RNA Biology 10:8, 1262–1265; August 2013; © 2013 Landes Bioscience

Long non-coding RNAs and human X-chromosome regulation

A coat for the active X chromosome

Céline Vallot and Claire Rougeulle*
Université Paris Diderot; Sorbonne Paris Cité; CNRS; UMR7216 Epigenetics and Cell Fate; Paris, France

In mammals, the genic disequilibrium between males (XY) and females (XX) is resolved through the inactivation of one of the X-chromosomes in females. X-chromosome inactivation (XCI) takes place in all mammalian species, but has mainly been studied in the mouse model where it was shown to be controlled by the interplay of several long non-coding RNA (lncRNA). However, recent data point toward the existence of species divergences among mammals in the strategies used to achieve XCI. The recent discovery of *XACT*, a novel lncRNA that coats the active X-chromosome specifically in human pluripotent cells, further highlights the existence of human-specific mechanisms of X-chromosome regulation. Here, we discuss the roles of lncRNAs in defining species-specific mechanisms controlling X-inactivation and explore the potential role of large lncRNAs in gene activation.

Keywords: long non-coding RNAs, X-chromosome inactivation, XIST, XACT, chromosome coating

Submitted: 05/06/13

Accepted: 07/18/13

http://dx.doi.org/10.4161/rna.25802

*Correspondence to: Claire Rougeulle; Email: claire.rougeulle@univ-paris-diderot.fr

X-chromosome inactivation (XCI) is established in the early stages of embryonic development. It switches in the mouse from an imprinted form, which characterizes pre-implantation development and extra-embryonic annexes, to a random process in the embryo proper after its implantation.[1] It was demonstrated in the mouse that the long non-coding RNA *Xist* is the master regulator of the process, as its expression triggers the initiation and spreading of X-chromosome silencing and the recruitment of repressive histone marks.[2] Mouse female embryonic stem cells are a useful tool to study XCI as they recapitulate ex vivo the X-inactivation process: displaying two active X chromosomes (Xa) in the pluripotent state, they undergo random XCI as they differentiate; *Xist* starts to be expressed at the onset of differentiation from the future inactive X chromosome. In humans, it has been possible to get insights into the early stages of XCI through a limited number of studies of human embryos, as well as the analysis of human embryonic stem (hES) cells. hES cells have been shown to be a more complex model system regarding XCI than their murine counterpart. Most undifferentiated hES cells have already undergone XCI and one X-chromosome has been inactivated; however, these cells tend to lose the expression of *XIST*.[3] hES cells, and their heterogeneity regarding XCI, constitute a spontaneously perturbed system to understand the early steps of X-chromosome repression.

LncRNAs in X-Chromosome Regulation: Is it Sequence or Function Conservation that Matters?

Xist, as well as the surrounding genomic region known as the X-inactivation center (*Xic*), is generally conserved in eutherians, and in particular between mouse and human. This region harbors several genes producing lncRNAs known or suspected to participate in the regulation of XCI (**Fig. 1**), and whose function has been mostly studied in the mouse. The extent to which these lncRNAs act similarly in other species remains to be thoroughly investigated. For instance, *Xist*, which is the most characterized of all, is commonly thought to control XCI in humans as it does in the mouse, but this has not been

