Available online at www.sciencedirect.com
ScienceDirect
Genomics 90 (2007) 453–463
GENOMICS
www.elsevier.com/locate/ygeno

A cross-species comparison of X-chromosome inactivation in Eutheria☆

Ziny C. Yen [a,1], Irmtraud M. Meyer [a,b], Sanja Karalic [a,2], Carolyn J. Brown [a,*]

[a] Department of Medical Genetics, University of British Columbia, Vancouver, Canada BC V6T 1Z3
[b] Bioinformatics Centre and Department of Computer Science, University of British Columbia, Vancouver, Canada BC V6T 1Z3

Received 8 November 2006; accepted 4 July 2007

Abstract

Mammalian X-chromosome inactivation achieves dosage compensation between the sexes by the silencing of one X chromosome in females. In Eutheria, X inactivation is initiated by the large noncoding RNA Xist; however, it is unknown how this RNA results in silencing of the chromosome or why, at least in humans, many genes escape silencing in somatic cells. We have sequenced the coast mole *Xist* gene and compared the Xist RNA sequence among seven eutherians to provide insight into the structure of the RNA and origins of the gene. Using DNA methylation of promoter sequences to assess whether genes are silenced in females we report the inactivation status of seven X-linked genes in humans and mice as well as two additional eutherians, the mole and the cow, providing evidence that escape from inactivation is common among Eutheria.

Keywords: X (inactive)-specific transcript; X-chromosome inactivation; Dosage compensation (genetics); DNA methylation; Gene silencing; RNA; Conserved secondary structure; Functional secondary structure

Most organisms with chromosomally determined sex have evolved a mechanism of dosage compensation to equalize expression of sex-linked genes between males and females. In mammals this involves inactivation of one X chromosome early in female development [1]. Inactivation of the same X is maintained through somatic cell divisions, resulting in mosaic females with cell populations containing one or the other X active. X-chromosome inactivation (X inactivation) is an excellent model for understanding epigenetic silencing and the emerging role of functional RNAs in the process of heterochromatinization. The inactive X acquires most of the features of heterochromatin, including late replication, peripheral nuclear localization, DNA methylation at the 5′ end of genes, modification of histones, and incorporation of a histone variant (reviewed in [2]). The initiating factor in the silencing of the X chromosome is the expression of *Xist* (Xi-specific transcript), which encodes a large (~17 kb) alternatively spliced and polyadenylated transcript that lacks any conserved open reading frame and is expressed solely from the inactive X in somatic cells. The spliced transcripts are seen to associate with the inactive X chromosome from which they are transcribed and the *Xist* gene itself is located within the X-inactivation center (*Xic*), the only region of the X that is required *in cis* for inactivation of an X. Knockout and transgene studies in mice have shown that *Xist* is necessary and sufficient for the initiation of inactivation. While the gene continues to be expressed, and contributes to the stable silencing of the X in somatic cells, it is not required for the maintenance of silencing (reviewed in [2]).

The *Xist* gene has been identified in all eutherian mammals analyzed [3,4]. Analysis of cow, human, and rodent *Xist* sequences has identified characteristic repetitive regions present at different copy numbers in each species [5–7]. We now extend such sequence comparisons by adding the sequence of an insectivore, the coast mole. We have also analyzed the dog (a carnivore) and the rat (an additional rodent) genomic sequences to generate theoretical *Xist* cDNAs for these species. In addition to comparing exon structure and repeats within the primary sequence, we also utilized two comparative RNA structure prediction methods to search for evolutionarily conserved

☆ Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. DQ845733 (coast mole *Xist* sequence).

* Corresponding author. Fax: +1 604 822 5348.
E-mail address: cbrown@interchange.ubc.ca (C.J. Brown).
[1] Current address: Faculty of Medicine, Queen's University, Kingston, ON, Canada.
[2] Current address: Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada.

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2007.07.002

putative Xist RNA structures within the seven eutherians examined—human, mouse, rat, vole, cow, dog, and mole. By employing a comparative rather than a noncomparative RNA secondary structure prediction approach, we are capable of finding evolutionarily conserved RNA structure elements that are likely to play a functional role.

Surprisingly, approximately 15% of human X-linked genes escape inactivation and another 10% show variable inactivation, whereby they are expressed from the inactive X in only a subset of females [8]. Far fewer genes have been reported to escape inactivation in mice, consistent with the lack of abnormal phenotype associated with 39,X mice that are missing the second sex chromosome. This contrasts with 45,X human females who rarely survive to term and result in Turner syndrome when viable (reviewed in [9]). By examining species differences in X inactivation clues to the regulation and maintenance of silencing may be obtained. Understanding why particular genes escape silencing in humans is of clinical importance for the potential differences in expression between males and females, as well as individuals with X-chromosome aneuploidies. It is also of considerable biological interest to understand how the *cis*-limited inactivation of the X chromosome skips over, or fails to be maintained upon, particular genes.

