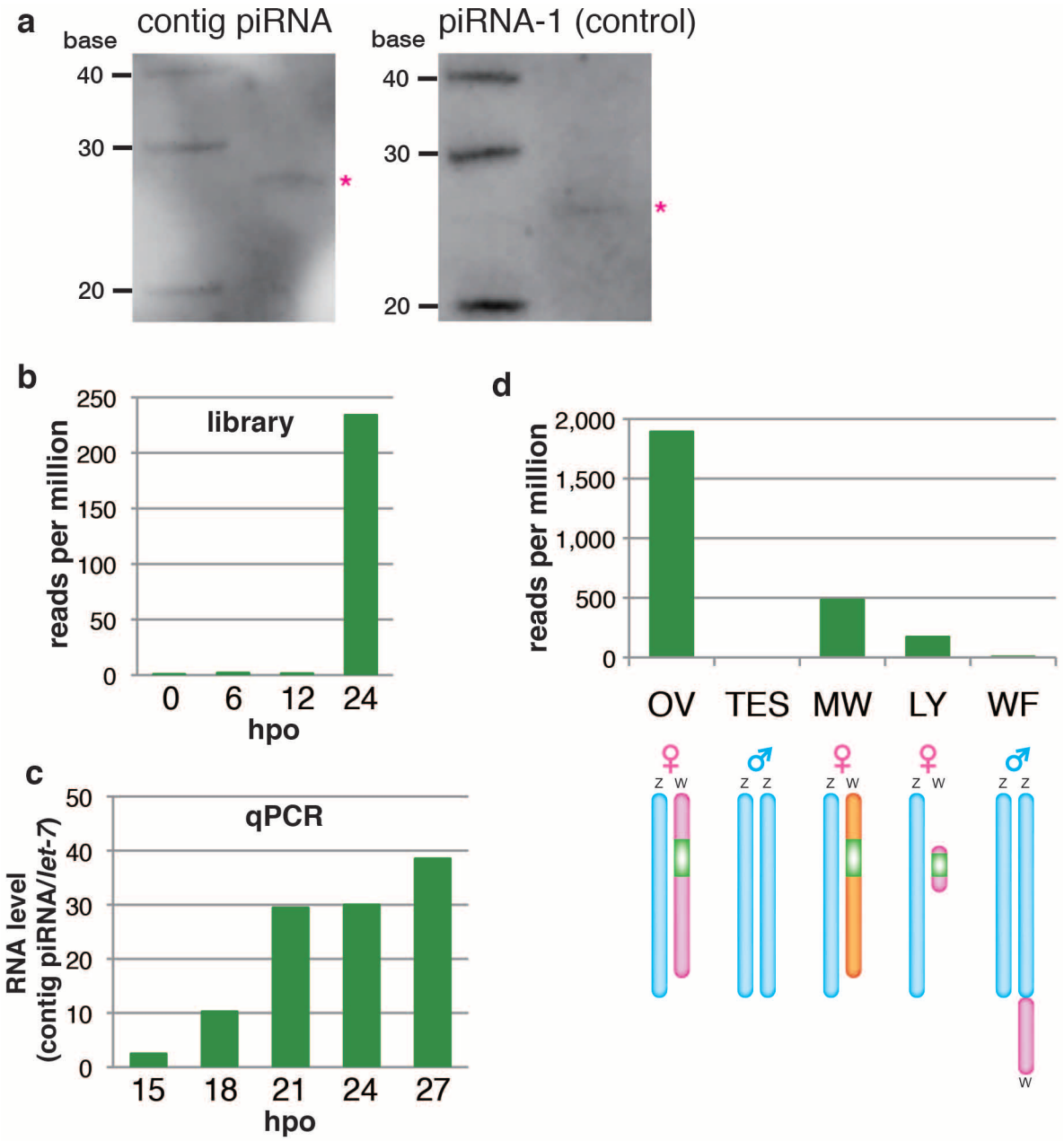
22. Sakai, H., Yokoyama, T., Abe, H., Fujii, T. & Suzuki, M. G. Appearance of differentiated cells derived from polar body nuclei in the silkworm, *Bombyx mori*. *Front. Physiol.* **4,** 235 (2013).
23. Sato, Y. *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nature Genet.* **45,** 860–867 (2013).
24. Sakudoh, T. *et al.* Diversity in copy number and structure of a silkworm morphogenetic gene as a result of domestication. *Genetics* **187,** 965–976 (2011).
25. Suzuki, M. G. *et al.* Establishment of a novel *in vivo* sex-specific splicing assay system to identify a trans-acting factor that negatively regulates splicing of *Bombyx mori dsx* female exons. *Mol. Cell. Biol.* **28,** 333–343 (2008).
26. Wang, L. *et al.* Mutation of a novel ABC transporter gene is responsible for the failure to incorporate uric acid in the epidermis of *ok* mutants of the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **43,** 562–571 (2013).
27. Yamaguchi, J., Mizoguchi, T. & Fujiwara, H. siRNAs induce efficient RNAi response in *Bombyx mori* embryos. *PLoS ONE* **6,** e25469 (2011).
28. Hutvágner, G., Simard, M. J., Mello, C. C. & Zamore, P. D. Sequence-specific inhibition of small RNA function. *PLoS Biol.* **2,** e98 (2004).
29. Shoji, K. *et al.* Characterization of a novel chromodomain-containing gene from the silkworm, *Bombyx mori*. *Gene* **527,** 649–654 (2013).
