Mammalian Genome 11, 220–224 (2000).

Mammalian Genome
Incorporating *Mouse Genome*

A revision of the human XIST gene organization and structural comparison with mouse *Xist*

Young-Kwon Hong, Sara D. Ontiveros, William M. Strauss

Harvard Institute of Human Genetics, 4 Blackfan Circle, Harvard Medical School and Beth Israel Deaconess Medical Center, Boston, Massachusetts 02115, USA

Received: 3 August 1999 / Accepted: 15 November 1999

Abstract. The XIST gene plays an essential role in X Chromosome (Chr) inactivation during the early development of female humans. It is believed that the XIST gene, not encoding a protein, functions as an RNA. The XIST cDNA is unusually long, as its full length is reported to be 16.5 kilobase pairs (kb). Here, comparison of sequences from the genomic interval downstream to the 3′ end of the human XIST gene against the human EST database brought to light a number of human EST sequences that are mapped to the region. Furthermore, PCR amplification of human cDNA libraries and RNA fluorescence in situ hybridization (RNA-FISH) demonstrate that the human XIST gene has additional 2.8 kb downstream sequences which have not been documented as a part of the gene. These data show that the full-length XIST cDNA is, in fact, 19.3 kb, not 16.5 kb as previously reported. The newly defined region contains an intron that may be alternatively spliced and seven polyadenylation signal sequences. Sequences in the newly defined region show overall sequence similarity with the 3′ terminal region of mouse *Xist*, and three subregions exhibit quite high sequence conservation. Interestingly, the new intron spans the first two subregions that are absent in one of the two isoforms of mouse *Xist.* Taken together, we revise the structure of human XIST cDNA and compare cDNA structures between human and mouse XIST/*Xist.*

Introduction

X chromosome inactivation is an early developmental process occurring in female mammals to compensate for differences between male and female mammals in dosage of genes residing on the X Chr (Brockdorff 1998; Brown et al. 1992; Lee and Jaenisch 1997; Lyon 1961). This mammalian dosage compensation is achieved by the transcriptional silencing of genes on one of the two X Chrs in females (Brockdorff 1998; Gartler et al. 1972; Lee and Jaenisch 1997). The inactivated X Chr can be microscopically observed during interphase as a condensed body at the nuclear periphery (Barr and Bertram 1949). Consistent with the chromosomal level of transcriptional silencing, the inactive X Chr is both hypermethylated on CpG islands and hypoacetylated on histone H4, compared with the active X Chr (Jeppesen and Turner 1993; Keohane et al. 1998; Miller et al. 1974).

From the study of chromosomal translocations, an interval called the XIC (X inactivation center) of the X Chr has been identified to control the process of inactivation. Translocations containing this segment to an autosome direct the inactivation of the autosome. A gene expressed exclusively from the inactivate X Chr has been cloned from human and mouse and mapped to the XIC region (Borsani et al. 1991; Brockdorff et al. 1992; Brown et al. 1992). This gene, called XIST/*Xist* (X inactive specific transcript), shows several interesting features. First, both human and mouse XIST/*Xist* cDNA are unusually long, reportedly 16.5 kb and 17.8 kb, respectively (Brown et al. 1992; Hong et al. 1999). Second, the transcript does not seem to encode a protein, on the basis of the lack of a significant open reading frame, absence of the *Xist* RNA from polysomes, and localization of the transcript in the nucleus (Brockdorff et al. 1992; Brown et al. 1992). Third, the XIST/*Xist* RNA physically associates with, or 'coats,' the inactive X Chr (Brown et al. 1992; Clemson et al. 1996). Fourth, XIST/*Xist* transcripts can be observed as early as the four-cell stage, and upon the initiation of X-inactivation, the steady-state level of the transcript rises dramatically, apparently by stabilization of the RNA (Panning et al. 1997; Sheardown et al. 1997). Although the function of XIST/*Xist* is not known, deletion of the gene leads to failure of X-inactivation, and knock-out mice die around the gastrulation stage (Marahrens et al.1997; Penny et al. 1996).

In this report, we revise the structure of the human XIST cDNA and discuss structural features of the newly defined region.

