more hypoxic environments. Two recent studies have implicated p53 in the control of glycolysis by its negative regulation of phosphoglycerate mutase (PGM) and Akt (*6*, *7*, *21*). Interestingly, despite the marked increase in glycolysis stimulated by Akt expression, cellular oxygen consumption remained stable, reflecting the tight homeostatic controls governing mitochondrial respiration (*22*). Additional functions of the *TP53* gene continue to be uncovered (*23*), and the effect of p53 status on exercise tolerance suggests that it has global functions that extend beyond processes related to cell birth and death (*24*). The COX complex is critical for aerobic eukaryotes, and *SCO2* and *MTCO2* (COX II) are ancient genes conserved in organisms as diverse as yeast and humans. The regulation of mitochondrial genome-encoded COX II by p53 suggests that our observations pertain to a fundamental control point in metabolism. Recent studies have implicated p53 as a mediator of senescence and aging (*25*–*27*), although specific aspects of this model remain controversial (*28*, *29*). Given the mounting evidence supporting a role for metabolism and oxidative stress in aging (*30*, *31*), the functional relationship between p53 and the COX complex assembly may underlie some aspects of organismal aging. In filamentous fungus, the genetic disruption of COX II markedly increases life-span (*32*). Future studies examining the role of p53 in mitochondrial regulation may clarify how a tumor suppressor gene can have such diverse and global effects on cellular and organismal functions.

References and Notes

1. B. Vogelstein, K. W. Kinzler, *Nat. Med.* **10**, 789 (2004).
2. O. Warburg, *Science* **124**, 269 (1956).
3. G. L. Semenza *et al.*, *Novartis Found. Symp.* **240**, 251 (2001).
4. N. C. Denko *et al.*, *Oncogene* **22**, 5907 (2003).
5. J. W. Kim *et al.*, *Mol. Cell. Biol.* **24**, 5923 (2004).
6. H. Kondoh *et al.*, *Cancer Res.* **65**, 177 (2005).
7. D. R. Plas, C. B. Thompson, *Oncogene* **24**, 7435 (2005).
8. P. C. Herrmann *et al.*, *Proteomics* **3**, 1801 (2003).
9. S. Zhou, S. Kachhap, K. K. Singh, *Mutagenesis* **18**, 287 (2003).
10. F. Bunz *et al.*, *Science* **282**, 1497 (1998).
11. S. Sariban-Sohraby, I. T. Magrath, R. S. Balaban, *Cancer Res.* **43**, 4662 (1983).
12. L. D. Attardi, L. A. Donehower, *Mutat. Res.* **576**, 4 (2005).
13. K. Matsumoto, K. Ishihara, K. Tanaka, K. Inoue, T. Fushiki, *J. Appl. Physiol.* **81**, 1843 (1996).
14. P. M. Hwang *et al.*, *Nat. Med.* **7**, 1111 (2001).
15. E. A. Shoubridge, *Am. J. Med. Genet.* **106**, 46 (2001).
16. P. Pecina, H. Houstkova, H. Hansikova, J. Zeman, J. Houstek, *Physiol. Res.* **53** (Suppl. 1), S213 (2004).
17. L. C. Papadopoulou *et al.*, *Nat. Genet.* **23**, 333 (1999).
18. J. Yu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14517 (1999).
19. B. Vogelstein, D. Lane, A. J. Levine, *Nature* **408**, 307 (2000).
20. M. Kohli, C. Rago, C. Lengauer, K. W. Kinzler, B. Vogelstein, *Nucleic Acids Res.* **32**, e3 (2004).
21. A. J. Levine, Z. Feng, T. W. Mak, H. You, S. Jin, *Genes Dev.* **20**, 267 (2006).
22. R. L. Elstrom *et al.*, *Cancer Res.* **64**, 3892 (2004).
23. L. J. Hofseth, S. P. Hussain, C. C. Harris, *Trends Pharmacol. Sci.* **25**, 177 (2004).
24. U. Wisloff *et al.*, *Science* **307**, 418 (2005).
25. S. D. Tyner *et al.*, *Nature* **415**, 45 (2002).
26. B. Maier *et al.*, *Genes Dev.* **18**, 306 (2004).
27. J. Campisi, *Cell* **120**, 513 (2005).
28. I. Garcia-Cao *et al.*, *EMBO J.* **21**, 6225 (2002).
29. M. V. Poyurovsky, C. Prives, *Genes Dev.* **20**, 125 (2006).
30. L. Bordone, L. Guarente, *Nat. Rev. Mol. Cell Biol.* **6**, 298 (2005).
31. R. S. Balaban, S. Nemoto, T. Finkel, *Cell* **120**, 483 (2005).
32. S. W. Stumpferl, O. Stephan, H. D. Osiewacz, *Eukaryot. Cell* **3**, 200 (2004).