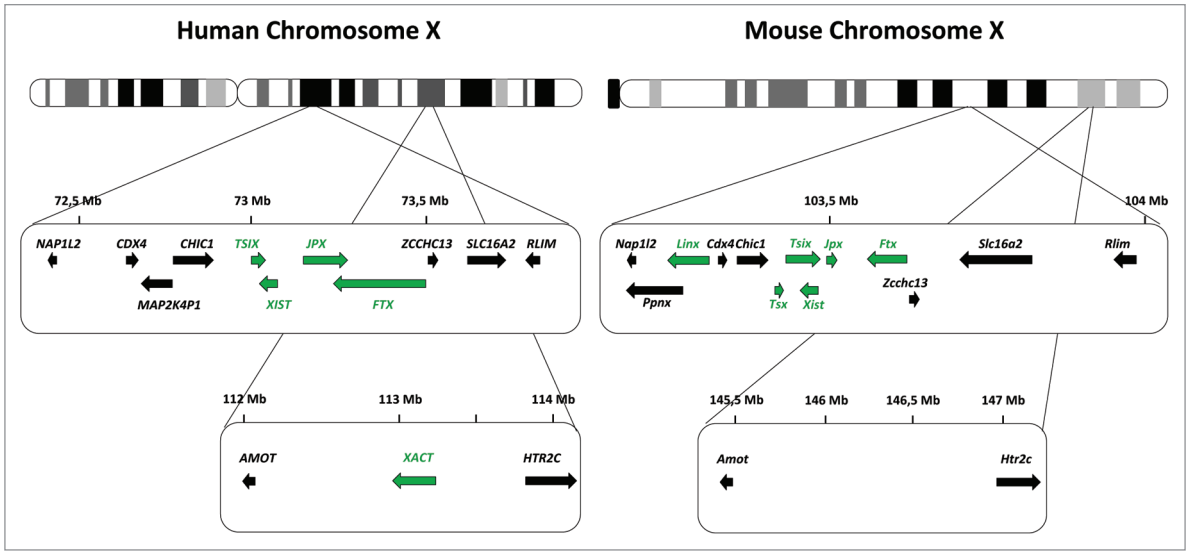

Figure 1. Comparative maps of the human and mouse X-chromosomes, including the X-inactivation center region (upper part), and the *XACT* region (lower part), according to the RefSeq annotation. Note that in the mouse, no transcript similar to *XACT* could be detected in the syntenic region. Protein-coding genes are indicated in black, genes producing lncRNAs in green.

formally proven; in fact, several differences in *Xist* expression profiles between human and mouse have recently been reported. *XIST* expression is highly variable in hES cells, in contrast to the mouse, where *Xist* upregulation is tightly associated with ES cell differentiation.[4] More strikingly, observations in human pre-implantation embryos, where X-chromosomes are active yet coated by *XIST*, further support some degree of uncoupling between *XIST* expression and X-chromosome repression in human.[5]

Among the other lncRNAs within the *Xic*, *Jpx*, and *Ftx* stand as potential activators of the X-inactivation process.[6,7] They have well-defined human orthologs, whose functions have not yet been investigated. In contrast, *Tsix* conservation is the subject of some debate. *Tsix* in the mouse is antisense to *Xist* and inhibits the accumulation of *Xist* transcripts on the future active X, protecting it from inactivation.[8] *Tsix* is also believed to be involved in the transition from imprinted to random X-inactivation,[9] an event that does not exist in the human. A *XIST* antisense transcript has been identified in humans, but it bears very little similarity to the mouse *Tsix* and its properties argue against a role in *XIST* repression.[10] In particular, *TSIX* and *XIST* can be concomitantly expressed from the same X-chromosome in human cells.[11] Although the hypothesis of an alternative antisense transcript that could assume *Tsix*-like function in humans remains formally possible, the differences in the dynamics of *XIST* expression during early development in mouse and human raises the question as to the necessity for conserving such a function. Finally, *Linx* is a recently identified lncRNAs that is suspected to participate in the control of *Tsix* expression.[12] Its conservation in humans is not yet documented, but the lack of *Tsix*-like function in humans would argue against *Linx* conservation.

The comparison of human and mouse XCI therefore reveals a high variability in the involvement of lncRNAs. Analysis of more distantly related species further uncovered the intriguing finding that non-homologous lncRNAs could harbor similar function in XCI; in marsupials, where *Xist* is not conserved, another lncRNA, *Rsx*, has indeed been shown to have *Xist*-like properties.[13] *Rsx* coats the inactive X-chromosome, and can induce some degree of gene silencing when inserted on an autosome in mouse ES cells. It appears altogether that the regulation of X-chromosome activity in mammals relies on a plastic interplay of rapid-evolving lncRNAs. It is therefore tempting to speculate that additional non-coding actors, yet to be discovered, might come into play and define species-specific mechanisms.

Identification of XACT, a Large lncRNA Coating the Human Xa

We recently discovered a novel long non-coding RNA, *XACT*, which displays the striking property of forming a cloud around active X-chromosome(s) (Xa) in human cells.[14] *XACT* is only the second example (after *XIST*) of a lncRNA coating a chromosome, and the first to coat an active one. *XACT* is expressed from a region of the human X-chromosome located some 40 Mb (megabases) from the *XIC* (**Fig. 1**). *XACT* expression and coating of the Xa is restricted to pluripotent and early differentiating cells, when XCI is still extremely dynamic. *XACT* repression appears to correlate with the establishment of stable, irreversible XCI. Whether *XACT* is strictly involved in XCI remains to be investigated. However, its expression pattern leads us to speculate that *XACT* might be involved in the regulation of the activity of the Xa during early human embryogenesis, potentially by protecting it from inactivation in the early steps of XCI. No *XACT*-like transcript could be identified from the syntenic region in the mouse. Although one cannot exclude the possibility that another lncRNA transcribed from a different mouse genomic location might share *XACT* features, our findings strongly suggest that *XACT* could participate in human-specific mechanisms for controlling X-chromosome activity.