Materials and methods

Reagents. Genomic DNA was purified from a human female placenta, as described (Strauss 1999), and used as genomic DNA control for all the PCR reaction. Human cDNA libraries were obtained as follows: female pancreas (CLONTECH Catalog No. HL 1163b), male colon (CLONTECH Catalog No. HL 1034b), male/female bone marrow (CLONTECH Catalog No. HL 5005b), male liver (CLONTECH Catalog No. HL 3006b), male/female pituitary gland (CLONTECH Catalog No. HL 1139b), male/female fetal brain (CLONTECH Catalog No. HL 5015b), and fetal heart (sex unknown, CLONTECH Catalog No. HL 1114b). DNA sequences of oligonucleotides used for this study are listed in Table 1.

Table 1. Names, locations and sequences of primers used for this study

Names of primers	Locations (bp)[a] A	B	Sequences
WS 925	−160	16321	ACCTTGACCTGGCCTACAGA
WS 926	520	17001	TTGTTCCTGTGTTTCCACCA
WS 927	339	16820	TTGCTCATTGGTCTGGCTTA
WS 928	1075	17556	CCATGCCCCTAACAAGAAAA
WS 929	1049	17530	TGGCTTGTTTTCTTGTTAGGG
WS 930	1759	18240	CCCACCCTCTGTGAGTGATT
WS 931	1659	18140	TTGGCCAAAATTGAAAGGAA
WS 932	2351	18832	CAGCTGAAGAAAGGGGTGTT
WS 933	2052	18533	AAAGCTGAAGCCAAAATATGC
WS 934	2754	19235	CCAACTCCCCAGTTTGTTTC
WS 935	1341	17822	TGAGCCACAATTGGTTTTGA
WS 936	2559	19040	AAGGACAATGACGAAGCCAC
GAPDH-F			GAAGGTGAAGGTCGGAGTC
GAPDH-R			GAAGATGGTGATGGGATTTC

[a] Locations of primers are shown as the distances from the previously published end (**A**) and from the transcriptional initiation site (**B**) of XIST (Brown et al. 1992).

Correspondence to: W.M. Strauss; e-mail: wstrauss@hihg.med.harvard.edu

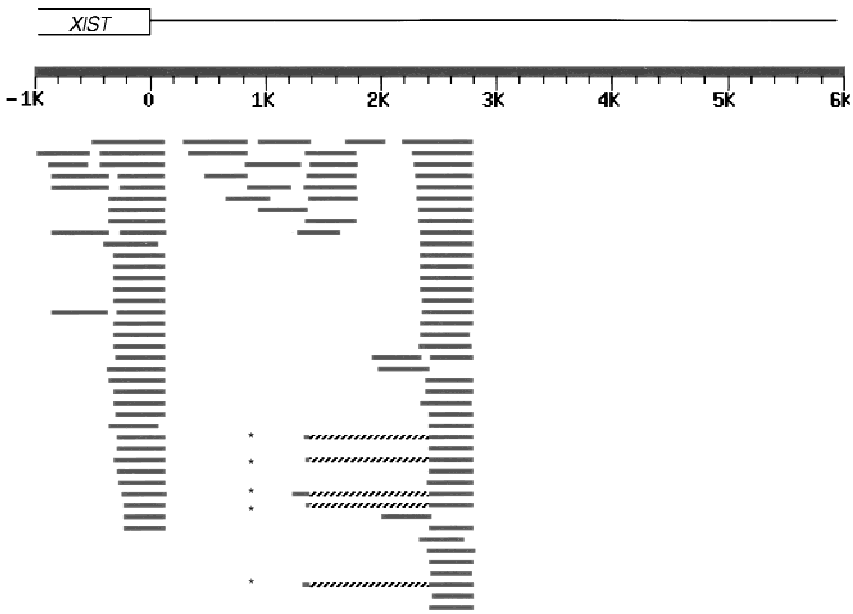


Fig. 1. Results of BLAST sequence similarity search of the 3′ end and downstream region of the human XIST gene with NCBI BLAST 2.0. Genomic sequences of 7000 bp from the human Xq13 region (GenBank Accession No. U80460) spanning the 3′ end of XIST and its downstream were searched against the human EST database with an Expect value, 0.0001, without filtering sequences of low compositional complexity, as of June 15, 1999 (www.ncbi.nlm.nih.gov). The previously reported end of the XIST gene is sketched on top, and relative locations from the end are shown in kilobases (K). Each bar represents an EST sequence registered in the human EST database. Gray bars represent segments in the XIST gene that have a similarity to the EST sequences above the threshold Expect value. A hatched region in the middle of a bar indicates a region of similarity below the Expect value. GenBank accession numbers of the EST sequences marked by asterisks are in order from the top N68599, AI268939, AI693632, N94964, and N93942.