33. We thank S. Nemoto, C. Mcleod, I. Pagel, C. Rago, and O. Mian for technical advice and assistance, and T. Finkel, M. N. Sack, N. Epstein, R. S. Balaban, and B. Vogelstein for helpful suggestions and critical reading of this manuscript. The p53-disrupted HCT116 cell lines, tetracycline-inducible p53 DLD1 cell lines, and plasmid constructs were gifts from B. Vogelstein and B. H. Park. This research was supported by the Division of Intramural Research, National Heart, Lung, and Blood Institute, NIH.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1126863/DC1
Materials and Methods
Figs. S1 and S2
References and Notes

1 March 2006; accepted 17 May 2006
Published online 25 May 2006;
10.1126/science.1126863
Include this information when citing this paper.

The *Xist* RNA Gene Evolved in Eutherians by Pseudogenization of a Protein-Coding Gene

Laurent Duret,[1]* Corinne Chureau,[2] Sylvie Samain,[3] Jean Weissenbach,[3] Philip Avner[2]

The *Xist* noncoding RNA is the key initiator of the process of X chromosome inactivation in eutherian mammals, but its precise function and origin remain unknown. Although *Xist* is well conserved among eutherians, until now, no homolog has been identified in other mammals. We show here that *Xist* evolved, at least partly, from a protein-coding gene and that the loss of protein-coding function of the proto-*Xist* coincides with the four flanking protein genes becoming pseudogenes. This event occurred after the divergence between eutherians and marsupials, which suggests that mechanisms of dosage compensation have evolved independently in both lineages.

Mammalian X and Y chromosomes evolved from a pair of autosomes shortly after the divergence of mammals from other amniotes (*1*). In eutherians and in marsupials, the desiquilibrium in gene dosage between XY males and XX females is compensated for by silencing one of the X chromosomes in females (*2*–*4*). In eutherians, this silencing involves the *Xist* gene, which is located in the X inactivation center (*Xic*) and encodes a long untranslated RNA (*5*). *Xic* is located on the long arm of the human X chromosome, which corresponded to the proto–X chromosome in the mammalian cenancestor (last common ancestor) (*6*, *7*). This observation is consistent with the hypothesis that X-chromosome inactivation might have emerged contemporaneously with the chromosomal sex-determining mechanism, early in mammalian evolution (*8*).

To study the evolution of X inactivation, we searched for homologs of *Xist* in 14 vertebrate genomes (*9*). We found *Xist* in all eutherians (Fig. 1), which demonstrates that *Xist* was already present in the eutherian cenancestor. With BLAST, we failed to detect significant sequence similarity to *Xist* in noneutherian vertebrates.

In humans, the genomic region surrounding the *Xist* gene contains three protein-coding genes (*Cdx4*, *Chic1*, and *Xpct*) that have orthologs in all vertebrate classes (table S1). The linkage between these genes is conserved in chicken and in *Xenopus* (Fig. 2A). We will hereater refer to the genomic interval between *Chic1* and *Xpct* in noneutherian species as the *XicHR* (*Xic* homologous region). In eutherians, besides *Xist*, the *Xic* region contains two RNA genes (*Jpx* and *Ftx*) and two protein-coding genes (*Tsx* and *Cnbp2*) (*10*) (Fig. 2A). *Cnbp2* is a retrotransposed gene that is specific to eutherians (*9*). We failed to detect any homolog of *Tsx*, *Jpx*, or *Ftx* genes in noneutherian vertebrates. In both chicken and *Xenopus*, the *XicHR* contains five protein genes (*Fip1l2*, *Lnx3*, *Rasl11c*, *UspL*, and *Wave4)* that have no detectable orthologs in eutherian genomes (table S1). The gene content, order, and orientation of the *XicHR* is perfectly conserved between chicken and *Xenopus* (Fig. 2A), which indicates that the chicken *XicHR* (on an autosome) corresponds to the ancestral state in the tetrapod cenancestor.

