Absence of dosage compensation at the transcription level of a sex-linked gene in a female heterogametic insect, *Bombyx mori*

MASATAKA G. SUZUKI, TORU SHIMADA* & MASAHIKO KOBAYASHI

Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113, Japan

To confirm whether sex-linked genes of the silkworm are dosage compensated or not, we investigated three RAPDs and one RFLP marker which have been mapped to the Z chromosome as candidate genes to test for dosage compensation. Northern blot analysis demonstrated that one of the RAPDs, T15.180a, is transcribed into several mRNAs in the embryo, midgut, Malpighian tubule, testis and ovary. Southern blot analysis revealed that T15.180a is present at one copy in the genome. Sequence data showed that T15.180a contained two open-reading frames (ORFs) separated by an intron of 605 bp. These ORFs encoded a possible polypeptide of 445 amino acids containing the proline cluster. These results suggested that T15.180a is part of a functional gene on the Z chromosome of the silkworm. We have called this gene *T15*, and compared its transcriptional level between males and females. Northern blot analysis showed that the level of *T15* mRNA in males was twice that of females. For a more precise analysis, the levels of transcription from both alleles of *T15* (T15^P, T15^C) were measured by the single-nucleotide primer extension (SNuPE) assay. The results indicated that the amount of transcription from each *T15* allele in both sexes (male, T15^P/T15^C; female, T15^P/W) was equivalent. All the results suggest that sex-linked genes in the silkworm are not dosage compensated.

Keywords: *Bombyx mori*, dosage compensation, female heterogamety, RAPD, sex-linked gene, Z chromosome.

Introduction

The heterogametic sex has morphologically and genetically distinct sex chromosomes. With male heterogamety, males carry X and Y sex chromosomes and females are XX. With female heterogamety, females are ZW and males are ZZ. The chromosome that is restricted to the heterogametic sex (Y or W) often lacks most of the genes carried on its partner (X or Z). In those cases, there is a two-fold difference in the doses of X (or Z)-linked genes between the two sexes. In male heterogametic organisms as diverse as nematodes, fruit flies, mice and humans, this difference in gene dosage is compensated so that males and females produce equal amounts of gene products from such sex-linked genes. In mammals, dosage compensation is achieved by the inactivation of one of the two female X chromosomes (Borsani & Ballabio, 1993). In *Caenorhabditis elegans*, dosage compensation is brought about by decreasing by half the rate of transcription of genes on both of the female X

chromosomes (Hsu & Meyer, 1993). In *Drosophila melanogaster*, in contrast, dosage compensation occurs by doubling the transcription rate of genes on the male X chromosome (Baker *et al*., 1994). Products of many individual X-linked genes are found in equal amounts in males and females, whether assayed by their genetic (Muller, 1950) or enzymatic function (Lucchesi & Manning, 1987), or by direct measurement of protein (Korge, 1975) or transcript levels (Birchler *et al*., 1982; Ganguly *et al*., 1985). Evidence that dosage compensation occurs at the level of RNA synthesis at most X-linked loci comes from comparison of [^{3}H]uridine incorporation into nascent RNA along the polytene X chromosomes in male and female salivary glands (Muckherjee & Beermann, 1965).

On the other hand, in female heterogametic organisms, there are very few studies reported on sex-linked gene dosage compensation. For a small number of species in birds and butterflies, it has been confirmed that sex-linked genes are not dosage compensated (Cook, 1964; Johnson & Turner, 1979; Baverstock *et al*., 1982). Cytological studies and data

---

*Correspondence. E-mail: shimada@ss.ab.a.u-tokyo.ac.jp

on sex-linked morphological mutants suggest that birds do not show dosage compensation for sex-linked genes (Ohno, 1967). An apparent lack of dosage compensation is also shown by the measurement of enzymatic activity. In the domestic fowl, house sparrow and spotted turtledove, cytoplasmic aconitase is Z-linked and its activity in males is twice that in females (Baverstock *et al.*, 1982). As for lepidopteran insects, there was no compensation for a sex-linked gene for melanism in the moth *Lymantria monacha*, but this observation was hard to interpret, as there were indications that the two sexes differ in their general sensitivity to melanization (Goldschmidt, 1921). Stehr (1959) reported that haemolymph colours in Choristoneura spp. resulted from the interaction of an autosomal locus and a sex-linked modifier locus without dosage compensation. In *Heliconius melpomene* and *H. elato*, the activity of 6-phosphogluconate dehydrogenase (6PGD) located on the Z chromosome is not compensated between the two sexes (Johnson & Turner, 1979). In these heterogametic animals, no direct evidence that the absence of dosage compensation occurs at the level of transcription has been published.

As in the case with other Lepidoptera, *Bombyx mori* has female heterogamety (Hashimoto, 1933), but there have been no reports on dosage compensation in this insect. If *B. mori* does not also show dosage compensation like other heterogametic animals, the level of transcription of Z-linked genes would show a twofold difference between males and females. To confirm whether this is the case, it would be most desirable to compare the levels of Z-linked gene transcripts in the two sexes. More than 400 visible mutations have been found in this species, and about 200 of them have been mapped on 28 linkage groups. Although 15 mutants have been localized on the Z chromosome, it is very difficult to quantify their gene activities because these mutations are characterized by morphological or physiological abnormalities in egg formation, choriogenesis, embryogenesis, larval growth rate and larval epidermal pigmentation (Doira *et al.*, 1992). In addition to this conventional linkage map, linkage maps of RAPDs and RFLPs have been constructed. Three RAPDs (M1.80a, N20.70b, T15.180a) and one anonymous cDNA marker (Rcf96) have been mapped on the Z chromosome (Promboon *et al.*, 1995; Shi *et al.*, 1995). To investigate whether these DNA markers contain sequences of functional genes, the DNAs were cloned and used as probes for Northern blot analysis. One of the RAPDs, T15.180a, detected transcripts in various kinds of tissue at all developmental stages. Southern blot and sequence analysis showed that T15.180a is part of a

single copy gene on the Z chromosome, suggesting that it would be suitable for an examination of dosage compensation. We then named the gene *T15*, and compared the levels of transcription from *T15* in both sexes. In this paper, we present the first direct evidence for the absence of dosage compensation at the level of transcription of a Z-linked gene in Lepidoptera.