30. Katsuma, S. *et al.* Novel macula-like virus identified in *Bombyx mori* cultured cells. *J. Virol.* **79,** 5577–5584 (2005).
31. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnol.* **29,** 644–652 (2011).
32. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12,** 323 (2011).
33. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11,** R106 (2010).
34. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
35. Abe, H. *et al.* Identification of the female-determining region of the W chromosome in *Bombyx mori*. *Genetica* **133,** 269–282 (2008).
36. Fujii, T. *et al.* The female killing chromosome of the silkworm, *Bombyx mori*, was generated by translocation between the Z and W chromosomes. *Genetica* **127,** 253–265 (2006).
37. Reuter, M. *et al.* Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature* **480,** 264–267 (2011).



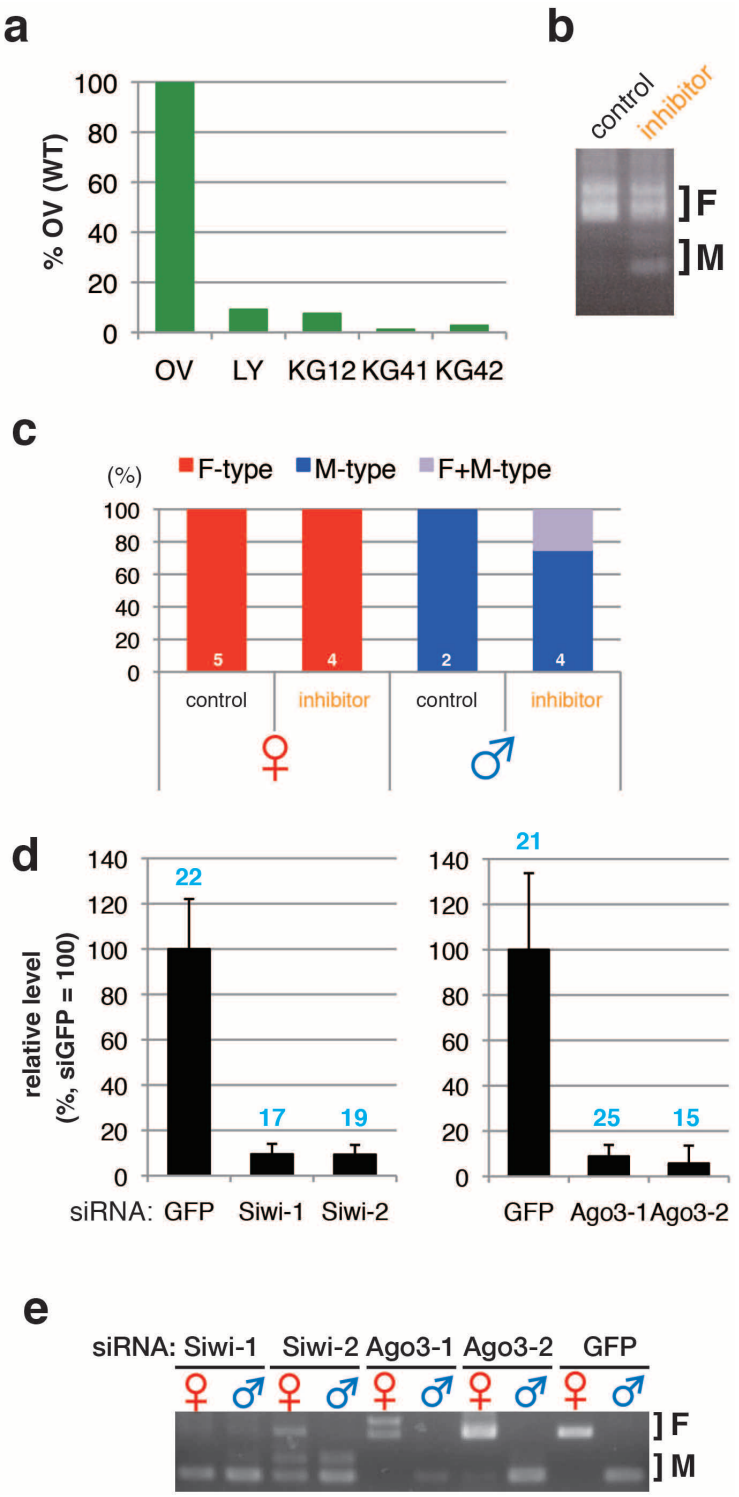
Extended Data Figure 1 | Molecular sexing and comparative transcriptome analysis of embryonic *B. mori*. **a**, Molecular sexing of individual embryos at 21 hpo. *Musashi*, *Sasuke* and *Bonsai* are W chromosome RAPD markers. 'Chr2' control bands are generated from a primer set that amplifies a sequence within the 2nd chromosome of *B. mori*. **b**, MA plots of RNA-seq data. The comp73859_c0 contig is indicated by red dots and highlighted by arrows. The axes show: *A* (*x*-axis) = ($\log_2$(transcripts per million in male) + $\log_2$(transcripts per million in female))/2. *M* (*y*-axis) = $\log_2$(transcripts per million in male) − $\log_2$(transcripts per million in female). **c**, Number of the comp73859_c0-derived transcripts in each RNA-seq library. Note that the comp73859_c0-derived transcripts detected in male libraries may be derived from incorrectly sexed embryos or RNA produced by polar bodies. Combined with RT–qPCR results of Fig. 1c, the expression level of this contig peaks around 18–21 hpo in the *B. mori* embryo.



