

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Investigation into the Molecular Mechanisms that Control Random X Chromosome Inactivation

Permalink

<https://escholarship.org/uc/item/7s0270gd>

Author

Royce-Tolland, Morgan Elizabeth

Publication Date

2009

Peer reviewed|Thesis/dissertation

Investigation into the Molecular Mechanisms that Control Random X Chromosome
Inactivation

by

Morgan Elizabeth Royce-Tolland

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Cell Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Copyright 2009
by
Morgan Elizabeth Royce-Tolland

For my Grandma

Acknowledgements

I must start by thanking my thesis advisor, Barbara Panning, for her support of me and my research over the past 7 years. I am especially grateful to Barbara for nurturing my life outside of lab in addition to my successes in it.

My thesis committee members, Hiten Madhani and Richard Locksley, have been instrumental to my growth as a scientist and I thank them for their genuine interest in my career and wellbeing. David Page was my first research advisor and I am appreciative to him and to my undergraduate mentor, Julie Bradley, for exposing me to the joys of scientific discovery.

The members of the Panning lab came from diverse scientific backgrounds to create a stimulating work environment. I thank all of them for sharing their time, expertise, ideas and reagents. I would especially like to acknowledge Angela for pioneering the A-repeat project, Susanna and Mary Kate for their contributions to the SIAR story, Katie for being a good friend through the ups and downs of graduate school and Jason for being a bottomless source of knowledge and ridicule.

Graduate school is a trying experience and I appreciate the friends who have brought levity to my life over the past 8 years. Erika, Andrew, Reba, Dani, Jonathan, Dennis, Jeannie, Sonya, Heather and Liz E—I don't what I would have done without you. Rich, Jon and Ben— creating art with you boys is always an experience. Wendy—I am comforted to know that no matter where life takes me, you will always be there.

Finally, I am thankful to my entire family for their unconditional love and encouragement. My parents have been my biggest cheerleaders my whole life and I thank them for reminding me to smile and to keep life in perspective.

Susanna Mlynarczyk-Evans, Mary Kate Alexander, Angela Andersen, Hannah R. Koyfman and Barbara Panning have all contributed to data in this thesis. The first chapter of this work and the figures contained within it were originally published in PLoS Biology (Mlynarczyk-Evans et al. 2006). The original publication is available at <http://biology.plosjournals.org/perlserv/?request=get-document&doi=10.1371/journal.pbio.0040159>

Investigation into the Molecular Mechanisms that Control Random X Chromosome Inactivation

Abstract

In eutherian mammals, dosage compensation between XX females and XY males is achieved by the transcriptional silencing of one X-chromosome in each female cell. The choice of which chromosome will be silenced is random in that both X chromosomes have an equal probability of being silenced. Despite nearly fifty years of research, the mechanisms that enable a cell to designate precisely one active and one inactive X chromosome remain elusive. What follows is an investigation into how X chromosomal fate is assigned. I present evidence that the two X chromosomes in a female cell adopt distinct states and that these states correlate with fate. I also explore the role of *Xist*'s A-repeat in this process. The A-repeat is a highly conserved *Xist* element that is necessary for Xist-dependent silencing. The work within this thesis shows that the A-repeat is also required for random choice. I demonstrate that the A-repeat is important for post-transcriptional processing of Xist RNA and that the A-repeat binds the essential splicing factor ASF/SF2. In combination, these findings provide the foundation for a model in which regulation of Xist RNA splicing in ES cells is part of the stochastic process that determines which X will be inactivated in wild-type cells.

Table of Contents

	Page
Dedication	iii
Acknowledgments	iv
Abstract	vi
List of Tables	viii
List of Figures and Illustrations	viii
General Introduction	1
Chapter 1	5
<i>X chromosomes alternate between two states prior to random X-inactivation</i>	
Chapter 2	45
<i>The A-repeat regulates Xist RNA processing and is required for random X-inactivation</i>	
Concluding Remarks	94
Appendix	98
References	100

List of Tables

	page
Table 1. BACs used as templates for FISH probes	33
Table 2. Primers used in Chapter 1	34
Table 3. Primers used in Chapter 2	75

List of Figures

Figure 1. X chromosomal loci display a high proportion of SD FISH signals in female ES cells.	35
Figure 2. SD FISH signals at X-chromosomal loci are independent of asynchronous DNA replication in ES cells.	36
Figure 3. X Chromosomes differ from one another in ES cells.	39
Figure 4. The future X _a and future X _i exhibit distinct frequencies of singlet FISH signals.	40
Figure 5. X chromosomes switch between states.	41
Figure 6. SIAR is specific to pluripotent cells <i>in vitro</i> and <i>in vivo</i> .	42
Figure 7. Models for achieving randomness in X-inactivation.	43
Figure 8. Features of SIAR are not revealed in MeOH samples.	44
Figure 9. Targeted deletion of the A-repeat in female ES cells.	76
Figure 10. X ^{AA} X cells undergo primary non-random X-inactivation.	77
Figure 11. X ^{AA} X cells do not show allelic enrichment for histone modifications.	80
Figure 12. Reduced levels of correctly spliced Xist transcripts in X ^{AA} Y cells.	82
Figure 13. The A-repeat binds ASF/SF2.	88

Figure 14. Binding of ASF/SF2 to the A-repeat depends on ASF/SF2 consensus sequences and not secondary structure.	91
Figure 15. ChIP analysis of the H3K27me3 hotspot.	93

General Introduction

Sex chromosomes have fascinated scientists since their discovery over a century ago. These chromosomes provide the genetic basis for sex determination and, unlike their autosomal counterparts, do not form a matched pair. In most cases, one sex chromosome is larger and gene-rich while the other sex chromosome is atrophied and gene-poor. This difference in size and gene content produces an imbalance between homogametic and heterogametic sexes. A variety of mechanisms are used by organisms to compensate for chromosome imbalances between sexes. *Drosophila melanogaster* males up-regulate gene expression two-fold from their single X chromosome to equalize gene expression with XX females (Lucchesi et al. 2005). *Caenorhabditis elegans* XX hermaphrodites down-regulate expression from both X chromosomes by 50% to balance expression with males that only have one X chromosome (Meyer 2000). Mammals use a third strategy and silence one X chromosome per female cell (Lucchesi et al. 2005). This mechanism acts to balance gene dosage between females that have two X chromosomes and males that have one X chromosome and one Y chromosome. It also introduces a feature that is unique to mammalian dosage compensation: the unequal treatment of X chromosomes within the same nucleus. How distinct expression profiles are established and maintained on the active and inactive X chromosomes (X_a and X_i) has been the focus of much research over the past fifty years.

X-inactivation occurs in two stages during mouse development (Goto and Monk 1998). Initially, all cells undergo imprinted X-inactivation to silence the paternally derived X chromosome (X_p) (Takagi and Sasaki 1975). As only female progeny receive an X_p, this method ensures that silencing occurs only in XX females and not in XY

males. In the cells that go on to form the extraembryonic tissues, Xp silencing is maintained. In the cells of the inner cell mass (ICM), which give rise to the embryo proper, Xp silencing is reversed (Okamoto et al. 2004). These cells then undergo a second wave of X-inactivation that coincides with the transition of the pluripotent ICM cells into more developmentally restricted lineages (Panning et al. 1997; Sheardown et al. 1997). This second round of X-inactivation differs from the initial silencing in that it is random-- either the maternally or paternally derived X-chromosome can be inactivated. Random X-inactivation also occurs upon differentiation of ICM-derived embryonic stem (ES) cells, making these cells a useful model system in which to study this process (Martin et al. 1978).

Both imprinted and random X-inactivation are initiated from a genetically-defined region on the X chromosome, the X-inactivation center (*Xic*). This region contains a pair of antisense genes, *Xist* and *Tsix*. The *Xist* gene encodes a 17kb spliced, polyadenylated non-coding nuclear RNA that is expressed from both X chromosomes in undifferentiated cells and exclusively from the Xi in differentiated cells. This transcript is both necessary and sufficient for transcriptional silencing of *cis*-linked sequences. *Xist*'s anti-sense partner, *Tsix*, also produces a non-coding RNA that is expressed from both X chromosomes in undifferentiated cells and functions to negatively regulate *Xist*.

Upon differentiation, *Xist* expression is upregulated on one X chromosome. On this chromosome, *Xist* RNA spreads from its site of transcription and triggers a cascade of epigenetic modifications including an increase in DNA methylation, hypoacetylation of histones H2A, H3 and H4, hypomethylation of histone H3K4, hypermethylation of H3K9, H3K27 and H4K20, ubiquitination of histone H2A and enrichment of the histone

variant macroH2A (Cohen et al. 2005). It has been proposed that Xist mediates silencing by coordinating chromatin modifying activities, however it appears that at least a subset of these chromatin alterations may not participate in transcriptional inactivation of the Xi. Instead, they likely act synergistically to maintain its silent state. Illustrating this point, mutant Xist RNA lacking a highly conserved element, the A-repeat, can coat *cis*-linked sequences and recruit both the H3K27 methyltransferase, Ezh2, and H3-3mK27 without causing silencing (Chaumeil et al. 2006).

The mechanisms that initially designate one Xi and one Xa per female cell are not understood. There is evidence that an epigenetic mark is established during oogenesis on the Xm which protects this chromosome from silencing during imprinted X-inactivation (Huynh and Lee 2001; Plath et al. 2002). How X chromosome fates are assigned during random X-inactivation is harder to explain. Specifically, how does a female cell ensure that the two X chromosomes take on mutually exclusive fates as the Xa and Xi? And what is the basis of randomness that allows for the maternally inherited X chromosome to be silenced in half of the cells of a population and for the paternally inherited X chromosome to be silenced in the remaining half? Analyses of *Xist* and *Tsix* mutant cell lines indicate that these genes are important for stochasticity during random X-inactivation as mutation of either *Xist* or *Tsix* in female cells results in non-random X-inactivation. In *Xist* mutant cells, the wild-type X chromosome is always silenced (Marahrens et al. 1998; Gribnau et al. 2005) and in *Tsix* mutant cells, the mutant X chromosome is always silenced (Lee et al. 1999; Luikenhuis et al. 2001; Sado et al. 2001). It is worth noting that in heterozygous *Xist* or *Tsix* mutants the two X chromosomes adopt opposite fates. So, while these genes are important for randomly

assigning X chromosome fates, they are not necessary to ensure that one X chromosome is silenced per female nucleus.

The following study is an investigation into how X_a and X_i fates are assigned during random X-inactivation. In chapter 1, I present evidence that prior to X-inactivation, the future X_a and future X_i chromosomes in a female cell differ from one another and that this difference correlates with their fates. In chapter 2, I explore the functions of the A-repeat. I show that deletion of the A-repeat from one X chromosome in female ES cells causes these cells to undergo non-random X-inactivation. I also demonstrate that the A-repeat is necessary for post-transcriptional processing of Xist RNA and binds the splicing factor ASF/SF2. These results suggest that Xist RNA metabolism may be an integral step in establishing a difference between the two X chromosomes. I finish with some future directions for this work, including several experiments to test aspects of the model presented in chapter 2.

Chapter 1

**X chromosomes alternate between two states
prior to random X-inactivation**

I. Summary

Early in the development of female mammals, one of the two X chromosomes is silenced in half of cells and the other X chromosome is silenced in the remaining half. The basis of this apparent randomness is not understood. We show that before X-inactivation, the two X chromosomes appear to exist in distinct states that correspond to their fates as the active and inactive X chromosomes. Xist and Tsix, noncoding RNAs that control X chromosome fates upon X-inactivation, also determine the states of the X chromosomes prior to X-inactivation. In wild-type ES cells, X chromosomes switch between states; among the progeny of a single cell, a given X chromosome exhibits each state with equal frequency. We propose a model in which the concerted switching of homologous X chromosomes between mutually exclusive future active and future inactive states provides the basis for the apparently random silencing of one X chromosome in female cells.

II. Introduction

At least 10% of mammalian genes are transcribed from only one allele that, in most instances, is chosen at random (Singh et al. 2003). The mechanisms for achieving differential regulation of homologous alleles in a stochastic manner are poorly understood (Ohlsson et al. 1998). X-chromosome inactivation in mammals is an example of random monoallelic expression; the majority of genes on one X chromosome in XX females is silenced to equalize expression of X-linked genes with XY males (Lyon 1961). Understanding the process by which the two X chromosomes are assigned active and

inactive fates in a stochastic manner will provide insight into how randomness can be achieved in other biological contexts.

X-inactivation is a well-defined system for studying the mechanisms that generate the randomness of monoallelic expression. In the earliest stages of female embryogenesis, X-linked genes exhibit biallelic expression in female cells. Upon receipt of a developmental cue, one X chromosome is silenced in half of the cells of the embryo and the other X chromosome is silenced in the remaining half. Once the identities of the active and inactive X chromosomes (X_a and X_i) are established, they are stably propagated throughout all subsequent cell divisions (Avner and Heard 2001). Female mouse embryonic stem (ES) cells provide a model system to study the mechanisms that control the initial, stochastic determination of the X_a and X_i *in vitro*. As in the pluripotent cells of the early embryo, genes are expressed from both X chromosomes in female ES cells. X-inactivation can be induced *in vitro*, recapitulating the random silencing process that occurs *in vivo* during differentiation (Martin et al. 1978). In addition, genetic elements have been identified that affect the randomness of X-inactivation (Lee 2002; Ogawa and Lee 2003) making female ES cells a useful tool for the study of the mechanisms that control random monoallelic expression.

The *X-inactivation center* (X_{ic}) is a master *cis*-regulatory element on the X chromosome that controls X-inactivation (Avner and Heard 2001). The X_{ic} contains a number of elements, including X_{ist} and T_{six} , a sense/antisense pair of noncoding RNAs that are transcribed from both X chromosomes prior to X-inactivation (Panning and Jaenisch 1996; Sheardown et al. 1997; Lee et al. 1999). When embryonic cells differentiate and X-inactivation is initiated, X_{ist} RNA spreads *in cis* from the X_{ic} to coat

and silence one X chromosome (Panning and Jaenisch 1996; Sheardown et al. 1997). *Xist* and *Tsix* play a role in the stochastic determination of which X chromosome will become the Xi and which will become the Xa. In cells heterozygous for either an *Xist* or a *Tsix* mutation, X-inactivation is nonrandom: an *Xist* mutant chromosome always becomes the Xa (Marahrens et al. 1998; Gribnau et al. 2005) and a *Tsix* mutant chromosome always becomes the Xi (Lee and Lu 1999; Luikenhuis et al. 2001; Sado et al. 2001).

Heterozygous mutations of *Xist* or *Tsix* also exhibit effects *in trans*: not only are the fates of the mutant X chromosomes fixed but also the fates of the wild-type X chromosomes are determined. Thus, information about the fate of one X chromosome must be transmitted to the other X chromosome, ensuring that decisions about their fates are coordinated and that there is random and exclusive silencing of one X chromosome in female cells. The manner in which the opposing activities of *Xist* and *Tsix* regulate the fates of both X chromosomes in female cells remains mysterious.

In this paper, we present evidence that in individual pluripotent embryonic cells that are poised for X-inactivation, the X chromosomes exist in two mutually exclusive states. In heterozygous *Xist* and *Tsix* mutant cells, these states predict the fates of the X chromosomes, indicating that one X chromosome adopts a future Xa state and the other X chromosome adopts a future Xi state. In wild-type cells, X chromosomes switch between these states such that, among the progeny of a single cell, a given X chromosome exhibits each state with equal frequency. Thus, the concerted switching of homologous X chromosomes between mutually exclusive future Xa and future Xi states may provide the basis for the apparent randomness of X-inactivation.

III. Materials and Methods

Cell lines and culture

Wild-type female ES cells (Marahrens et al. 1997), $\Delta Xist/+$ ES cells (Csankovszki et al. 1999; Gribnau et al. 2005), *Tsix-pA/+* ES cells (Luikenhuis et al. 2001), X.2/Xwt ES cells (Tada et al. 1993), and MEFs were cultured according to standard practice. Wild-type and $\Delta Xist/+$ (Csankovszki et al. 1999) 3.5-d postconception blastocysts were harvested by standard procedures and cultured overnight in ES medium without LIF. To identify cells in S phase, cells were cultured with BrdU (Amersham Biosciences, Little Chalfont, United Kingdom) for 15 to 30 min.

Cell cycle fractionation

Flow cytometry was performed on live ES cells labeled with BrdU (Amersham) and stained with 40 μ g/ml Hoechst 33342 (Molecular Probes, Eugene, OR, USA) for 45 min prior to harvesting. Cells were resuspended in ES medium plus Hoechst, 7% Cell Dissociation Buffer (GIBCO, San Diego, CA, USA), and 10 mM EDTA for sorting and were cooled during the procedure. Cytometry was performed using a FACSDiVa Cell Sorter (Becton-Dickinson, Palo Alto, CA, USA). Cell cycle profiles were generated by excitation with a violet laser; Hoechst emission was measured with a HQ445/50 bandpass filter (Chroma Technology, Rockingham, VT, USA). Cells were gated to exclude debris and double cells. For microscopy, fractions were sorted into PBS in multiwell slides pretreated with 1 mg/ml poly-L-lysine and allowed to settle and adhere.

Sample preparation

ES cells and MEFs were fixed for FISH using PFA (Marahrens et al. 1998) or MeOH (Gribnau et al. 2003). Blastocysts were treated with acid tyrode to remove zona pellucidae, applied to 2% gelatin-coated slides using a Cytospin apparatus (Shandon, Pittsburgh, PA, USA), PFA-fixed for 10 min, and permeabilized with a 5-min incubation in PBS plus 0.5% Triton X-100.

FISH

BACs used for genomic probes are listed in Table 1. All pairwise DNA FISH was performed on loci separated by 40 Mb or less; linked sequences can be reliably scored as being on the same chromosome over distances of up to 50 Mb (Ensminger and Chess 2004). A collection of PCR products spanning exon 1 was used to make *Xist* probes. Probes were generated using a BioPrime kit (Invitrogen, Carlsbad, CA, USA), or using cy3-dCTP or FITC-dUTP (Amersham; Enzo Life Sciences, Farmingdale, NY, USA) with kit reagents. Strand-specific probes to detect *Xist* and *Tsix* RNA were generated by *in vitro* transcription with FITC-UTP or bio-CTP (Enzo Life Sciences) from an *Xist* exon 7 template.

FISH for genomic DNA was performed as described (Gribnau et al. 2003). Biotinylated probes were detected with FITC-avidin (Vector Laboratories, Burlingame, CA, USA) or cy3-streptavidin (Amersham). Combined DNA and RNA FISH was performed as described (van Raamsdonk and Tilghman 2001), with the addition of a pepsin

pretreatment prior to the initial step (Gribnau et al. 2003). BrdU detection was performed as described (Gribnau et al. 2003) using mouse monoclonal α -BrdU antibody (Becton Dickinson) and α -mouse FITC (Vector Laboratories).

Samples were scored on an Olympus BX60 microscope. Images were collected with a Hamamatsu ORCA-ER digital camera using Openlab 4.0.1 software, assembled using Adobe Photoshop 7.0, and levels adjusted to enhance contrast. ICM cells of blastocysts were scored from three-dimensional images collected using a DeltaVision system as described below.

FISH signal intensity quantification

Images were collected as 0.1- μ m optical section stacks using an Olympus IX70 microscope with a motorized stage controlled by DeltaVision 2.10 software (Applied Precision, LLC, Issaquah, WA, USA) and a MicroMax CCD camera (Roper Scientific, Tucson, AZ, USA). SoftWorx 2.50 software was employed to deconvolve three-dimensional images and sum pixel intensity through relevant sections of image stacks to generate two-dimensional projections representing total intensity of each FISH signal. The software was allowed to delineate the signal circumference and to integrate pixel intensities to generate an overall intensity value in arbitrary units.

Mimosine arrest and release replication timing assay

S-phase time-course experiments were performed using a mimosine arrest-release protocol (Gribnau et al. 2003). DNA from each time point was isolated and sonicated as described (Hansen et al. 1993). BrdU-labeled human DNA (0.5 μ g) was mixed with 10

μg of DNA from each time point for normalization. Labeled DNA was immunopurified using an α-BrdU monoclonal antibody (Becton Dickinson) and Protein G Sepharose 4 Fast Flow beads (Amersham) and resuspended in 1 ml of 1 mM Tris, 0.1 mM EDTA (pH 8.0). PCR primers are listed in Table 2. For FACS analysis of mimosine arrest/release time points, samples were fixed overnight in 70% ethanol, treated with 0.2 μg/ml RNase, stained with 20 μg/ml propidium iodide (Molecular Probes), and analyzed using a FACSCalibur (Becton Dickinson).

Statistics

All p-values were determined by comparing the observed distribution of signal patterns at each allele to a random, 50/50 distribution (null hypothesis) using a χ^2 distribution test with one degree of freedom.

IV. Results

X-Chromosomal Loci Show a High Frequency of Singlet/Doublet Fluorescence In Situ Hybridization Signals in ES Cells

While using fluorescence *in situ* hybridization (FISH) to visualize the *Xic* in female ES cells fixed with paraformaldehyde (PFA), we observed that a high frequency of cells displayed a single pinpoint FISH signal at one allele and a double pinpoint at the other. To determine whether this feature was unique to the *Xic*, we analyzed a number of other X-chromosomal loci. Cells were scored as showing a singlet signal for each allele (SS), a doublet signal for each allele (DD), or a pattern in which one allele appeared as a singlet and the other as a doublet (SD) (Figure 1A). For all X-chromosomal loci examined, the

SD signal class was the most abundant, comprising 40% to 50% of the population (Figures 1B and 1C). This proportion is significantly greater than the fraction of SD signals observed for two autosomal loci, which exhibited the SD pattern in fewer than 20% of cells (Figures 1B and 1C). Thus, the high proportion of cells displaying SD signals is a unique feature of X-linked sequences.

A high frequency of SD FISH signals can be indicative of asynchronous replication of the two alleles, as singlet and doublet signals can reflect unreplicated and replicated loci, respectively (Selig et al. 1992). However, a singlet FISH signal can also occur at a replicated locus (Azuara et al. 2003). To determine whether the high proportion of cells exhibiting SD signals was due to asynchronous replication, we directly measured the replication timing of X-linked sequences. ES cells were blocked in G1 and released into S phase (Figure 2A). At hourly intervals, cells were pulsed with 5-bromo-2-deoxyuridine (BrdU) to label replicating DNA. BrdU-containing DNA was immunopurified from each time point and assayed by PCR for X-linked sequences. The *Xic* and *Pgk1* each showed a single peak of BrdU incorporation early in S phase (Figure 2B). Therefore, neither X-linked locus was subject to highly asynchronous DNA replication. We next performed FISH for the *Xic* and *Pgk1* in female ES cells sorted by DNA content (Figure 2C). The proportion of cells exhibiting SD signals increased at the beginning of S phase, remained fairly constant throughout S phase, and decreased at the end of S phase (Figure 2D). Taken together, these data demonstrate that even though both alleles of each of these X-chromosomal loci replicated early, SD signals persisted throughout S phase. Furthermore, they suggest that the singlets in cells exhibiting SD FISH signals are replicated alleles.

