Seminars in Cell & Developmental Biology 56 (2016) 48–57

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

Enlightening the contribution of the dark matter to the X chromosome inactivation process in mammals

Miguel Casanova, Tharvesh Moideen Liyakat Ali, Claire Rougeulle*

Sorbonne Paris Cité, Epigenetics and Cell Fate, UMR 7216 CNRS, Université Paris Diderot, Paris, France

ARTICLE INFO

Article history:
Received 21 December 2015
Received in revised form 4 May 2016
Accepted 4 May 2016
Available online 9 May 2016

Keywords:
X chromosome inactivation
Long non-coding RNAs
Transposable elements
Stem cell
Pluripotency

ABSTRACT

X-chromosome inactivation (XCI) in mammals represents an exceptional example of transcriptional co-regulation occurring at the level of an entire chromosome. XCI is considered as a means to compensate for gene dosage imbalance between sexes, yet the largest part of the chromosome is composed of repeated elements of different nature and origins. Here we consider XCI from a repeat point of view, interrogating the mechanisms for inactivating X chromosome-derived repeated sequences and discussing the contribution of repetitive elements to the silencing process itself and to its evolution.

Contents

1. Introduction

The evolution of dimorphic sex chromosomes, X and Y, from an ancestral pair of autosomes in some Bilateria introduced a disequilibrium of gene dosage between XY males and XX females. In therian mammals, this imbalance is compensated, at least in part, by silencing one of the two X chromosomes in females [1,2], in a process commonly known as X chromosome inactivation (XCI).

XCI is a paradigm of epigenetic regulation that is developmentally regulated and during which the two X homologs are differentially treated within the same nucleus, resulting in a full chromosome being transcriptionally shut down and converted to facultative heterochromatin. XCI is tightly linked to the cellular context and is generally established concomitantly to cell differentiation. In most cases XCI is random, affecting either the paternal or the maternal X chromosome. However imprinted XCI, with selective inactivation of the paternal X chromosome, occurs in marsupials and in extraembryonic lineages of some eutherian species including rodents [3,4]. A key player in the XCI process in eutherians is the long noncoding RNA (lncRNA) *Xist* [5–7], which is produced from an X-linked control region, the X-inactivation center (*Xic*). The *Xic* region contains several additional non-coding genes, including *Tsix*, *Ftx* and *Jpx* that are believed to orchestrate *Xist* expression and XCI [8]. The *Xist* lncRNA has the peculiar property of covering the chromosome from which it originates, forming a nuclear RNA domain and triggering transcriptional and chromatin changes

* Corresponding author.
E-mail address: claire.rougeulle@univ-paris-diderot.fr (C. Rougeulle).

http://dx.doi.org/10.1016/j.semcdb.2016.05.003
1084-9521/

throughout the chromosome [2]. An initial cascade of events during the establishment of XCI includes exclusion of RNA polymerase II (RNA PolII) and modifications of core histone tails (hypoacetylation of histone H3 and H4, hypomethylation of H3K4, hypermethylation of H3K9, H3K27and H4K20). In later stages, additional changes are observed that contribute to lock in the inactive state, with a switch to late replication timing, histone variant exchange and DNA methylation of CpG islands [9].

XCI is usually seen as a way to compensate for gene dosage imbalance between sexes. However the genic fraction of mammalian chromosomes is relatively low, with repetitive sequences accounting for approximately half of the genomes (Fig. 1A). The X in particular is among the chromosomes displaying the lowest gene density (5.2 genes per Mb of DNA, compared to an average of 7.6 for autosomes), and an excess of repetitive elements [10]. This excess is mostly explained by an enrichment of long interspersed nuclear elements (LINEs) of the L1 subclass [10,11], and also, albeit to a lower extent, of long terminal repeat (LTR) elements, inverted repeats and retrogenes (Fig. 1B, C, Glossary) [12]. These repeat elements have often been referred to as "selfish DNA", or "junk DNA", mostly because they were thought to spread in the genome without conferring any advantages to the host organism [13]. However, mounting evidence shows that repeat elements may play important functional roles in genomic regulation [14] and one might wonder whether this holds true for XCI. More specifically, do X-linked repetitive sequences play a role in XCI (and if so, how?) or are they neutral to the process? Is silencing of X-linked repetitive sequences achieved through equivalent mechanisms to that of genes? In this review, we will describe the silencing of the X chromosome repeat fraction and focus on the role of those repeats in XCI, revisiting the possible function of classical players, such as LINEs and discussing potential roles for different classes of repeat elements.