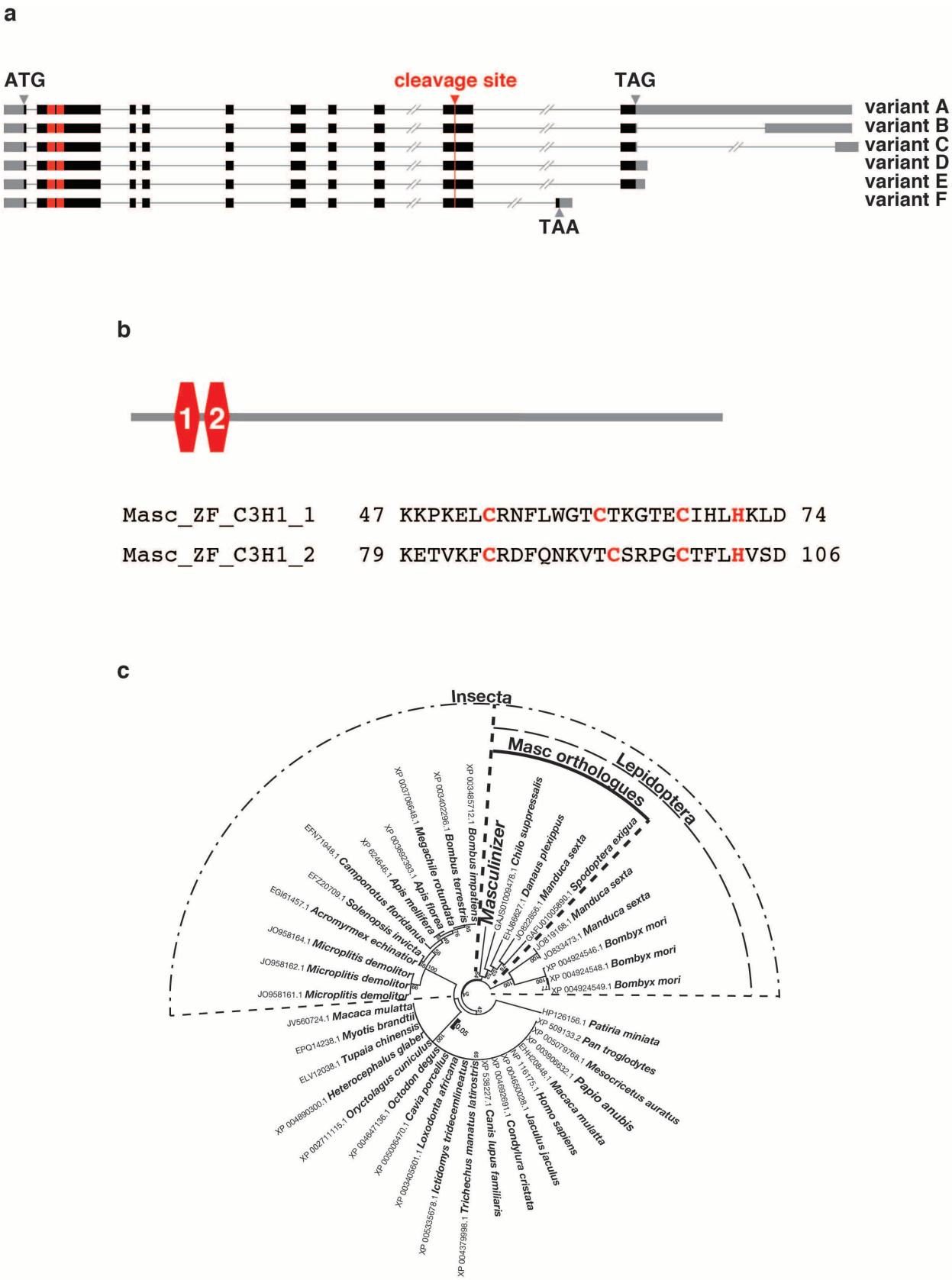
Extended Data Figure 2 | Expression profile of the female-specific comp73859_c0 contig. **a**, Developmental expression profile of the female-specific contig in ovary during the larval (4th and 5th instars) and pupal stages. RT–qPCR was performed using total RNA that was isolated from ovary of 4th and 5th instar larvae, and pupae (p50T). This contig was detected in the ovary of 4th and 5th instar larvae, and pupae of *B. mori* with a strong peak expression at an early pupal stage. *rp49* was used as an internal control. Data shown are mean + s.d. of three individuals, except for day 0 of 5th instar ($n = 2$). **b**, The mRNA expression in 17 different tissues from day 3, 5th instar larvae (p50T). RT–qPCR was performed using total RNA from brain (BR), prothoracic gland (PG), salivary gland (SG), fat body (FB), trachea (TR), haemocyte (HC), testis (TES), ovary (OV), anterior silkgland (ASG), middle silkgland (MSG), posterior silkgland (PSG), foregut (FG), midgut (MG), hindgut (HG), Malpighian tubules (MT), integument (IG) of male and female larvae (except for testis and ovary) or BmN4 cells (BmN). *rp49* was used as an internal control. **c**, Amplification of the female-specific transcript. Long PCR using female gDNA and cDNA as templates was performed with primers 1F and 1R. Black arrows show bands corresponding to single or multiple units of this transcript. The predicted structure of each unit was also indicated. **d**, Northern blot analysis of total RNA that was prepared from embryos (24 hpo) and tissues from day 3 5th instar F1 hybrid Kinshu × Showa larvae (ovary, testis, fat body, and silk gland), and BmN cells. The asterisks show major transcripts.