If a singlet signal in a cell showing an SD signal pattern represents a replicated allele, then it should contain the same amount of DNA as the two pinpoints comprising the doublet signal. To test this hypothesis, we determined the relative fluorescence intensity of singlet and doublet signals for the Xic and Pgk1. We validated this assay on X-linked sequences in female fibroblasts, in which singlet FISH signals correspond to unreplicated loci and doublet FISH signals correspond to replicated alleles (Gartler et al. 1999). In female mouse embryo fibroblasts (MEFs) exhibiting DD signals for Pgk1, the distribution of D/D intensity ratios centered around one (Figure 2E), demonstrating that two replicated alleles display equal signal intensities. We next compared the intensity of the singlet to the sum of the two pinpoints in the doublet (S/D) in female MEFs exhibiting SD signals for Pgk1. The S/D intensity ratios centered around 0.5 (Figure 2E), indicating that relative fluorescence intensity can be used to measure a two-fold difference in DNA content. Quantification of FISH signals can therefore be used to assay differences in DNA content at individual loci in single cells. In ES cells, the ranges of S/D intensity ratios for both the Xic and Pgk1 were very similar to the ranges of D/D ratios: in both cases, the distributions centered approximately around one (Figure 2E). These results indicated that the singlet and doublet FISH signals in ES cells exhibiting an SD pattern for the Xic and Pgk1 contained the same amount of DNA. In combination, these analyses demonstrate that the unusually high proportion of ES cells displaying SD signals for X-chromosomal loci reflects something other than asynchronous DNA replication. We refer to the high frequency of PFA-fixed cells exhibiting SD signals that are Independent of Asynchronous DNA Replication as SIAR.

In a recent study, it was suggested that nuclear organization was important for replicated sequences to appear as singlet FISH signals (Azuara et al. 2003). To determine whether SIAR required an intact nucleus, we compared the proportion of ES cells exhibiting SD signals upon PFA fixation to that seen upon methanol:acetic acid (MeOH) fixation. While PFA fixation preserves the three-dimensional organization of the nucleus by cross-linking nucleic acids and proteins, MeOH fixation destroys nuclear architecture by extracting the bulk of histones and other chromatin proteins (Hendzel and Bazett-Jones 1997). The distribution of SS, SD, and DD signals for an autosomal gene, *Fnl*, did not differ significantly between fixation methods (Figure 2F). In contrast, these two fixation conditions resulted in different distributions of FISH signals for the X-chromosomal loci *Mecp2*, *Pgk1*, and the *Xic*. For these loci, the proportion of cells displaying SD signals decreased while the proportion of cells exhibiting DD signals increased in MeOH-fixed samples compared to PFA-fixed samples (Figure 2F). The changes in the relative proportions of cells exhibiting SD and DD signals indicate that when nuclear structure is disrupted, some replicated X-chromosomal loci that would appear as singlets in an intact nucleus resolve into doublets. In addition, these results suggest that native chromatin structure and/or nuclear organization is necessary for some replicated alleles of X-chromosomal loci to appear as singlets.

X Chromosomes Differ Prior to X-Inactivation

All X-chromosomal loci examined (Figure 3A) exhibited SIAR in female ES cells. We tested whether the singlet FISH signals for multiple loci appeared on the same chromosome or were randomly distributed between the two X chromosomes. Closely

linked loci were analyzed pairwise and, among the cells that exhibited SD signals for both probes, the proportion in which singlet signals occurred on the same chromosome (concordant signals) was determined (Figure 3B). *Ccnb3* and *Hprt*, *Hprt* and *Mecp2*, and *Mecp2* and *Pgkl* each exhibited approximately 65% concordant signals (Figure 3C). This 65% concordance is significantly higher than the 50% concordance that would be expected for a random distribution of singlet signals between chromosomes ($p < 0.02$), indicating that the behavior of loci on each X chromosome is coordinated. One X chromosome exhibits a higher frequency of singlet signals along its length, and the other X chromosome exhibits a higher frequency of doublet signals. These results suggest that, even though X-inactivation has not yet occurred and X-linked genes are biallelically expressed, the two X chromosomes in female ES cells already differ from each other.

When carrying out pairwise analysis of X-chromosomal loci, we found that the *Xic* was unusual in that it was oppositely coordinated with adjacent genes. *Mecp2* and the *Xic*, and the *Xic* and *Pgkl* exhibited a bias against concordant signals, displaying concordance in only 38% of cells scored (Figure 3C). The *Xic* contains *Xist*, which is unusual in that it is the only gene expressed exclusively from the Xi. The opposite behavior of the *Xic* relative to other X-chromosomal loci in ES cells therefore parallels the opposite expression patterns of *Xist* and X-linked genes after X-inactivation. This parallel suggests a relationship between the appearance of the X chromosomes by FISH prior to X-inactivation and their fates as the Xa and Xi.

Xist and Tsix Control SIAR and X Chromosome Fate

To determine if there was a correlation between the appearance of the X chromosomes by FISH prior to X-inactivation and their fates after X-inactivation, we analyzed SIAR in ES cell lines that will undergo nonrandom X-inactivation. If such a correlation exists, then the identities of the X chromosomes displaying singlet and doublet FISH signals should be nonrandom in ES cell lines that are poised for nonrandom X-inactivation. In ES cells heterozygous for an *Xist* mutant chromosome ($\Delta Xist/+$), the wild-type X chromosome is always silenced and the mutant chromosome always remains active upon X-inactivation (Csankovszki et al. 1999; Gribnau et al. 2005). In ES cells bearing a *Tsix* mutant chromosome (*Tsix-pA/+*), the wild-type X chromosome remains active and the mutant chromosome is inactivated upon X-inactivation (Luikenhuis et al. 2001).

We performed allele-specific FISH for X-chromosomal loci in $\Delta Xist/+$ ES cells (Figure 4A) or *Tsix-pA/+* ES cells (Figure 4B). The *Xic*, *Pgk1*, and *Mecp2* were scored individually in both cell lines. While the overall proportions of SS, SD, and DD signal patterns for these loci did not differ between wild-type ES cells and *Xist* and *Tsix* mutant ES cells (data not shown), the identities of the alleles displaying the singlet and doublet FISH signals in cells with SD signals were nonrandom in the mutant cell lines. In both $\Delta Xist/+$ and *Tsix-pA/+* cell lines, the X chromosome that will become the Xi exhibited singlet signals for the *Xic* at a high frequency (~70% of SD cells; Figure 4C) and singlet signals for *Pgk1* or *Mecp2* at a low frequency (~30% of SD cells; Figure 4C). The future Xa showed the opposite patterns (Figure 4C). These results demonstrate that *Xist* and *Tsix* mutations affect SIAR and that, prior to nonrandom X-inactivation, the future Xi and future Xa show different probabilities of exhibiting a singlet signal for the *Xic* and other X-chromosomal loci. X-chromosomal loci showed distinct frequencies of FISH signal

patterns on the future Xa and future Xi in *Xist* and *Tsix* mutant ES cells, supporting the idea that the two X chromosomes adopt distinct states prior to nonrandom X-inactivation. One X chromosome exists in a future Xi state, which causes singlet FISH signals to occur at a high frequency at the *Xic* and at a low frequency at other X-linked sequences. The second X chromosome adopts the future Xa state, which causes singlet FISH signals to occur at a low frequency at the *Xic* and at a high frequency at other X-chromosomal loci.

X Chromosomes Alternate between Two States prior to Random X-Inactivation

The coordination of SIAR in wild-type ES cells (Figure 4C) indicated that the two X chromosomes exist in two distinct states similar to those observed in *Xist* and *Tsix* mutants. We hypothesized that in cells poised for random X-inactivation, these X chromosome states might also be indicative of X chromosome fates. Two predictions arise from this hypothesis. First, each X chromosome in wild-type ES cells should exist in the future Xi state in the same proportion of cells in which that chromosome will be inactivated. Second, because the fate of each X chromosome is not yet fixed in wild-type ES cells, the state of each X chromosome must not be fixed either.

In an ES cell line that will undergo random X-inactivation, one X chromosome should exist in the future Xi state in half of the cells in the population and the other X chromosome should exist in the future Xa state in the remaining half. By FISH, an X-linked gene should appear as a singlet on one X chromosome in 50% of SD cells and on the other X chromosome in the remaining 50%. To test this prediction, we performed allele-specific FISH in X.2/Xwt ES cells, which are poised to undergo random X-inactivation (Tada et al. 1993). In these cells, one X chromosome is marked by a

centromeric fusion to Chromosome 2 (Tada et al. 1993). *Ccnb3*, an X-chromosomal locus that exhibited SIAR (Figure 1B) and is closely linked to the fusion point, was analyzed in combination with a probe proximal to the centromere of Chromosome 2, which identified the X.2 chromosome (Figure 5A). The X.2 chromosome and the wild-type X chromosome each exhibited a singlet signal for *Ccnb3* in approximately 50% of SD cells (Figure 5B), consistent with the marked X chromosome existing in the future Xi state in half of the cells and the wild-type X chromosome adopting this state in the remaining half.

X-inactivation is partially skewed in cells that are heterozygous at the *Xce*, an X-linked control element that influences randomness of X-inactivation (Cattanach 1975). We analyzed ES cell lines containing X chromosomes that carry different *Xce* alleles to determine whether the frequency with which a chromosome adopts each state correlates with the degree of skewing observed upon X-inactivation. In ES cells heterozygous for *M. musculus 129* and *M. castaneus Ei* X chromosomes (*129/cas*), the *129* X chromosome will be inactivated in approximately 80% of cells, and the *cas* X chromosome will be inactivated in the remaining 20% (Cattanach 1975; Ogawa and Lee 2003). Allele-specific FISH for the *Xic* (Figure 5C) was performed in two independent *129-tet/cas* ES cell lines in which the *Xic* allele on the *129* X chromosome is marked by a tet-operator array integration that does not disrupt the *Xce* effect (Gribnau et al. 2005). Based on the frequency with which the future Xi exhibited a singlet FISH signal in mutant cells destined for nonrandom X-inactivation, we calculated that in *129/cas* ES cells, the *Xic* on the *129* allele should appear as a singlet in approximately 62% of SD cells (data not shown). Consistent with this prediction, the singlet appeared on the *129* allele in 60% of

SD cells (Figure 5D). Together, analysis of ES cell lines that undergo completely random, completely nonrandom, and skewed X-inactivation suggests that there is a relationship between the state of an X chromosome in ES cells and its fate as the Xa or Xi.

The fates of the X chromosomes are not fixed in wild-type ES cells, as a population of ES cells derived from a single progenitor cell undergoes random X-inactivation upon differentiation (Penny et al. 1996; Gribnau et al. 2005). If the states of the X chromosomes in ES cells reflect their fates upon X-inactivation, the states should not be fixed in ES cells that undergo random or skewed X-inactivation. To test this hypothesis, we first analyzed four clonal derivatives of the X.2/Xwt ES cell line, each of which is subject to random X-inactivation (data not shown). In each clone, the X.2 chromosome exhibited a singlet signal for *Ccnb3* in approximately half of cells displaying SD signals (Figure 5B). Thus, each clone recapitulated the pattern observed in the parental cell line. In the two clonally derived *129-tet/cas* ES cell lines, the *Xic* on the *129* chromosome also appeared as a singlet in some cells exhibiting SD signals and a doublet in others (Figure 5D). Thus, the marked chromosome in each cell line, which would have existed in either the future Xa or the future Xi state in the founding cell of each clone, assumed the future Xa state in some daughter cells and the future Xi state in others. Therefore, the states of the X chromosomes cannot be fixed; rather, the two X chromosomes must switch between states in cycling ES cells.

SIAR Is Restricted to Cells Poised for Random X-Inactivation

If future Xi and future Xa states underlie random X-inactivation, then these states should be observed *in vivo*. We performed FISH for the *Xic* in blastocyst-stage female mouse embryos. Cells from the inner cell mass (ICM) of these embryos are poised for random X-inactivation (Sugawara et al. 1983). The percentage of ICM cells exhibiting SD signals for the *Xic* was comparable to that seen in ES cells (Figure 6A), suggesting that ICM cells also display SIAR. Allele-specific FISH for the *Xic* in $\Delta Xist/+$ blastocysts (Figure 6B) indicated that the wild-type X chromosome, which will become the Xi, more frequently exhibited a singlet signal, and the mutant X chromosome, which will become the Xa, more frequently appeared as a doublet (12 of 15 SD cells, $p = 0.02$). These observations suggest that, as in ES cells, the two *Xic* loci in cells of early embryos adopt distinct configurations that are regulated by *Xist* and correlate with fate. Thus, the chromosome states reflected by FISH may play a role in the initial random designation of the Xa and Xi *in vivo*.

After one X chromosome is silenced, it remains the Xi throughout all subsequent cell divisions. We analyzed differentiated cells to determine whether SIAR persists after X-inactivation is established. In MEFs, the X-chromosomal loci *Mecp2* and the *Xic* exhibited significantly reduced proportions of SD signals when compared to ES cells (Figures 6A and 6C), and the proportions of SD signals were not altered upon MeOH fixation (Figures 6D-F), indicating that MEFs do not exhibit SIAR. A second differentiated cell population, trophectoderm cells from blastocyst-stage embryos, showed a proportion of SD *Xic* signals comparable to that seen in MEFs and significantly

lower than that seen in ES cells (Figure 6A). Taken together, these observations indicate that SIAR occurs only in cells that are poised for X-inactivation. Furthermore, they suggest that SIAR does not reflect a mechanism that maintains the identity of the X_a and X_i in differentiated cells but instead reflects a process that is involved in the initial designation of X chromosome fates.

V. Discussion

In this study, we used FISH to demonstrate that the two X chromosomes in female mouse ES cells differ prior to X-inactivation. On one X chromosome, the *Xic* tended to appear as a singlet signal when assayed by FISH and other X-linked genes more often appeared as doublets. The second X chromosome was more likely to show the opposite pattern. In ES cell lines that are destined for nonrandom X-inactivation, the future X_i exhibited a high frequency of singlet signals for the *Xic* and of doublet signals for other X-chromosomal loci, while the future X_a showed the opposite pattern. Taken together, these data suggest that, prior to X-inactivation, the two X chromosomes in a female cell exist in distinct future X_i and future X_a states.

We propose that an absolute difference between the X chromosomes underlies the future X_i and future X_a states. The state of the X chromosome in turn affects the probability that a locus will appear as a singlet or doublet FISH signal. In *Tsix-pA/+* and $\Delta Xist/+$ ES cells, where a given X chromosome will become the X_i in 100% of cells, the *Xic* appeared as a singlet on that chromosome in 70% of SD cells. A simple explanation for this observation is that FISH reflects dynamic behavior of loci, and the future X_i and future X_a differ in their dynamics. For example, the *Xic* may fluctuate between appearing

as a singlet or doublet on both the future Xi and the future Xa, appearing more frequently as a singlet on the future Xi and more frequently as a doublet on the future Xa. In this model, the frequency with which a locus on the future Xi or Xa appears as a singlet or doublet in the population, and not its appearance in a single cell at any given point in time, reveals the underlying state of the chromosome.

In cell lines destined for random and completely nonrandom X-inactivation, the frequency with which a given X chromosome adopts the future Xi state correlates with the frequency with which it will be inactivated. The same relationship was observed in cell lines that will display skewed X-inactivation due to the *Xce* effect. Thus, the *Xce* effect is manifested prior to X-inactivation. This result is consistent with the suggestion that the *Xce* effect influences the initial assignment of X chromosome fates (Percec et al. 2002).

Our observation that X chromosomes adopt distinct future Xi and future Xa states suggests that X chromosomes know their fates prior to silencing. These states exist even in cell lines that will undergo random X-inactivation, raising the question of how randomness is achieved. In clonal populations of ES cells that will undergo random X-inactivation, a marked X chromosome showed a singlet signal for an X-linked gene in 50% of cells, suggesting that X chromosomes can switch states.

We propose a model in which switching of X chromosomes between mutually exclusive future Xi and future Xa states is the source of the apparent randomness of X-inactivation (Figure 7A). In each cell, one X chromosome adopts the future Xi state and the other X chromosome adopts the future Xa state. When a cell receives the cue to initiate X-inactivation, the chromosome that exists in the future Xi state in that cell will

be silenced. However, as long as that cell remains pluripotent, the states of the chromosomes are not fixed and can switch in a concerted fashion. Because of this switching, the two X chromosomes each assume the future Xi state in half of the cells in a population. This randomization of states provides the basis for silencing to occur in an apparently random manner when X-inactivation is triggered. In this model, the mechanisms that determine the randomness of X-inactivation function prior to the receipt of the differentiation signal that initiates X-inactivation. In contrast, the prevailing model for randomness of X-inactivation posits that the two X chromosomes are equivalent in pluripotent cells, and that designation of which chromosome will be silenced occurs stochastically upon receipt of the signal that triggers X-inactivation (Figure 7B) (Rastan 1983; Avner and Heard 2001). Instead, our data support a model similar to the class of models proposed by Williams and Wu (2004), who speculate that a switching-based mechanism, like that regulating mating-type switching in fission yeast, may underlie the randomness of X-inactivation.

X-chromosomal loci exhibited SD signals in a significant fraction of MeOH-fixed and of PFA-fixed ES cells. In a recent study, Gribnau et al. (2005) also observed a high frequency of SD FISH signals for X-chromosomal loci in MeOH-fixed ES cells. Based on the large fraction of cells exhibiting SD signals, it was suggested that X-linked genes are subject to asynchronous replication in ES cells. We directly assayed replication timing of two X-chromosomal loci, *Pgkl* and the *Xic*, in ES cells. Each exhibited a single peak of replication. Cells showing SD FISH signals for the *Xic* or *Pgkl* occurred at a high frequency even after both alleles of these two X-chromosomal loci had replicated. In addition, a FISH signal intensity quantification assay revealed that in individual ES cells,

the singlet and doublet signals for the *Xic* or for *Pgkl* contained comparable amounts of DNA (Figure 9A). Together, these data indicate that the high frequency of SD FISH signals for X-chromosomal loci in ES cells is not due to highly asynchronous DNA replication.

In PFA-fixed ES cells, the *Xic* was unusual in that it tended to exhibit a singlet FISH signal while other genes on the same X chromosome more often showed doublet signals. In contrast, the *Xic* and adjacent genes exhibited concordant FISH signals when ES cells were fixed with MeOH (Figure 8B) (Gribnau et al. 2005). Thus, the opposite behavior of the *Xic* was observed only in PFA-fixed ES cells. In addition, when ES cells were fixed with PFA, the future Xi showed a higher frequency of singlet signals for the *Xic* or of doublet signals for other X-linked genes. In contrast, when ES cells were fixed with MeOH, the future Xi and future Xa showed equal probabilities of exhibiting singlet signals for the *Xic* or other X-linked genes (Figure 8C) (Gribnau et al. 2005). PFA fixation maintains nuclear structure while MeOH extracts histones and other chromatin proteins and perturbs nuclear organization (Hendzel and Bazett-Jones 1997). Therefore, the differences between PFA- and MeOH-fixed samples suggest that some aspect of chromatin structure or nuclear organization is required for replicated alleles on the future Xi and future Xa to appear as singlets with different probabilities. One possible explanation is that there is differential cohesion between sister chromatids on the two X chromosomes. Whatever the physical basis of the singlet and doublet FISH signals, they demonstrate that the X chromosomes differ prior to X-inactivation in a manner that is predictive of Xi and Xa fates.

The *Xic* behaved oppositely to other loci on both the future Xi and the future Xa, mirroring the opposite expression patterns of *Xist* and most X-linked genes after X-inactivation and suggesting a correlation between the appearance of a locus by FISH and its future expression status. Azuara et al. (2003) reported a correlation between FISH signal appearance and current expression status. In that study, a replicated transgene tended to exhibit a doublet FISH signal when it was expressed and a singlet FISH signal when it was silenced. In our study, the replicated allele that more often appeared as a singlet signal in ES cells was the one that would be expressed after X-inactivation is triggered by differentiation. This result suggests that organization of sequences within the nucleus of a pluripotent embryonic cell may impact their expression after differentiation.

In wild-type ES cells, X chromosomes appear to switch between states in a concerted manner: when one X chromosome assumes the future Xi state, the other adopts the future Xa state. In heterozygous *Xist* and *Tsix* mutant ES cells, it appears that the X chromosomes no longer switch between states, suggesting that these noncoding RNAs are either required for switching to occur or affect the likelihood that a chromosome will adopt the future Xi or future Xa state each time a switch occurs. The probability that a chromosome will become the Xi (or the Xa) is determined by the opposing activities of *Xist* and *Tsix* on that chromosome (Plath et al. 2002). These observations suggest that *Xist* and *Tsix* influence the fates of both X chromosomes by determining how effectively each chromosome competes to adopt the future Xi (or future Xa) state prior to X-inactivation. Both *Xist* and *Tsix* RNAs have been implicated in the regulation of chromatin structure (Heard 2004; Navarro et al. 2005; Sado et al. 2005). Prior to X-inactivation, these noncoding RNAs can be detected exclusively at the *Xic* on both

transcriptionally active X chromosomes (Lee et al. 1999). Perhaps Xist and Tsix RNA mediate changes in chromatin structure at the Xic prior to X-inactivation. Such changes may in turn modulate the state of the entire X chromosome and direct its fate.

Our data show that prior to random X-inactivation, X chromosomal loci differ in a manner that is predictive of future expression status. In addition to X-linked genes, several thousand autosomal genes are also subject to random monoallelic expression. It has been suggested that random monoallelic loci on autosomes and the X chromosome may share regulatory features (Singh et al. 2003; Ensminger and Chess 2004). It will be interesting to determine whether a common mechanism of concerted switching between two states underlies the randomness of monoallelic expression throughout the genome.