Materials and methods

Insects

Silkworm strains p50 and C108 were originally supplied by the Institute of Genetic Resources, Faculty of Agriculture, Kyushu University, Fukuoka, Japan and the National Institute of Genetics, Mishima, Japan, respectively. Individuals in the F_1 hybrid generation between C108 males and p50 females were also used in these experiments. Larvae were reared on fresh mulberry leaves.

Midguts were dissected out in PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4) from day-4 fifth instar larvae, frozen in liquid nitrogen, and stored at −80°C. In some experiments, other tissues, such as fat body, Malpighian tubule, ovary and testis were obtained from Day-4 and Day-5 fifth instar larvae and Day-0 and Day-3 pupae.

Cloning of sex-linked DNA markers

Sex-linked RAPDs M1.80a, N20.70b and T15.180a were amplified as described in Promboon *et al.* (1995). Amplified products were separated on a 2% agarose gel. The appropriate size of bands (M1.80a, 0.8 kbp; N20.70b, 0.7 kbp; T15.180a, 1.8 kbp) were cut out and DNAs were purified using a Geneclean kit (Bio101, La Jolla, CA), and cloned into a TA cloning site of pGEM-T vector, (Promega, Madison, WI). These clones were denoted as pGZM1, pGZN20 and pGZT15, respectively. The amplification product of Rcf96 (Shi *et al.*, 1995) was provided by M. R. Goldsmith (University of Rhode Island, Kingston, RI), and cloned as described above.

RNA purification and Northern analysis

Total RNA was isolated by the acid guanidinium thiocyanate phenol-chloroform (AGPC) method (Chomcyzynski & Sacchi, 1987). Poly(A)+RNA was isolated using a poly(A)+RNA extraction kit Quick Prep Micro (Pharmacia Biotech, Uppsala, Sweden) according to the protocol provided by the manufacturer. The electrophoretically separated and blotted RNAs were hybridized with the ^{32}P-labelled pGZT15 DNA which had been prepared by random primer labelling or with the [^{32}P]cRNA transcribed

from the pGZT15 DNA template using Riboprobe *in vitro* transcription systems (Promega, Madison, WI) according to the manufacturer's instructions at 42°C overnight in a hybridization solution ($5 \times$ SSC, 50% formamide, $5 \times$ Denhardt's reagent, 0.5% SDS and 100 μg mL^{-1} sheared salmon sperm DNA). The filter was washed in $0.1 \times$ SSC containing 0.1% SDS at room temperature, and $2 \times$ SSC containing 0.1% SDS at 50°C, and then in $0.1 \times$ SSC, 0.1% SDS at the same temperature. Exposure, analysis and printing were performed using a BAS 2000 Bioimage Analyser (Fuji Photo Film, Tokyo, Japan). As a control probe to ensure equal RNA loading, the ribosomal DNA and alkaline phosphatase gene of *B. mori* were used. The ribosomal DNA was obtained from pBmR161 (Fujiwara *et al.*, 1984) provided by H. Maekawa, National Institute of Infectious Diseases, Japan. The alkaline phoshatase gene was synthesized by PCR amplification with primers AF1 (5′-ATGGCAGAACGATAAGGA-GTC-3′) and AR1 (5′-GAAAGGCATTCTGTTG-CTGTC-3′), which were designed using the cDNA sequence (Itoh *et al.*, 1991), and cloned into a pGEM-T vector as described above.

Southern blot analysis

Genomic DNA was extracted from the posterior silk glands of mid-fifth instar larvae. A pair of posterior silk glands was disrupted in 1.0 mL of proteinase K solution (20 mM Tris-HCl, 100 mM EDTA, 1% Sarcosyl, 1 mg mL^{-1} proteinase K), and incubated at 50°C for 2 h with occasional swirling. One extraction was performed with phenol : chloroform : isoamyl alcohol (25 : 24 : 1), and one with chloroform : isoamyl alcohol (24 : 1) at room temperature. The resulting product was incubated with 10 μg mL^{-1} RNase A (Takara Shuzo Co. Ltd, Kyoto, Japan) for 30 min at 37°C, followed by phenol extraction as described above. Digested DNAs were separated by electrophoresis on a 1.5% agarose gel and transferred to a Hybond-N+ nylon membrane (Amersham International plc, Buckinghamshire, UK). Hybridization and washing were performed according to the manufacturer's recommendations. Radioactivity was visualized as described above.

DNA sequence and computer analysis

To determine the nucleotide sequence of T15.180a, the insert of pGZT15 was excised by digestion with *Apa*I and *Sac*I, and separated into 0.4 kb and 1.1 kb *Sac*I fragments and a 0.7 kb *Apa*I–*Sac*I fragment. These fragments were subcloned into pBluescript II SK+ (Stratagene Cloning System, La Jolla, USA) and sequenced as double-stranded DNA by dideoxy chain termination using fluorescence-labelled dideoxynucleotides (Perkin Elmer, Applied Biosystems Division, Foster City, CA). Sequence data were analysed with the sequence analysis software package, DNASIS v. 3 (Hitachi Software Engineering Co., Ltd, Yokohama, Japan). The deduced amino acid sequence was compared with the protein databases (Swiss-Prot and Pir) and the nucleotide sequence databases (GenBank, EMBL, DDBJ and PDB) using the programs BLASTP and TBLASTN, respectively (Altschul *et al.*, 1990, 1994). The DUST program was used to mask compositionally biased segments in protein sequence that tend to produce spurious hits in database searches (Hancock & Armstrong, 1994).

