Biochimica et Biophysica Acta xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Review

X-marks the spot: X-chromosome identification during dosage compensation ☆

Jessica Chery, Erica Larschan *

Department of Molecular Biology, Cellular Biology, and Biochemistry, Brown University, 185 Meeting Street, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 3 August 2013
Received in revised form 29 December 2013
Accepted 30 December 2013
Available online xxxx

Keywords:
Dosage compensation
Drosophila
Mammals
C. elegans

ABSTRACT

Dosage compensation is the essential process that equalizes the dosage of X-linked genes between the sexes in heterogametic species. Because all of the genes along the length of a single chromosome are co-regulated, dosage compensation serves as a model system for understanding how domains of coordinate gene regulation are established. Dosage compensation has been best studied in mammals, flies and worms. Although dosage compensation systems are seemingly diverse across species, there are key shared principles of nucleation and spreading that are critical for accurate targeting of the dosage compensation complex to the X-chromosome(s). We will highlight the mechanisms by which long non-coding RNAs function together with DNA sequence elements to tether dosage compensation complexes to the X-chromosome. This article is part of a Special Issue entitled: Chromatin and epigenetic regulation of animal development.

1. Introduction

Assuring that all genes are transcribed at the correct gene dosage across the genome is a key process because protein and RNA components of large macromolecular complexes are encoded on different chromosomes. Altering the stoichiometry of macromolecular complexes causes severe phenotypes [1]. The most commonly known disease of gene dosage is Down's syndrome, which is caused by trisomy of chromosome 21 [2–4]. Furthermore, copy number variation throughout the genome has been linked to diverse diseases from autism to schizophrenia, and neurons are particularly sensitive to copy number changes [5].

There is a naturally occurring situation in which there is a seeming imbalance in gene dosage: the sex chromosomes in heterogametic species. Sex-chromosome dosage compensation is an essential process that evolved to regulate the levels of transcription of X-linked genes in heterogametic species [6–8]. In the eutherian mammals (humans and mice) and *Drosophila*, the male is heterogametic (XY) and the female is homogametic (XX) [1]. In contrast, *Caenorhabditis elegans* (*C. elegans*) hermaphrodites are XX and males are XY. In mammals, dosage compensation occurs by silencing one of the two female X chromosomes [9]. In *Drosophila*, the genes along the length of the single male X-chromososome are upregulated two-fold to equalize transcript levels between males and females [10,11]. In *C. elegans*, the levels of transcription from each of the two X chromosomes in hermaphrodites are downregulated two-fold [12]. Despite these diverse mechanisms, there are commonalities including nucleation of dosage compensation on the X-chromosome followed by spreading along its length [13–15] (*Drosophila*); [16] (mammals); [17] (worm). Furthermore, recent work has revealed a key commonality among dosage compensation mechanisms in all organisms, which is the mechanism by which transcription levels of genes on the single active X-chromosome are equalized to those on autosomes: the acetylation of histone H4 at lysine 16 [18,19].

In mammals, *C. elegans* and *Drosophila*, the process of dosage compensation is essential for proper development [10,12,20,21]. In humans, imbalance in sex chromosome number has strong phenotypic consequences including infertility [22]. Humans that lack one X chromosome (XO) develop Turner syndrome and show symptoms of ovarian failure, infertility, short stature, and similar congenital malformations [23]. Males with an extra X chromosome (XXY), termed Klinefelter syndrome, are sterile [22]. Similarly males with two Y chromosomes (YYX), Double Y Syndrome, are sterile if the extra Y is not lost early in sperm cell development. In addition, neurons are very sensitive to gene dosage and autism and schizophrenia have been linked to deletions of a single copy of several genetic loci that each contain many genes [9]. In *C. elegans* and *Drosophila*, dosage compensation is essential for viability.

2. Sex determination

The process of sex determination occurs upstream of dosage compensation and is directly linked to dosage compensation in *Drosophila* and *C. elegans*. In contrast, in mammals, it is not known how dosage compensation is linked to sex-determination. Master regulators such as *Sry* in mammals, *sex-lethal* (*sxl*) in *Drosophila*, *xol* in *C. elegans*, and *Dmrt1* conserved in flies, mammals, and worms regulate the process

☆ This article is part of a Special Issue entitled: Chromatin and epigenetic regulation of animal development.