PCR amplification and Southern blotting. Each PCR reaction was performed at 94°C 30 s, 58°C 30 s, and 72°C 30 s for 40 cycles (except 35 cycle for GAPDH). All the PCR products were analyzed in 2% agarose gels with the 100-bp size marker (Promega). The PCR products derived from the female pancreas cDNA library were cloned in pBSIISK+ (Stratagene) and sequenced with both T3 and T7 primers on an Applied Biosystems 377 sequencer in the Beth Israel Deaconess Medical Center sequencing facility. These clones were subsequently radiolabeled by the random hexamer method (Feinberg and Vogelstein 1983) in the presence of [^{32}P] dCTP and used as probes for Southern blot analysis as follows. The PCR products derived from either genomic DNA or cDNA libraries were electrophoresed on an 2% agarose gel and transferred to nylon membrane and hybridized with each probe for 12–24 h at 65°C in a buffer containing 7% SDS, 2 mM EDTA (pH 7.6), and 0.5 M sodium phosphate (pH 7.5).

RNA fluorescence in situ hybridization (FISH). Male and female human fibroblast cell lines were obtained from the NIGMS Repositories (Catalog numbers GM04033 and GM04281, respectively), grown on chamber slides, and fixed as described (Trask 1991). Two sets of DNA probes were used: the previously published XIST G1A (Clemson et al. 1998), ~10 kb genomic plasmid spanning from the fourth intron to the 3′ end of the human XIST, was labeled with biotin, and the plasmids containing the PCR fragments, WS927/928, WS929/930, WS931/932, and WS933/934 (Fig. 2), were pooled and labeled with digoxygenin. Probes were hybridized either separately or simultaneously to the interphase spreads of human fibroblasts. After washing, the slides were labeled with antidigoxygenin rhodamine or avidin-fluorescein. Images were collected with a NIKON E-800 microscope equipped with a Sensys (Photometrics, AZ) digital CCD camera. Grayscale images for either FITC, rhodamine, or DAPI filter sets were pseudocolored, and images were merged in the 12-bit format. The 12-bit data were compressed at eight-bit data during export to Photoshop 5.0 (Adobe) for final figure preparation.

Results

We reviewed the human XIST gene structure by comparing the sequences encompassing the 3′ end and downstream of the gene against the human Expressed Sequence Tag (EST) databases. Surprisingly, this examination brought to light a number of ESTs (~100 sequences) mapped to, not only the 3′ end of the gene, but also further downstream of the gene (Fig. 1; Brown et al. 1992). This implied that the human XIST gene may encode additional information at the 3′ end. In order to test this, we designed DNA oligonucleotide primers spanning both the 3′ end and downstream