Reverse transcription and polymerase chain reaction (RT-PCR)

RT-PCR was performed using the LA RNA PCR kit (Takara Shuzo Co. Ltd, Kyoto, Japan) following the manufacturer's instructions. First-strand cDNA was synthesized from 100 ng of poly (A)+RNA from larval midgut with four T15.180a-specific antisense primers across the T15.180a, and the whole sample was used for PCR amplification with several primer pairs that covered the entire T15.180a sequence. All of the RT-PCR products were cloned into pGEM-T and sequenced as above.

Single nucleotide primer extension (SNuPE) assay

RT-PCR was carried out as described above. The first-strand cDNA was synthesized from 100 ng poly-(A)+RNA from larval midgut of individuals in the F_1 hybrid generation between a C108 mother and a p50 father with a random 9-mer primer. The resulting product was PCR-amplified with T15.180a specific primer pairs (p2, nucleotide numbers 344–364 in Fig. 3; and p3, nucleotide numbers 586–606 in Fig. 3) common to the target RNA and a standard DNA. The insert of pGMut464, which was derived from pGZT15 to possess one base substitution at the position next to the 3′ end of the SNuPE primer (A→X, nucleotide number 464 in Fig. 3), was used as a standard DNA. Ten ng of amplified DNA was used as a template for the SNuPE assay as described by Singer (Singer-Sam & Riggs, 1993). The SNuPE primer was localized at nucleotide numbers 484–465 in Fig. 3. The 3′ end of the SNuPE primer is just 5′ to the position of a polymorphism between strains p50 and C108 and a standard DNA (p50, G; C108, A; standard, C, nucleotide number 464 in Fig. 3). In separate tubes, the [^{32}P]dNTP appropriate for extension of each of the amplified products was added (p50, [^{32}P]dATP; C108, [^{32}P]dCTP; standard, [^{32}P]dGTP). Radio-

activity of the SNuPE products was normalized using a standard DNA.

Results

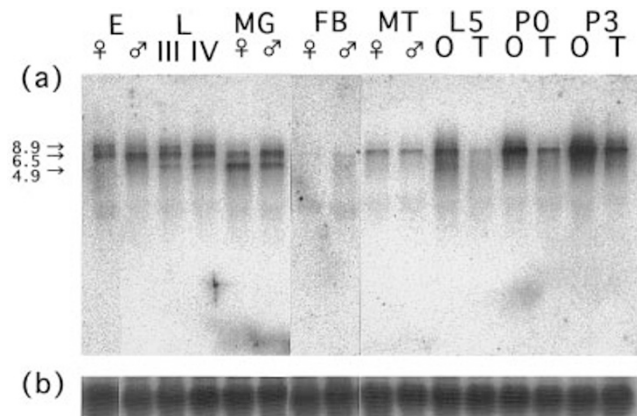### Expression of sex-linked DNA markers

Full length fragments of M1.80a, N20.70b, T15.180a and Rcf96 were used as hybridization probes for Northern analysis of RNAs from fertilized eggs, whole bodies and various tissues of larvae and pupae. M1.80a and N20.70b did not hybridize to RNA from any tissues. Rcf96 detected consistently a very low level of a 4.7 kb transcript in larval and pupal gonads (data not shown). T15.180a clearly hybridized to RNA of ≈ 6.5 kb from all tissues at any stages (Fig. 1). The level of this mRNA species was very weak in fat body, but increased in ovary and testis in accordance with their development. In addition to the 6.5 kb species, the probe also hybridized to RNA of 8.9 kb from fertilized eggs and whole bodies of third and fourth instar larvae. The probe also detected a 4.9 kb transcript in RNA from larval midgut and whole bodies of third and fourth instar larvae.

These results of Northern analysis indicate that T15.180a detects several transcripts in RNAs from various kinds of tissues at all developmental stages, from embryogenesis to the pupal stage. The expression patterns of these transcripts were different

Fig. 1 Northern blot analysis of T15.180a mRNA. (a) 10 μg of total RNA was loaded on each lane and hybridized with radiolabelled full-length T15.180a. E, embryos 33 h after oviposition; LIII, LIV, whole bodies of third and fourth instar larvae; MG, midgut; FB, fat body; MT, Malpighian tubule; O, ovary; T, testis. L5, P0 and P3 represent developmental stage, day-5 fifth instar larvae and day-0 and day-3 pupae, respectively. (b) The same blot was hybridized with *Bombyx* ribosomal DNA as a control. The molecular size in kilobases is indicated to the left of the figure.

among tissues. No sex-specific transcripts were detected in the examined tissues. However, the amount of 6.5 kb transcript was different between ovary and testis. This difference may be caused by the morphological and functional difference between these two tissues.
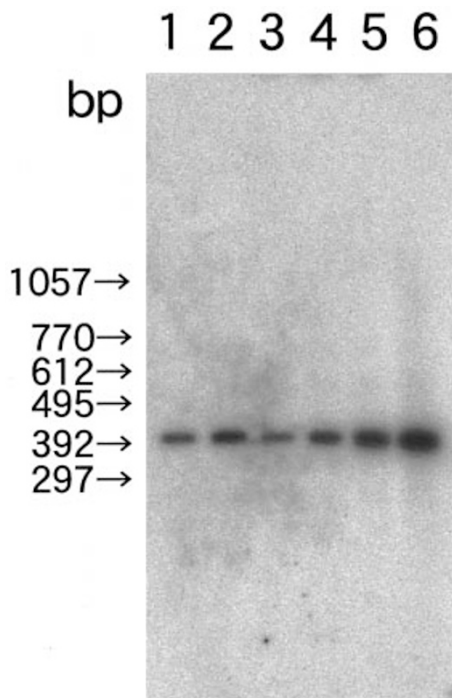
Copy number of T15.180a