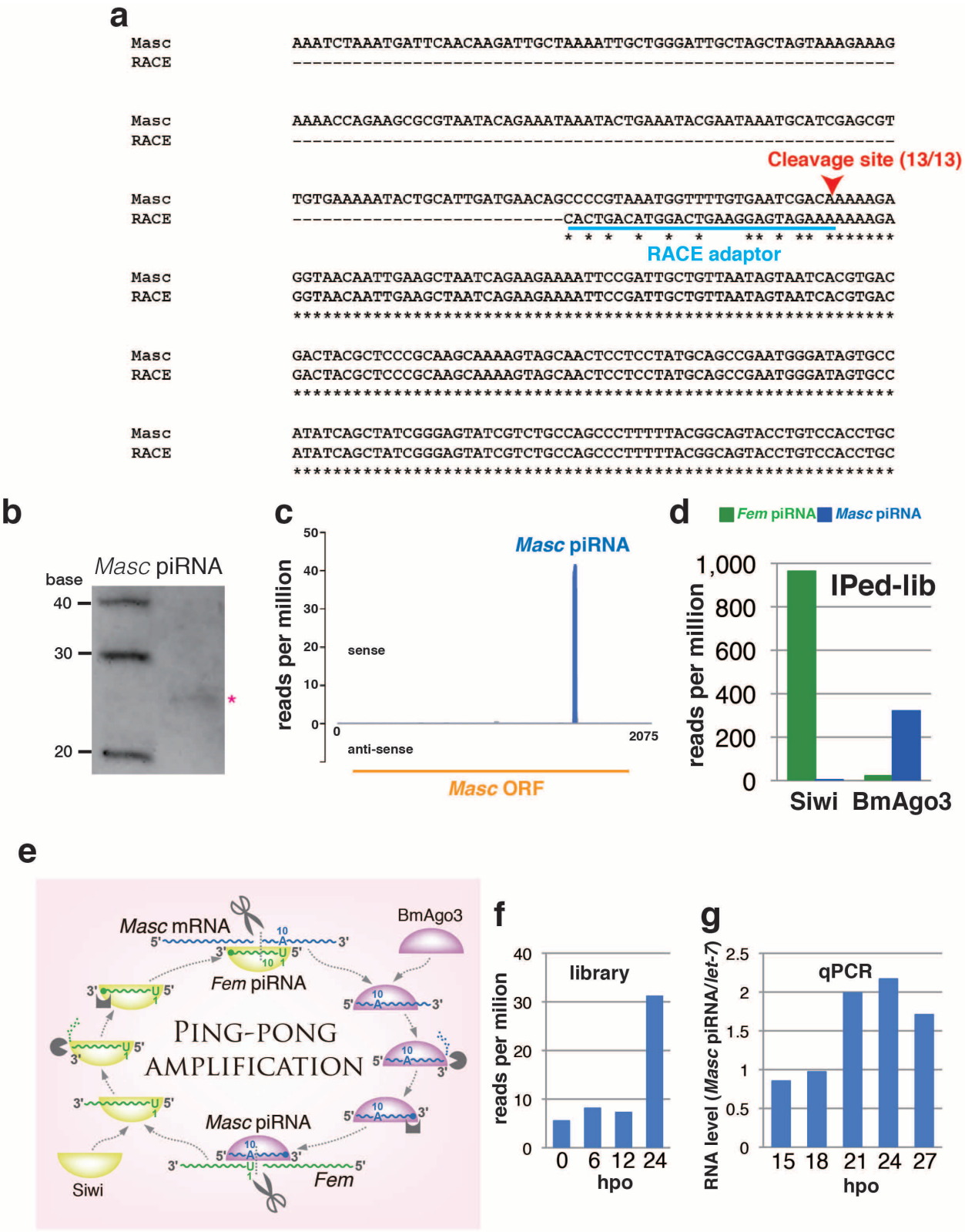
Extended Data Figure 3 | Characterization of the female-specific piRNA. **a**, Detection of contig-derived piRNA and piRNA-1 (control). Northern blot analysis was performed using total RNA prepared from early embryos. The asterisks show the location of each piRNA. **b**, Normalized reads of the female-specific piRNA in embryonic piRNA libraries of *B. mori*[12] generated at 0, 6, 12, and 24 hpo. Reads of 26–29 nucleotides that showed 2 or fewer mismatches to the corresponding piRNA sequence were scored as a positive match. **c**, RT–qPCR estimation of the female-specific piRNA levels in early embryos. The piRNA level was normalized to that of *let-7*. **d**, Normalized reads of the female-specific piRNA in piRNA libraries[6] from ovary and testis of wild-type *B. mori* or W chromosome mutants. Schematic representation of sex chromosomes of each strain is shown below the panel. The putative sex-determining region is represented by the green box. The orange bar represents the W chromosome derived from *B. mandarina*. OV, ovary from wild-type; TES, testis from wild-type; MW, ovary from MW strain; LY, ovary from LY strain; WF, testis from WF strain.