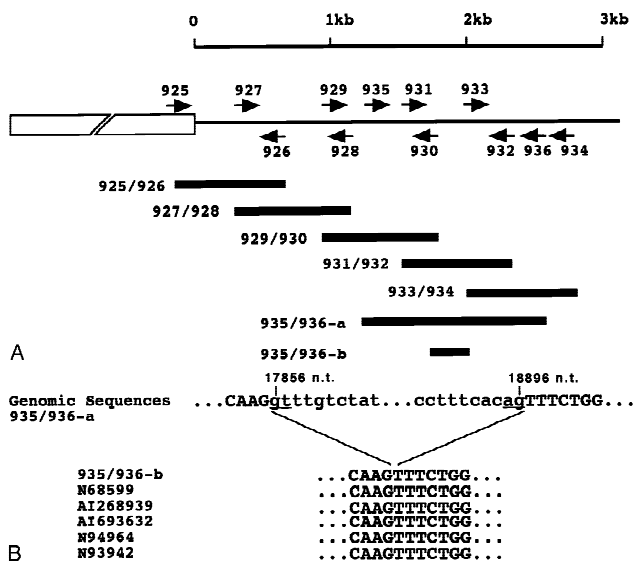


Fig. 2. (A) Relative locations of PCR primers (arrows) against the previously reported end of XIST and the overlapping PCR products (black bars) resulting from PCR reactions with each primer set. Names of primers used for the PCR reactions are shown next to the black bars. PCR amplification reaction with WS935 and WS936 yields products of two sizes, 935/936-a and 935/936-b. DNA sequences of the primers are listed in Table 1. **(B)** Sequence alignment of genomic DNA and 935/936-a to the PCR product 935/936-b, as well as the five human ESTs shown in Fig. 1 (GenBank Accession No. N68599, AI268939, AI693632, N94964, and N93942). Intron sequences are lower-cased, and sequences of splice donor and receptor are underlined. Nucleotide sequence positions of the 5′ and the 3′ ends of the intron are shown based on the numbering by Brown et al. (1992).

of the gene (Table 1, Fig. 2A). Using these primers, we PCR-amplified the corresponding regions from several human cDNA libraries as well as human genomic DNA. The human cDNA libraries used for this purpose were derived either from male, female, or male/female pooled RNAs. We cloned all the PCR fragments amplified from the human female pancreas cDNA library into the pBSIISK+ vector. Sequencing analyses of the cloned fragments confirmed that the PCR-amplified products were indeed

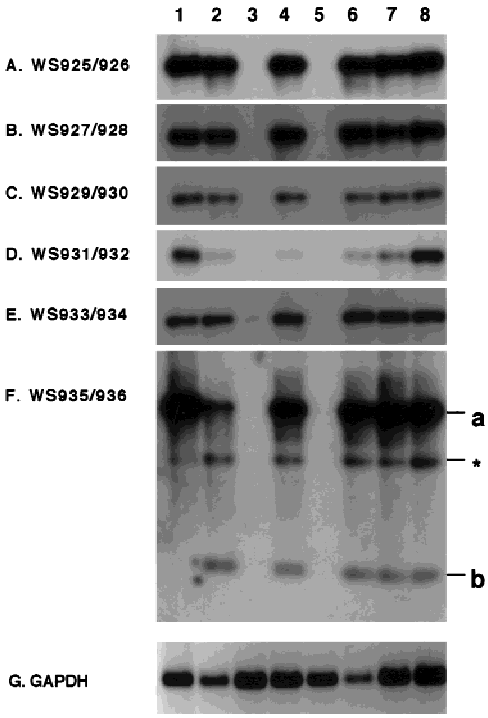


Fig. 3. Southern blot analyses of the PCR products. DNA fragments were PCR-amplified with each primer shown in Fig. 2A from either human genomic DNA or cDNA libraries (lane 1, female genomic DNA from placenta; lane 2, human female pancreas cDNA library; lane 3, human male colon cDNA library; lane 4, human male/female bone marrow cDNA library; lane 5, human male liver cDNA library; lane 6, human male/female pituitary gland cDNA library; lane 7, human male/female fetal brain cDNA library; lane 8, sex-unknown human fetal heart cDNA library). The PCR primer sets used for the reaction are shown next to each panel (A≈F). The PCR products were electrophoresed, transferred to a nylon membrane, and hybridized with probes. The probes were prepared from the cloned and sequenced PCR products, which were amplified with the corresponding primer set from the human female pancreas cDNA library (panels A through E) or from genomic DNA (panel F). In panel F, a and b indicate the PCR products, 1219 bp and 178 bp, amplified from unspliced and spliced cDNA templates, respectively, and *, nonspecific PCR product. The PCR amplification reactions are controlled by amplifying the human GAPDH gene (panel G).