The result of Northern blot analysis indicated that T15.180a detected at least three species of transcripts. This suggested that these multiple transcripts were produced by either alternative splicing of mRNA originated from the same gene or transcription from the several closely related genes. If the latter is the case, the levels of these transcripts could not represent the transcriptional levels of Z-linked genes, because there is the possibility that some of such closely related genes are located on autosomal chromosomes. We therefore examined the average copy number of the gene hybridized with T15.180a in the genome of *B. mori* in order to confirm whether or not there are genes which share sequence similarity with T15.180a. Genomic DNA from larval midgut was digested with *Sac*I and blotted after gel electrophoresis. The Southern blot also included DNA from pGZT15 digested with *Sac*I (the positions cleaved with *Sac*I are indicated by arrowheads in Fig. 3), containing T15.180a equivalent to 1, 2, 4 and 8 copies per haploid male genome (i.e. one set of autosomes and one Z chromosome). When the Southern blot was probed with a ^{32}P-labelled 0.4 kb *Sac*I fragment of T15.180a, the only genomic DNA band that hybridized was 0.4 kb long (Fig. 2). When the intensities of the hybridization in the genomic DNAs are compared with those of the various equivalents in the cloned DNA, it is apparent that the copy number of T15.180a is two in the male genome but one in the female genome. Because T15.180a is located on the Z chromosome, this result indicates that both males and females have one copy of the T15.180a sequence per Z chromosome. Therefore, the multiple hybridizing bands observed in Northern blot analysis of T15.180a expression (Fig. 1) represent differentially spliced forms of mRNA originated from the same gene, and do not correspond to the transcripts from several closely related genes. Taken together with the size difference between T15.180a (1.8 kb) and three transcripts (4.9 kb, 6.5 kb, 8.9 kb) observed in Northern hybridization (Fig. 1), these observations show that T15.180a encodes part of a Z-linked single copy gene likely to be at least 8.9 kb long based on the Northern blot data. As a matter of convenience, we denote this gene as *T15*.

T15.180a structure and sequence analysis

To obtain insight into the function of the *T15* gene through the identification of genes that share significant sequence similarity with *T15*, the nucleotide sequence of T15.180a was determined. Because T15.180a was derived from genomic DNA, it is desirable to confirm which strand is transcribed into mRNA observed in Northern blot analysis of T15.180a expression (Fig. 1). pGZT15 was transcribed by SP6 or T7 RNA polymerase in the presence of [^{32}P]GTP to prepare labelled RNA probe complementary to each strand of T15.180a, and these labelled cRNAs were hybridized with RNAs from the tissues indicated in Fig. 1. The hybridization pattern probed with the cRNA synthesized by T7 RNA polymerase was completely identical to that observed in Fig. 1 (data not shown). From this experiment, it was found that the sequence presented in Fig. 3 was the sense strand of T15.180a. (The nucleotide sequence data will appear in the DDBJ, EMBL and GenBank nucleotide

Fig. 2 Determination of T15.180a copy number in male and female genomic DNA. Genomic DNAs from female (lane 1) and male (lane 2) larvae digested with *Sac*I and electrophoresed on a 1.5% agarose gel. Included in the gel was DNA from pGZT15 digested with *Sac*I, 1, 2, 4 and 8 male haploid genome equivalents (lanes 3–6, respectively). A Southern blot of the gel was hybridized with a radiolabelled 0.4 kb *Sac*I fragment of T15.180a. Molecular sizes in basepairs are indicated to the left of the figure.

sequence databases with accession number AB006912.) To isolate T15.180a cDNA, RT-PCR was performed using poly(A) + RNA from midgut as a template with several different primers along T15.180a. This experiment was also expected to provide the molecular basis for the size difference of the *T15* messages because the RNA from midgut contained two species of transcripts produced by alternative splicing of *T15* mRNA as shown in Fig. 1. The sequences of the respective primers were follows. p1: 5′-TAGCTTTGACCTAAACCTAAC-3′ (nucleotide numbers 14–34 in Fig. 3); p2: 5′-CGAG-CAATGCGATCGTTCCATGA-3′ (nucleotide numbers 344–364 in Fig. 3); p3: 5′-CAATTCGGGA GTTGGGCTTAC-3′ (nucleotide numbers 586–606 in Fig. 3); p4: 5′-ACCTTCCACAGATATAA-TTCC-3′ (nucleotide numbers 802–822 in Fig. 3); p5: 5′-TGCTCTTCATTATGCCCTTGC-3′ (nucleotide numbers 1205–1225 in Fig. 3); p6: 5′-GCTATGCTTACGTGCCACGTC-3′ (nucleotide number 1891–1911 in Fig. 3). The couples of primers p1/p3, p2/p3, p2/p5 and p4/p5, respectively, amplified products of 600 bp, 260 bp, 880 bp and 430 bp, respectively. The sizes of these products corresponded to those of the products expected from the sequence data of T15.180a. The primer pair p4/p6 amplified a product of 520 bp, which was shorter than the predicted size (data are not shown). The sequences of these RT-PCR products were compared with the T15.180a sequence and the smaller product generated by the p4 and p6 primer pair was found to lack a segment of 605 bp found in the T15.180a sequence (nucleotide numbers 1267–1871 in Fig. 3). This segment started with GT and ended with AG (Fig. 3). Therefore, we conclude that T15.180a retains a 605 bp intron. From these experiments, we did not detect any transcripts produced by alternative splicing of *T15* mRNA. This suggests that alternative spliced sites are located at the flanking region of T15.180a. Knowledge of the full extent of alternative splicing awaits isolation and analysis of the entire gene sequence.