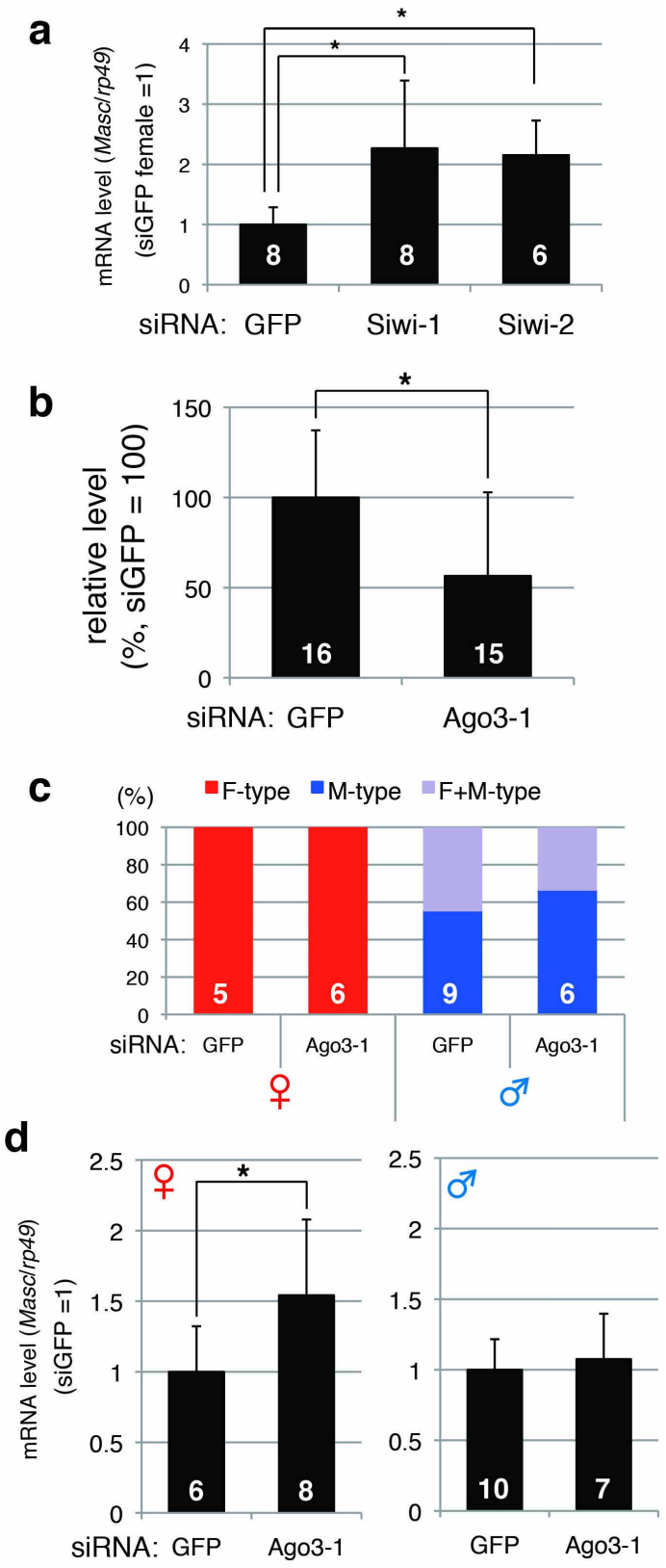
Extended Data Figure 4 | Effect of inhibition of the piRNA pathway on the splicing of *Bmdsx* transcripts. **a**, Abundance of the female-specific piRNA in three piRNA libraries constructed from three KG individual ovaries (KG12, KG41 and KG42)[13]. Of these, two (KG41 and KG42) showed a severe masculinized phenotype, and the rest (KG12) showed a weak phenotype. KG12 expressed a slightly lower amount of this piRNA than that of LY (82.4% of LY), whereas its expression in the ovary of severe masculinized individuals (KG41 and KG42) were markedly lower than LY's (12.1 and 29.7% of LY, respectively). The abbreviations are the same as in Extended Data Fig. 3d. **b**, Effect of the inhibitor RNA on the *Bmdsx* splicing. BmN4 cells were transfected with the inhibitor RNA or control RNA (that is, inhibitor for *GFP* piRNA), and the splicing patterns of *Bmdsx* were examined by RT–PCR. The F and M indicate female- and male-type splicing of *Bmdsx*, respectively. Similar results were obtained in three independent experiments. **c**, Effect of the RNA inhibitor on the *Bmdsx* splicing. The *Bmdsx* splicing patterns were examined at about 240 h post-injection (immediately after hatching). The abbreviations are the same as in Fig. 2c. The number indicates the sample size. **d**, Knockdown of *Siwi* or *BmAgo3* mRNAs in female and male embryos. The embryos were injected with two types of siRNAs that target *Siwi* (Siwi-1 and Siwi-2) or *BmAgo3* (Ago3-1 and Ago3-2) or a control siRNA that targets *GFP*. Total RNA was isolated from female or male siRNA-injected embryos at 72 h post-injection and RT–qPCR was performed. The data shown are mean + s.d. The number above each bar indicates the sample size of each group. **e**, Representative patterns of the *Bmdsx* splicing in siRNA-injected embryos. The F and M indicate female- and male-type splicing of *Bmdsx*, respectively.