To search for possible vestiges of *XicHR* genes in eutherians, we compared the chicken genomic sequence to its counterpart from four species representative of different eutherian orders

[1]Laboratoire de Biométrie et Biologie Evolutive (UMR 5558), CNRS and Université Lyon 1, 16 rue Raphaël Dubois, 69622 Villeurbanne Cedex, France. [2]Unité de Génétique Moléculaire Murine, URA CNRS 1947, Institut Pasteur, 75015 Paris, France. [3]Genoscope, Centre National de Séquençage and CNRS UMR8030, Case Postale 5706, 91057 Evry Cedex, France.

*To whom correspondence should be addressed: E-mail: duret@biomserv.univ-lyon1.fr

(human, mouse, dog, and cow) (*9*). The *XicHR* covers 162 kb in chicken (998 kb in human), of which 5% (2%) consists of exons and 3% (59%) of repeat sequences. Comparison of human and chicken sequences revealed 22 alignments in nonrepeated sequences. Although these alignments are short (on average 62 base pairs with 72% identity), eight of them overlap with known exons in chicken, of which five also correspond to exons in human. The probability that such alignments would occur by chance is extremely low ($P < 10^{-7}$ in each species) (*9*), indicating that they correspond to homologous regions, conserved between human and chicken. Overall, we detected 63 distinct fragments in chicken, covering 3.4 kb, that align with at least one of the four eutherian species, with 12 of these alignments overlapping with chicken exons (from *Fip1l2*, *Lnx3*, and *Rasl11c*) (table S2).

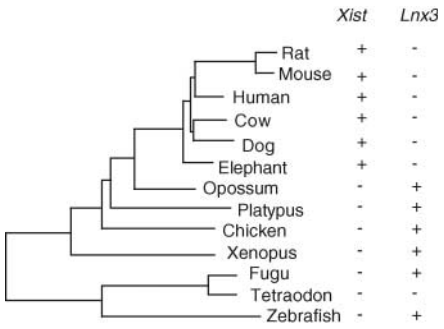

Fig. 1. Phylogenetic distribution of *Xist* and *Lnx3* within vertebrate species for which whole-genome sequence data are available. The phylogenetic tree was adapted from references (*16*, *17*). We searched for homologs of *Xist* within genomic sequences with BLASTN. The presence of a significant hit homologous to *Xist* is indicated by a plus sign (+). We searched for homologs of *Lnx3* with BLASTP against Ensembl protein predictions and with TBLASTN against genome assemblies, except for the platypus, for which we used whole-genome shotgun sequences. Phylogenetic analyses were conducted to distinguish *Lnx3* from its paralogs (fig. S2). The presence of a *Lnx3* ortholog is indicated by a plus sign (+).

There are six exons of *Fip1l2* that show homology with the human or mouse sequence. Three of them correspond to exons of *Tsx* (Fig. 2B). The protein alignment revealed that the mouse *Tsx* is a truncated gene, encoding a protein orthologous to the N-terminal end of *Fip1l2*. *Tsx* is functional (transcribed and translated) in mouse and rat, but is evolving very rapidly (*11*). *Tsx* is a pseudogene in human (*10*), as it is in dog and cow. We identified the four exons of the *Rasl11c* gene in cow and dog (one in human, none in mouse), but in all these species, *Rasl11c* has become a pseudogene. Two exons of *Lnx3* are homologous to *Xist* (Fig. 2C). The first corresponds to *Xist* exon h4/m4, which is well conserved in eutherians (Fig. 3A and fig. S1). The second corresponds to the exon h5/m6, which, although conserved in human and mouse, is more divergent in dog and cow (Fig. 3B). The probability that, by chance, two independent alignments overlap exons in both species is extremely low (5×10^{-5}), which indicates that these exons of the *Xist* RNA gene are homologous to the *Lnx3* protein gene.