To analyse the coding potential of T15.180a, we searched for ORFs in its sequence. T15.180a contains two ORFs separated by a 605 nt intron. The first ORF is 1266 nt and the second one is 69 nt in length. A feature of potential interest of the T15.180a protein sequence is the presence of a proline-rich region between positions 358 and 384 (Fig. 3). Database searches indicated that the T15.180a protein shared partial homology with many proteins, most of which have proline-rich domains, but none of these homologies extended beyond the proline cluster into the surrounding regions. In addition, when the database searches using BLASTP and

TBLASTN (Altschul *et al.*, 1990, 1994) were combined with the DUST program (Hancock & Armstrong, 1994) to mask compositionally biased regions of the polypeptide (e.g. proline-rich regions), T15 did not show significant sequence similarity to proteins in current databases. Therefore, we were not able to estimate the function of T15 protein. The knowledge of the full length amino acid sequence awaits isolation and analysis of the entire cDNA sequence.

Comparison of T15 *mRNA level between males and females*

As shown above, *T15* is a single-copy gene located on the Z chromosome of the silkworm. *T15* is transcribed in various kinds of tissues at all developmental stages, and its expression pattern is the same between males and females. These suggest that *T15* can be used to investigate the dosage compensation

of Z-linked genes in the silkworm by comparing the transcriptional level between males and females.

Initially, we examined the steady-state level of *T15* mRNA transcripts in males and females by Northern blot hybridization. In this case, it is necessary to use RNAs from tissues which do not show sexual dimorphism. Among such tissues, midgut contained the largest amount of *T15* mRNA (Fig. 1). Therefore, poly(A)+RNA from larval midgut was subjected to Northern blot analysis using T15.180a as a probe. In addition to the 4.9 kb and 6.5 kb hybridizing bands, an 8.9 kb band which was observed in RNAs only from embryos and larval whole bodies (Fig. 1) was also detected (Fig. 4). Moreover, a very weak but detectable band of ≈3.4 kb was seen, probably as a result of *T15* mRNA being concentrated by the use of poly-(A)+RNA instead of total RNA. The hybridization patterns were the same between males and females.

```
   1 GATGCCACTACCCTAGCTTTGACCGAAACCGAACTGTCGGACTGGGCGGCCGAGAGTGGAGTATTAGACGATTGTGGT
   1  D  A  T  T  L  A  L  T  E  T  E  L  S  D  W  A  A  E  S  G  V  L  D  D  C  G
  79 TTTGACGAGAAGGACGACAGGAAAAGAAGCAAGAATCCCAGGACATTGAGCGGGCCAAAGCTGGTGCACGAAGCCAAG
  27  F  D  E  K  D  D  R  K  R  S  K  N  P  R  T  L  S  G  P  K  L  V  H  E  A  K
 157 AACATATCAGCTGTAGCGTCGCACGTATGCGGCAAGACGAGCCCCACCGAACCGGTAGTTTATTCGAACGCGTTAGAA
  52  N  I  S  A  V  A  S  H  V  C  G  K  T  S  P  T  E  P  V  V  Y  S  N  A  L  E
 235 CATTTCGAGTTCGCTGACGAGGGAGAACAAGATCCGTCCCTCGAAAGCCCCCTCACGCCCAAGAACTACGGCTACATG
  79  H  F  E  F  A  D  E  G  E  Q  D  P  S  L  E  S  P  L  T  P  K  N  Y  G  Y  M
 313 GAGCTGGTCGACGACGACTACGCGTATTCGCCGAGCAACGATCGTTCCATGAATTTCATCGAAAAAACCTTTTCGGAG
 105  E  L  V  D  D  D  Y  A  Y  S  P  S  N  D  R  S  M  N  F  I  E  K  T  F  S  E
 391 ACCGTTTTGAAGCCCGCTAGCCAAGAAGTTAGCGTCCGTCAAGAGGAGCCGATCACCGAAACGAGCGCTGAAACAGAA
 131  T  V  L  K  P  A  S  Q  E  V  S  V  R  Q  E  E  P  I  T  E  T  S  A  E  T  E
 469 GCCGTATCAGAACCTCAAGGCGAGTTAAAGTCGCCGCCGCACTGCTTGAAGTATACAGACACCGATCCAAGCGACAGG
 157  A  V  S  E  P  Q  G  E  L  K  S  P  P  H  C  L  K  Y  T  D  T  D  P  S  D  R
 547 TCGACGAGCGAGAAACAAACCAGTGATAGCATAACAAACGTAAGCTCAACTCCCGAATTGAAAATAGGGGAGACTGAGT
 182  S  T  S  E  K  Q  T  S  D  S  I  T  N  V  S  S  T  P  E  L  K  I  G  R  L  S
 625 CTATCGGAGGTATCTTCCCCGATAGATAAAGTAGACAAGACTGAGTCTTTAGAGGAGGCCTCGCCGCCGCTGATCTCC
 208  L  S  E  V  S  S  P  I  D  K  V  D  K  T  E  S  L  E  E  A  S  P  P  L  I  S
 703 GAGTCAACCATGTCCAGTATAGTTCCGAATTCCATACCGTTCTCTGGGTCTTGTTCGATAAGGCTCTATTCTCCGGCG
 235  E  S  T  M  S  S  I  V  P  N  S  I  P  F  S  G  S  C  S  I  R  L  Y  S  P  A
 781 ATATGCAGGTCCGCAAGCGAAACCTTCCACAGATCGAATTCCCGCAACTTGGAGTCCCCACGAAGCGGTGACGTATCG
 261  I  C  R  S  A  S  E  T  F  H  R  S  N  R  N  L  E  S  P  R  S  G  D  V  S
 859 ATGTCGATCAACTTGAGCTCGGCCAAGTTCCATATCACCAATGAGTCCCGTGCTGGCGCAGGACACAACTGATAAAGTA
 287  M  S  I  N  L  S  S  A  S  S  I  S  P  M  S  P  V  L  A  Q  D  T  T  D  K  V
 936 CAAGAGATAAAGCGAGAAAGAGAGGAACAGACCGAGCTGGTGAGACGACTGGTCCTGGAGAGACTCGGCAGCGGTCCT
 313  Q  E  I  K  R  E  R  E  E  Q  T  E  L  V  R  R  L  V  L  E  R  L  G  S  G  P
1014 CGTGCGAATAGGAAATGCAACCGAAGAATACGTGCATCGCCTACATGCAGCATGTCTCCACCCGTACCTCCTCCGCCG
 339  R  A  N  R  K  C  N  R  R  I  R  A  S  P  T  C  S  M  S [P  P  V  P  P  P  P
1093 GTATTGCCGGAGCCTCCACCGCCGCCACCACCCCCGAGACCTTCTCCTCCCCTTATGCCTATACAAGTCACCCCTTCA
 365  V  L  P  E  P  P  P  P  P  P  P  P  R  P  S  P  P  L  M  P] I  Q  V  T  P  S
1171 TTCTCCGACCCAGAACTCGCGAGGGAAAGGCGAAGCAAGGGCATAATGAAGAGCATATCGAGTTATCTAAATCGTAGA
 391  F  S  D  P  E  L  A  R  E  R  R  S  K  G  I  M  K  S  I  S  S  Y  L  N  R  R
1249 CTCGGACCGCGACATAAGGTGAGCTCATAATATCCTCGTCTAATCATGTAAACATATTCACTAAAAGCAAGCAGAGGG
 417  L  G  P  R  H  K
1326 TACCTGACATCAATTGTGTGATAATTGTCACTGCAATGCGCATCGTTCATTGCATAGATGCACAAGAACATTTCCATT
1405 ACATTAAAAAATTGTGTCGAATATTAAGATTGATTGTTAAAAACTAATATACACTAAGATTCCATTCAACAAATGAAA
1483 AAATCACACTGAAAATGTGAATATTGTCAAATTATAACCCAAACGGTATTGGAGTTTAACAAAAAGAAGGGTGTCCAC
1561 GATTTGGCTTTAATTTTAGTTAAACACGGAGTTGGAATAGTTGTTGGCACTGTTTTAATTTTCTCTATAAAACATTTG
1638 GAGGAGAGTAACATTGAAATTGTTAAGAAAAACACGTAAATCACAATTATGTTTACATTTGAAATTGAACTTCAAAAT
1717 TTTTGTTACTAGTAACAATTTAATTATACACTTGCATAATATGTAAAAGAGTTACCCGTTTCCAAATTTTCAAAATTT
1794 CGTGTATTTTTTTTTCTATTTATGAAATTTTTCGTAGATTAGGACACCGGTAGACTAACAAAAGTTGAATTTTCAG
1872 TGGGTTTCCGAGCCGGAGCTGACGTGGCACGTAAGCATAGCATTTGAATCTGTCGTGAGTGGCATCCAA  3'
 423  W  V  S  E  P  E  L  T  W  H  V  S  I  A  F  E  S  V  V  S  G  I  Q
```