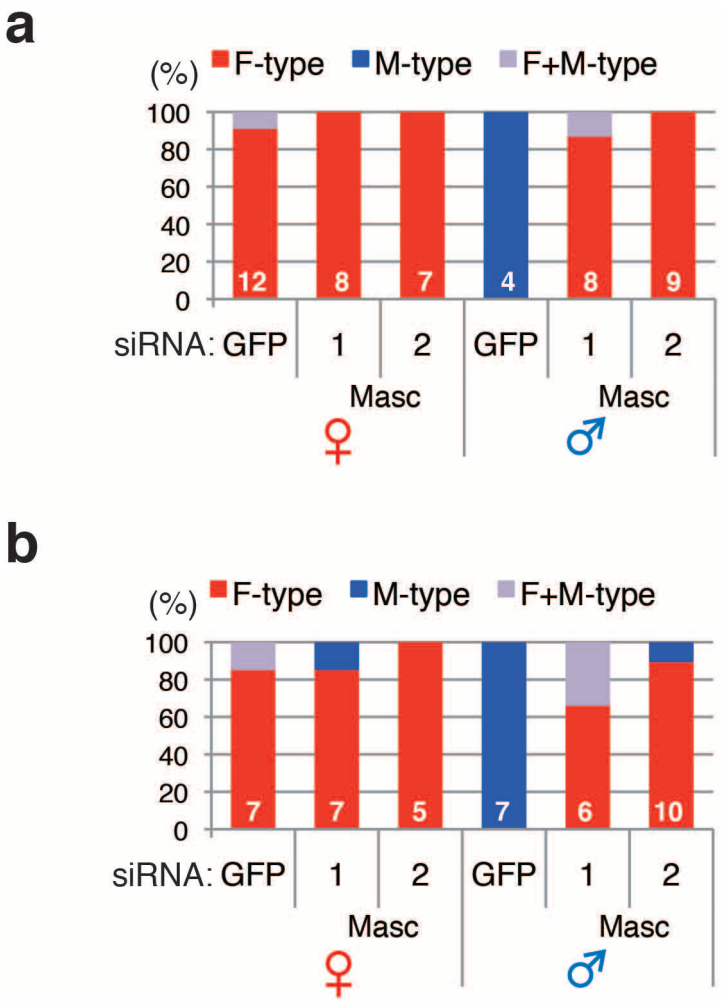
Extended Data Figure 5 | Characterization of *Masc*. **a**, Structure of *Masc* mRNA. Five *Masc* transcripts (A–E) that encode full-length Masc proteins but show unique splicing patterns in the 3′-untranslated region as well as one transcript (F) that encodes a truncated Masc protein are found. **b**, Domain structure of the Masc protein. The hexagons show the location of two CCCH-type zinc finger domains. The amino acid sequences of these domains are shown below. The conserved CCCH residues are shown in red. **c**, Phylogenetic analysis of Masc proteins. The neighbour-joining tree was generated using the amino acid sequences of zinc finger domains from proteins showing homology to Masc. The numbers on the internal branches represent the support value in the bootstraps of 1,000 replicates.



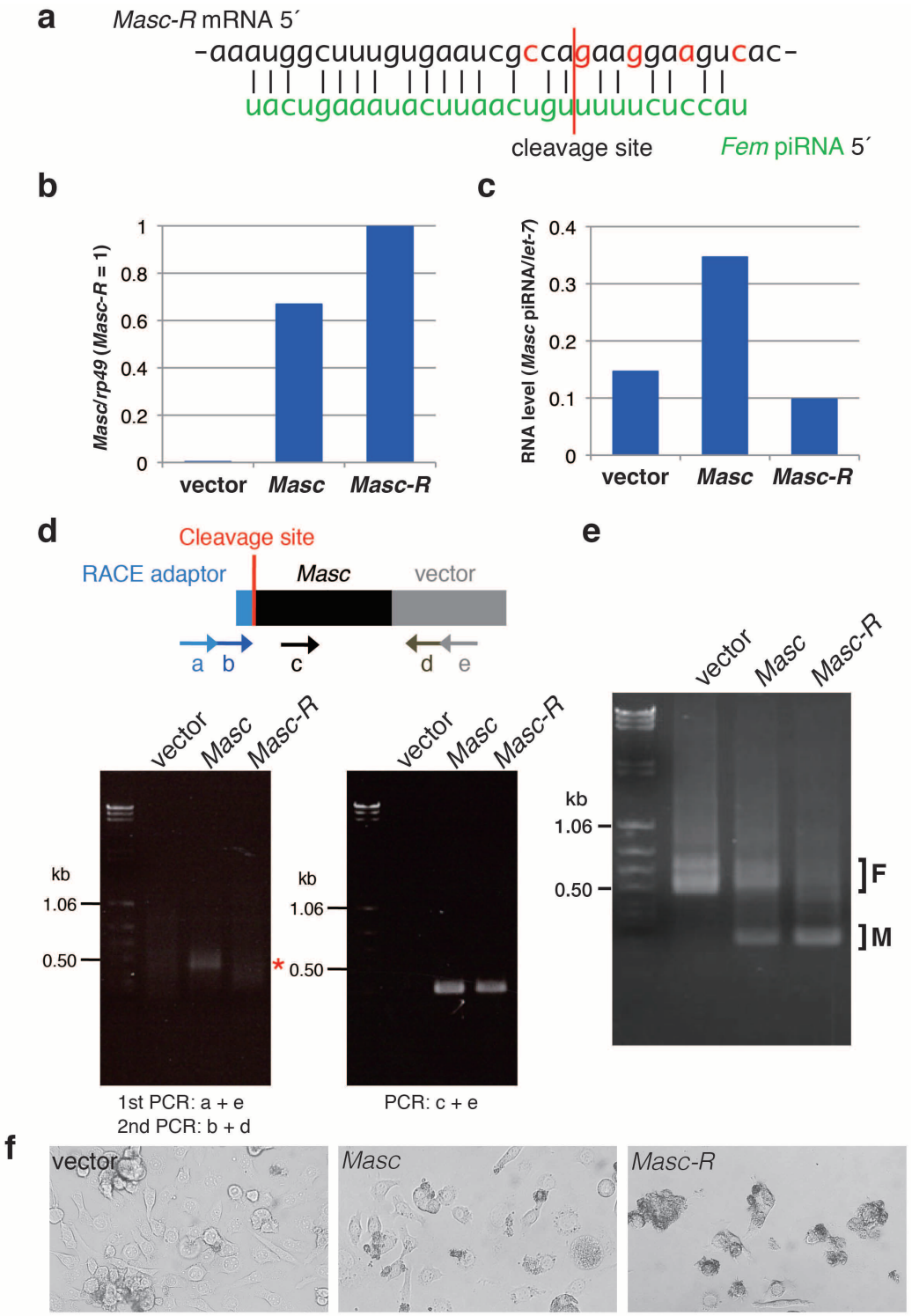
Extended Data Figure 6 | Cleavage of *Masc* mRNA. **a**, Identification of the cleavage site of *Masc* mRNA. The *Masc* mRNA-derived RNA fragments were amplified by a modified RACE method, cloned, and sequenced. The RACE adaptor and the cloned 5′-end are indicated. Thirteen 5′-ends were determined and all showed identical sequences. Nucleotides identical to the top sequence are represented by asterisks. **b**, Detection of *Masc* piRNA. Northern blot analysis was performed using total RNA prepared from early embryos. The asterisk shows the location of *Masc* piRNA. **c**, Mapping of embryonic piRNAs (24 hpo) onto *Masc* mRNA. The relative location of ORF of *Masc* is shown below. **d**, Normalized reads of *Fem* piRNA and *Masc* piRNA in Siwi- or BmAgo3-immunoprecipitated libraries from BmN4 cells[14]. **e**, A ping-pong amplification model of *Fem* piRNA/*Masc* piRNA. **f**, Normalized reads of *Masc* piRNA in embryonic piRNA libraries. Reads of 26–29 nucleotides that showed 2 or fewer mismatches to the *Masc* piRNA sequence were scored as positive. **g**, RT–qPCR estimation of *Masc* piRNA in early embryos. The *Masc* piRNA level was normalized to that of *let-7*.