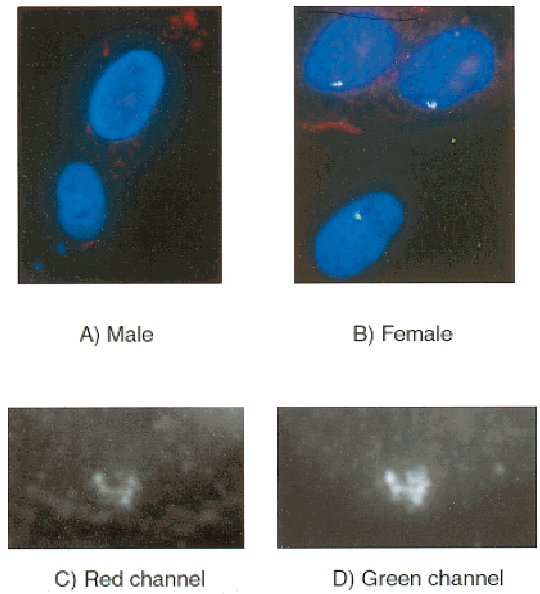


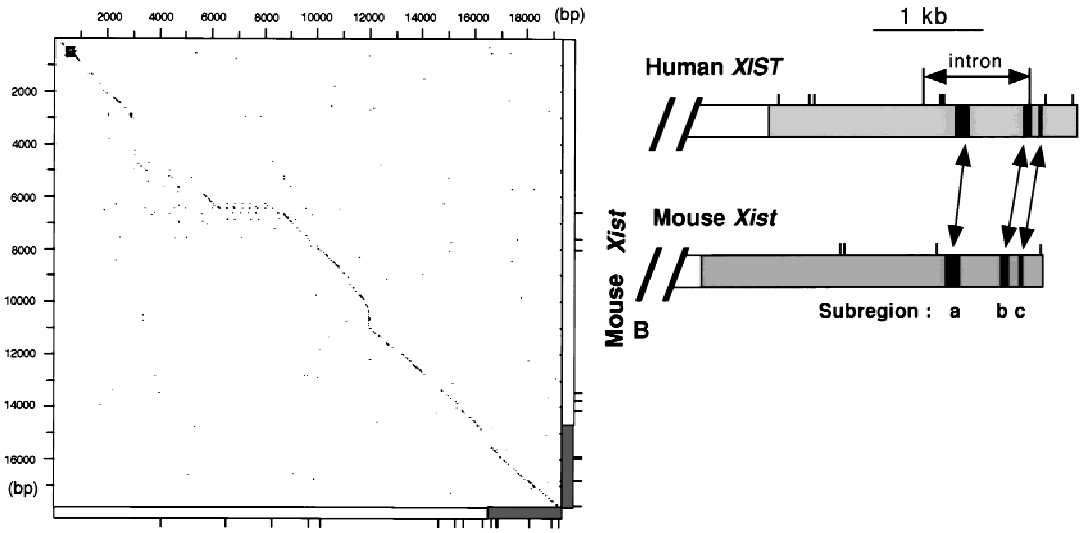
Fig. 4. Two-color RNA-FISH photography of human male **(A)** and female **(B)** interphase nuclei. XIST G1A (green), which spans the genomic sequence from the fourth intron to the previously documented end of the gene, and the newly defined region (red) are colocalized on the inactive X Chr in female (yellow). Enlarged gray-scale images of rhodamine (C) and fluorescein (D) signals of the upper right cell in panel (B) are shown.

mapped to the 3′ end of XIST and its downstream region (data not shown). Subsequently, the PCR products amplified from both human genomic DNA and human cDNA libraries were subjected to Southern blot analysis by using the cloned and sequenced PCR fragments as probes (Fig. 3, A–E). This Southern analysis demonstrates that only female-specific cDNA libraries yield the PCR products. Furthermore, the PCR products amplified from the female-specific cDNA libraries and human genomic DNA are of the same size and hybridize equally to the sequenced probes (Fig. 3A–E). These data imply that the corresponding sequences are expressed only in females, consistent with the female-specific expression of XIST. More importantly, the sequences downstream from the documented human XIST end, which have not been considered as a part of the gene, are transcribed and contiguous with the XIST RNA (Brown et al. 1992). Therefore, these data lead us to conclude that the full length of the human XIST RNA is at least 19.3 kb, not 16.5 kb as previously reported (Brown et al. 1992).