XACT has the intriguing property of being a 251.8 kbp-long mostly unspliced RNA. Apart from *XACT*, few large and mostly unspliced lncRNAs have previously been identified in mammals, including Kcnq1ot1 (91.5 kbp), Airn (108 kbp), and Nespas (32 kbp).[15] None of them broadly accumulates around a chromosome. These large unspliced lncRNAs have previously been overlooked in RNA-seq data sets; *XACT* was for example not listed in a reference catalog of more than 8,000 human lncRNAs recently characterized through an integrative analysis of various RNA-seq data sets.[16] One reason for this is that workflows of genome-wide lncRNA identification have limited their analysis to spliced RNA species. This is likely due to the fact that, in the absence of splicing event, it remains hazardous to delineate accurately unspliced lncRNAs with current RNA-seq technologies. The low expression levels of lncRNAs moreover contribute to their uncertain identification; lncRNAs species might be difficult to distinguish from pervasive transcription. In addition, it was recently shown that RNA-seq of total RNA depleted of rRNAs, which was used for the identification of *XACT*, is advantageous for detecting these large lncRNAs over the widely used polyA+ RNA fractions.[17] Indeed, large lncRNAs might be easily degraded and generate fragments lacking polyA tails. Numerous ENCODE RNA-seq data sets have been generated using polyA RNA fractions,[18] large unspliced lncRNAs are likely under-represented or absent from these reference data sets. One important question that remains regarding these peculiar RNA species is whether it is the transcripts themselves or their transcription which is essential to their function. In the case of *Airn,* it has recently been shown that it is the transcription of the locus, rather than the transcript, which controls the imprinting of the locus.[19] In the case of *XACT*, however, the fact that it is located in a gene desert region and the ability of the transcript to coat the chromosome argue for a role of the transcript itself rather than its transcription.

LncRNAs and Gene Activation

The three large lncRNAs mentioned above are involved in the control of imprinted gene expression, and more precisely, in the silencing of gene clusters in *cis*.[15] More generally, lncRNAs are often associated with gene repression. Various lncRNAs, among which *Xist*, *HOTAIR*, *Kcnq1ot1*, *COLDAIR*, and others associate for example to the Polycomb complex PRC2 and target this complex in *cis* and/or *trans* to mediate H3K27 methylation and gene silencing.[20] In mammalian cells, however, some examples of lncRNAs involved in gene activation have also been characterized. *HOTTIP*, which is located at the 5' end of the *HOXA* locus, coordinates the activation of HOXA genes through its binding to a H3K4 methylase MLL.[21] Similarly, NeST recruits the MLL complex to activate the interferon-γ gene, thus modifying susceptibility to pathogens.[22] It has also been shown that ncRNA-a, ncRNA-activating, can activate their neighboring genes using a cis-mediated mechanism implicating Mediator in order to enhance transcription.[23,24] In all cases, however, the activating transcriptional effect remains local, with only close-by neighbor genes being affected. Although the potential role of *XACT* in the regulation of the active X-chromosome activity remains to be determined, it seems that *XACT* could have a role in regulating a large fraction of the active X-chromosome as it co-localizes with a significant portion of the chromosome territory. It remains to be addressed whether *XACT* actively promotes the activity of the Xa or simply protects the Xa from inactivation. In this line of thought, instructive examples are provided by the dosage compensation systems in fruit flies, where the male-specific-lethal complex increases the transcription levels of the male X-chromosome to compensate for sex chromosome asymmetry.[25] The MSL complex includes two ncRNAs, *roX1* and *roX2*, which are transcribed from the male X-chromosome. *roX* RNAs appear to be important for correct targeting and spreading of the chromatin modifying MSL complex to the male X-chromosome. They provide an example where lncRNAs participate in chromosome-wide activating activities. Moreover, considering this dosage compensation system, it appears that the determination of *XACT* interaction partners, as well as its fixation sites on the human Xa, will be crucial to gain insight into its function and potential mode of action.

Lessons from mammals and flies are teaching us how central lncRNAs are to dosage-compensation processes, no matter the underlying mechanisms. Their RNA nature and the weak evolutionary constraints exerted on them could allow them to evolve more rapidly than protein-coding genes and to adapt to species specificities. X-chromosome inactivation in mammals in particular relies on the interplay of rapid-evolving lncRNAs regulating the activity of the X-chromosomes. The two faces of this RNA-orchestrated symphony are provided by non-orthologous lncRNAs playing similar function in different species (*XIST* and *Rsx*), as well as by the existence of species-specific long, non-coding transcripts such as *XACT*. The function of the latter in human development, and more specifically in the peculiar regulation of XCI in early embryos, remains to be understood.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E. Epigenetic dynamics of imprinted X inactivation during early mouse development. Science 2004; 303:644-9; PMID:14671313; http://dx.doi.org/10.1126/science.1092727
2. Chow JC, Heard E. Nuclear organization and dosage compensation. Cold Spring Harb Perspect Biol 2010; 2:a000604; PMID:20943757; http://dx.doi.org/10.1101/cshperspect.a000604
3. Shen Y, Matsuno Y, Fouse SD, Rao N, Root S, Xu R, et al. X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. Proc Natl Acad Sci USA 2008; 105:4709-14; PMID:18339804; http://dx.doi.org/10.1073/pnas.0712018105
4. Makhlouf M, Rougeulle C. Linking X chromosome inactivation to pluripotency: Necessity or fate? Trends Mol Med 2011; 17:329-36; PMID:21411371; http://dx.doi.org/10.1016/j.molmed.2011.02.001
5. Okamoto I, Patrat C, Thépot D, Peynot N, Fauque P, Daniel N, et al. Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature 2011; 472:370-4; PMID:21471966; http://dx.doi.org/10.1038/nature09872
6. Chureau C, Chantalat S, Romito A, Galvani A, Duret L, Avner P, et al. Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. Hum Mol Genet 2011; 20:705-18; PMID:21118898; http://dx.doi.org/10.1093/hmg/ddq516
7. Tian D, Sun S, Lee JT. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell 2010; 143:390-403; PMID:21029862; http://dx.doi.org/10.1016/j.cell.2010.09.049