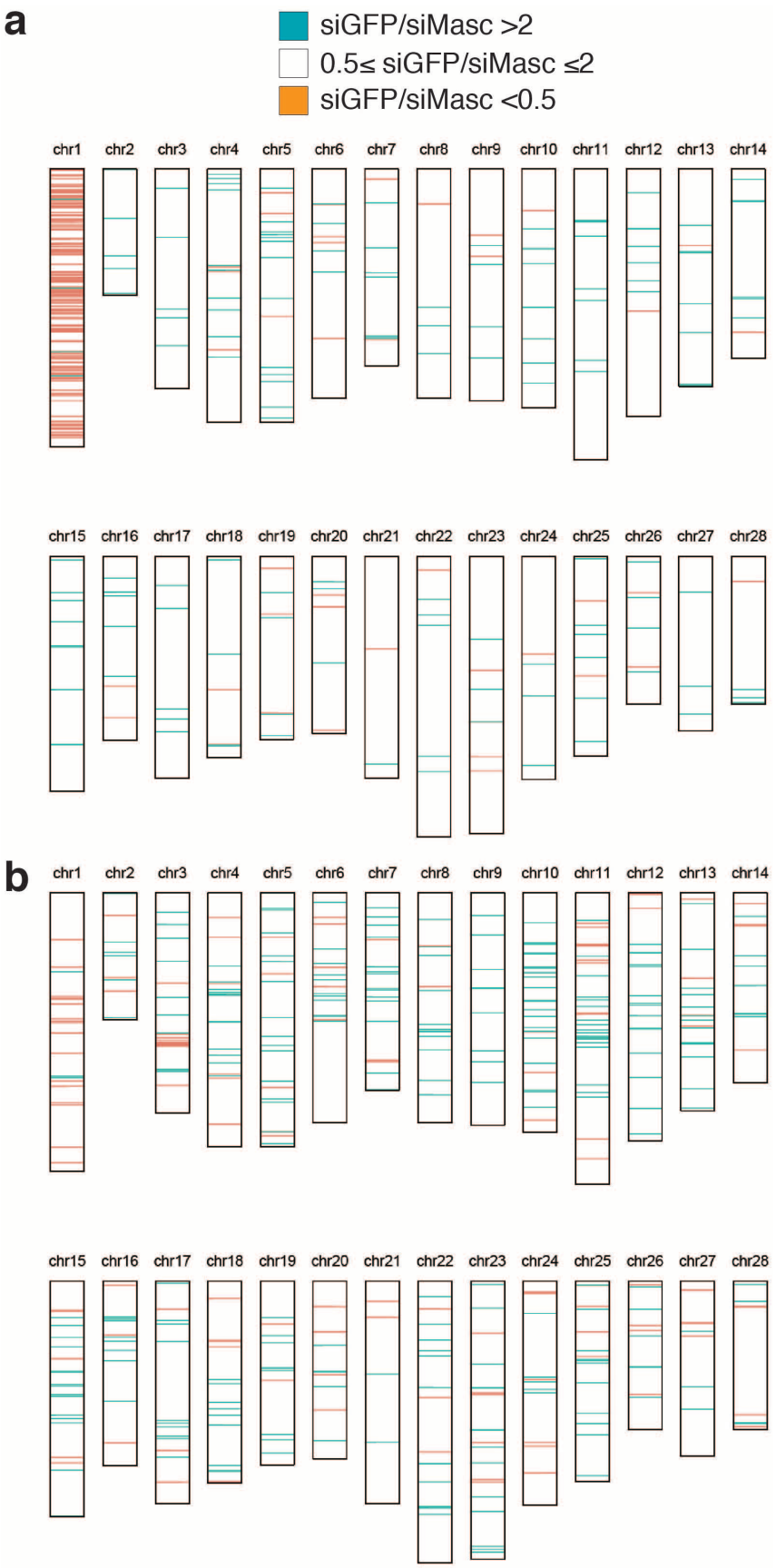
Extended Data Figure 7 | Effects of *Siwi* or *BmAgo3* knockdown on the *Bmdsx* splicing and *Masc* expression. **a**, *Masc* expression in female embryos injected with two types of siRNAs that target *Siwi* (Siwi-1 and Siwi-2) or a control siRNA that targets GFP. Total RNA was isolated from female siRNA-injected embryos at 18 hpo and RT–qPCR was performed. The data shown are mean + s.d. The number at the base of each bar indicates the sample size of each group. Data were subjected to Kruskal–Wallis analysis with post hoc Dunn's test. $*P < 0.05$. The expression levels of *Siwi* mRNA decreased to 23 and 44% after injecting Siwi-1 and Siwi-2 siRNAs, respectively, compared with that in GFP-siRNA-injected embryos. **b**, Knockdown of *BmAgo3* mRNA in newly hatched larvae. The embryos were injected with *BmAgo3* or *GFP* (control) siRNA. Total RNA was isolated from newly hatched larvae (at about 240 h post-injection) and RT–qPCR was performed. The data shown are mean + s.d. The number indicates the sample size of each group. $*P < 0.05$, one-sided Mann–Whitney test. **c**, Splicing of *Bmdsx* in newly hatched larvae that were injected with *BmAgo3* siRNA. The *Bmdsx* splicing patterns were examined at about 240 h post-injection. The number indicates the sample size. The abbreviations are the same as in Fig. 2c. **d**, *Masc* expression in newly hatched larvae that were injected with *BmAgo3* siRNA. Total RNA was isolated from siRNA-injected newly hatched larvae (at about 240 h post-injection) and RT–qPCR was performed. The data shown are mean + s.d. The number indicates the sample size of each group. $*P < 0.05$, one-sided Mann–Whitney test.



Extended Data Figure 8 | Splicing of *Bmdsx* in *Masc* siRNA-injected embryos. **a**, **b**, The *Bmdsx* splicing pattern was determined at 144 h (**a**) and 216 h (**b**) post-injection. The abbreviations are the same as in Fig. 2c. The number indicates the sample size.