In marsupials, the *XicHR* is located on the X chromosome (*7*, *12*). We have sequenced an opossum genomic clone including *Rasl11c* and the 5′ end of *Lnx3*, and we have sequenced the *Lnx3* mRNA (*9*). We have identified *Wave4* in sequence databases (Fig. 2A and table S1). Phylogenetic analyses indicate that these three genes are functional in the opossum (see fig. S2 for details on *Lnx3*). Thus, the loss of protein-coding function of *Lnx3* occurred in the eutherian lineage and was concomitant with the pseudogenization of at least two of the four other *XicHR* genes.

Lnx3 is conserved in all vertebrate classes and is highly similar to its paralogs *Lnx1* and *Lnx2* (*13*). The exons conserved in *Xist* correspond to two PDZ motifs, and both contain frameshift mutations (Fig. 3). By screening databases of expressed sequence tags, we found that in both chicken and *Xenopus*, *Lnx3* is transcribed in varied tissues and developmental stages. In the

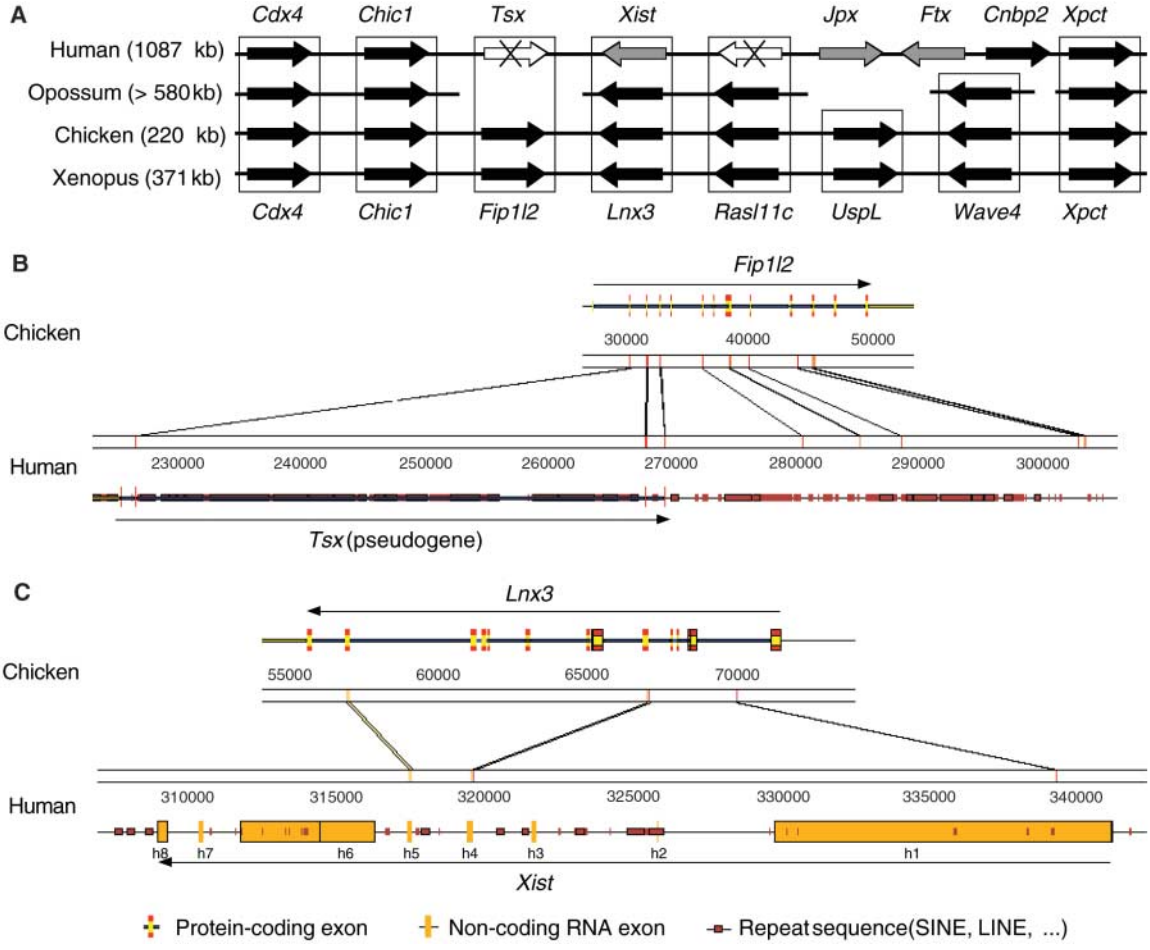

Fig. 2. Comparison of the human *Xic* region and of its orthologous region in opossum, chicken, and *Xenopus*. (**A**) Genomic map. Protein-coding genes are indicated in black, RNA-genes in gray, and pseudogenes in white. Groups of genes for which there is evidence of homology are surrounded by a rectangle. The assembly of the opossum genome is incomplete, and the order and orientation of contigs (thick black lines) is therefore not known. (**B**) Alignment of the chicken *Fip1l2* genomic region with the human *Tsx* region. (**C**) Alignment of the chicken *Lnx3* genomic region with the human *Xist* region. The numbering of *Xist* exons corresponds to the nomenclature proposed by reference (*10*). Positions are indicated as base pairs.

opossum, *Lnx3* is expressed both in males and females, a behavior very different from that of *Xist* in eutherians. In mice, although *Xist* exon h4/m4 (which is homologous to *Lnx3*) is dispensable for X inactivation (*14*), the exon has been shown to affect the transcription and/or processing of *Xist* RNA (*14*). This suggests that *Xist* might have retained some regulatory elements of the *Lnx3* transcription unit.