* Corresponding author. Tel.: +1 401 863 1070; fax: +1 401 863 1201.
E-mail address: Erica_Larschan@Brown.edu (E. Larschan).

1874-9399/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagrm.2013.12.007

of sex determination. Because sex determination has been reviewed elsewhere [24–27], we will provide only a brief summary below.

In mammals the *Sry* gene on the Y-chromosome encodes the SRY protein that determines the male fate by acting on the Sox9 gene. A positive feedback mechanism ensures continual expression of Sox9 leading to testis differentiation. In the absence of SRY or temporal mis-expression of SRY, Sox9 is silenced and follicle cell and ovary formation occurs.

In *Drosophila*, the ratio of the sex chromosomes to autosomes determines sex. The gene dosage of the *sisterless* genes that are encoded on the X-chromosome results in the differential regulation of the master regulator of sex determination, the *sex-lethal* gene. *sex-lethal* (*sxl*) determines sexual fate by using different promoters to produce sex-specific products at precise developmental stages, and the SXL protein to control alternative splicing of the female specific *tra* protein. The Tra protein, in turn, regulates sexual dimorphism [26]. Once initiated, the female fate is continually maintained through autoregulation of the *sxl* gene by the SXL protein. Furthermore, in females SXL translationally represses the *male-specific lethal-2* (MSL2) protein, a core component of the male specific lethal (MSL) complex that upregulates expression of the single X chromosome in males [13]. In this way, sex determination is linked to dosage compensation.

Like *Drosophila*, in *Caenorhabditis elegans* (*C. elegans*), it is the ratio of sex chromosomes to autosomes that determines sex. The ratio of 2X:2A and 3X:4A defines the hermaphrodite, while the male is defined by 1X:2A and 2X:3A ratios [27]. The key sex-determining factor is the XO-lethal gene (*xol*), which activates the cascade for the male fate. In hermaphrodites (2X:2A), *xol* is repressed by X-signal elements (XSEs) encoded on the X chromosome. In hermaphrodites, repression of *xol* allows activation of the sex determination and dosage compensation complex (SDC-2), which is involved both in establishing hermaphrodite sexual development as well as dosage compensating the XX state [27,28].

Dmrt1, a conserved transcription factor in worms, flies, and mammals, is critical for maintaining the male fate. In mice, *Dmrt1* inhibits activation of the female fate genetic network through suppression of key female fate factors such as the FOXL2 transcription factor and WNT4/β-catenin [24]. In the absence of Dmrt1 and upregulation of female fate genes, male specific cells such as the Sertoli cells, differentiate into female-specific granulosa cells producing estrogen [24]. Subsequent to the initial steps of the sex determination process, dosage compensation mechanisms begin.

3. Identification of the X-chromosome during dosage compensation

Due to the haploinsufficiency of many loci throughout the genome, the cell is sensitive to dosage of a large number of genes including those that encode critical transcription and translation factors [1,5]. However, there is a widespread naturally occurring haploinsufficiency in all heterogametic species: genes that are encoded on the sex chromosomes. Dosage compensation mechanisms have evolved to specifically distinguish the X chromosome from autosomes for subsequent transcriptional modulation. X chromosome identification involves the following factors: 1) long non-coding RNAs that are likely to serve as nucleation centers [29–34]; 2) DNA sequence elements; 3) chromatin marks that alter packed chromatin; and 4) proteins that read and write the chromatin marks [30,35–39]. Below we discuss how each of these factors contributes to specifically target the dosage compensation machinery to the X-chromosome.