Extended Data Figure 9 | Functional analysis of the *Fem* piRNA-resistant *Masc* transcript. **a**, Sequence of the *Fem* piRNA-resistant *Masc* (*Masc-R*) mRNA. Five nucleotide mutations that do not result in amino acid substitutions for the Masc protein are shown by red letters. The putative cleavage site by the *Fem* piRNA–Siwi complex is shown by the red line. **b**, RT–qPCR of *Masc* mRNA in cDNA-transfected BmN4 cells. BmN4 cells were transfected with *Masc* expression vectors or control vector. The *Masc* mRNA level was normalized to that of *rp49*. Data shown are means of duplicates. **c**, RT–qPCR of *Masc* piRNA in BmN4 cells transfected with *Masc* expression vectors or control vector. The *Masc* piRNA level was normalized to that of *let-7*. Data shown are means of duplicates. **d**, Identification of the cleavage site of exogenously introduced *Masc*. BmN4 cells were transfected with *Masc* expression vectors or control vector. Three days after transfection, zeocin (final concentration, 500 μg ml^{-1}) was added to the medium. Six days after drug selection, the *Masc* mRNA-derived RNA fragment (shown by the red asterisk) expressed from the transfected plasmids was amplified by a modified RACE method. The fragment was cloned, sequenced, and identified as the *Masc* mRNA-derived one. The locations of the primers are shown by arrows. **e**, Effect of *Masc* transfection on the *Bmdsx* splicing in BmN4 cells. The splicing patterns of *Bmdsx* in stably transfected BmN4 cells (six days after drug selection) were examined by RT–PCR. The F and M indicate female- and male-type splicing of *Bmdsx*, respectively. Similar results were obtained in two independent experiments. **f**, Light microscopic observations of BmN4 cells stably transfected with *Masc* expression vectors or control vector (2 weeks after drug selection).



Extended Data Figure 10 | Distribution of *Masc*-regulated genes throughout the silkworm genome. **a**, **b**, The genome loci where *Masc*-regulated genes are located were identified using RNA-seq data from male (**a**) and female (**b**) embryos injected with control (siGFP) and *Masc* (siMasc) siRNAs (72 h post-injection).