Sequence comparison of the 3′ downstream region of XIST to the EST databases also suggested the presence of an intron in the newly defined 2.8-kb region (Fig. 1). In order to prove the presence of an intron, we used a primer set (WS935 and WS936) that spans the potential intron and then PCR-amplified the region from both human genomic DNA and the human cDNA libraries. PCR products from human genomic DNA or cDNA library from female pancreas RNA were cloned and sequenced. Southern blot analysis shows two PCR fragments amplified from the female RNA-derived cDNA libraries, whereas only one fragment was amplified from the human genomic DNA (Fig. 3F). Sequence alignment between the XIST genomic DNA and the PCR product from the pancreas RNA-derived cDNA library, as well as the five ESTs marked in Fig. 1, demonstrate that the sequence (1041 bp) between nucleotide positions 17856 and 18896 is processed (Fig. 2B). The splice junctions match the consensus splice donor and acceptor sequences for vertebrates (Shapiro and Senapathy 1987). Therefore, we conclude that the newly identified region of the human XIST gene contains an intron.

RNA-fluorescence in situ hybridization (FISH) was employed to determine whether the newly identified region colocalized with the established sequences of XIST on the inactive X Chr in human female fibroblast cells. We used the previously characterized XIST genomic DNA probe, XIST G1A, which spans from the fourth intron to the formerly documented end of the gene and has been shown to paint the inactive X Chr in human females (Clemson et al. 1998). In Fig. 4, the DNA probes derived from the newly defined region colocalize with the XIST G1A probe on the inactive X Chr in a female-specific manner. This evidence clearly demonstrates that the newly defined region is a part of the human XIST transcript and correctly localizes on the inactive X Chr in human female cells.

Recently, we showed that the murine *Xist* gene has an additional 3.1 kb in sequence than was previously documented (Hong et al. 1999). In an attempt to explore the possible structure/functional conservation between the two homologs, we compared sequences of the human and murine XIST/*Xist* cDNA (Fig. 5A). The dot matrix sequence comparison of the two genes with a threshold of 80% sequence identity in a 21-base pair window sliding along the genes shows a substantial conservation between the two species. Importantly, the newly defined 3′ regions are



A

Human *XIST*

Subregion a

```
XIST:  18199 ctgggctttgacctttccgtttttaaatcactcacagagggtgggacaggaggaagagt 18258
             ||| |||||||||| |||||  |||||   ||||||||||||||||||||||||||  ||||
Xist:  16957 ctgtgctttgacccttccatttttctttcactcacagagggtgggacaggaggccgagt 17016

XIST:  18259 gaaggaaaaggtcaaacctgt 18279
             |||||||| |||| | |||||
Xist:  17017 gaaggaaagggtccagcctgt 17037
```

Subregion b

```
XIST:  18856 tggaaaatggaaagaagaaagtaactaaaagccttcctttcacagtttctggcatca 18912
             |||||  || |||| | ||| ||||| ||||||||||||||||||| ||| |||||
Xist:  17494 tggaatttgcaaaggaaaaaataactgaaagccttcctttcacagttactgccatca 17550
```

Subregion c

```
XIST:  18953 catcatctgctttattcacataaatgaa 18980
             |||||||||||||  |||||||||||||
Xist:  17579 catcatctgctttgatcacataaatgaa 17606
```

C

Fig. 5. (A) Dot matrix sequences comparison between the human XIST and mouse *Xist* cDNA sequences. With the DNA STRIDER 1.2 program, human and mouse XIST/*Xist* cDNA were compared by scoring sequence identity of at least 17 bp out of a 21-bp window moving along the genes. Therefore, a dot represents location for each gene that shares the similarity. Horizontal and vertical line diagrams at the bottom and the right side of the matrix depict the human and mouse XIST/*Xist* cDNA sequences, respectively. Shaded region of each line diagram shows the newly defined sequences of human XIST and mouse *Xist*, respectively (Hong et al. 1999). Short line ticks in the each diagram indicate locations of the consensus polyadenylation signal sequence. **(B)** Schematic representation of homologous regions between the 3′ ends of human and mouse XIST/*Xist.* Three subregions, a, b and c, are depicted that show a high sequence identity (a, 85% over 81 bp; b, 84% over 57 bp; c, 93% over 28 bp). Locations of polyadenylation signal sequences (ticks) and of the intron are shown. **(C)** Alignments of sequences from the three subregions. Nucleotide numbers are shown for both human XIST and mouse *Xist* cDNA sequences.

among regions that show the most significant similarity between the two genes. The newly defined region of human XIST also contains seven polyadenylation consensus signal sequences and three subregions with relatively high sequence similarity to mouse *Xist* (Fig. 5B, C). The newly identified intron in human XIST spans the first two subregions. A further examination of regions downstream of the new 3′ sequences did not show any significant sequence similarity.