VI. Figure Legends

Figure 1. X chromosomal loci display a high proportion of SD FISH signals in female ES cells. (A) FISH for Xic genomic sequences (red) demonstrates the three classes of signals in PFA-fixed female ES cells. DNA was stained with DAPI (blue). (B) ES cells display an elevated proportion of SD signals at X-chromosomal loci. Average data from two to four experiments ($n > 150$), scored by two independent scorers, are presented. Error bars indicate one standard deviation. (C) Proportions of PFA-fixed ES cells exhibiting SS (white), SD (black) and DD (grey) signals at autosomal (*Hba1*, *Fn1*) and X-chromosomal (*Ccnb3*, *Hprt*, *Mecp2*, *Xic*, *Pgk1*, *Grpr*) loci. Average data from the samples scored in Figure 1B are shown.

Figure 2. SD FISH signals at X-chromosomal loci are independent of asynchronous DNA replication in ES cells. (A) FACS analysis of female ES cell samples from a mimosine arrest/release timecourse. Samples collected at two-hour intervals, starting at 1 hour after washout of mimosine and continuing until 11 hours after washout were analyzed by propidium iodide (PI) staining for DNA content. Histograms illustrate release of the majority of cells from G1 arrest and synchrony of progression through S phase. Co-cultured replication –incompetent feeder cells exhibiting 2nDNA content contribute to the G1 peak in each sample. (B) Analysis of replication of X-linked loci in the samples collected in A. Samples were pulsed with BrdU at hourly intervals after release from G1 arrest and DNA was isolated. An equal amount of BrdU-labeled human DNA was added to each time point. BrdU-labeled DNA was immunoprecipitated, and sequences present in each fraction were assessed by PCR. Standard shows amplification of a human sequence, as a control for variability in immunoprecipitations. IgG represents PCR analysis from labeled DNA purified with mouse IgG instead of anti-BrdU antiserum. –BrdU indicates analysis of an anti-BrdU immunoprecipitation from unlabeled DNA. Pre-IP depicts PCR analysis of the input DNA. Analysis of *Xic* and *Pgkl* reveals a single peak of replication for each locus in female ES cells. (C) Live female ES cells, pulse-labeled with BrdU, were sorted into six fractions by Hoechst staining for DNA content. (D) FISH for the *Xic* (upper panel), and *Pgkl* (lower panel) in the fractions delineated in C shows constant, high proportions of SD signals (red triangles) throughout S-phase. Proportions of SS (black circles) and DD (gray squares) are also shown. The high proportion of BrdU-positive cells in all fractions (bold black line) shows that a substantial proportion of cells in all six fractions are in S phase. Over 80% of cycling ES

cells are in S phase and fewer than 10% are in G1 (Savatier et al. 1994), inevitably leading to the inclusion of early S phase cells in fraction 1. Data are representative of two or three independent experiments. (E) In ES cells, singlet and doublet FISH signals for X-chromosomal loci exhibit equivalent fluorescence intensity. Plots show the ratios of S/D (solid symbols) or D/D (open symbols) FISH signal intensities in individual MEF or ES cell nuclei displaying an SD or DD pattern for *Pgk1* or the *Xic* as indicated. The intensity of both pinpoints in each doublet was summed to calculate the total intensity of doublet signals. When calculating the D/D intensity ratios, the two doublets in a cell were randomly assigned to the numerator or denominator. Mean ratio values and 95% confidence intervals for the means are indicated. (F) Comparison of the proportions of cells displaying SS (white), SD (black), and DD (gray) signals for an autosomal locus (*Fnl1*) and three X-chromosomal loci (*Mecp2*, *Xic*, and *Pgk1*) in S phase ES cells upon PFA or MeOH fixation.

Figure 3. X Chromosomes differ from one another in ES cells. (A) Map of the X chromosome showing positions (Mb) of loci assayed for SIAR. (B) Concordant *Mecp2* (red) and *Hprt* (green) (left) and discordant *Xic* (red) and *Mecp2* (green) (right) FISH signals. (C) Frequencies of concordance and discordance for specified locus pairs in ES cells. p-values, determined using a χ^2 test, reflect the probability that the observed distributions are random.

Figure 4. The future Xa and future Xi exhibit distinct frequencies of singlet FISH signals. (A) Allele-specific FISH for the *Xic* (red) in $\Delta Xist/+$ ES cells. An *Xist* probe

(green) identifies the wild-type allele. White arrowhead indicates the $\Delta Xist$ allele. (B) Allele-specific FISH for the *Xic* (red) in *Tsix-pA/+* ES cells. *Tsix* RNA (green) identifies the wild-type allele. Grey arrowhead indicates the *Tsix-pA* allele. (C) Table summarizing scoring of allele-specific FISH in $\Delta Xist/+$ and *Tsix-pA/+* ES cells. For three X-chromosomal loci, SD cells were scored for identity of the allele displaying the singlet signal. The X chromosome indicated in black always becomes the X_a, and that in gray and marked with an asterisk always becomes the X_i. The allele indicated in green will be the expressed allele after X-inactivation and the allele indicated in red will be the silent allele. p-values reflect the probability that the observed distributions are random.

Figure 5. X chromosomes switch between states. (A) The two SD signal configurations observed for *Ccnb3* (red) by allele-specific FISH in X.2/Xwt ES cells, in which one X chromosome is fused to Chromosome 2. The marked allele (asterisk) is scored by its proximity to a CEN-2 probe (green). The CEN-2 signal on wild-type Chromosome 2 is indicated by parentheses. (B) Allele-specific scoring of *Ccnb3* in X.2/Xwt ES and four single cell-derived clones. Non-significant p-values indicate a random distribution. (C) Allele-specific FISH for the *Xic* (red) in *129-tet/cas* ES cells. The *129-tet* allele is identified by a tet-operator probe (green; yellow in overlap; indicated by white arrowheads) and can be found in both the singlet (left) or doublet configuration in SD cells. (D) Allele-specific scoring of the *Xic* in two independently derived *129-tet/cas* ES cell lines. p-values indicate that the *Xic* on the *129* chromosome exhibits a singlet signal at a higher frequency than would be expected by random chance.

Figure 6. SIAR is specific to pluripotent cells *in vitro* and *in vivo*. (A) Percent SD signals observed by FISH for the *Xic* in cell types that are poised for X-inactivation (ES and ICM), and in trophectoderm (TE) cells and MEFs, which have completed X-inactivation. (B) Allele-specific FISH for the *Xic* (red) in $\Delta Xist/+$ ICM cells. An *Xist* probe (green) identifies the wild-type allele; white arrowhead indicates the $\Delta Xist$ allele. In 12 of 15 SD cells ($p = 0.02$), the $\Delta Xist$ allele exhibited the doublet signal. (C) Comparison of the proportion of PFA-fixed S phase (BrdU+) ES cells (dark grey) and MEFs (light grey) exhibiting SD signals for an autosomal locus (*Hba1*) and two X-chromosomal loci. (D) Percent SD signals observed for an autosomal biallelic locus (*Hba1*) and two X chromosomal loci (*Mecp2* and the *Xic*) in MEFs fixed with PFA (dark gray) or MeOH (light gray). Average data from two experiments ($n \geq 150$) are presented; error bars indicate one standard deviation. (E) Full presentation of the proportions of S phase MEFs fixed with PFA (left) or MeOH (right) exhibiting SS (white), SD (black) and DD (grey) signals at *Hba1*, the *Xic* and *Mecp2*.

Figure 7. Models for achieving randomness in X-inactivation. (A) The data in the present study suggest a model in which the X chromosomes in pluripotent cells (m indicates maternal X chromosome, p indicates paternal X chromosome) coordinately switch between future Xa (light blue with dark blue *Xic*) and future Xi (dark blue with light blue *Xic*) states in cycling cells. The fates of the X chromosomes as the Xa (green with red *Xic*) or the Xi (red with green *Xic*) are determined by their states at the time that the cell receives the cue to initiate X-inactivation. (B) The prevailing model holds that the two X chromosomes in pluripotent cells (black, *Xic* indicated in white) are equivalent until the

cue that initiates X-inactivation causes differential marking of the two X chromosomes (gray cross), thus designating the X_a and the X_i.

Figure 8. Features of SIAR are not revealed in MeOH samples. (A) *Xic* FISH signal intensity quantification as in Figure 2E for MeOH-fixed MEFs and ES cells. S/D FISH signal intensity ratios in individual nuclei displaying SD signals are plotted. The signal intensity ratio distributions are consistent with the interpretations that (1) in MEFs exhibiting an SD pattern, most singlet and doublet signals represent unreplicated and replicated alleles, respectively, and (2) in ES cells exhibiting an SD pattern, the majority of the singlet and doublet FISH signals reflect equal amounts of DNA at the two alleles. (B) Coordination of FISH signals for X-chromosomal loci in MeOH-fixed ES cells. Frequencies of concordance and discordance for specified locus pairs were scored as in Figure 2B. p-values reflect the probability that the observed distributions are random. A high p-value for *Mecp2* and the *Xic* reflects the small sample size. As previously reported, the *Xic* exhibited a high proportion of concordant FISH signals with the linked *Mecp2* gene in samples fixed with MeOH (Gribnau et al. 2005), whereas in PFA-fixed samples this locus pair exhibited predominantly discordant FISH signals (Figure 3C). (B) Scoring of allele-specific FISH in MeOH-fixed *Tsix-pA/+* ES cells. Cells displaying an SD pattern for the *Xic* were scored for identity of the allele displaying the singlet signal as in Figure 4B. p-values reflect the probability that the observed distributions are random. This analysis supports the conclusion that in MeOH-fixed ES cells, the *Xic* on the future X_i does not display an increased likelihood of exhibiting a singlet signal (Gribnau et al. 2005), in contrast to what was seen in PFA-fixed ES cells (Figure 4C).

Table 1. BACs used for FISH probes

<u>Probe</u>	<u>BAC</u>
Hba1	RP23-71G18
Fn1 (B)	RP23-27O4
Ccnb3	RP24-270D24
Hprt (B)	RP23-412J16
Mecp2 (A)	RP23-77L16
Xic	RP23-309B17
Pgk1	R24-90H17
Grpr (B)	RP23-231H22
2CEN (A)	RP23-382P22

Table 2. Primers used in Chapter 1

<u>Gene</u>	<u>Sequence</u>
<i>ZIP1</i> (human standard)	ATCTCCAGTCAGTGGCTAGTCC CACGCTTGGTCCACGTTGGGATTT
<i>Xic</i>	AAGTCAATAAAGCACTCCCCATCTC TTGGCTCAGTGCTTATGGTG
<i>Pgk1</i>	TGCAACTGTTAGACCTGAGGAACCTTG TTGCCAGCAGAGATTTGAGTTCAGC

Figure 1

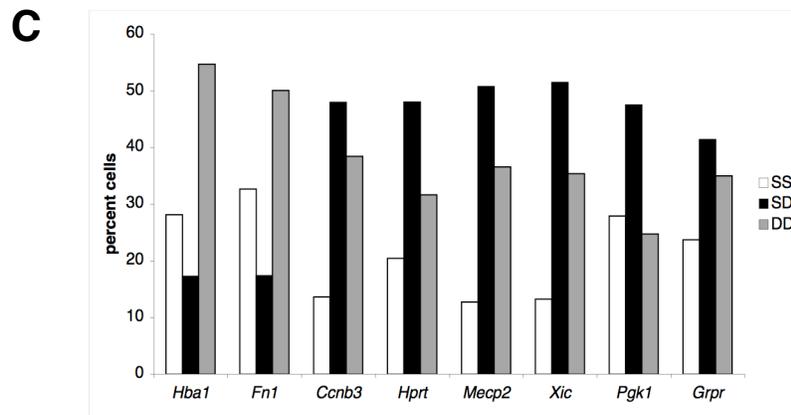
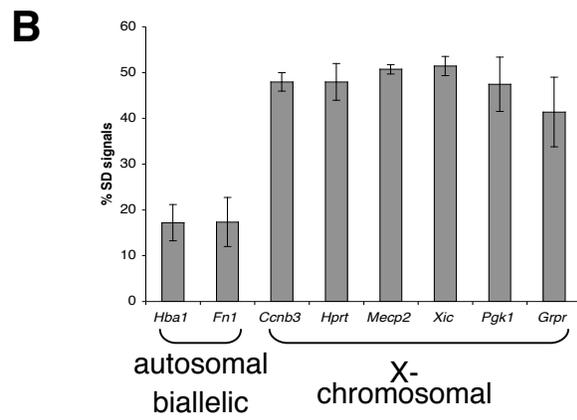
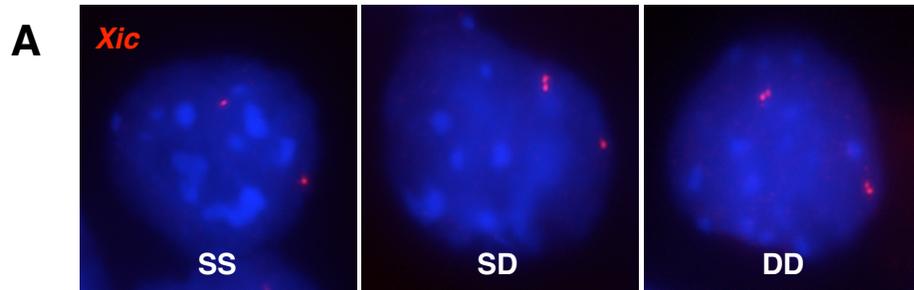


Figure 2

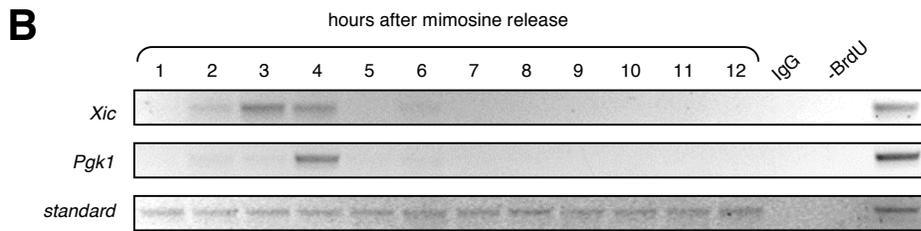
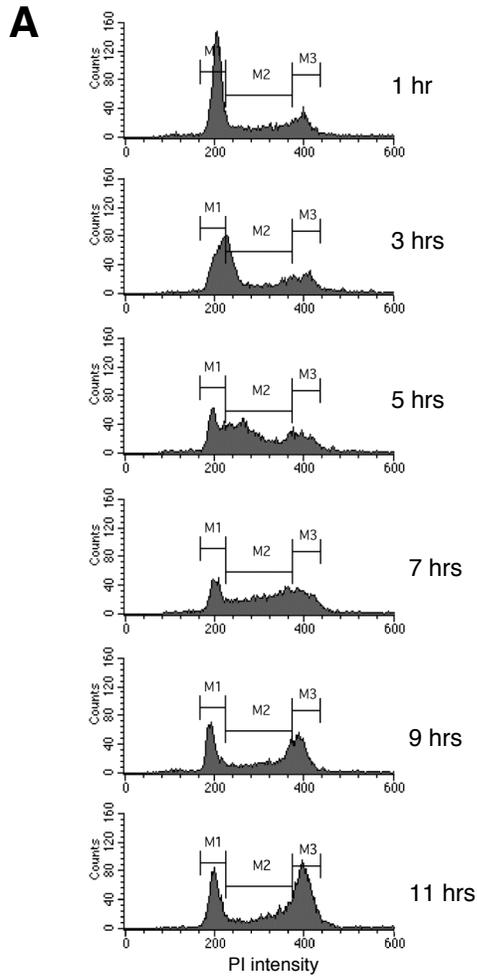


Figure 2 continued

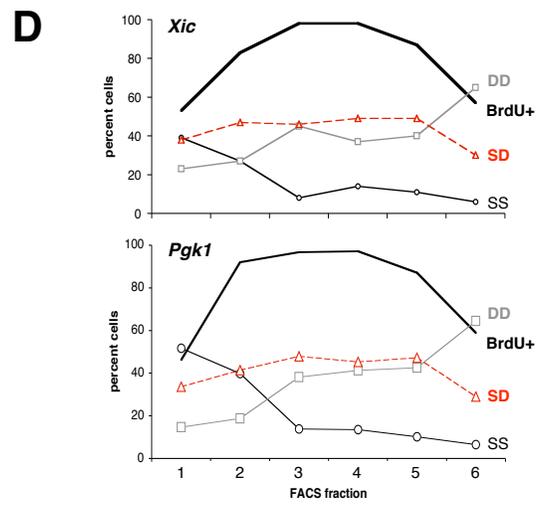
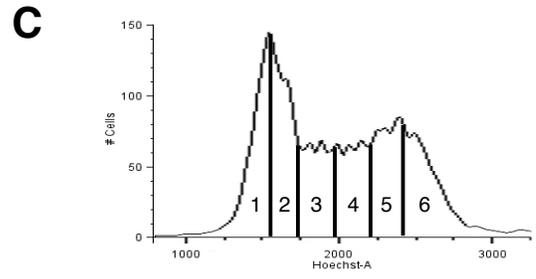
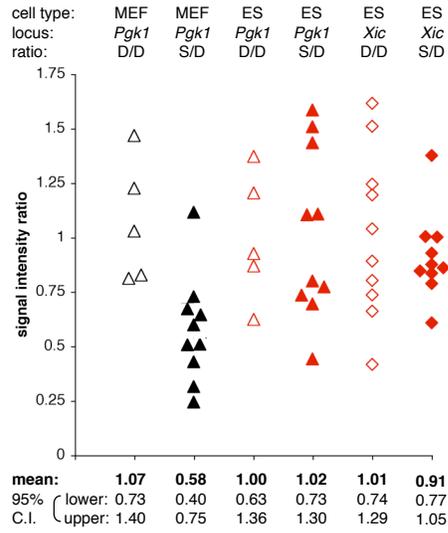


Figure 2 continued

E



F

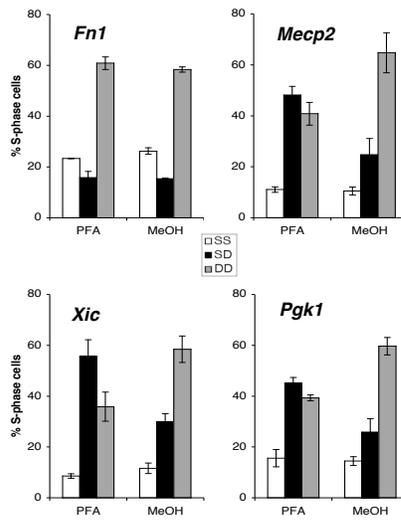
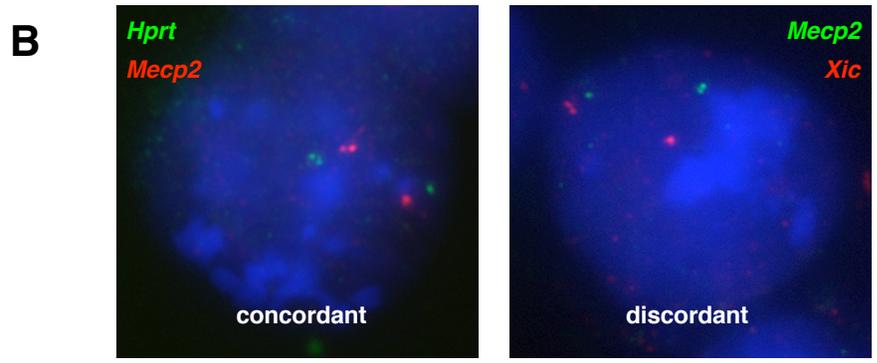
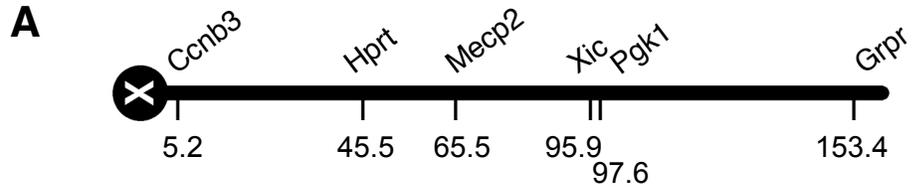


Figure 3



C

locus pair	SD SD cells:		p =	(n)
	concordant	discordant		
<i>Ccnb3</i> & <i>Hpvt</i>	64%	36%	0.009	(84)
<i>Hpvt</i> & <i>Mecp2</i>	65%	35%	0.014	(66)
<i>Mecp2</i> & <i>Pgk1</i>	65%	35%	0.001	(115)
<i>Mecp2</i> & <i>Xic</i>	38%	62%	0.006	(134)
<i>Xic</i> & <i>Pgk1</i>	38%	62%	0.019	(104)

Figure 4

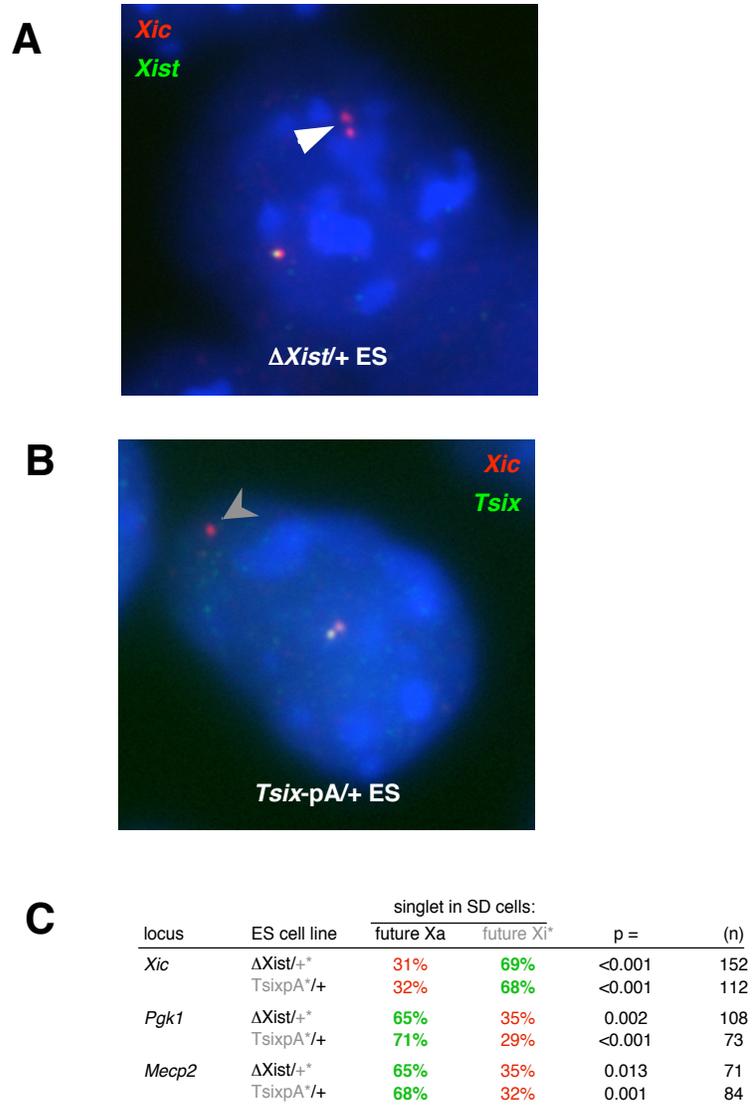
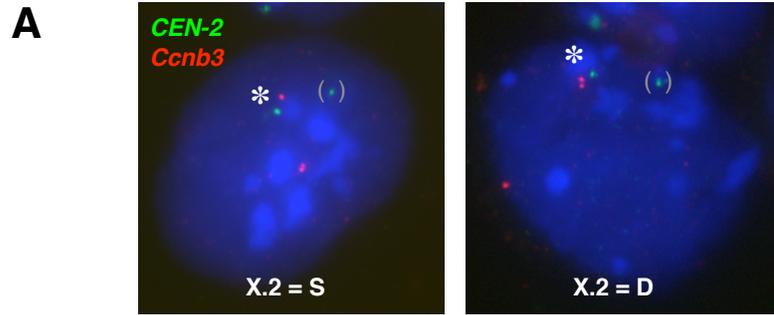