Our results show that two exons of *Xist* derive from *Lnx3*. However, *Lnx3* and *Xist* contain, respectively, 11 and 6 other exons, for which we failed to detect significant similarities. Notably, we did not detect homology to the *Xist* A-repeat, which is a discrete sequence element implicated in X-silencing function (*15*). This lack may be because RNA genes and protein genes are subject to very different selective constraints and may rapidly diverge. It is also possible that the first exons of *Xist* are not homologous to *Lnx3* and derive from the insertion of a sequence (e.g., a transposable element) that was recruited to form a proto-*Xist* gene. We analyzed the opossum genomic interval between *Rasl11c* and *Lnx3* to search for hallmarks of a potential proto-*Xist* gene, but failed to detect any significant similarity with *Xist,* even using the most accurate alignment software (*9*). Given that *Xist* exons are highly conserved among eutherians, the lack of similarity with the opossum strongly suggests that marsupials do not contain any proto-*Xist* gene at this locus and, hence, that *Xist* is specific of eutherians.

The mechanisms of dosage compensation in marsupials and eutherians both involve chromosomewide X inactivation (XCI), but with some significant differences. In marsupials, it is always the paternal X-chromosome that is inactivated, and the inactivation, which is incomplete and tissue-specific, does not seem to involve DNA methylation (*4*). Our results, moreover, indicate that in marsupials, XCI does not involve *Xist*. In monotremes, the *XicHR* has been translocated to an autosome, which indicates that dosage compensation does not require this locus (*6*). There is, therefore, no evidence that the processes of dosage compensation in eutherian, marsupial, and monotremes are homologous. It is possible that *Xist*-independent XCI existed in the mammalian cenancestor and that *Xist* overtook this mechanism in eutherians. However, it should be stressed that in the earliest stages of the divergence of the X and Y chromosomes, most of the X-linked genes still had an active Y homolog and so did not need dosage compensation. It is only after the Y chromosome had lost a large number of genes that it might become advantageous to achieve dosage compensation by inactivation of the whole X chromosome. We, therefore, propose that the emergence of XCI might be a late event in the evolution of sexual chromosomes.