Several different hypotheses have been proposed to explain the differences in degree of silencing between the human and the mouse chromosomes. As the genes escaping silencing in humans are enriched in the X added region on the short arm of the X, their presence has been proposed to reflect the evolutionarily more recent addition of this region to the eutherian X [8], while mice may show more complete silencing due to greater decay of the Y [9]. A second hypothesis suggests that the spread of silencing is impeded by the presence of centric heterochromatin, such that the acrocentric X chromosome of mouse shows more complete inactivation [10]. Supporting this argument, it has been shown that in rodents in which an X/autosome translocation has occurred during sex-chromosome evolution, a block of heterochromatin is present, which is proposed to prevent the spread of X-chromosomal silencing into the autosomal material [11]. A third hypothesis is that the inactivation signal may diminish with longer distance from the *Xic,* as observed during early development in mouse embryos before X inactivation becomes stable [12] and in unbalanced X/autosome translocations that show a gradient effect to autosomal silencing [13]. To address these alternatives we have examined the inactivation status of seven different genes in human, cow, coast mole, and mouse, using DNA methylation as an assay for silencing, as was done previously by Jegalian and Page in their study of three X-linked genes in multiple eutherians [14].

Results

Sequence of the coast mole (Scapanus orarius) Xist gene

As diagrammed in Fig. 1A, the sequence of the mole *Xist* was derived from a combination of genomic DNA and cDNA by PCR amplification using degenerate primers designed from existing *Xist* sequences and subsequently using primers from mole sequence for inverse or traditional PCR (primers in Supplementary Table 1). The acquired mole *Xist* sequences aligned best to other *Xist* sequences using Blastn [15,16], confirming their likely origin from the mole *Xist* gene. Furthermore, amplification in cDNA was observed only in cDNA from female, not male, moles, as anticipated for an ortholog of *Xist* (Fig. 1B). Although the sequence including and

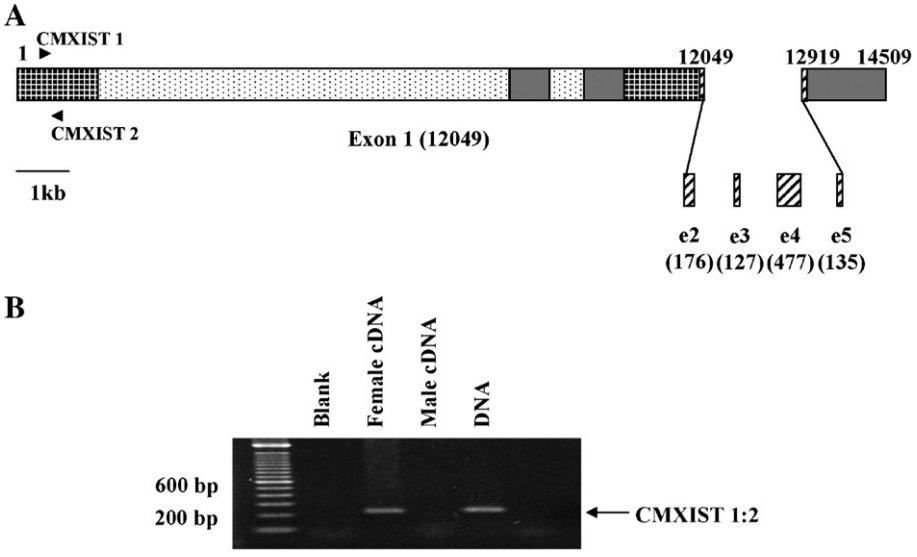

Fig. 1. Coast mole *Xist* cDNA. (A) Sequencing of the coast mole *Xist* was based on direct sequencing of PCR products (solid boxes) or sequencing of cloned PCR products (spotted regions). Amplification of coast mole regions was by degenerate or coast-mole-specific PCR of genomic DNA or cDNA (diagonals) or inverse PCR with mole-derived primers (cross-hatch). Numbers along the top indicate the position in the sequence (Accession No. DQ845733), while the numbers in parentheses under the exons signify the sizes of exons in base pairs. (B) RT-PCR shows female-specific expression. Mole-specific primers (CMXIST1 and CMXIST2) at the locations marked by the arrows in (A) were used to amplify cDNA derived from male and female mole RNA, with amplification detected only in female cDNA.