8. Lee JT, Lu N. Targeted mutagenesis of *Tsix* leads to nonrandom X inactivation. Cell 1999; 99:47-57; PMID:10520993; http://dx.doi.org/10.1016/S0092-8674(00)80061-6
9. Navarro P, Pichard S, Ciaudo C, Avner P, Rougeulle C. Tsix transcription across the Xist gene alters chromatin conformation without affecting Xist transcription: implications for X-chromosome inactivation. Genes Dev 2005; 19:1474-84; PMID:15964997; http://dx.doi.org/10.1101/gad.341105
10. Migeon BR, Chowdhury AK, Dunston JA, McIntosh I. Identification of *TSIX*, encoding an RNA antisense to human *XIST*, reveals differences from its murine counterpart: implications for X inactivation. Am J Hum Genet 2001; 69:951-60; PMID:11555794; http://dx.doi.org/10.1086/324022
11. Migeon BR, Lee CH, Chowdhury AK, Carpenter H. Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation. Am J Hum Genet 2002; 71:286-93; PMID:12023758; http://dx.doi.org/10.1086/341605
12. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 2012; 485:381-5; PMID:22495304; http://dx.doi.org/10.1038/nature11049
13. Grant J, Mahadevaiah SK, Khil P, Sangrithi MN, Royo H, Duckworth J, et al. Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. Nature 2012; 487:254-8; PMID:22722828; http://dx.doi.org/10.1038/nature11171
14. Vallot C, Huret C, Lesecque Y, Resch A, Oudrhiri N, Bennaceur-Griscelli A, et al. XACT, a long noncoding transcript coating the active X chromosome in human pluripotent cells. Nat Genet 2013; 45:239-41; PMID:23334669; http://dx.doi.org/10.1038/ng.2530
15. Guenzl PM, Barlow DP. Macro lncRNAs: a new layer of cis-regulatory information in the mammalian genome. RNA Biol 2012; 9:731-41; PMID:22617879; http://dx.doi.org/10.4161/rna.19985
16. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011; 25:1915-27; PMID:21890647; http://dx.doi.org/10.1101/gad.17446611
17. Huang R, Jaritz M, Guenzl P, Vlatkovic I, Sommer A, Tamir IM, et al. An RNA-Seq strategy to detect the complete coding and non-coding transcriptome including full-length imprinted macro ncRNAs. PLoS One 2011; 6:e27288; PMID:22102886; http://dx.doi.org/10.1371/journal.pone.0027288
18. Consortium EP; ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol 2011; 9:e1001046; PMID:21526222; http://dx.doi.org/10.1371/journal.pbio.1001046
19. Latos PA, Pauler FM, Koerner MV, Şenergin HB, Hudson QJ, Stocsits RR, et al. Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. Science 2012; 338:1469-72; PMID:23239737; http://dx.doi.org/10.1126/science.1228110
20. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem 2012; 81:145-66; PMID:22663078; http://dx.doi.org/10.1146/annurev-biochem-051410-092902
21. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 2011; 472:120-4; PMID:21423168; http://dx.doi.org/10.1038/nature09819
22. Gomez JA, Wapinski OL, Yang YW, Bureau JF, Gopinath S, Monack DM, et al. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-γ locus. Cell 2013; 152:743-54; PMID:23415224; http://dx.doi.org/10.1016/j.cell.2013.01.015
23. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, et al. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature 2013; 494:497-501; PMID:23417068; http://dx.doi.org/10.1038/nature11884
24. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell 2010; 143:46-58; PMID:20887892; http://dx.doi.org/10.1016/j.cell.2010.09.001
25. Conrad T, Akhtar A. Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. Nat Rev Genet 2011; 13:123-34; PMID:22251873; http://dx.doi.org/10.1038/nrg3124