Fig. 3 The nucleotide sequence of T15.180a. Deduced amino acid sequences encoded by the putative open-reading frames (ORFs) are shown below the nucleotide sequence. The positions cleaved with *Sac*I are indicated by arrowheads. The invariant bases in the consensus sequences of splice acceptor and donor sites are underlined. The proline-rich region is boxed.
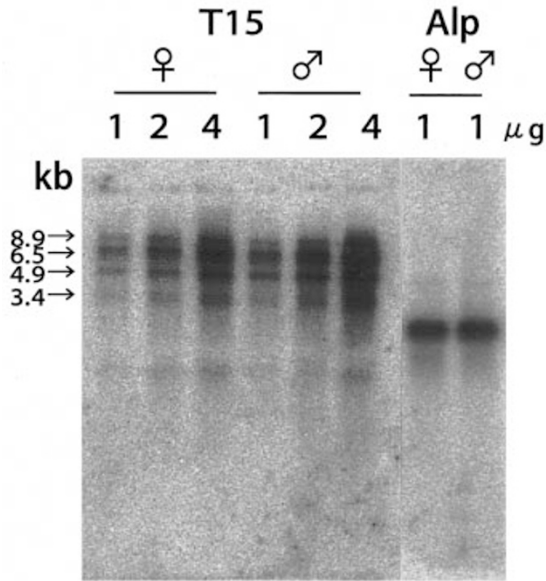
Fig. 4 Determination of the relative steady-state level of the *T15* mRNA in male and female poly(A)+RNA extracted from the midgut of day-3 fifth instar larvae. Northern blots were hybridized with ³²P-labelled T15.180a DNA and *Bombyx* Alp (alkaline phosphatase) cDNA. Numbers at the top of the lanes represent the amount of poly(A)+RNA loaded in each lane. Molecular size in kilobases is indicated to the left of the figure.

As shown in Table 1, the level of each *T15* mRNA transcript in male RNA was approximately twice as great as that of the female RNA. To determine whether the observed difference in the steady-state levels of *T15* mRNAs in males and females is a result of inaccuracy in RNA quantification, the steady-state level of *Bombyx* alkaline phosphatase (Alp) mRNA, which is encoded by an autosomal gene, was measured in these two RNA samples. As shown in Table 1, the ratio of the steady-state levels of *T15* mRNAs over Alp mRNA in males was approximately twice that in females.