2. Inactivating the X chromosome—the silence of the repeats

In interphase nuclei of differentiated cells, the inactive X chromosome (Xi) is found as a highly compacted heterochromatic structure, initially described as the Barr body in humans and usually located at the nuclear or nucleolar periphery [15–17]. Classically the Barr body was considered as a chromosomal structure promoting the silencing of embedded genes. However it has been shown, at least in human, that the core of the Barr body is enriched in inactive repetitive sequences, including satellite DNA and Cot-1 sequences (Glossary) [18]. Assessment of the kinetics of XCI in the mouse further revealed that silencing of the repetitive portion of the X-chromosome is one of the first observable events, occurring prior to gene inactivation [19,20]. At the onset of XCI, *Xist* expression and accumulation leads to the quick exclusion of the transcription machinery from the X chromosome and to the rapid silencing of the repeat fraction [19,21]. Genes initially remain located outside of the *Xist* RNA domain and stay temporarily active; they are subsequently relocalized closer or into this domain and inactivated (Fig. 2) [18,19]. Alternative approaches aiming at capturing chromosome conformation in an allelic-specific manner produced high-resolution topology maps of the active and inactive X chromosomes, which further support the hypothesis of a random organization of silenced genes within the Xi territory, while escapees appear to cluster and locate at its periphery [22].

The differential behavior of genes and repeats during the establishment and maintenance of random XCI in mouse and humans suggests that inactivation of the two fractions could rely on different mechanisms and/or on different portions of the *Xist* RNA. The *Xist* transcript comprises 8 regions (A–H) made of relatively conserved tandem repeat elements [23] (see Section 4). Only one, the A-repeat, located at the 5′ end of the transcript, is critical for gene silencing [24]. The A-repeat is required for the relocalization of genes within the *Xist* repressive compartment [19] and for spreading of the *Xist* RNA through gene dense regions [25]. Intriguingly however, the A-repeat is dispensable for the initial formation of the repressive compartment and for silencing of the repeat fraction of the chromosome [19]. Random X-inactivation is thus a two-step process, initiating with an A-repeat independent repression of repetitive elements, followed by an A-repeat dependent gene silencing.

Inactivation of the X chromosome in two mechanistically independent steps also appears to take place during early mouse embryonic development [20], where imprinted X inactivation of the paternal X occurs. Repetitive elements on the Xp are silenced very early, from the two-cell stage onwards, in a process that is independent of *Xist* and might rely on paternal imprints acquired during gamete maturation and meiotic sex chromosome inactivation (MSCI); later, at the morula-to-blastocyst transition, genic silencing on the Xp is initiated, this time in a *Xist*-dependent manner [20]. A distinct study suggested that initiation of genic repression could also be *Xist*-independent; in this model *Xist* would rather be required to stabilize silencing [26]. Altogether and independently of the discrepancies, these observations reveal a bimodal silencing of the X, with repeat and genic fractions following different kinetics and relying on separate mechanisms, of which some appear independent of *Xist*. It is even conceivable that the initial formation of a repressed repeat compartment is essential to prepare for subsequent and efficient silencing of the entire X, through pan-chromosomal structural and nuclear reorganization. In this scenario, X-linked gene silencing would have a stronger dependency on the repeat fraction of the X chromosome than previously anticipated.

In agreement with this hypothesis, a recent study pointed to a surprising role for repeat RNAs in promoting chromatin decompaction [27]. Cot-1 RNAs are indeed highly abundant and tightly associated with the chromosomes from which they are expressed in several primary and cancer mouse and human cell lines as well as in different tissue sections. Cot-1 RNAs distribute mostly across euchromatin regions and are excluded from heterochromatin foci, including the inactive X (as discussed above). Releasing Cot-1 RNAs from chromatin resulted in aberrant chromatin distribution and condensation, suggesting a role for repeat transcripts in preventing higher-order chromatin packaging. In this context, and given the abundance of repeats on the X chromosome, it can be speculated that the X-linked repeat fraction has to be rapidly silenced to allow proper reorganization and efficient silencing of the Xi.

3. The role of repeats in promoting silencing of the X chromosome

Several studies on X/autosome translocations reported less efficient spreading of *Xist* RNA and XCI on autosomes compared to the X chromosome [28,29]. Similar observations led Gartler and Riggs to propose the existence of X-specific way stations that would act as boosters for the efficient spread of inactivation along the X chromosome [30]. The concept of way stations and the particular repeat composition of the X chromosome suggested that repeat sequences could be involved in X-chromosome silencing. Mary Lyon later proposed the repeat hypothesis, in which LINE-1 (L1) elements would serve as the way stations postulated by Gartler and Riggs [31].

L1s are autonomous, non-LTR retrotransposons, about 6–7 kb in length (Fig. 1C). Most L1 elements are truncated at their 5′ end and therefore inactive. However, a small cohort of full-length L1s, that still have retro-transposition potential, can be found both in human