3.1. Long noncoding RNAs that function during dosage compensation in mammals and flies

Sex chromosome dosage compensation in mammals, flies, and worms is achieved either by upregulation or downregulation of a single or two sex chromosomes. While there are differences between the dosage compensation mechanisms among these species, the establishment of nucleation centers is required to initially identify the whole X-chromosome for dosage compensation. In mammals and flies, a key structural scaffold of these nucleation centers is the use of long noncoding RNAs (lncRNAs) as targeting platforms [40,41].

lncRNAs, are defined as RNAs that are larger than 200 nucleotides and do not code for functional protein [30,31,42]. The ability for the secondary structure and location of lncRNAs to be preserved in the absence of sequence conservation [30,33] makes them attractive targeting platforms because this is likely to decrease the frequency of spatial temporal error in the establishment of a domain of coordinate gene regulation. lncRNAs also have the capacity to bring different factors in close proximity [30]. Furthermore, lncRNAs can function in *cis* and *trans*, interact with multiple proteins, and shuttle between the nucleus and cytoplasm [30,42,43]. For example, the *roX* lncRNAs (RNA on X) can induce transcription of nearby genes when ectopically inserted onto autosomes [44,45]. Using the transcription of lncRNAs as a means of regulating expression of other genes is also supported by the observation that promoters of lncRNAs are the regions of lncRNAs under the greatest selective pressure [30,33]. In summary, lncRNAs are an ideal system by which to nucleate chromatin domains in a spatially and temporally regulated manner (Fig. 1).

The dosage compensation system in mammals uses lncRNAs to control X-chromosome identification. In mammals, dosage compensation is achieved by silencing one of two X chromosomes in female (XX), a process called X-chromosome Inactivation (XCI) [35]. XCI is controlled by the X inactivation center (Xic), which is a 100–500 kb region found on the chromosome to be silenced, Xi, and the active chromosome, Xa [9]. The Xic is a rich source of lncRNAs that are involved in X chromosome choice, counting, pairing, and silencing [9]. Xic regulation involves an intricate network of RNA and protein interactions. In the first steps of XCI, the X-specific-inactive-transcript (*Xist*) is transcribed only from the Xic on Xi into a 17–20 kb noncoding RNA and coats Xi in *cis*. Silencing of Xi requires *Xist* expression as *Xist* utilizes its conserved repeat motif (Rep A) to bind directly to Polycomb repressive complex 2 (PRC2) to target PRC2 to Xi. PRC2 deposits the silencing mark: trimethylation of histone H3 at lysine 27 (H3K27me3), facilitating chromatin compaction of DNA from active transcription [35]. *Xist* function however is regulated by its antisense transcript, *Tsix* another lncRNA. *Xist* upregulation requires *Tsix* downregulation [46]. *Tsix* inhibits *Xist* by silencing *Xist* through recruitment of a DNA methyltransferase (Dnmt3a) and inhibiting *Xist* binding to the PRC2 repressor protein [47,35,9]. In addition to *Tsix*, other lncRNAs encoded at the Xic repress *Xist* function. These include the X-inactivation intergenic transcription element (*Xite*) and the testis-specific X-linked gene (*Tsx*). A recently identified lncRNA called *Linx* (large intervening transcript in the Xic) may also have a role in *Tsix* regulation via modulating *Xite* [41]. There are also a number of lncRNAs encoded at the Xic that activate *Xist*. These include the RepA RNA and the *Jpx* RNA that can function both in *cis* and *trans* to activate *Xist* [46]. Overall, an intricate network of inter-dependent lncRNAs nucleates mammalian dosage compensation.

Drosophila sex chromosome dosage compensation also utilizes lncRNAs to establish local neighborhoods of regulation. In contrast to mammals, *Drosophila* sex chromosome dosage compensation is achieved by upregulating transcription of the single X-chromosome two-fold in males [11,48]. The male specific lethal (MSL) complex, which releases paused RNA polymerase into active elongation, mediates transcriptional upregulation by increasing histone H4 acetylation [49,50]. As described earlier, the MSL2 core complex component is translationally repressed in females by SXL and is therefore only expressed in males [13,51]. The MSL complex is a ribonucleoprotein complex composed of five proteins: MSL1, MSL2, MSL3 (male-specific lethal 1, 2, and 3), MOF (males absent on the first), MLE (maleless), and two lncRNAs: *roX1* and *roX2*, which are required for dosage compensation [10,52,53].

Although *roX1* and *roX2* differ in sequence and size (*roX1* is 3.7 kb and *roX2* is 0.5–1.2 kb), they are functionally redundant [54–56]. The

Please cite this article as: J. Chery, E. Larschan, X-marks the spot: X-chromosome identification during dosage compensation, Biochim. Biophys. Acta (2014), http://dx.doi.org/10.1016/j.bbagrm.2013.12.007