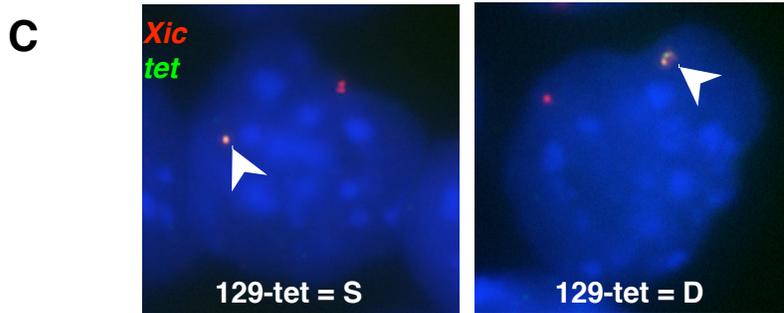
Figure 5



B

Ccnb3 singlet in SD cells:

cell line	marked	unmarked	(n)	p =
X.2/X ^{wt}	47%	53%	179	0.501
X.2/X ^{wt} clone 6G	46%	54%	52	0.579
X.2/X ^{wt} clone 7C	50%	50%	50	1.000
X.2/X ^{wt} clone 7D	47%	53%	64	0.617
X.2/X ^{wt} clone 7E	52%	48%	50	0.777



D

Xic singlet in SD cells:

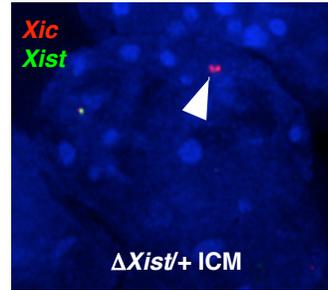
	129-tet	cas	(n)	p =
129-tet/ <i>cas</i> clone A	60%	40%	209	0.003
129-tet/ <i>cas</i> clone B	60%	40%	237	0.002

Figure 6

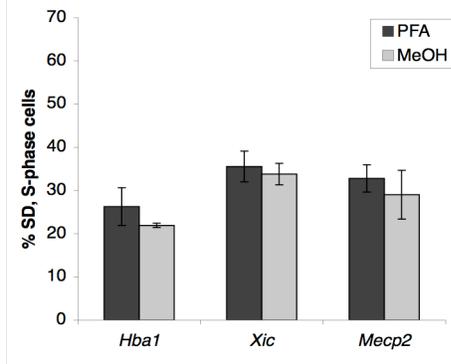
A

cell type	<i>Xic</i>	
	%SD	(n)
ES	51	(182)
ICM	53	(40)
TE	31	(36)
MEF	28	(141)

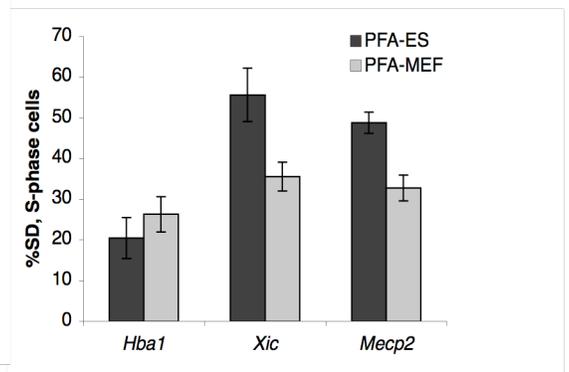
B



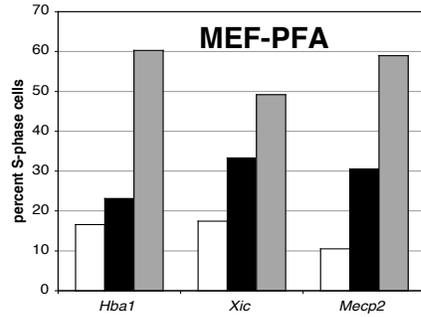
C



D



E



F

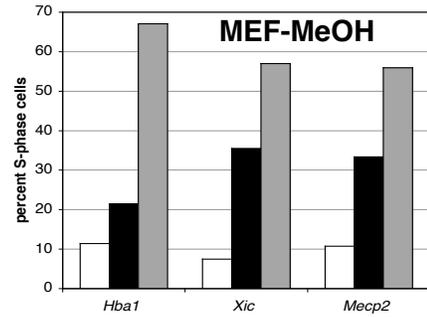


Figure 7

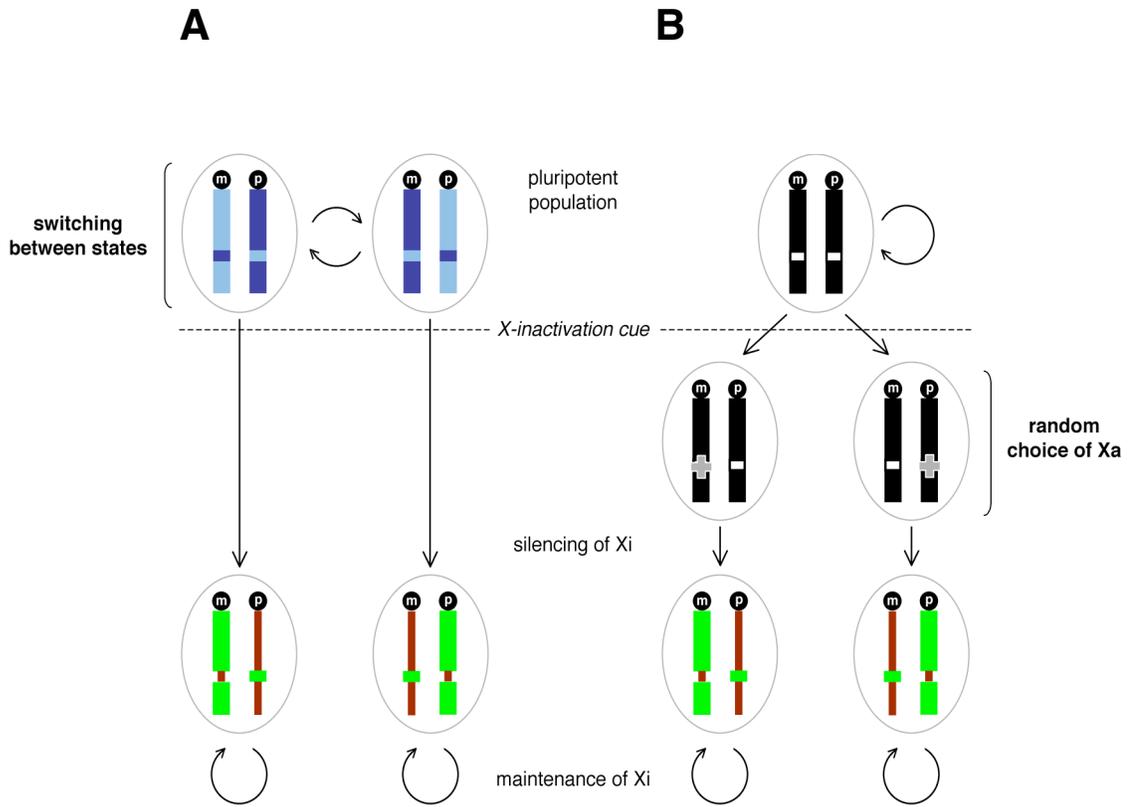
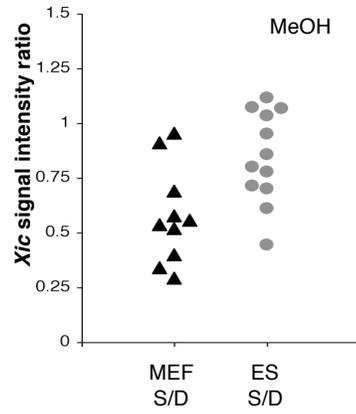


Figure 8

A



B

	% SD SD cells		p =	(n)
	concordant	discordant		
<i>Mecp2</i> & <i>Cnbp2</i>	77	23	0.003	(30)
<i>Mecp2</i> & <i>Xic</i>	62	38	0.275	(21)

C

	locus	% singlet in SD cells		p =	(n)
		marked	unmarked		
<i>Tsix</i> -pA/+	<i>Xic</i>	47	53	0.475	(159)

Chapter 2

**The A-repeat regulates Xist RNA processing and is required for
random X-inactivation**

I. Summary

One X chromosome is silenced in each female mammalian cell. *Xist* plays two roles in this process. It is essential for X-inactivation to occur randomly and to establish X chromosome silencing. A conserved element, the A-repeat, is necessary for *Xist* RNA-dependent silencing. We deleted this element from one copy of *Xist* in female cells to determine if the A-repeat is also necessary for random choice. Here we show that the A-repeat mutant X always remains active. We also find that the A-repeat is required for correct *Xist* RNA metabolism and that it binds the splicing factor ASF/SF2. In combination, our data suggest that normal *Xist* RNA processing is important for X-inactivation to occur randomly. We propose that regulation of *Xist* RNA processing may be part of the stochastic process that determines which X will be inactivated.

II. Introduction

In order to balance X-linked gene dosage between XX females and XY males, female mammals silence one X chromosome in each cell early during embryogenesis. In eutherian embryos, X-inactivation is random in that each X chromosome has an equal probability of being inactivated (Lyon 1961). This process can be recapitulated in embryonic stem (ES) cells. Female ES cells have two active X chromosomes and undergo random X-inactivation when differentiated (Martin et al. 1978), making them a useful model system in which to study the factors that influence which X will be silenced and which will remain active.

The choice of which chromosome will be silenced is regulated *in cis* by the X-inactivation center (*Xic*). The *Xic* encodes an antisense pair of genes, *Xist* and *Tsix*, each

of which produces a non-coding RNA (Avner and Heard 2001). Loss-of-function mutations in either *Xist* or *Tsix* cause non-random X-inactivation (Marahrens et al. 1998; Gribnau et al. 2005). *Xist* is also necessary and sufficient to silence the inactive X chromosome (Xi) (Wutz and Jaenisch 2000). When X-inactivation is triggered, *Xist* RNA coats and silences the X that will become the Xi (Panning and Jaenisch 1996; Sheardown et al. 1997). A conserved repeat element at the 5' end of *Xist*, the A-repeat, is necessary for establishment of silencing when *Xist* RNA is ectopically expressed from an inducible cDNA transgene (Wutz et al. 2002). This element is also necessary for X chromosome silencing during imprinted X-inactivation in mouse extraembryonic tissues, where the paternally-inherited X chromosome is silenced in every cell (Hoki et al. 2009).

To determine if the A-repeat also plays a role in random X-inactivation, we deleted this element from one chromosome in female ES cells. Our results show that female A-repeat mutant ($X^{AA}X$) cells undergo non-random X-inactivation, demonstrating that the A-repeat is necessary for random choice during X-inactivation. In addition, we identify two new functions of the A-repeat that may explain why X-inactivation is non-random in $X^{AA}X$ cells. First, the A-repeat is important for post-transcriptional processing of *Xist* RNA and second, the A-repeat binds the essential splicing factor ASF/SF2. In combination, our data suggest a model in which *Xist* RNA splicing is a regulatory step used to ensure that X-inactivation occurs randomly.

III. Materials and Methods

ES cells

Culture and differentiation of wild-type female ES cell line ES 2-1 (Panning et al. 1997) and $X^{\text{tet}^O}X$ (Gribnau et al. 2005), $X^{\text{Tp}^A}X$ (Fa2L (Luikenhuis et al. 2001)), $X^{\Delta\text{Xist}}X$ ($X^{\text{Ilox}}X$ (Csankovszki et al. 1999)), $X^{\Delta A}X$, XY, $X^{\Delta A}Y$ (Blewitt et al. 2008) and XY-GFP (Fazzio et al. 2008) ES cells were carried out using standard conditions.

Genomic targeting and generation of $X^{\Delta A}X$ cells

To produce the ΔA targeting construct a puromycin cassette flanked by *loxP* sites was cloned between 5' and 3' homology arms that directly flank the A-repeats. 5' and 3' homology arms were generated by PCR using primers indicated in Table 3. A P_{gk}-promoter upstream of the diphtheria toxin A gene was inserted at the 3' end of the 3' homology arm to allow for counter selection against random integrants. Linearized targeting constructs were introduced into ES 2-1 female cells by electroporation and clones were selected with 5 $\mu\text{g/ml}$ puromycin starting 48 hours after electroporation.

Fluorescent *in situ* hybridization

Xist exon1 or A-repeat concatamer PCR products were used as templates for Xist and A-repeat probes, respectively. The tetO sequences were detected as described previously (Gribnau et al. 2005). Random priming (Bioprime kit reagents; Invitrogen) was used to incorporate cy3-dCTP (Amersham), FITC-dUTP (Roche), or biotin-dCTP (Invitrogen). RNA and DNA FISH were performed as previously described (Mlynarczyk-Evans et al. 2006). Samples were scored on an Olympus BX60 microscope. Images were collected

with a Hamamatsu ORCA-ER digital camera using Openlab 4.0.1 software, assembled using Adobe Photoshop 7.0, and levels adjusted to enhance contrast.

Chromatin immunoprecipitation

For histone immunoprecipitations, confluent 10cm plates ($\sim 2 \times 10^7$ cells) were trypsinized and resuspended in 10 ml fresh growth medium. Formaldehyde was added to a final concentration of 1%, and after 10 min, cross-linking was quenched with 125 mM glycine. Cells were washed once with PBS and pellets were flash frozen and stored at -80°C . Pellets were allowed to thaw on ice and 1 ml swelling buffer (10 mM Tris-HCl pH8, 85 mM KCl, 0.5% NP40, with protease inhibitors: 5 $\mu\text{g/ml}$ leupeptin, 0.7 $\mu\text{g/ml}$ pepstatin, 1 $\mu\text{g/ml}$ aprotinin, 100 μM AEBSF) was added, followed by 60 strokes of Dounce homogenization. Nuclei were collected and resuspended in *micrococcal* nuclease buffer (50 mM Tris-HCl pH7.5, 60 mM KCl, 3 mM CaCl_2 , 340 mM sucrose plus protease inhibitors). One hundred units of *micrococcal* nuclease (Worthington) were added and after 5 min at room temperature, reactions were quenched with 50 mM EGTA. Nuclei were sonicated twice for 20 sec each with a Branson 250 sonicator, power setting 3, resulting in DNA fragments 200-1000 bp in length. Insoluble material was removed by centrifugation and for each immunoprecipitation, 400-500 μg (protein) of soluble chromatin extract was diluted in RIPA140 (10 mM Tris-HCl pH8, 1 mM EDTA, 0.5 mM EGTA, 140 mM NaCl, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS plus protease inhibitors) to a final volume of 2.5 ml. Either 20 μl magnetic protein A or protein G beads (Dynal) were washed with PBS and pre-bound with H3K27me3 (5 μl Abcam 6002), H3K4me2 (10 μl , Upstate Biotechnology 07-030), H4ac (10 μl Upstate

Biotechnology 06-598) or rabbit IgG (10 μ l Jackson ImmunoResearch 011-000-C03) at 4°C for at least 1 hour. These beads were added to the extract and rotated at 4°C overnight. The beads were washed 3X with RIPA140, 1X with RIPA500 (RIPA140 with 500mM NaCl total), and once more with RIPA140 for 5 minutes each at room temperature. Chromatin was eluted and cross-links were reversed with 200 μ l elution buffer (50 mM Tris-HCl pH8, 1 mM EDTA, 0.5% SDS, 100 mM NaCl, 100 μ g/ml proteinase K), incubated at 65°C overnight. RNA was degraded with RNase and DNA was purified by phenol-chloroform extraction. 1/50th of the chromatin immunoprecipitate was added to each 20 μ l PCR reaction.

Allele-specific PCR

Allele-specific PCR analyses were based on single nucleotide polymorphisms that resulted in restriction fragment length polymorphisms. To eliminate the effect of heteroduplex DNA on digestion results, PCR reactions from RT-PCR or ChIP were diluted 1:2 into a new PCR mixture containing a single radiolabeled primer (indicated by *) and one final cycle of PCR was performed prior to digestion with the appropriate restriction enzyme (Table 3). Digested products were separated on an 8% polyacrylamide gel. Products were visualized on a PhosphorImager (Molecular Dynamics) and ImageQuant (Molecular Dynamics) was used to calculate the relative proportion of *I29* and *cas* PCR products in each reaction.

RT-PCR

RT was performed on total RNA isolated with Trizol reagent (Invitrogen) and treated with RNase-free DNase (Promega). For RT-PCR of spliced transcripts, first strand cDNA was synthesized from 0.5 µg of total RNA using random primers and 40 cycle PCR was performed. To detect unspliced Xist transcripts, cDNA was synthesized from 5 µg of total RNA using a strand-specific primer containing a 5' T3 anchor sequence. +RT and – RT reactions were performed for all samples and minus primer controls were included for all strand-specific RT reactions. 1/20th of the first strand cDNA was amplified by PCR. All primers are indicated in Table 3.

Gel mobility shift and UV crosslinking assays

A-repeat RNA (5' GGGCCCAUCGGGGCUGCGGAUACCUGCC 3', Dharmacon) was labeled at the 3' end with 5'-[³²P] pCp and RNA ligase. End-labeled RNA (10 nM) was incubated with 10-20 µg of ES or HeLa cell (Chemicon) nuclear extract and 10 µg tRNA in 20 mM HEPES (pH 7.9), 50 mM KCl, 10% glycerol and 1 mM DTT in a volume of 20 µl at 4 °C for 30 min. A 100-fold molar excess of wild-type A-repeat or non-specific competitor RNA (MS2 hairpin, 5' AGCUGAGGAUCACCCAGCA 3', Dharmacon) was included where indicated. The mixture was separated by nondenaturing 6% PAGE at 4°C, or UV crosslinked (800 µJ on ice 10-15 min). The crosslinked reaction was added to 2 volumes sample buffer (20% glycerol, 5% SDS, 2 M Tris HCl (pH 6.7), 0.2 mg/ml bromophenol blue, 5%β-mercaptoethanol), boiled for 10 min, placed on ice 5 minutes and separated by 12% SDS-PAGE. Gels were fixed in MeOH/acetic acid, dried and visualized using a PhosphorImager (Molecular Dynamics). For protein purification,

unlabeled A-repeat RNA was incubated with HeLa cellnuclear extract and partially purified (Figure 13D-F). The mixture was separated by nondenaturing 10% PAGE with a 4% stack and stained with silver.

RNA IP

293 cells were transiently transfected with a plasmid directing production of GFP-tagged ASF/SF2. Immunoprecipitation was performed as described (Huang et al. 2004), except that phosphatase inhibitors were omitted and Dynal magnetic beads were used, with rabbit anti-GFP antibody (Abcam) and control rabbit IgG (Jackson ImmunoResearch).

Statistics

All p-values were determined by comparing the observed distribution of signal patterns at each allele to the expected 25/75 distribution (null hypothesis) using a chi-squared distribution test with one degree of freedom.

IV. Results

Targeted deletion of the A-repeat

To investigate the role of the A-repeat in the regulation of X chromosome inactivation, we generated a female ES cell line bearing a deletion of the A-repeat from one copy of *Xist*. The murine A-repeat consists of 7.5 copies of a 24-mer repeat unit connected by linkers of variable size. Our targeting vector was designed to replace the A-repeat with a puromycin cassette flanked by *loxP* sites (Figure 9A) without removing any adjacent sequences. We introduced this vector into female ES cells carrying one X chromosome

of *Mus musculus 129/Sv (129)* origin and one of *Mus castaneus (cas)* origin. Correctly targeted clones ($X^{\Delta A_{\text{puro}}}X$) were identified by Southern blot hybridization (Figure 9B). The puromycin cassette was then deleted by transient expression of Cre recombinase and deleted clones ($X^{\Delta A}X$) were identified by Southern blot hybridization (Figure 9C). Allele-specific PCR was used to confirm that the A-repeat was deleted from the *129* chromosome in $X^{\Delta A}X$ ES cells and sequencing was performed to ensure that no additional mutations were introduced within the homology arms (data not shown). The resulting line replaces the 400 bp A-repeat with a *loxP* site on the *129* X chromosome.

Deletion of the A-repeat causes primary non-random X-inactivation

To determine the effect of the A-repeat deletion on X-inactivation, we differentiated $X^{\Delta A}X$ and parental XX cells and used allele-specific reverse transcriptase (RT)-PCR to determine the ratio of *129* to *cas* Xist RNA. In wild-type *129/cas* cell lines, X-inactivation is skewed away from a 1:1 ratio because the *129* and *cas* X chromosomes contain different *X controlling element (Xce)* alleles (Cattanach and Rasberry 1994). Differentiated parental XX cells exhibited the expected ratio of *129:cas* transcripts, while differentiated $X^{\Delta A}X$ cells expressed only *cas* Xist transcripts (Figure 10A). This result indicates that the ΔA chromosome never becomes the Xi.

Next, we used allele-specific fluorescence *in situ* hybridization (FISH) as an independent assay to determine the frequency with which Xist RNA coats each X chromosome in differentiated $X^{\Delta A}X$ and control cells. A control line, $X^{\text{tetO}}X$ (Gribnau et al. 2005), which is derived from the same parental female ES cell line as $X^{\Delta A}X$, carries an insertion of a tet operator (*tetO*) array that marks the *129* X chromosome (Figure 10B).