Discussion

In this report, we demonstrate that the human XIST transcript is, in fact, 2.8 kb longer than the 16.5 kb previously documented (Brown et al. 1992). This finding is based on three independent approaches; the EST database analysis with the BLAST search algorithms, female specific-PCR amplification of XIST cDNA, and finally, co-localization of the newly defined region with XIST sequences previously shown to be associated with the inactive X Chr in female cells. Sequence comparisons of the 3′ end and the region downstream of XIST against the human EST databases reveal that a number of EST sequences are reported to map not only to the established 3′ end, but also further downstream of the gene (Fig. 1). Female-specific PCR-amplification of the 3′ end as well as the newly defined region of XIST from human cDNA library demonstrates that the additional 2.8-kb sequence is co-linear to the XIST transcript (Figs. 2, 3). Finally, the DNA probes containing the downstream region co-localize with the established XIST sequence on the inactive X Chr in human female cells (Fig. 4). On the basis of these data, we conclude the human XIST full-length cDNA is at least 19.3 kb long.

It is an interesting coincidence that the cDNA structures of both human and mouse XIST/*Xist* have been incorrectly characterized and reported to be shorter (Brockdorff et al. 1992; Brown et al. 1992). One possible explanation is the presence of adenosine-rich sequences in the genome sequence. We found a stretch of sequence with a high adenosine content downstream of the previously reported ends of both human XIST (25 A residues out of 27 nucleotides, GenBank Accession No. U80460) and mouse *Xist* (15

out of 17, GenBank Accession No. X99946). The adenosine-rich stretch is located 118 bp downstream from the previously reported ends of human XIST and immediately follows the previously defined end of the mouse *Xist* gene. This may have induced misprimings of the oligo dT primers for the first-strand synthesis during the construction of the cDNA libraries, from which the authors collected the XIST/*Xist* cDNA clones (Brockdorff et al. 1992; Brown et al. 1992). Consistent with this idea, the ends of a number of human EST sequences are mapped at the adenosine-rich stretch (Fig. 1). Nevertheless, we do not rule out the possible presence of shorter isoforms of XIST that may be formed by alternative usage of polyadenylation signals at the 3′ end. In fact, we have observed two major species of the *Xist* RNA that differ at their 3′ ends in the mouse (Hong et al. 1999).

It is our belief that the human XIST RNA does not extend further downstream than 2.8 kb, based on the following three lines of observation. First, even after we reduced the stringency of the BLAST search (Expected value = 0.01), we did not find any human EST sequences with a significant similarity to the region up to 10 kb downstream from the newly defined end (Fig. 1, data not shown, apart from a number of EST sequences with partial sequence similarity to the Alu repeats). Second, all the EST sequences that mapped close to the new 3′ end unanimously terminate at one location, 2.8 kb downstream from the previously documented end, which is not followed by any adenosine-rich sequence string (Fig. 1). Third, there is no significant sequence similarity between human and mouse sequences further downstream of the newly defined regions of the two genes (Fig. 5).

Two lines of data suggest that the intron in the newly defined region is likely to be alternatively spliced: PCR reactions using primers (WS935 and WS936) flanking the intron yielded DNA products amplified from unprocessed cDNA template as well as from spliced template, and PCR reactions using primers priming at the intron (WS929/930, WS931/932, WS933/934) also generated products, implying the presence of unprocessed cDNA template (Fig. 2). Interestingly, while the region shares a considerable sequence similarity between human and mouse, no intron has been found in the corresponding region of mouse *Xist*, to date. One implication may be that mouse *Xist* may depend only on differential polyadenylation to generate isoforms. Human XIST may utilize alternative splicing in addition to differential polyadenylation. Within the same context, it is noteworthy that the new intron in XIST spans the first two out of three subregions of high sequence similarity between mouse and human (Fig. 5B, C). Thus, it is formally possible that murine splice variants may be produced, but as of yet are undetected. We discovered that one of two isoforms of *Xist* ends before the first similarity subregion (Hong et al. 1999, data unpublished). Taken together, it is tempting to think that the two similarity subregions may provide important functional distinctions to the XIST/*Xist* isoforms. In order to validate this idea, additional study will be required.