A *Xist* exon h4/m4:

```
Chicken-Lnx3    TTTTTTCTCTTTGATTTCCACCATCAGATGTTCCTC-AGAGAGCTAGCCCTGTGGCTGAA
Human-Xist      ---TTTTTCTTTTATCTCTTTT--CAGATCTTCCTC-AGAAGAATAGGCTTGTTGTTTTA
Mouse-Xist      TGGCTTTTCTTTCACCTCTTTT-CCAGATCTCCCCCCAGAATTGTGGGCTTGCTGCTTTG
Dog-Xist        TTATTTTTCTTTGATCTCTTCT-CCAGATGTTCCTC-AGAAGAATAGGCTTGTTGCTCTA
Cow-Xist        TTATTTTTCTTTTATCTCTTTT-CCAGATCTACCTC-AAAAGAATAGGCCTGTTGCTTTA
                :  :** ***** *: **      ***** * **:* *:*    *:* * **: *:*  :
                                        <- exon start

Chicken-Lnx3    ACGAGCACAATAGAAATTCACCGAGAAGATCCAGAGGAAGAGCTTGGG-ATGAGAATAGT
Human-Xist      CAGTGTTAGTGATCCATTCCCTTTGACGATCCCTAGGTGGAGATGGGGCATGAGGATCCT
Mouse-Xist      CAGTGCTGGCGACCTATTCCCTTTGACGATCCCTAGGTGGAGATGGGGCATGAGGATCCT
Dog-Xist        CAGTGTTAGTGACCCGTTCCCTTTGACGATCCCTAGGTGGAGATGGGGCATGAGGATCCT
Cow-Xist        CAGTGTTAGTGACCCATTCCCTTTGACGATTCCTAGGTGGAGATGGGGCATGAGGATCCT
                  * *      *   :*** *   ** ***:*  ***  *** * *** ***** **  *

Chicken-Lnx3    T--GGGGGTAAGGACACG-CCACTAGGAAACA
Human-Xist      CCAGGGGAAAAGCTCACTACCACTGGGCAACA
Mouse-Xist      CCAGGGGAATAGCTCACCACCACTGGGCAACA
Dog-xist        CCAGGGGAAAAGCTCACTACCACTAGGCAACA
Cow-Xist        CCAGGGGAAAGATTCACTACCACTGAGCAACA
                   ****  :::  ***  ***** :* ****
```

B *Xist* exon h5/m6:

```
Chicken-Lnx3    TCTTTGCAGCTACCTTCACTG----------CTGTCAAGATATTGTCCTTAGCAAAGGCA
Human-Xist      TTTTTATAGCTCTTCATTGTTCC-----TATCTGCCAAATCATTATACTTCCTACAAGCA
Mouse-Xist      --TTTGTAGTGCC---------------ATCTACCAAAT-ATTACCCTTCCCCAAAGCA
Dog-Xist        TTTTTATAGCTCCTGATTGTTCCCTTTTTATCTACCAAATCATTGTC--TCCCAAAAGCA
Cow-Xist        CTTTTGTAGCTCCTGGTTGTTCCCTTCATATTTGCCAAATCATTATCTTTCCCTGAAGTA
                  ***  **                      :*  ***   *** :: :*  :  * *:*
                       <- exon start

Chicken-Lnx3    ACCAGGAAAGCTGGGGCTTCAGCATTGTTGGAGGCTTTGAGGAGAGCAAAGGAAACCAGC
Human-Xist      GTGCAGAGAGCTGAGTCTTCAGCAGGTCCAAGAAATTTGAACACACTGAAGGAAGTCAGC
Mouse-Xist      GCACAGAAAACTGGGTCTTCAGCGTGATCAAGCAATGTGAACACACAAAAGGAAGGCAGC
Dog-Xist        GTGCAGAGGGC-----------------AAGAAAGTGGAACATATCAAAG---------
Cow-Xist        GTGCAAAGAGC-----------------AAGAAATGTGAACACACC-AAG---------
                     :*  :*                    : :**  * *   ***

Chicken-Lnx3    CCTTCT----TCATCAAAACCATCGTGCCTGGGACGCCTGCC-TGCCGCAGCGCAAGGCT
Human-Xist      CTTCCC----ACCTGAAGATCAACATGCCTGGCACTCTAGCA-CTTGAGGATAGCTGAAT
Mouse-Xist      TTTATAAATGACCCGAGGATCAACATGCCTG--ACTGCAGCATCTTAAAAGCAATAGAAT
Dog-Xist        --------------GAAGATCAACATGCCTGGCATGCTAGCA-TTTTAGAACATCAGAAT
Cow-Xist        --------------GAAGATCAACATGCCTGCAATGCTAGCA-TTTTAGAATAGCAGAAT
                              *: * ** * ******  *  :  **        :      *  *

Chicken-Lnx3    GAAGTGAGT
Human-Xist      GAAGTAAGT
Mouse-Xist      GAGGTAAGT
Dog-Xist        GAAGTAAGT
Cow-Xist        GAAGTAAGT
                **:** ***
          exon end ->
```