We used DNA FISH to detect the tetO sequences and RNA FISH to detect Xist transcripts in differentiated $X^{\text{tetO}}X$ cells. Xist RNA coated the unmarked *cas* X chromosome in ~25% of $X^{\text{tetO}}X$ cells (9 of 40), consistent with the expected frequency of *cas* silencing in a *129/cas* cross ($p=0.72$) (Cattanach and Rasberry 1994; Gribnau et al. 2005). Two RNA FISH probes, one within the A-repeat and a second downstream of the A-repeat in exon 1 were used to identify the wild-type and ΔA alleles in $X^{\Delta A}X$ cells (Figure 10B). Wild-type Xist RNA is detected by both probes while mutant Xist RNA is only detected by the exon 1 probe (Figure 10C). In 100% of differentiated $X^{\Delta A}X$ cells (55 of 55), wild-type Xist RNA coated the Xi (Figure 10D). This result is significantly different from the 25% of cells expected to silence the *cas* X chromosome ($p<0.0001$). In combination, these allele-specific RT-PCR and FISH data indicate that deletion of the A-repeat changes the frequency with which the *129* X chromosome is inactivated from 75% to 0%.

Non-random X-inactivation can be due to a primary effect on choice or a secondary effect caused by the death of cells that silence too many or too few X chromosomes (McMahon and Monk 1983). To test for secondary effects of the ΔA mutation on choice, we compared the viability of differentiating wild-type and $X^{\Delta A}X$ cells. When wild-type or $X^{\Delta A}X$ female ES cells were co-differentiated with green fluorescent protein (GFP)-expressing wild-type male ES cells, the percentage of $X^{\Delta A}X$ cells with Xist coating at each timepoint was similar to that observed in wild-type cells (Figure 10E), indicating normal X-inactivation kinetics in $X^{\Delta A}X$ cells. In addition, there was no change in the ratio of GFP+/GFP- cells over time (Figure 10E). Therefore, differentiating $X^{\Delta A}X$ cells were not at a proliferative disadvantage, consistent with

primary non-random X-inactivation. Differentiating $X^{\Delta A}X$ cells do not undergo more cell death than wild-type cells upon differentiation (data not shown), also consistent with the deletion of the A-repeat having a primary affect on choice.

To further distinguish between primary and secondary non-random X-inactivation, we examined *Xist* and *Tsix* expression on each X chromosome in cells in the early stages of X-inactivation. Shortly after ES cells are induced to differentiate and X-inactivation is triggered, *Xist* RNA coats the X chromosome that becomes the Xi while *Tsix/Xist* expression persists transiently on the active X (*Xa*), appearing as a pinpoint FISH signal (Panning et al. 1997; Sheardown et al. 1997; Lee et al. 1999). We used allele-specific RNA FISH to determine which X was silenced in this early stage of X-inactivation in $X^{\Delta A}X$ and control $X^{\text{tetO}}X$ cells. In differentiating $X^{\text{tetO}}X$ cells, the pinpoint *Tsix/Xist* RNA signal was associated with the tetO marked 129 X chromosome in 4 of 24 cells (Figure 10F), consistent with the fact that the 129 chromosome becomes the *Xa* in ~25% of differentiated wild-type cells ($p=0.35$) (Gribnau et al. 2005). In $X^{\Delta A}X$ cells, the pinpoint *Tsix/Xist* signal arose from the ΔA chromosome in 18 of 18 cells ($p<0.0001$)(Figure 10F). These results further support the conclusion that $X^{\Delta A}X$ cells undergo primary non-random X-inactivation.

The results of a third assay were also consistent with primary non-random X-inactivation in $X^{\Delta A}X$ cells. Before female ES cells are differentiated and X-inactivation is initiated, the X chromosome that will remain active (future *Xa*) and the X chromosome that will be silenced (future Xi) differ from each other when analyzed by FISH. This conclusion is based on the observation that X-linked genes in female ES cells show an unusually high frequency of singlet:doublet (SD) FISH signals. In *Xist* and *Tsix* mutant

ES cells, which undergo primary non-random X-inactivation, the future Xa shows predominantly doublet FISH signals at *Xist* and the future Xi exhibits predominantly singlet FISH signals (Mlynarczyk-Evans et al. 2006). We compared the *Xist* loci in $X^{\Delta A}X$ ES cells to those in *Xist* ($X^{\Delta Xist}X$ (Gribnau et al. 2005)) and *Tsix* ($X^{TpA}X$ (Luikenhuis et al. 2001)) mutant ES cells. The A-repeat mutant chromosome showed a high frequency of singlet FISH signals at *Xist*, a pattern similar to that seen on the future Xa in $X^{\Delta Xist}X$ and $X^{TpA}X$ ES cells (Figures 10G-I). Thus, $X^{\Delta A}X$ cells exhibit the FISH signature characteristic of cells that will undergo primary non-random X-inactivation. This observation in combination with our results above indicate that deletion of the A-repeat from one X chromosome in female ES cells results in primary non-random X-inactivation upon differentiation.

Histone modifications do not correlate with fixed chromosomal fate

In an ES cell line heterozygous for a deletion that removes the major *Tsix* promoter there is primary non-random X-inactivation such that the mutant X chromosome always becomes the Xi (Lee and Lu 1999). In addition, histone modifications that are associated with silent chromatin, such as enrichment of histone H3 trimethylated on lysine 27 (H3K27me3) and depletion of histone H3 dimethylated on lysine 4 (H3K4me2) and tetraacetylated histone H4 (H4ac), were reported to mark the *Tsix* mutant chromosome in these cells (Sun et al. 2006). These observations led to the suggestion that chromatin marks designate the *Tsix* mutant chromosome as the X that will become the Xi. We performed chromatin immunoprecipitation (ChIP) for H3K27me3, H3K4me2, and H4ac in wild-type, $X^{\Delta A}X$ and $X^{TpA}X$ ES cells to determine if primary non-random X-

inactivation caused by an *Xist* mutation also resulted in chromatin modifications on the X destined to be the Xi.

We used restriction enzyme polymorphisms to determine the relative abundance of H3K27me3, H3K4me2, and H4ac in the *Xist/Tsix* region (Figure 11A) on the *I29* and *cas* chromosomes in undifferentiated XX, X^{ΔA}X and X^{TPA}X ES cells. In wild-type cells, the distribution of all three histone modifications was roughly equal between the *I29* and *cas* chromosomes at the *Xist* promoter region, in the *Xist* gene body, and at the *Tsix* promoter region (Figures 11B, C). Insertion of a polyadenylation signal downstream of the *Tsix* promoter (X^{TPA}X) caused a skewed distribution of H3K27me3 at all three sites, with an increased proportion of this modification occurring on the *Tsix* mutant chromosome. The distribution of H3K4me2 and H4ac did not differ dramatically from wild-type cells at any of the three sites. While the allelic enrichment of H3K27me3 in X^{TPA}X cells is consistent with the analysis of the *Tsix* promoter deletion ES cells, our H3K4me2 and H4ac results differ. One possible reason for this difference is that our analysis was carried out using a technique that eliminates heteroduplex DNA and allows more accurate allele-specific quantitation.

In contrast to *Tsix* mutant ES cells, X^{ΔA}X ES cells showed an equal distribution of H3K27me3, H3K4me2, and H4ac between the *I29* and *cas* alleles at all three sites analyzed. Thus, the non-random X-inactivation that occurs in differentiated X^{ΔA}X ES cells does not correlate with altered distribution of H3K27me3, H3K4me2, or H4ac in undifferentiated ES cells. These data indicate that the unequal distribution of these chromatin marks at *Xist/Tsix* is not necessary to preemptively mark X chromosomes for primary non-random X-inactivation.

Properly spliced ΔA Xist transcripts fail to accumulate

To investigate the effects of an A-repeat deletion on Xist RNA accumulation, we employed an independently-generated male $X^{\Delta A}Y$ ES cell line which contains a deletion that removes the A-repeat and approximately 400 bp of flanking sequence from the sole *Xist* locus (Blewitt et al. 2008). A-repeat mutant male ES cells were used for this analysis because populations of ES cells contain a low percentage of differentiated cells even when grown under conditions that promote self-renewal. The amount of Xist RNA in differentiated female cells is several orders of magnitude greater than in undifferentiated ES cells (Norris et al. 1994). As a result, a substantial proportion of the Xist RNA in a population of female ES cells arises from the low percentage of differentiated cells within that population (Figure 12A). Because the wild-type *cas X* always becomes the Xi in $X^{\Delta A}X$ ES cells, differentiated cells may mask any bias in the *cas:129 Xist* RNA ratio among the undifferentiated cells in the population. *Xist* is silenced in differentiated male cells (Brockdorff et al. 1991; Brown et al. 1992), so only ES cell-specific Xist transcripts are assayed. Also, there are a limited number of polymorphisms that allow us to distinguish between Xist transcripts arising from the wild-type *cas X* and the mutant *129 X* chromosomes in $X^{\Delta A}X$ ES cells; using male cells simplifies the analysis.

We compared the levels of spliced Xist and Tsix RNAs in parental XY and $X^{\Delta A}Y$ ES cells. Two spliced Xist products were examined- exon 1 to exon 3 (1-3) and exon 3 to exon 6 (3-6) (Figure 12B)- and both were significantly reduced in $X^{\Delta A}Y$ ES cells. In contrast, elevated Tsix RNA levels were detected from the mutant chromosome in $X^{\Delta A}Y$

ES cells (Figure 12C). RT-qPCR for spliced Tsix (exon 3 to exon 4) revealed that steady state levels of spliced Tsix are approximately 6-fold higher in $X^{\Delta A}Y$ cells to XY control cells.

We next examined whether the absence of correctly spliced Xist transcripts in $X^{\Delta A}Y$ ES cells could be attributed to reduced levels of primary transcripts. Strand-specific RT reactions were performed to ensure that Xist, not Tsix, transcript levels were measured. We found that control RT reactions carried out without primers could yield RT-PCR products (Figure 12D) suggesting that Xist and Tsix RNA can prime each other, as has been reported for other sense/antisense RNA pairs (Perocchi et al. 2007). As a result, it is unclear whether PCR products from strand-specific RT-PCR reactions arose from the Xist or Tsix strand. Therefore, we used a primer containing a 5' anchor sequence for the RT reaction and a primer within the anchor sequence to subsequently amplify the cDNA (Figure 12E). This method ensured that only cDNAs produced from the RT primer were amplified in the PCR step. Across both the first (exon 1 to intron 1) and last (intron 6 to exon 7) exon-intron junctions, similar amounts of correctly sized products were detected in wild-type and $X^{\Delta A}Y$ cells (Figure 12F, G). Therefore, the reduced levels of correctly spliced Xist RNA cannot be attributed to a reduction of primary transcripts from the ΔA allele. In contrast, strand-specific RT-PCR within exon 1 and exon 7, each of which measure the combined levels of spliced and unspliced transcripts, revealed approximately 4-fold fewer Xist transcripts in ΔA mutant cells relative to wild-type cells (Figure 12F, G). We believe that the decrease in transcript abundance within exons 1 and 7 reflects the absence of correctly processed Xist RNA in the ΔA mutant.

To determine if Xist RNA splicing was perturbed in ΔA cells, we performed strand-specific RT-PCR across intron 4 and measured the ratio of spliced:unspliced Xist transcripts. Only 20% of the transcripts in $X^{\Delta A}Y$ cells were spliced across this junction compared to 96% in wild-type cells (Figure 12H). Further RT-PCR analysis also revealed several mis-processed transcripts. At 50 cycles of PCR, we occasionally observed a smaller than expected exon3-6 product in ΔA cells (Figure 12I). Sequencing revealed that this product corresponds to an aberrantly spliced transcript that skips exons 4 and 5 of Xist and partially includes introns 3 and 5 (Figures 12J, K). 50 cycles of PCR amplification across exon 1 and exon 4 or exon 1 and exon 5 also amplified incorrectly sized products in the mutant cell line (Figure 12L). These findings support the conclusion that Xist RNA splicing is defective in $X^{\Delta A}Y$ cells.

We next assessed the accumulation of spliced Xist RNA in female $X^{\Delta A}X$ cells. To assess whether spliced Xist RNA was produced from the mutant *I29* X chromosome, we employed a polymorphism in the ex 1-3 RT-PCR product that generates a restriction enzyme cleavage site in the *I29* cDNA. While digestion of the ex1-3 cDNA from wild-type female ES cells revealed both *I29* and *cas* bands, no *I29* product was detected in $X^{\Delta A}X$ cells (Figure 12M). Thus, spliced Xist RNA is also absent from the mutant X chromosome in $X^{\Delta A}X$ ES cells. Within exon 1, *I29* transcripts were detected from the mutant *I29* chromosome (Figure 12N), suggesting that the lack of spliced Xist RNA from the ΔA X chromosome does not arise from a transcriptional defect. Together, the analyses of male and female ΔA cells indicate that deletion of the A-repeat impairs production or accumulation of spliced Xist RNA.

The A-repeat binds ASF/SF2

To gain insight into the molecular role of the A-repeat, we identified proteins that interact with this RNA sequence. Radiolabeled RNA corresponding to a consensus repeat unit (Figure 13A) (Wutz et al. 2002) was incubated with HeLa or ES cell nuclear extracts and bound complexes were separated from unbound RNA by nondenaturing polyacrylamide gel electrophoresis (PAGE) (Figure 13B) or by UV-crosslinking followed by SDS-PAGE (Figure 13C). Two prominent band shifted complexes were observed by both methods. Both complexes could be competed away with unlabeled A-repeat RNA, but not with a non-specific competitor, indicating that both complexes were specific to the A-repeat (Figure 13B). A three-step purification enriched the larger complex (complex II) approximately 50-fold (Figures 13D-F). To determine the protein composition of this complex, this partially purified HeLa nuclear extract was separated by SDS-PAGE. Mass spectrometry identified the most prominent bands that migrated near the molecular weight of the larger complex (Figure 13G) as aldolase A and Alternative Splicing Factor/Splicing Factor 2 (ASF/SF2).

To determine whether either ASF/SF2 or aldolase A bound the A-repeat, radiolabeled A-repeat RNA was incubated with HeLa cell nuclear extract and UV cross-linked. Immunoprecipitation of UV-crosslinked complex II with either ASF/SF2 or aldolase antibodies was performed. ASF/SF2 but not aldolase antibodies precipitated the radiolabeled A-repeat RNA-protein complex (Figure 13H), indicating that ASF/SF2 binds an A-repeat of *Xist* RNA *in vitro*. To confirm that ASF/SF2 interacts with XIST RNA *in vivo*, we transiently expressed GFP-tagged ASF/SF2 protein in human female cells. Whole cell extracts from these cells were immunoprecipitated using a GFP

antibody. RT-qPCR revealed that 100-fold more XIST RNA was pulled down in cells expressing ASF/SF2-GFP than in cells expressing the GFP alone (Figure 13I). This result suggests that ASF/SF2 interacts with XIST RNA in cells.

We next performed mutational analysis on an A-repeat unit to identify sequences necessary for ASF/SF2 binding. Each A-repeat unit is predicted to form a pair of stackable stem-loop structures (Figure 13A) by the free energy minimization program MFOLD (Zuker and Jacobson 1995; Mathews et al. 1999). In addition, the A-repeat consensus unit contains two sites, identified by ESEfinder (Cartegni et al. 2003), that conform to the consensus motif for ASF/SF2 (Figure 14A, B). Complex II did not form with RNAs in which either site was mutated such that it was no longer predicted to bind ASF/SF2 (Figure 14C, D, mutations C, E, F, G, and H). Mutations that were predicted to bind ASF/SF2 but that affected base pairing within either or both stems did not affect complex II formation (Figure 14C, D, mutations B and D). In combination, these data argue that the primary sequence, and not the secondary structure, is important for ASF/SF2 to interact with the A-repeat. Because band-shifted complexes required the presence of both predicted ASF/SF2 binding sites within the A-repeat unit, each unit may form a bipartite binding site for ASF/SF2, or ASF/SF2 may cooperatively bind within the A-repeat unit.

V. Discussion

The A-repeat is a highly conserved, multi-functional element

Our studies identify three new activities for the A-repeat. This highly conserved sequence is necessary to ensure that X-inactivation is random, it binds the splicing factor ASF/SF2 and it is required for correct Xist RNA processing. These observations suggest that regulation of Xist at the level of RNA processing, possibly splicing, may play a role in the stochastic choice of which X chromosome is silenced.

X^{ΔA}X cells undergo primary non-random X-inactivation, similar to what is observed in other heterozygous loss-of-function *Xist* mutant ES cells (Marahrens et al. 1997) (Marahrens et al. 1997; Csankovszki et al. 1999; Sado et al. 2005). While other mutations abolish or prematurely terminate *Xist* transcription, deletion of the A-repeat skews choice without eliminating *Xist* transcription. Therefore, events downstream of *Xist* transcription must be important for random choice.

The A-repeat constitutes one of the most conserved regions of *Xist*. This element is conserved in primary sequence, relative position and repeat unit copy number across species (Wutz et al. 2002). Mutational analysis within this element indicates that its predicted secondary structure, and not its primary sequence, is important for silencing (Wutz et al. 2002). Our data show that ASF/SF2's interaction with the A-repeat is sequence-dependent and appears to be independent of predicted secondary structure. In contrast, a second A-repeat binding protein Ezh2 appears to recognize the A-repeat via its predicted secondary structure (Zhao et al. 2008). Perhaps the A-repeat is so highly

conserved because the silencing, Ezh2 binding and ASF/SF2 binding activities need to be simultaneously maintained.

Chromatin modifications at Xist and Tsix do not correlate with fate in ΔA cells

In heterozygous *Tsix* mutant ES cells, the *Xist/Tsix* locus on the mutant chromosome was reported to be enriched for H3K27me3 and depleted for H3K4me2 and H4ac (Sun et al. 2006). Because the *Tsix* mutant X is always silenced upon differentiation, it was proposed that the acquisition of these modifications provides a preemptive mark that directs non-random X-inactivation. Deletion of the A-repeat also causes non-random X-inactivation, but our data indicate that the X chromosome that will be silenced in these cells does not acquire H3K27me3 enrichment, depletion of H3K4me2 or histone H4 hypoacetylation. Thus, these epigenetic alterations on the future Xi are not required to predetermine the fates of the X chromosomes. While differential modification of the *Xist/Tsix* region is not necessary for primary non-random X-inactivation in $X^{\Delta A}X$ cells, modifications at other sites may correlate with fate in both *Xist* and *Tsix* mutant cells. One candidate for such a site is the H3K27me3 hotspot upstream of the *Xist* promoter (Heard et al. 2001; Rougeulle et al. 2004). However, analysis of this site did not reveal any allelic bias in H3K27me3, H3K4me2 or H4ac in $X^{\Delta A}X$ or $X^{TpA}X$ mutant ES cells (Figures 15A, B)

Why is H3K27me3 enriched on the future Xi in *Tsix* mutants but not in the ΔA mutant? H3K27me3 is mediated by Polycomb Repressive Complex 2 (PRC2) (Cao and Zhang 2004), which is recruited by Xist RNA (Plath et al. 2003; Silva et al. 2003; Kohlmaier et al. 2004; Zhao et al. 2008). It is possible that the increase in Xist RNA that

occurs on *Tsix* mutant chromosomes in ES cells (Lee and Lu 1999; Luikenhuis et al. 2001; Morey et al. 2001; Shibata and Lee 2003) results in increased recruitment of PRC2. Consistent with this possibility, male *Tsix* mutant ES cells exhibit increased Xist RNA levels and an increase in H3K27me3 at the *Xist/Tsix* locus (Navarro et al. 2005; Shibata et al. 2008).

Regulation of Xist and Tsix RNA levels by the A-repeat

The lack of spliced Xist RNA and elevated levels of spliced Tsix RNA from A-repeat mutant chromosomes suggest that this element is necessary for normal metabolism of both non-coding RNAs. *Xist* and *Tsix* produce convergent and overlapping transcripts that can recruit chromatin-modifying activities and, therefore, have the potential to influence the other's transcription, processing and turnover. Because *Xist*, including the A-repeat, is encompassed within *Tsix*, any mutation of *Xist* also alters *Tsix*. As a result, it is difficult to definitively establish whether the alterations in Xist and Tsix RNA levels seen upon deletion of the A-repeat are a consequence of the absence of this sequence in *Xist*, *Tsix*, or both. Below we discuss two possibilities: that the A-repeat regulates *Xist* expression directly and *Tsix* expression indirectly and vice versa.

The possibility that Xist RNA processing is directly regulated by the A-repeat is supported by the fact that aberrantly spliced Xist transcripts can be detected in A-repeat mutant cells and that the A-repeat binds ASF/SF2. In other instances of splicing regulation by ASF/SF2, mutation of the ASF/SF2 binding site affects splice site selection at the exon containing the ASF/SF2 binding site and processing of other exons is unaffected (Cartegni et al. 2002). In contrast, deletion of the A-repeat disrupts processing

of the entire *Xist* primary transcript, suggesting a novel *cis*-regulatory role for this highly conserved element. Why might such regulation have evolved? The production of spliced *Xist* transcripts faces two hurdles: the very large size of *Xist*'s first exon and the presence of an antisense transcript. The position of the A-repeat at the 5' end of the first exon may aid proper processing of this unusual RNA.

It is also possible that the A-repeat does not directly participate in *Xist* RNA processing, but rather regulates *Xist* RNA metabolism indirectly through *Tsix*. There is evidence that *Tsix* regulates *Xist* as loss-of-function *Tsix* mutations cause a five-to-ten-fold increase in spliced *Xist* RNA abundance (Morey et al. 2001; Shibata and Lee 2003; Navarro et al. 2005; Sado et al. 2006). Perhaps the increase in *Tsix* transcripts in ΔA cells blocks accumulation of spliced *Xist* RNA *in cis*. Elucidating the mechanisms by which *Xist* and *Tsix* RNA regulate each other will be important for understanding how these non-coding transcripts ensure that X-inactivation is random.

Model for stochastically assigning chromosomal fates during random X-inactivation

Any model accurately describing how *Xa* and *Xi* fates are assigned in wild-type female cells must meet two criteria. It should account for how the two X chromosomes adopt mutually exclusive fates and it should explain why each X chromosome has an equal probability of becoming the *Xi*. Most models for stochasticity and mutual exclusivity posit that *trans*-acting factors asymmetrically assemble on the two X chromosomes in each female cell. As a result of this asymmetric distribution one X is marked to become the *Xa* and the other to become the *Xi*. Transgenic experiments have identified regions of the *Xic* that, when introduced onto an autosome, trick male ES cells into behaving as if

they have two X chromosomes. The minimal transgenes that function in this manner contain *Xist* and flanking sequences, suggesting that *trans*-acting factors asymmetrically assemble at a site within or closely linked to *Xist*. Deletion of the sequences contained in these transgenes does not perturb mutual exclusivity in female cells, suggesting either redundancy or an alternate mechanism. Here, we propose an alternative mechanism in which *Xist* RNA, and not a DNA element, is central to establishing randomness and mutual exclusivity.