*Measurement of transcripts from both alleles of* T15

If absence of dosage compensation is the case in the silkworm, the rate of transcription of genes on the Z chromosome will not be altered under both ZZ (male) condition and ZW (female) condition. We therefore analysed the levels of transcription from both alleles of *T15*. In order to quantify the allelic transcripts separately, we used single nucleotide primer extension (SNuPE) assay (Singer-Sam & Riggs, 1993) by taking advantage of a single-base polymorphism in an exon of *T15* between strains p50 and C108. We denote the p50 form of *T15* as $T15^P$ and that of C108 as $T15^C$. cDNAs were prepared from midgut samples dissected from individuals in the F₁ hybrid generation between a C108 mother

Table 1 Levels of *T15* and Alp mRNA in male and female *Bombyx mori*

Probe	μg RNA	Transcript no.†	PSL value‡		*T15*/Alp ratio	
			Male	Female	Male	Female
T15.180a	I	I	37.9	18.9	0.07	0.04
		II	63.7	34.3	0.11	0.07
		III	52.1	23.5	0.09	0.05
		IV	22.1	9.2	0.04	0.02
	2	I	74.9	43.3	0.13	0.09
		II	136.2	70.1	0.24	0.14
		III	109.8	48.8	0.2	0.1
		IV	43.6	22	0.08	0.04
	4	I	144.2	82.4	0.26	0.17
		II	265.3	133.2	0.48	0.27
		III	217.1	92.5	0.39	0.19
		IV	91.2	46.2	0.16	0.09
Alp DNA	I		557.4	495.3	—	—

†8.9 kb, 6.5 kb, 4.9 kb and 3.4 kb transcripts observed in Fig. 4 are denoted as I, II, III and IV, respectively.
‡PSL value indicates the radioactivity of respective area of Northern blot showing hybridization measured using Bioimage Analyser BAS2000.

C108 × p50 F₁
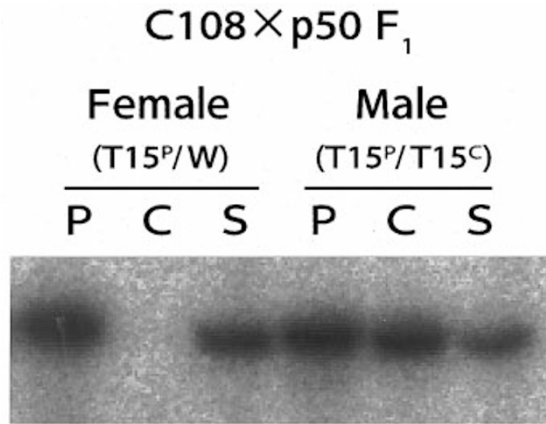
Fig. 5 Determination of the levels of transcription from both *T15* alleles. The genotype of *T15* in male and female is bracketed (p50 form of *T15* allele, T15^P; C108 form of *T15* allele, T15^C; W indicates the W chromosome). Upper case at the top of each lane shows the SNuPE product from each *T15* allele (P, *T15^P*; C, *T15^C*; S, standard).

and a p50 father. These cDNAs were mixed separately with a fixed amount of standard DNA prepared from pGMut464 (see Materials and methods). We then compared the amount of allelic transcript of *T15* between males (*T15^P*/*T15^C*) and females (*T15^P*/W) (Fig. 5).

In males, which possess two Z chromosomes, the level of transcript from *T15^P* was equal to that from *T15^C* (Table 2). On the other hand, females that possess only one Z chromosome derived from a p50 father, had the transcript produced only by *T15^P*, and the level was almost the same as that produced by each of the *T15* alleles in males. These results indicate that the transcription rate of *T15* alleles is not altered in both sexes, showing that there is an apparent lack of dosage compensation for *T15* in the silkworm.

Table 2 Levels of SNuPE products from the *T15^P* and *T15^C* alleles of *T15* in male and female *Bombyx mori*

	Male		Female	
	T15^P	*T15^C*	*T15^P*	*T15^C*
PSL value†	2.49	1.97	2.12	0.04
Relative level‡	1.17	0.93	1.00	0.02

†PSL value indicates the amount of radioactivity measured by use of Bioimage Analyser BAS2000.
‡Relative level shows the ratio of SNuPE obtained to the amount resulting from female *T15^P* (1.00).

Discussion

In order to study dosage compensation of sex-linked genes, it is critically important to quantify the levels of sex-linked gene transcripts in males and females. By Northern blot analysis using T15.180a, which is one of the known Z-linked RAPDs in *B. mori*, as a probe, we found a gene expressed in various kinds of tissues at all developmental stages, from embryogenesis to pupal stage (Fig. 1), and named this gene *T15*. *T15* was found to be transcribed into at least four variant forms produced by alternative splicing (Fig. 4). As the expression patterns of *T15* are different among tissues, it seems that *T15* mRNA would undergo tissue-specific splicing. We did not detect any sex-specific *T15* transcripts. The predicted amino acid sequence of T15.180a contained a proline-rich region. These regions are found, for example, in many proteins involved in the regulation of intracellular signalling, such as 3BPI, formin, WBP-1, 2 and Nef (Ren *et al.*, 1993; Chen & Sudol, 1995; Saksela *et al.*, 1995).

The level of *T15* mRNA in males is approximately twice that in females. The respective amounts of transcripts from *T15* alleles in both sexes were equal. These observations strongly suggest that the transcription rate of the *T15* allele is not altered in both sexes.

Our findings indicate that sex-linked genes in the silkworm are not dosage compensated. When taken in conjunction with the morphological and enzymatic data (Goldschmidt, 1921; Ohno, 1967; Johnson & Turner, 1979), the present data suggest that lack of dosage compensation is a general rule in lepidopterans. Previous studies in birds and butterflies have shown the absence of dosage compensation at the phenotypic level, such as morphological or enzymatic characteristics. Our results suggest that the lack of dosage compensation occurs not only at the phenotypic level but also at the level of transcription. This observation indicates that the approximately twofold difference in sex-linked enzyme activity of the two sexes in birds and butterflies (Johnson & Turner, 1979; Baverstock *et al.*, 1982) results from the steady-state amount of those mRNAs. To confirm the validity of our conclusion, further study will be required to determine whether or not the other Z-linked genes in the silkworm are compensated.