Fig. 3. Alignment of the two homologous regions of chicken *Lnx3* and eutherian *Xist*. (**A**) Sequence alignment [computed with MUSCLE (*18*)] of the 5′ part of *Xist* exon h4/m4 and of *Lnx3* exon 3 (Ensembl transcript ENSGALT00000012483). (**B**) Alignment of the entire *Xist* exon h5/m6 and *Lnx3* exon 9. Exon boundaries are indicated in bold. Sites that are conserved in all species or in four out of five species are indicated respectively by an asterisk (*) and a colon (:).

References and Notes

1. M. Kohn, H. Kehrer-Sawatzki, W. Vogel, J. A. Graves, H. Hameister, *Trends Genet.* **20**, 598 (2004).
2. B. Charlesworth, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5618 (1978).
3. M. T. Ross *et al.*, *Nature* **434**, 325 (2005).
4. F. Grutzner, J. A. Graves, *Curr. Opin. Genet. Dev.* **14**, 642 (2004).
5. J. C. Chow, Z. Yen, S. M. Ziesche, C. J. Brown, *Annu. Rev. Genomics Hum. Genet.* **6**, 69 (2005).
6. P. D. Waters *et al.*, *Chromosome Res.* **13**, 401 (2005).
7. J. A. Graves, M. Westerman, *Trends Genet.* **18**, 517 (2002).
8. H. S. Chandra, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6947 (1985).
9. Methods and other details are described in Supporting Online Material.
10. C. Chureau *et al.*, *Genome Res.* **12**, 894 (2002).
11. D. B. Cunningham, D. Segretain, D. Arnaud, U. C. Rogner, P. Avner, *Dev. Biol.* **204**, 345 (1998).
12. E. Koina *et al.*, *Chromosome Res.* **13**, 687 (2005).
13. D. S. Rice, G. M. Northcutt, C. Kurschner, *Mol. Cell. Neurosci.* **18**, 525 (2001).
14. M. L. Caparros, M. Alexiou, Z. Webster, N. Brockdorff, *Cytogenet. Genome Res.* **99**, 99 (2002).
15. A. Wutz, T. P. Rasmussen, R. Jaenisch, *Nat. Genet.* **30**, 167 (2002).
16. W. J. Murphy *et al.*, *Nature* **409**, 614 (2001).
17. E. H. Margulies *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3354 (2005).
18. R. C. Edgar, *Nucleic Acids Res.* **32**, 1792 (2004).

19. This work was supported by the CNRS, the Institut Pasteur, the Genoscope, the European Union Epigenome Network of Excellence, and the Mission de la Recherche et de la Technologie. We thank the Broad Institute for making the opossum sequences publicly available. Sequences have been deposited in EMBL (accession numbers AM230659 and AM230660).

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5780/1653/DC1
Materials and Methods
SOM Text
Figs. S1 and S2
Tables S1 and S2
References and Notes

16 February 2006; accepted 17 May 2006
10.1126/science.1126316