In our model, stochastic differences in *Xist* RNA levels between the two X chromosomes in each female ES cell underlie stochastic choice during random X-inactivation. *Xist* RNA is an attractive candidate as a factor that underlies randomness for two reasons. First, the *Xist* gene and its promoter are contained in all transgenes that are capable of tricking male cells into ectopically inactivating their sole X chromosome. In addition, there are approximately 10 copies of *Xist* RNA per female ES cell (Zhao et al. 2008). At such low levels, variability in *Xist* RNA production from the two X chromosomes in each cell is likely. Such differences could be introduced if *Xist* has a slight transcriptional advantage on one X chromosome. Stochasticity could also arise at the level of splicing or formation of a silencing competent *Xist* ribonucleoprotein (RNP) complex.

Once established, any stochastic difference could be amplified by a feedback loop that depends on the mutual regulation of *Xist* and *Tsix*. The feedback between *Xist* and *Tsix* ensures that the two X chromosomes robustly adopt mutually exclusive fates by amplifying any small difference in the amount of *Xist* RNA or RNP produced from each chromosome. In our model, the X chromosome that produces less *Xist* RNA or RNP

becomes the X_a upon differentiation. Assessing which chromosome produces less Xist RNA could occur when the two X chromosomes undergo homologous pairing (Bacher et al. 2006; Xu et al. 2006; Augui et al. 2007). Pairing may also be a mechanism that restricts X-inactivation to cells with more than one *Xic* (Wutz and Gribnau 2007). In such an RNA-based model, stochasticity of Xist production, feedback and pairing ensure that the X chromosomes in XX cells randomly and reliably adopt mutually exclusive fates.

RNA-based models can also explain how heterozygous *Xist* and *Tsix* mutations cause primary non-random X-inactivation without perturbing a cell's ability to designate one X_a and one X_i. Heterozygous loss-of-function mutations of *Xist* do not produce normal Xist RNA and will by default produce less Xist RNA/RNP. Likewise, Xist transcripts expressed from *Tsix* mutant chromosomes have a competitive advantage over wild-type transcripts within the same cell because the broken feedback loop on the mutant chromosome prevents positive feedback. Our model may also explain why homozygous *Tsix* mutant ES cells silence both X chromosomes at high frequency (Lee 2002). When the feedback loops on both chromosomes are broken due to the absence of *Tsix* RNA, small differences in Xist RNA levels are not amplified, impairing the mechanism that ensures mutual exclusivity.

While the roles of *Xist* and *Tsix* in randomness are well established, the molecular framework in which they function remains elusive. There are many possible ways that random choice could be achieved. We propose that stochastic regulation of the production of Xist RNA provides a mechanism that couples randomness with mutual exclusivity.

VI. Figure Legends

Figure 9. Targeted deletion of the A-repeat in female ES cells. (A) Targeting scheme for generating the $X^{\Delta A}$ allele. Positions of the relevant restriction enzyme sites and probes used in B and C are indicated. Triangles represent *loxP* sites. H, HindIII; D, DraI. (B) Homologous recombination in ES cells confirmed by Southern blot hybridization. Probes and restriction enzymes used are shown below and to the left of the blot, respectively. Correct band sizes are indicated on the right. (C) Excision of the puromycin-resistance cassette by Cre-recombinase was confirmed by Southern blot hybridization. Probes and restriction enzymes used are shown below and to the left of the blot, respectively. Correct band sizes are indicated on the right.

Figure 10. $X^{\Delta A}X$ cells undergo primary non-random X-inactivation. (A) Allele-specific RT-PCR for spliced Xist RNA (exon1-3) in wild-type and $X^{\Delta A}X$ cells at 0, 6 and 12 days of differentiation. (B) Genomic structure of *Xist*, indicating positions of the A-repeat and the tetO arrays in $X^{\text{tetO}}X$. FISH probes for A-repeat (red), tetO (red) and *Xist* exon 1 (green) are shown below. (C) RNA FISH for exon 1 (green) and the A-repeat (red) in male ES cells expressing inducible *Xist* (Tet-*Xist*) or inducible *Xist* lacking the A-repeat (Tet- ΔA) showing that the A-repeat probe only detects wild-type Xist. DAPI stained nuclei are blue. (D) Allele-specific FISH in differentiated $X^{\text{tetO}}X$ (left) and $X^{\Delta A}X$ (right) ES cells. tetO DNA FISH probes and A-repeat RNA FISH probes were used to identify the tetO and wild-type alleles, respectively, while exon1 probes identified all Xist transcripts. DAPI stained nuclei are blue, the A-repeat or tetO array is in red, and exon 1

is in green. Inset numbers indicate the percentage of cells with the pattern shown for each cell type. (E) Allele-specific FISH in $X^{\text{tetO}}X$ (*left*) and $X^{\Delta A}X$ (*right*) cells showing both Xist RNA coating and Xist/Tsix pinpoint expression. In the merged image, nuclei are blue, exon 1 is in green, and tetO or the A-repeat is in red. Numbers inset in the merged panel indicate the percentage of cells with the pattern shown. (F) Survival assay measuring the competitiveness of wild-type (*left*) and $X^{\Delta A}X$ (*right*) cells when co-differentiated with GFP-expressing male ES cells. Green bars indicate the % GFP negative cells at 4, 6, 8, 10 and 12 days of differentiation. Purple bars indicate the % of female cells exhibiting Xist RNA coating at each time point. At least 100 cells were counted for each time point in each replicate. Bars indicate one standard deviation. (G) FISH probes used for allele-specific FISH in $X^{\text{TpA}}X$, $X^{\Delta Xist}X$ and $X^{\Delta A}X$ ES cells. *Xist* is indicated in black, the A-repeats in purple, and *Tsix* in gray. Position and names of probes are indicated beneath the map, and primers used to generate probes are indicated in Table 3. (H) Allele-specific FISH for Xist/Tsix RNA in $X^{\Delta A}X$ cells, using A-repeat (red) and exon 1 (green) probes. The wild-type allele exhibits either a singlet (*left*) or doublet (*right*) FISH signal in the percentage of cells indicated. (I) Table summarizing results of allele-specific FISH in $X^{\Delta Xist}X$, $X^{\text{TpA}}X$ and $X^{\Delta A}X$ ES cells. The proportion of cells with SD FISH signals in which the future Xa and future Xi display the singlet signal are indicated. (n) indicates number of nuclei scored. p-values indicate that the frequency with which the mutant X chromosome in $X^{\Delta A}X$ ES cells exhibits a singlet FISH signal is comparable to that of the future active X in *Xist* mutant ES cells.

Figure 11. $X^{\Delta A}X$ cells do not show allelic enrichment for histone modifications. (A) Genomic structure of *Xist* (above line) and *Tsix* (below line) with location of PCR amplicons at *Xist* promoter (1), *Xist/Tsix* gene body (2) and *Tsix* promoter (3) indicated. (B) Allele-specific ChIP showing distribution of H3K27me3, H3K4me2, H4ac between the *129* and *cas* chromosomes in wild-type, $X^{TpA}X$ and $X^{\Delta A}X$ cells. Both *TpA* and ΔA mutations are on the *129* chromosome. (C) Data from (B) plotted as the fraction of products arising from the *129* allele. Fractions represent the average of at least three ChIP experiments and error bars indicate one standard deviation.

Figure 12. Reduced levels of correctly spliced *Xist* transcripts in $X^{\Delta A}Y$ cells. (A) RT-qPCR for *Xist* in male cells mixed with 0%, 0.2%, 6%, or 100% of differentiated female cells. All values are normalized to male cells alone which was set to one. (B) RT-PCR for spliced *Xist*, *Tsix* and β -actin in wild-type (XY) and ΔA ($X^{\Delta A}Y$) cells. For *Xist*, the amplicon is indicated (C) RT-qPCR for *Tsix* in XY and $X^{\Delta A}Y$ cells. Error bars indicate one standard deviation. (D) Strand-specific RT-PCR for *Xist* with and without primers indicate that *Xist* and *Tsix* can prime each other. (E) Diagram outlining strand-specific RT-PCR procedure. First strand synthesis is carried out using a specific primer with a T3 anchor, and PCR is carried out using a second specific primer and a T3 primer. (F) Strand-specific RT-PCR for *Xist* and β -actin in XY(wt) and $X^{\Delta A}Y$ (ΔA) ES cells. Numbers at right indicate the ratio of wt/ ΔA transcripts calculated from (G). (G) Strand-specific RT-qPCR for *Xist* and β -actin in XY and $X^{\Delta A}Y$ ES cells. (H) RT-PCR across *Xist* intron 4. Spliced and unspliced transcripts are indicated. (I) RT-PCR for *Xist* (3-6) showing a smaller than expected product. (J) Diagram showing splicing patterns from

Xist exons 3-6 in XY (*black*) and X^{ΔA}Y (*gray*) ES cells. 3^{ΔA} and 4^{ΔA} indicate the aberrant splicing in X^{ΔA}Y cells. (K) Sequence of the ΔA exon 3-6 RT-PCR product. *Xist* exons are in upper case and intronic sequences are lower case. Black sequences are absent in the ΔA RT-PCR product. (L) RT-PCR using primers in *Xist* exons 1 and 4 or exons 1 and 5 showing aberrantly sized products. (M) Allele-specific RT-PCR for spliced (exon 3-6) *Xist* RNA in undifferentiated XX and X^{ΔA}X ES cells and control *129* or *cas* cells. Size of *129* and *cas* bands indicated on the right. (N) Allele-specific RT-PCR for total (exon 1) *Xist* RNA in undifferentiated XX and X^{ΔA}X ES cells. Sizes of wild-type and ΔA bands are indicated on the right.

Figure 13. The A-repeat binds ASF/SF2. (A) Predicted hairpin-loop structure of the A-repeat unit consensus (Wutz et al. 2002). (B) Gel shift assay using ES cell or HeLa cell nuclear extracts identified two A-repeat interacting complexes, I and II. Both complexes were competed away by excess cold A-repeat RNA, but not control MS2 hairpin RNA. * indicates unbound radiolabeled A-repeat RNA. (C) SDS-PAGE gel of material UV cross-linked with radiolabeled A-repeat RNA. Sizes of molecular weight markers are indicated on the right. (D-F) Partial purification of A-repeat binding activities followed by gel mobility shift. (D) HeLa cell nuclear extracts were fractionated using ammonium sulfate ((NH₄)₂SO₄) at percentages indicated. The 60-70% and 70-80% fractions had a 7-fold increase in specific activity. (E) 60-70% and 70-80% (NH₄)₂SO₄ fractions were each incubated at 80°C for 15 min, cooled on ice, and pelleted. There was a roughly 8-fold enrichment of binding activity in the resulting supernatant. (F) Supernatants from (E) were subject to Q-Sepharose batch affinity chromatography, and material eluted with KCl

at increasing concentrations. Binding activity was enriched approximately 50-fold in the flow through. (G) Silver-stained SDS-PAGE gel of partially purified extract. Molecular weight markers are indicated on the right. The bands indicated were excised and identified by mass spectrometry. The two bands indicated both contained ASF/SF2, which migrates as two bands because a fraction is phosphorylated (Xiao and Manley 1997). (H) Immunoprecipitation of material cross-linked with radiolabeled A-repeat RNA using ASF/SF2 or aldolase A antibodies. Supernatant is indicated by sup. (I) RT-PCR for Xist on RNA that co-IPd with GFP-tagged ASF/SF2 or GFP alone. Input (I), Supernatant (S) and Pellet (P) are indicated.

Figure 14. Binding of ASF/SF2 to the A-repeat depends on ASF/SF2 consensus sequences and not secondary structure. (A and B) The sequences of the 7.5 repeat units from mouse (A) and human (B) A-repeat. The consensus repeat unit is indicated by *. The sequences highlighted in red were identified by ESEfinder as high score motifs corresponding to an ASF/SF2 consensus (Cartegni et al. 2003). (C) Summary of A-repeat mutations assayed by gel mobility shift. The line on top indicates predicted stems (straight lines) and loops (arcs). Stem-loop 1 is purple and stem-loop 2 is orange. The WT A-repeat consensus sequence is written in full. Blue (site 1) and pink (site 2) bases delineate the regions with similarity to the ASF/SF2 consensus sequence. Mutated sequences (A-H) are shown. All sequences are relative to the WT consensus sequence. Dashes represent unchanged bases, letters indicate substitutions and spaces represent deletions. Columns on the right indicate whether each A-repeat mutation maintains (+) or disrupts (-) complex II formation (binds), Xist RNA mediated silencing (Wutz et al.

2002), base pairing within the predicted stems (stem1, stem2), or the ASF/SF2 consensus sites (site1 or site2). ND indicates not done. * indicates mutations that do not alter the sequence away from the consensus, as predicted by ESEfinder (Cartegni et al. 2003). (-) indicates a mutation that is predicted to perturb silencing. (D-F) Autoradiograms of (D, E) UV-crosslinking or (F) gel mobility shift assays for A-repeat mutantions. * indicates unbound RNA. Complex I and/or II are indicated.

Figure 15. ChIP analysis of the H3K27me3 hotspot. (A) Allele-specific ChIP in female ES cells showing distribution of H3K27me3, H3K4me2, H4ac between the *129* and *cas* chromosomes in wild-type, $X^{TpA}X$ and $X^{\Delta A}X$ cells at the H3K27me3 hotspot. Both TpA and ΔA mutations are on the *129* chromosome. (B) Data from (A) plotted as the fraction of products arising from the *cas* allele.

Table 3. Primers used in Chapter 2

assay	amplicon	name	sequence	enzyme	
Allele-specific ChIP	Xist promoter	XP-3F	CGTCATGTCACTGAGCTTAC	Tsp509I	
		*X1-17R	GAGAAACCACGGAAGAACCG		
	Xist exon 7	*X7-7F+2	TCCTTTCTGTTCACTTTGAGC	MnII	
		X7-10R	GCGGTTCACTTCAGAGCCACTTG		
	Tsix promoter	*E-f	CTTCAAACCTCGCAAAGCTCT	AseI	
		E-r	GTCTGCCTACTAACACAGGT		
	Hotspot	*HotspotP-6f	GGGAAAGGTTCTCCATCTT	MseI	
		HotspotP-6r	GGAAAACCACATCTTGGTTGA		
RT-PCR	Xist (ex1-3)	NS66	GCTGGTTCGTCTATCTTGTGGG	ScrFI	
		NS33	CAGAGTAGCGAGGACTTGAAGAG		
	Xist (ex3-6)	mx23	ATGTTGATCCTCGGGTCATTTAT		
		mx20	ACTGCCAGCAGCCTATACAG		
	Tsix (ex3-4)	21b80F	CCTGCAAGCGCTACACACTT	ScrFI	
		AA51	TCCATAAGGACGTGAGTTTCGC		
	β-actin	β-actin F	GGCCAGAGCAAGAGAGGTATCC		
		β-actin R	ACGCACGATTTCCCTCTCAGC		
	Xist (ex1-4)	NS66	GCTGGTTCGTCTATCTTGTGGG		
		mrt63	CCCAGTGGTGGTGAGCTATT		
	Xist (ex 1-5)	NS66	GCTGGTTCGTCTATCTTGTGGG		
		mrt64	AGAATGGCTTCCTCGAAGGT		
strand-specific RT-PCR	Xist (exon 1)	*T3	AATTAACCCCTCACTAAA		
		mrt45	CCTGCAAGGGATACCGTTTA		
		mrt46/T3	AATTAACCCCTCACTAAAGGGTTCAAGTGCACAGAGCAGGT		
	Xist (ex1-in1)	NS66	GCTGGTTCGTCTATCTTGTGGG	ScrFI	
		AA34/T3	AATTAACCCCTCACTAAAGGGGTGGGTGGCAGACAAAAGAACTCGAATG		
	Xist (in6-ex7)	mrt70	TTGCATGACTAGGCCATTTG		
		mrt71/T3	AATTAACCCCTCACTAAAGGGGAAATAGCACACCCACAATACACA		
	Xist (exon 7)	mrt72	GCATACTGCCATCCTCCCTA		
		mrt73/T3	AATTAACCCCTCACTAAAGGGCTGCTCACCTAAGCCCAAAG		
	Xist (ex4-5)	mrt119	ATCCTCCAGGGGAATAGCTC		
		mrt120/T3	AATTAACCCCTCACTAAAGGGGTGGCATGAGTAGGGTAGCAG		
RNA IP	Xist A-repeat	AA19	CATCGCCCATCGGTGCTTTTTATGG		
		AA20	CTAAGCCGAGTTATGCGGCAAGTCT		
Southern	probe 1	DP2a	CATGGGTGCTATCGCCCCAGGTCAC		
	probe 2	DP2b	AAGGCTAGCCTGGGTTATATGCTAA TATTAATAGTAATCAATTAC GCGGCCGAGATTATATAAAC		
No primer RT	Xist (ex1-in1)	AA47	ATGAGAAAAAGATAGCTAG		
		AA34	GTGGGTGGCAGACAAAAGAACTCGAATG		
		NS66	GCTGGTTCGTCTATCTTGTGGG		
Cloning	5' targeting arm	5'HR-forward	AACTATTTATGTGAATGTCATTAG		
		5'HR-reverse	AGATTATATAAAACAATGAAAGAAAGG		
	3' targeting arm	3'HR-forward	GTGGATGAAAATATGGTTTGTGAG		
		3'HR-reverse	AGCAATAGCAGCAGCACTATTTGC		
qPCR	tsix	mrt104	CCAAGCAGCAGAAAAGATTCC		
		mrt105	AAGGACGTGAGTTTCGCTTG		
	β-actin	β-actin F	GGCCAGAGCAAGAGAGGTATCC		
		β-actin R	ACGCACGATTTCCCTCTCAGC		
FISH probes	Xist ex1-pooled	ex1-5a	GTCCATGGACAAGTAAACAAAGAAC		
		ex1-5b	TATGAGGGTATGGGATCTTGGTTA		
		ex1-6a	GATCCCATACCCTCATACCCTAAT		
		ex1-6b	CTTGAAGGACCATTGACCGTATT		
		ex1-7a	TGCTTTATGGAATTATGTATGTGC		
		ex1-7b	GGTCCGAAAAGTAATAAGGTTGTG		
		ex1-8a	ACTTTTCGGACCATTGTATCTCTT		
		ex1-8b	GAGAGCAGGTCATTCTGTCAGAG		
		ex1-9a	TCCCCTGCTAGTTTCCCAATGT		
		ex1-9b	TTTCCACAGACTCATCACCCCTCAG		
		ex1-10a	TTTTAAAAGGTGACTGGATGGTT		
		ex1-10b	TGATGTAACGGAGGAGCAGTAG		
		Tsix probe	X3'1F	AACCACTGCCACATTCCCCTTTTC	
			X3'1R	CCCTCCCGCCCTGGCCAGCACCCCT	

Figure 9

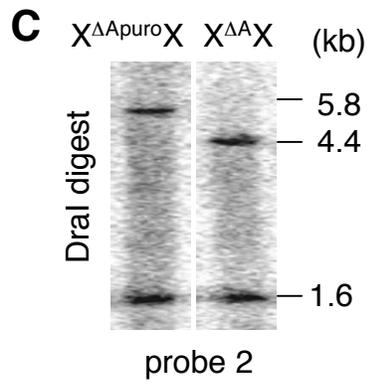
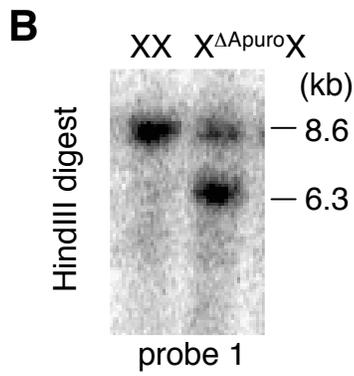
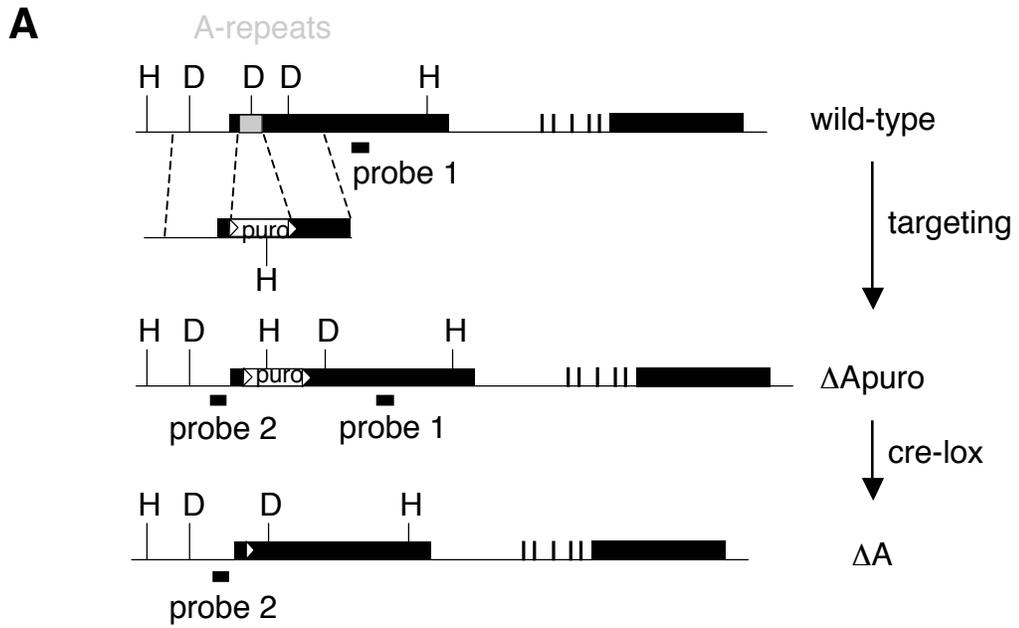