Acknowledgements

the Ministry of Education, Science, Sports and Culture, Japan.

References

ALTSCHUL, S. F., BOGUSKI, M. S., GISH, W. AND WOOTTON, J. C. 1994. Issues in searching molecular sequence databases. *Nature Genet.*, **6**, 119–129.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. AND LIPMAN, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.

BAKER, B. S., GORMAN, M. AND MARIN, I. 1994. Dosage compensation in *Drosophila*. *Ann. Rev. Genet.*, **28**, 491–521.

BAVERSTOCK, P. R., ADAMS, M., POLKINGHORNE, R. W. AND GELDER, M. 1982. A sex-linked enzyme in birds– Z-chromosome conservation but no dosage compensation. *Nature*, **296**, 763–766.

BIRCHLER, J. A., OWENBY, R. D. AND JACOBSON, K. B. 1982. Dosage compensation of serine-4 transfer RNA in *Drosophila melanogaster*. *Genetics*, **102**, 525–537.

BORSANI, G. AND BALLABIO, A. 1993. X chromosome dosage compensation in female mammals. *Semin. Dev. Biol.*, **4**, 129–139.

CHEN, H. I. AND SUDOL, M. 1995. The WW domain of yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 7819–7823.

CHOMCZYNSKI, P. AND SACCHI, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem.*, **162**, 156–159.

COOK, A. G. 1964. Dosage compensation and sex-chromatin in non-mammals. *Genet. Res.*, **5**, 354–365.

DOIRA, H., FUJII, H., KAWAGUCHI, Y., KIHARA, H. AND BANNO, Y. 1992. *Genetical Stocks and Mutations of* Bombyx mori: *Important Genetic Resources.* Institute of Genetic Resources, Faculty of Agriculture, Kyusyu University.

FUJIWARA, H., OGURA, T., TAKADA, N., MIYAJIMA, N., ISHIKAWA, H. AND MAEKAWA, H. 1984. Introns and their flanking sequences of *Bombyx mori* rDNA. *Nucl. Acids Res.*, **12**, 6861–6869.

GANGULY, R., GANGULY, N. AND MANNING, J. E. 1985. Isolation and characterization of the glucose-6-phosphate dehydrogenase gene of *Drosophila melanogaster*. *Gene*, **35**, 91–101.

GOLDSCHMIDT, R. B. 1921. Erblichskeitstudien an Schmetterlingen III. Der Melanismus der Noune, *Lymandria monacha* L. *Z. Indukt. Abst. U. Vererb.*, **25**, 89–163.

HANCOCK, J. M. AND ARMSTRONG, J. S. 1994. SIMPLE34: an improved and enhanced implementation for VAX and Sun computers of the SIMPLE algorithm for analysis of clustered repetitive motifs in nucleotide sequences. *Comput. Appl. Biosci.*, **10**, 67–70.

HASHIMOTO, H. 1933. The role of the W chromosome for sex determination in the silkworm, *Bombyx mori*. *Jap. J. Genet.*, **8**, 245–258.

HSU, D. R. AND MEYER, B. J. 1993. X chromosome dosage compensation its relationship to sex determination in *C. elegans*. *Semin. Dev. Biol.*, **4**, 93–106.

ITOH, M., TAKEDA, S., YAMAMOTO, H., IZUMI, S., TOMINO, S. AND EGUCHI, M. 1991. Cloning and sequence analysis of membrane-bound alkaline phosphatase cDNA of the silkworm, *Bombyx mori*. *Biochim. biophys. Acta*, **1129**, 135–138.

JOHNSON, M. S. AND TURNER, J. R. G. 1979. Absence of dosage compensation for a sex-linked enzyme in butterflies (*Heliconius*). *Heredity*, **43**, 71–77.

KORGE, G. 1975. Chromosome puff activity and protein synthesis in larval salivary glands of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 4550–4554.

LUCCHESI, J. C. AND MANNING, J. E. 1987. Gene dosage compensation in *Drosophila melanogaster*. *Adv. Genet.*, **24**, 371–429.

MUCKHERJEE, A. S. AND BEERMANN, W. 1965. Synthesis of ribonucleic acid by the X-chromosomes of *Drosophila melanogaster* and the problem of dosage compensation. *Nature*, **207**, 785–786.

MULLER, H. J. 1950. Evidence of the precision of genetic adaptation. *Harvey Lect.*, **43**, 165–229.

OHNO, S. 1967. *Sex-Chromosomes and Sex Linked Genes*. Springer, Berlin.

PROMBOON, A., SHIMADA, T., FUJIWARA, H. AND KOBAYASHI, M. 1995. Linkage analysis of random amplified polymorphic DNAs (RAPDs) in the silkworm, *Bombyx mori*. *Genet. Res.*, **66**, 1–7.

REN, R., MAYER, B. J., CICCHETTI, P. AND BALTIMORE, D. 1993. Identification of a ten-amino acid proline-rich SH3 binding site. *Science*, **259**, 1157–1161.

SAKSELA, K., CHENG, G. AND BALTIMORE, D. 1995. Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef$^+$ viruses but not for downregulation of CD4. *EMBO J.*, **14**, 484–491.

SHI, J., HECKEL, D. G. AND GOLDSMITH, M. R. 1995. A genetic linkage map for the domesticated silkworm, *Bombyx mori*, based on restriction fragment length polymorphisms. *Genet. Res.*, **66**, 109–126.

SINGER-SAM, J. AND RIGGS, A. D. 1993. Quantitative analysis of messenger RNA levels: Reverse transcription-polymerase chain reaction single nucleotide primer extension assay. *Methods Enzymol.*, **225**, 344–351.

STEHR, G. 1959. Haemolymph polymorphism in a moth and the nature of sex-controlled inheritance. *Evolution*, **13**, 537–560.