Acknowledgments. The authors acknowledge the extraordinary support of Dr. Fred Rosen and the Center for Blood Research, Harvard Medical School. We also thank Dr. Woo-Joo Song for his generous assistance and Dr. J.B. Lawrence for the XIST G1A probe. This work was supported by a US Army Prostate Cancer Research Award (PC970479), the Harvard Nathan Shock Center for Biology of Aging Core C (5 P30-AG13314), the Harvard Milton Fund, and a Beth Israel Deaconess New Investigator Award, all awarded to W.M. Strauss.

References

Barr ML, Bertram EG (1949) A morphological distinction between neurons of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. Nature 163, 676–677

Borsani G, Tonlorenzi R, Simmler MC, Dandolo L, Arnaud D et al. (1991) Characterization of a murine gene expressed from the inactive X chromosome. Nature 351, 325–329

Brockdorff N (1998) The role of Xist in X-inactivation. Curr Opin Genet Dev 8, 328–333

Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP et al. (1992) The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell 71, 515–526

Brown CJ, Hendrich BD, Rupert JL, Lafreneire RG, Xing Y et al. (1992) The human *XIST* gene: analysis of a 17kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 71, 527–542

Clemson CM, McNeil JA, Willard HF, Lawrence JB (1996) *XIST* RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J Cell Biol 132, 259–275

Clemson CM, Chow JC, Brown CJ, Lawrence JB (1998) Stabilization and localization of xist RNA are controlled by separate mechanisms and are not sufficient for X inactivation [In Process Citation]. J Cell Biol 142, 13–23

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132, 6–13

Gartler SM, Chen SH, Fialkow PJ, Giblett ER, Singh S (1972) X chromosome inactivation in cells from an individual heterozygous for two X-linked genes. Nat New Biol 236, 149–150

Hong YK, Ontiveros SD, Chen C, Strauss WM (1999) A new structure for the murine xist gene and its relationship to chromosome choice/counting during X-chromosome inactivation [in Process Citation]. Proc Natl Acad Sci USA 96, 6829–6834

Jeppesen P, Turner BM (1993) The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. Cell 74, 281–289

Keohane AM, Lavender JS, O'Neill LP, Turner BM (1998) Histone acetylation and X inactivation. Dev Genet 22, 65–73

Lee JT, Jaenisch R (1997) The (epi)genetic control of mammalian X-chromosome inactivation. Curr Opin Genet Dev 7, 274–280

Lyon MF (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). Nature 190, 372–373

Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R (1997) Xist-deficient mice are defective in dosage compensation but not spermatogenesis [see comments]. Genes Dev 11, 156–166

Miller OJ, Schnedl W, Allen J, Erlanger BF (1974) 5-Methylcytosine localised in mammalian constitutive heterochromatin. Nature 251, 636–637

Panning B, Dausman J, Jaenisch R (1997) X chromosome inactivation is mediated by Xist RNA stabilization. Cell 90, 907–916

Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N (1996) Requirement for Xist in X chromosome inactivation. Nature 379, 131–137

Shapiro MB, Senapathy P (1987) RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res 15, 7155–7174

Sheardown SA, Duthie SM, Johnston CM, Newall AE, Formstone EJ et al. (1997) Stabilization of Xist RNA mediates initiation of X chromosome inactivation [see comments]. Cell 91, 99–107

Strauss WM (1999) Preparation of genomic DNA from mammalian tissue. In *Current Protocols in Neuroscience*, vol Appendix 1H. (New York: John Wiley & Sons, Inc.) pp A.1H.1

Trask B (1991) Fluorescence in situ hybridization: applications in cytogenetics and gene mapping. Trends Genet 7, 149–154