Figure 10

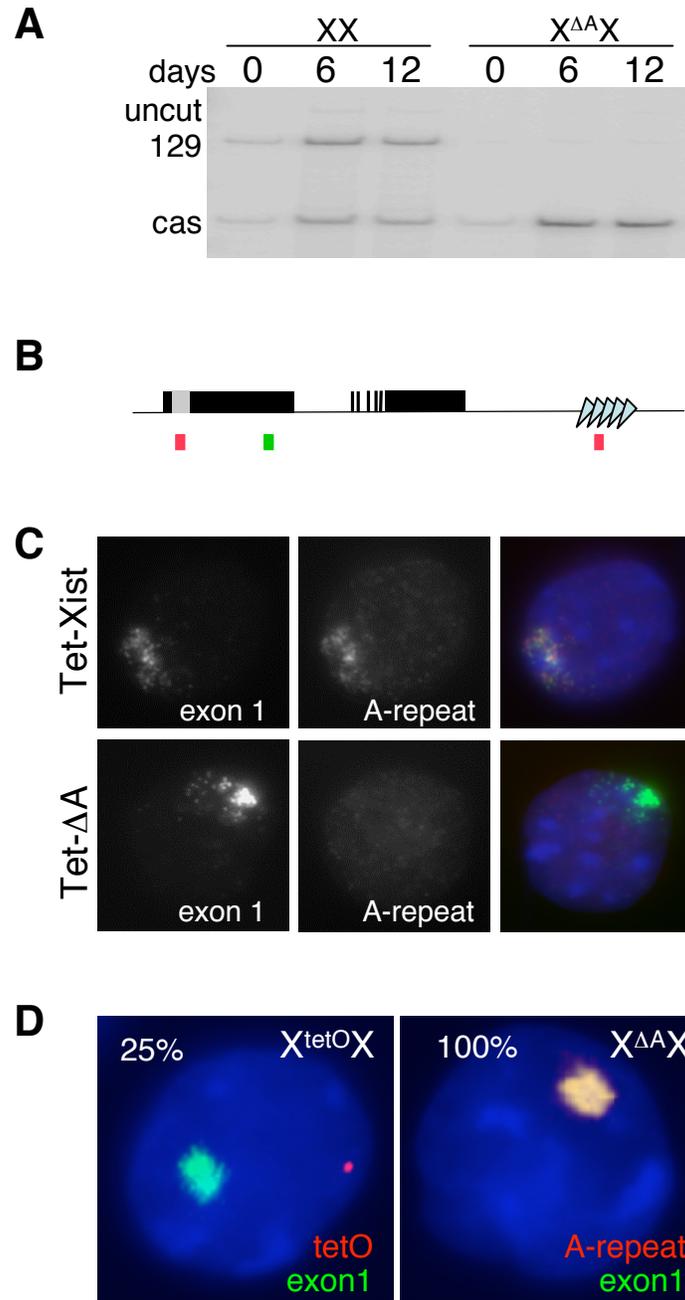
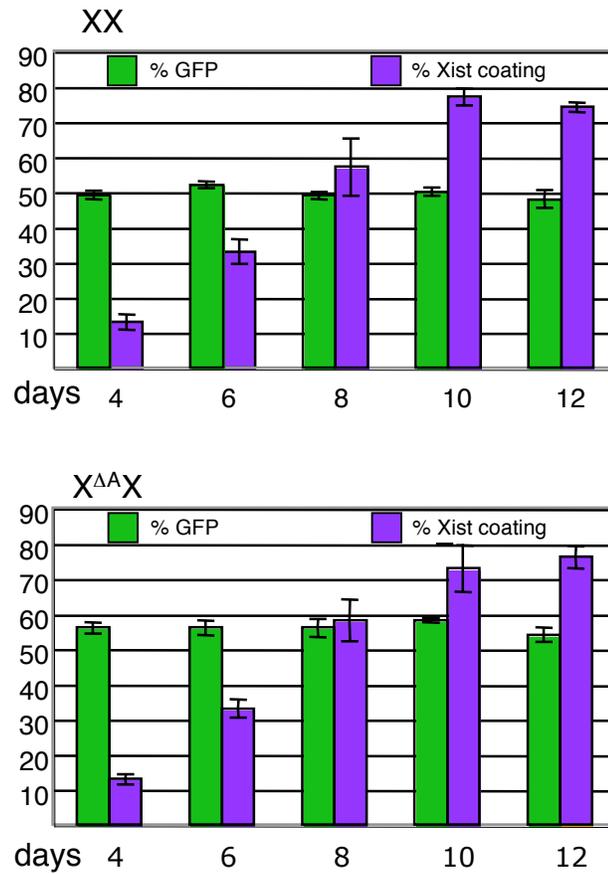


Figure 10 continued

E



F

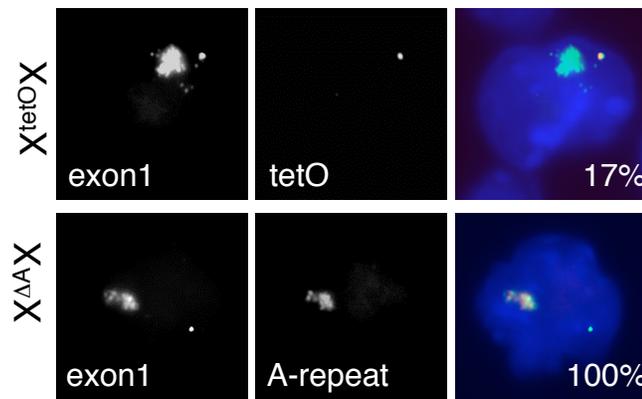
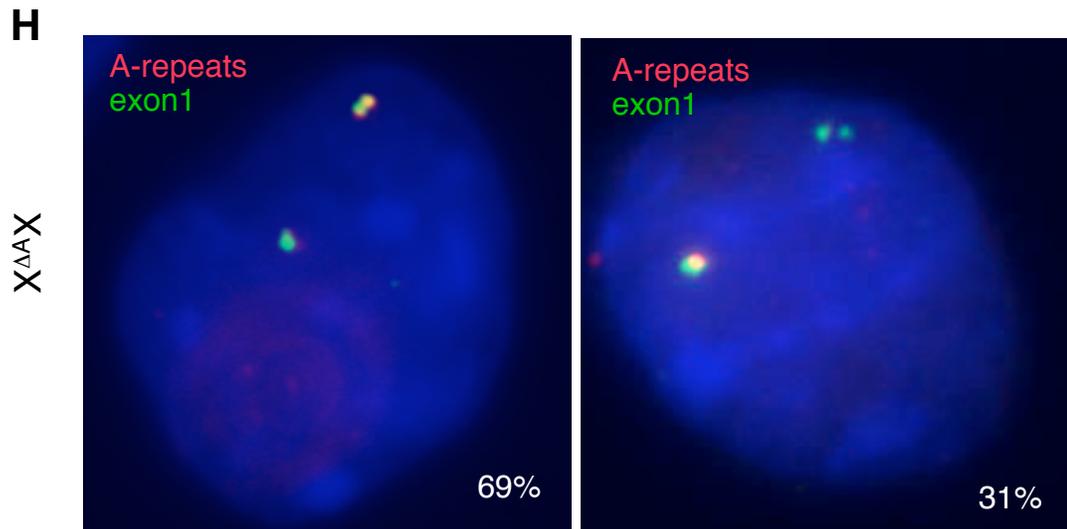
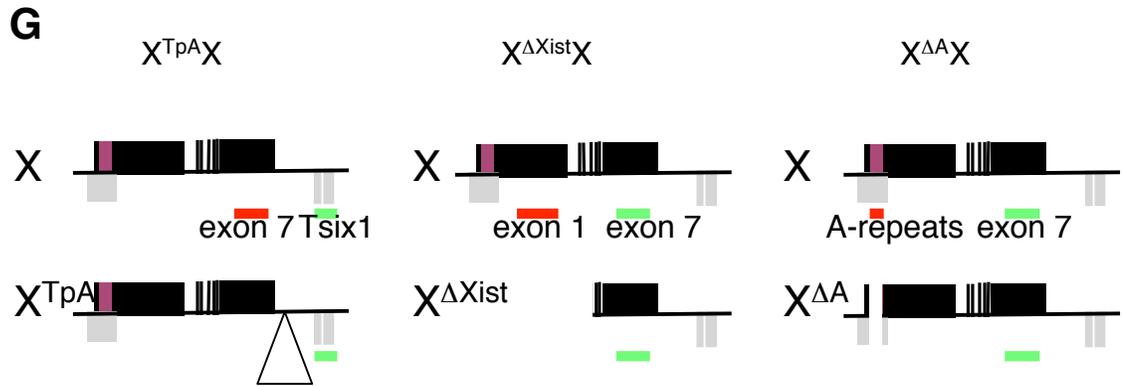


Figure 10 continued



I

cell line	locus	% singlet in SD cells		(n)	p
		future Xa	future Xi		
$X^{TpA}X$	<i>Xist/Tsix</i>	70	30	120	<0.01
	<i>Pgk1</i>	32	68	125	<0.01
$X^{\Delta Xist}X$	<i>Xist/Tsix</i>	69	31	135	<0.01
	<i>Pgk1</i>	33	67	122	<0.01
$X^{\Delta A}X$	<i>Xist/Tsix</i>	69	31	128	<0.01
	<i>Pgk1</i>	34	66	113	<0.01

Figure 11

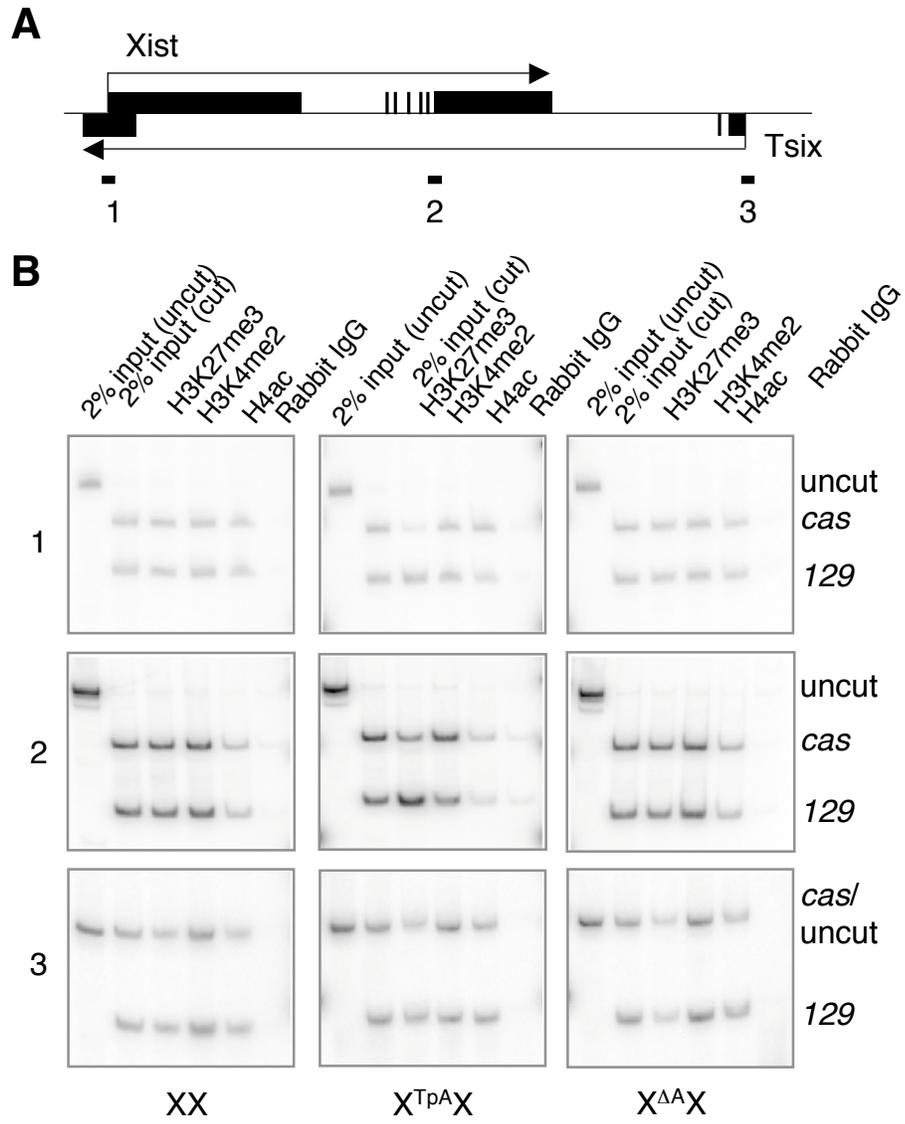
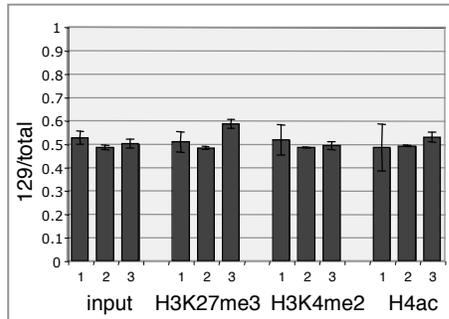


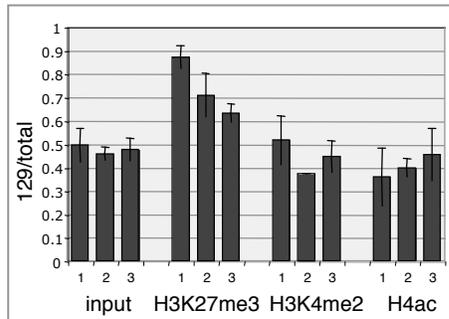
Figure 11 continued

C

XX



X^{Tr}A^X



X^{ΔA}X

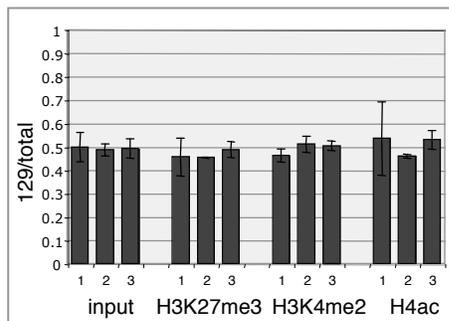


Figure 12

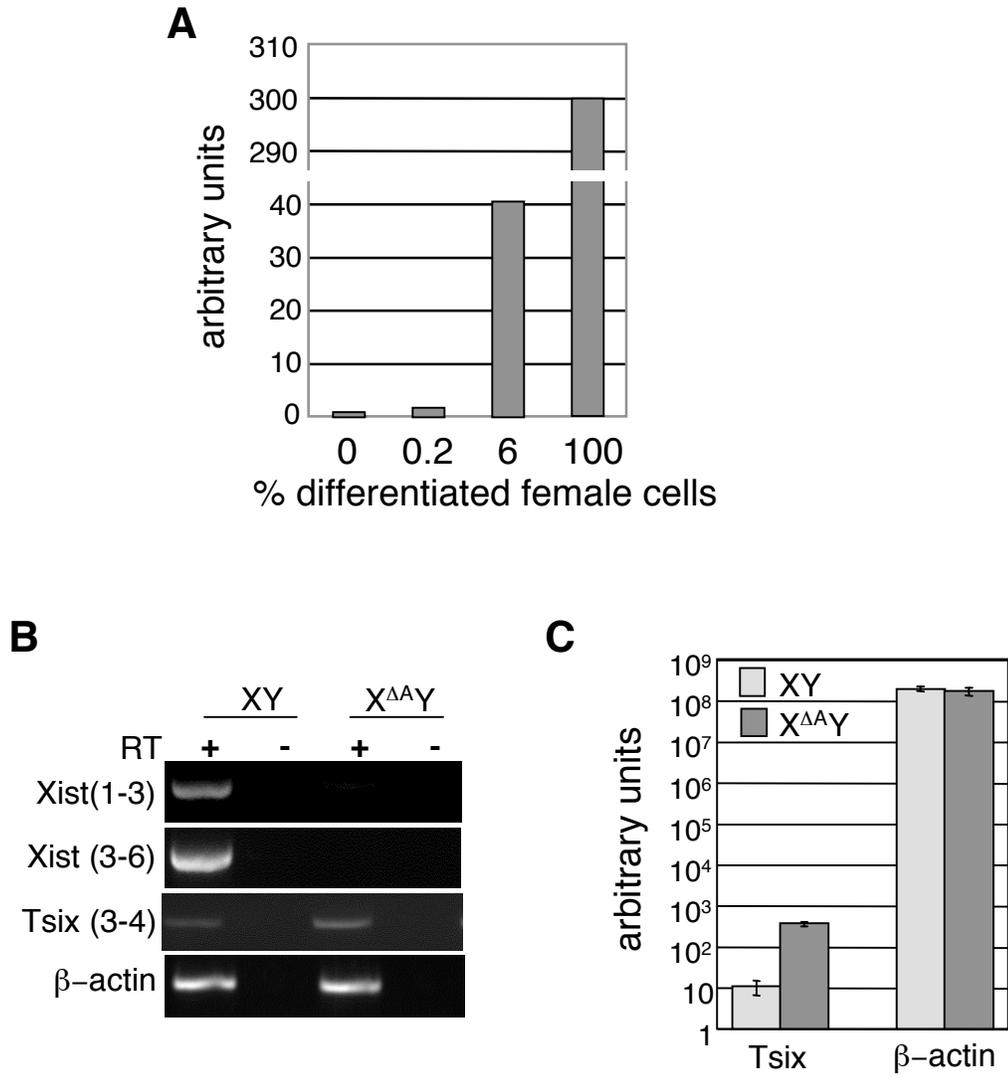


Figure 12 continued

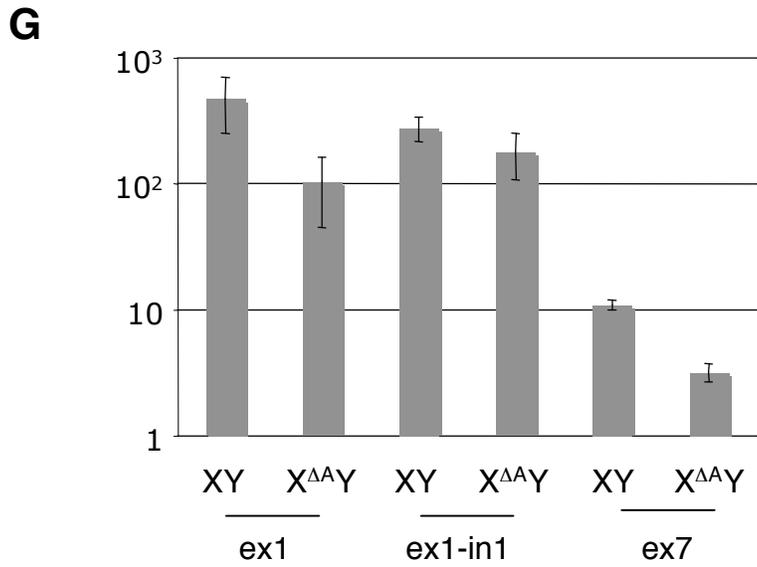
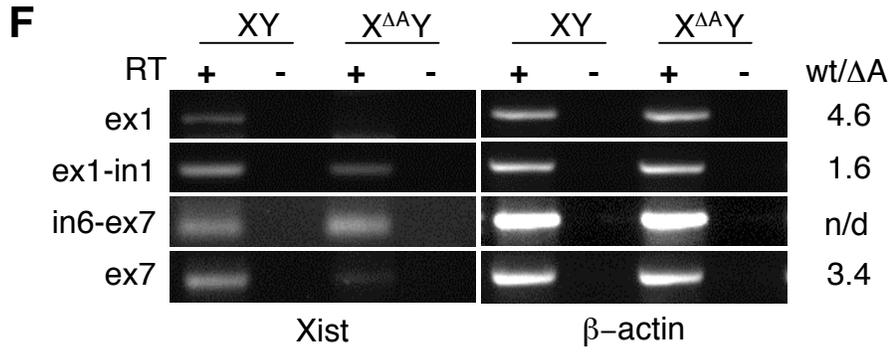
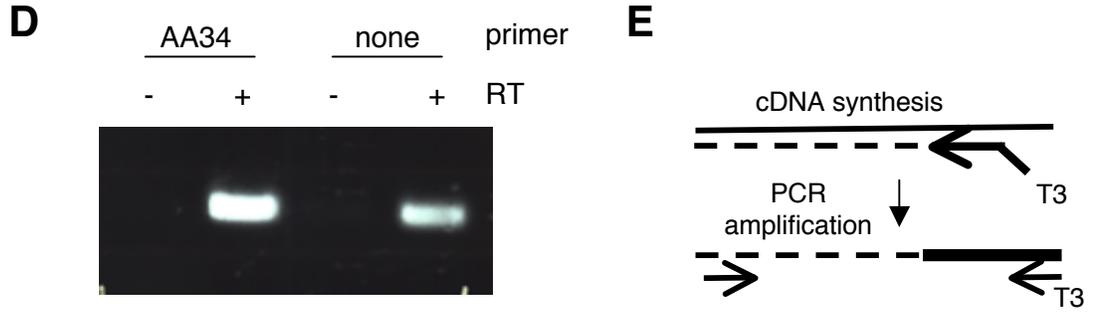


Figure 12 continued

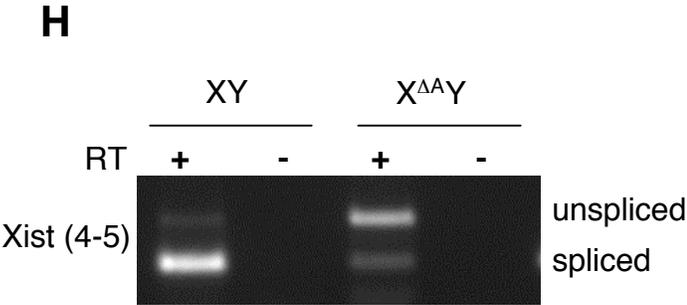
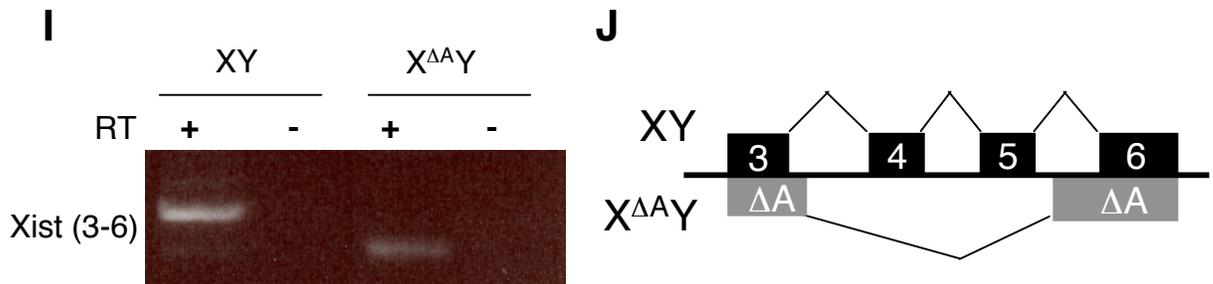


Figure 12 continued



K

ACTGCCAGCAGCCTATACAGCTACATTACATCTCAGCAGAACTTCTCTTC
AAGTCTCGCTACTCTGAACAAAAGCTTACAGGCCACATGGAGAAAAA
AAGgtactttggagaatttgcatttgggggttctatctgatctccagttcatcagagaatgagggggcaag
tcaataaagcactccccatctctgtcccttggcatagtcagttagccaatgagatccttgccttatgatga
ggccttgaatcatcccaccatgttatgaccaatctgtaacatttctgctctgaaatgcctaaaaggtagggt
atcagaggaatgaaggccagtttctgacaccctagctctttgtacttgatagagtaagtagatatggctgtgt
cacattctgtaggcctgtcataattctggtctctgaatatctcaaactcagaggccaatgctacaccagttc
tagggacatatgcaccataagcactgagccaaacaatgctattcacctgccagtggtctggcatggtattgt
ttaaagcacaataaaagaagaaaataatthaacccttattcccagagtattgaagtacaaggacttta
acacaacctataatctaaaggatctctcattggtagggtgacagaataaaagccagaaagaacttgatgg
agcaattctcatgtcttaagttaacactaagatctacaaatgtgaagggtcactctccaagcccctcccca
tgatttaacacaactgctggttactacataatgagacaacttcccagattcttgattcaagaattggcttttctt
tcaccttttccagATCTCCCCCAGAATTGTGGGCTTGCTGCTTTGCAGTGCT
GGCGACCTATTCCCTTTGACGATCCCTAGGTGGAGATGGGGCATGAGG
ATCCTCCAGGGGAATAGCTCACCACCACTGGGCAACAGGCCTAGCCCA
GATTCAGTGAGACGCTTTCCTGAACCCAGCAAGGAAGACAAAGGCTCA
AAGAATGCCACCCTACATCAAAGTAGgtaagtttctgcaacagtgtaataatthtaacttgg
atcttgtttccattaaagtcagcctttacttgagatatatacataaatttatcttttccacatctgggaaatga
actaaacagtgcaaatgtttcttctagGAGAAAAGCTGCTGCAATAGTGGCACTGACC
TTCGAGGAAGCCATTCTGCTCTATTTGGTTCTCTCTCCAGAAGCTAGGA
AAGCTTTGCCAGCTGTTTACATACTTCAAGATGCACTGCTACCCTACTCA
TGCCATATAATACACAAGtgagtagcttattacctacctagctatttacgagtagactgttctgtc
ctcagacacaccagaagagagcaccagatctcattacagatggttctgcccaccatgtggttctgtggg
attgaactcaggacctctggaagagcaatagggtctcttaaccactgagccatcactaacattactgatgat
gaatgcctctaactacaaatggaaacctggctggaattgatacagaagtagtatcattagggcagaatttaa
agagccaggatgttctagacttcttctctgcatcatacagcttctttgttcttttagTGCCATCTACC
AAATATTACCCTTCCCCAAAGCAGCACAGAAAAGTGGGTCTTCAGCGTG
ATCAAGCAATGTGAACACACAAAAGGAAGGCAGCTTTATAAATGACCCG

Figure 12 continued

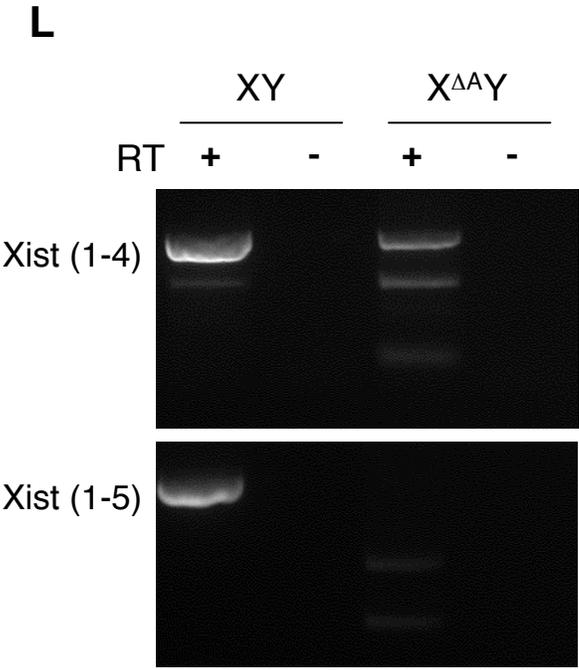
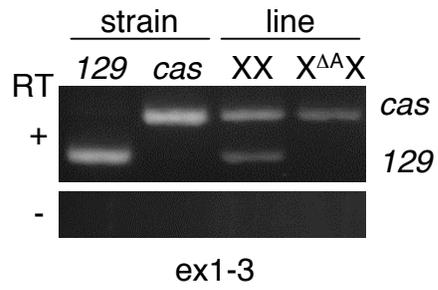


Figure 12 continued

M



N

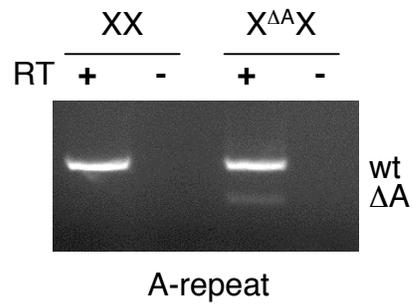


Figure 13

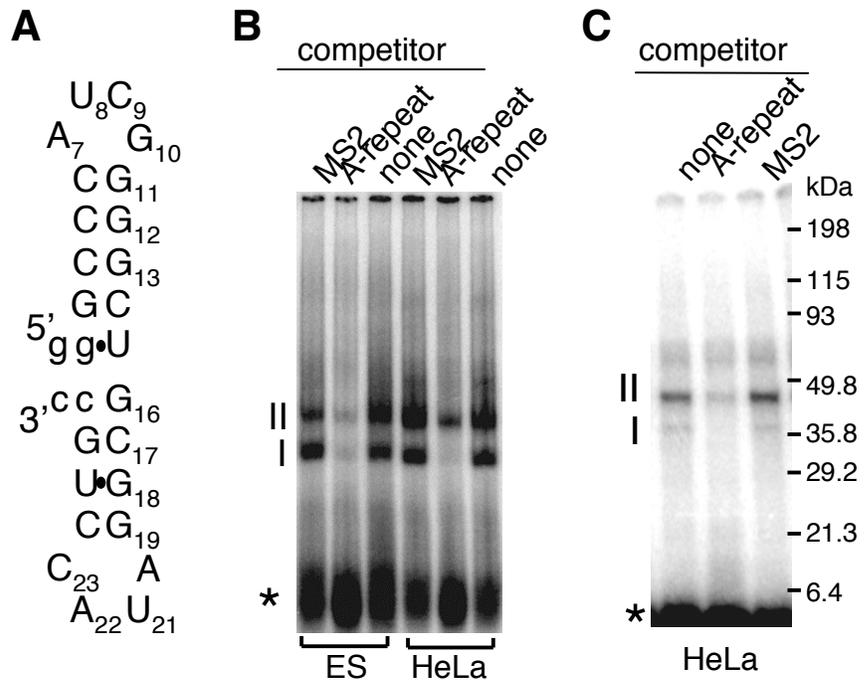


Figure 13 continued

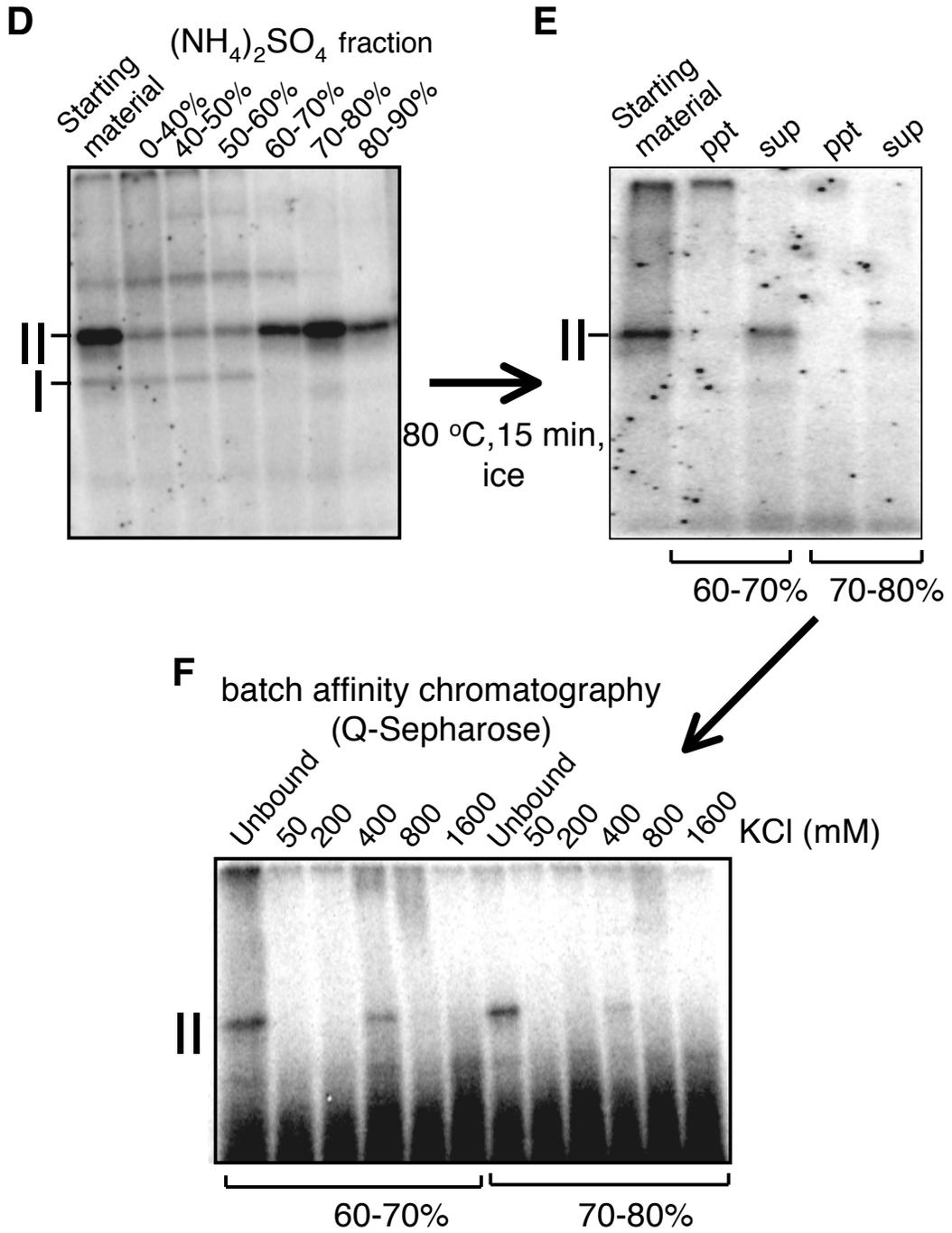


Figure 13 continued

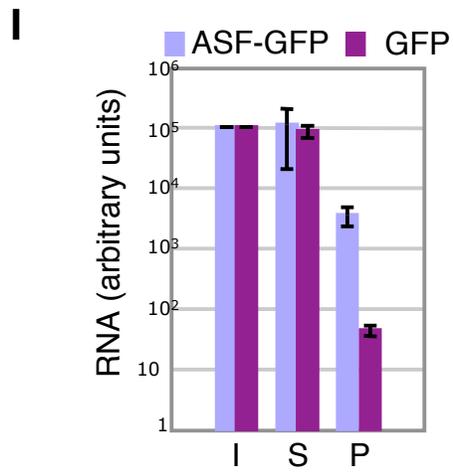
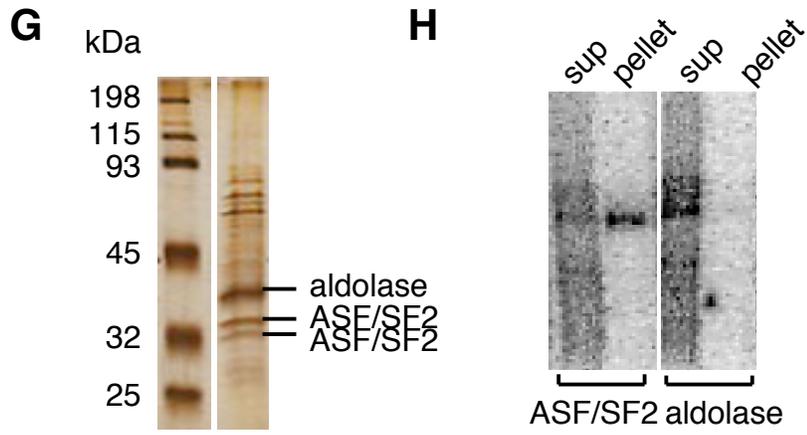


Figure 14

A

```
1 GCCCAUCGGGGCCACGGAUACCUG
2 GCCCAACGGGGUUUUGGAUACUUA
3 GCCCAUCGGGGCUGUGGAUACCUG
4 GCCCAUCGGGGCCAUGGAUACCUG
5 GUCCAUCGGGGACCUCGGAUACCUG
6 GCCCAACGGGGGCCUCGGAUACCUG
7 GCCCAUCGGGGCUGUGGAUACCUG
8 GCCCAACGGGGGCUU
```

B

```
*1 GCCCAUCGGGGCUGCGGAUACCUG
2 GCCCAACGGGGCCGUGGAUACCUG
3 GCCCAUCGGGGCCGCGGAUACCUG
4 GCCCAUCGGGGUAUCGGAUACCUG
5 GCCCAUCGGGGUGACGGAUAUCUG
6 GCCCAUCGGGGCUUCGGAUACCUG
7 GCCCAUCGGGGGCCUCGGAUACCUG
8 GCCCAUCGGGGC
```

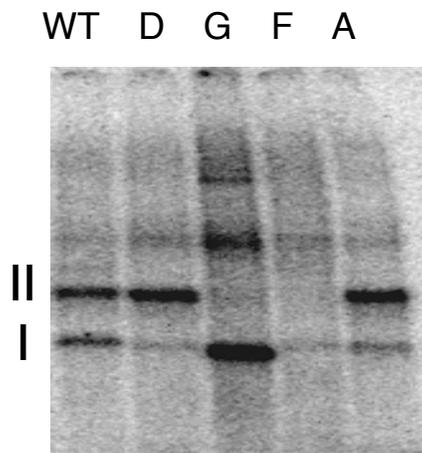
Figure 14 continued

C

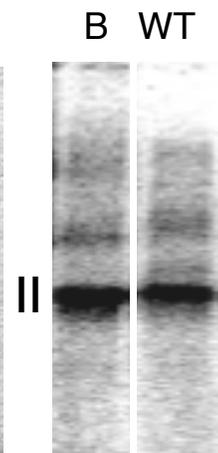


	WT	A	B	C	D	E	F	G	H	I	bindings	site1	site2	stem1	stem2
WT	GCCCAUCGGGGCUGCGGAUACCUGC	--G-----A-----	--G-----A--CU-----	---GAAA-----	-----A-A-----	-----A-A-----	-----	-----AAAA-----	-----AAAA---U-UU	-----GAGA-----	+	+	+	+	+
A											+	-	+*	+	-
B											+	(-)	+*	+	-
C											-	ND	-	+	+
D											+	ND	+*	+	-
E											-	ND	+*	-	-
F											-	-	+	-	+
G											-	ND	+	-	+
H											-	ND	+	-	+
I											+	ND	+	+*	+

D



E



F

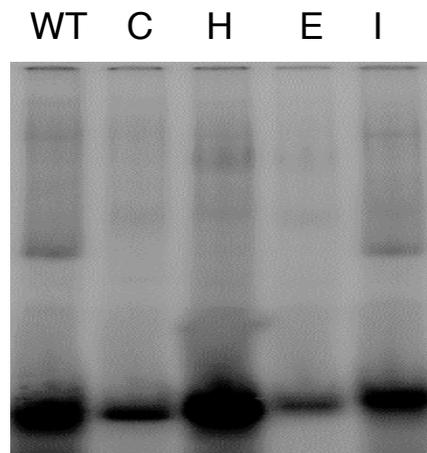
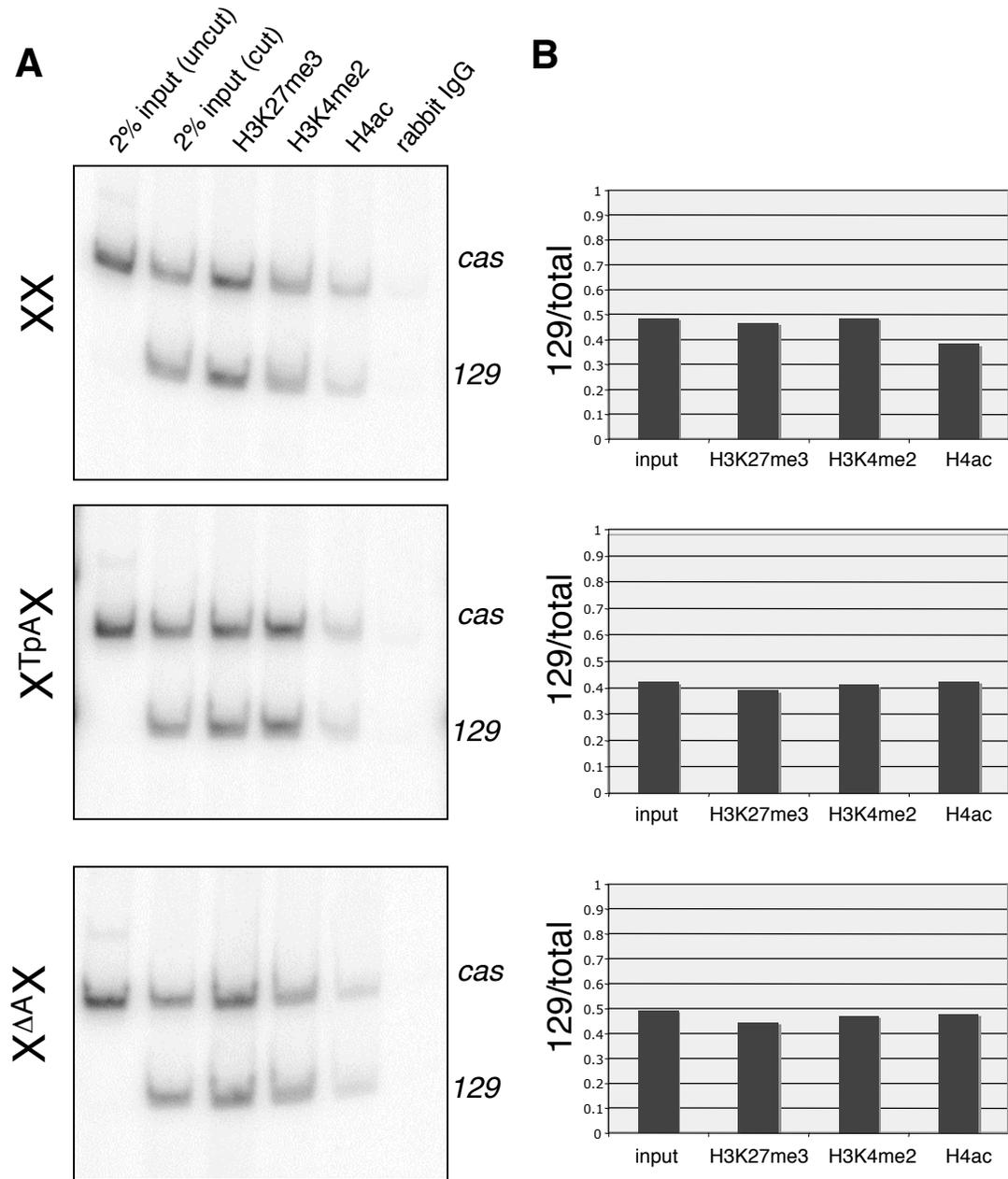


Figure 15



Concluding Remarks

Rethinking X-inactivation

Many models have been proposed to explain random choice during X-inactivation. The foundation for the majority of these models is the appearance of a differentiating “mark” on one X chromosome that designates that X as the active X. In most cases, it is suggested that this “mark” is a *trans*-acting factor such as a protein, however to-date, no proteins have been identified that function in this way. The idea that Xist RNA, an RNA already implicated in the silencing of the Xi, could establish a difference between the two X chromosomes in a female ES cell is intriguing and should be further explored.

Xist Metabolism in ES cells

If a difference in the levels of Xist RNA arising from each X are responsible for random choice of the Xa and Xi, determining how Xist RNA metabolism is regulated in ES cells will help us better understand this mysterious process. Xist transcripts, like most mRNAs, are capped, polyadenylated and spliced (Borsani et al. 1991; Brockdorff et al. 1991) and each of these steps could be subject to regulation. Why Xist transcripts are post-transcriptionally processed is unclear, as they do not contain any open reading frames and never leave the nucleus (Cohen et al. 2005). Perhaps these modifications exist to fine-tune the levels of functional Xist RNA within ES cells. *Xist*'s gene structure may also act to limit its expression. The 17.5 kb mature Xist transcript consists of two very large exons flanking five much smaller exons (Brockdorff et al. 1991). While the significance of the exceptionally large first exon is not clear, this organization is conserved in all species for which *Xist* has been sequenced. This degree of conservation suggests that this

unusual genomic organization has been maintained by evolutionary pressure. It is possible that this large first exon facilitates regulation of transcriptional elongation, which may be developmentally regulated. In support of this idea, fewer *Xist* transcripts are detected at the 3' end of the gene than at the 5' end in ES cells. (Figure 12G).

The data presented in chapter 2 argue that the A-repeat regulates *Xist* RNA production or stability. The exact nature of this regulation is unknown. Is it transcriptional or post-transcriptional? Is it *Tsix*-dependent? Follow-up experiments to address these questions may provide some insight into the molecular mechanisms that regulate *Xist* expression in ES cells.

Does the A-repeat deletion primarily affect Xist or Tsix?

Spliced *Xist* transcripts fail to accumulate in A-repeat mutant cells. In addition, *Tsix* RNA levels in these cells go up. As both *Xist* and *Tsix* contain the A-repeat and these genes have the potential to negatively regulate one another, it is unclear whether the A-repeat directly controls *Xist*, *Tsix* or both. To determine whether the increase in *Tsix* RNA abundance is responsible for the absence of spliced *Xist* transcripts, it would be useful to examine ΔA and wild-type *Xist* RNA accumulation in a *Tsix* mutant background. Male cells should be used for this experiment to simplify the analysis and could easily be made by introducing a polyadenylation signal immediately downstream of the *Tsix* promoter in XY and X ^{ΔA} Y cells. If the A-repeat regulates spliced *Xist* RNA accumulation through *Tsix*, *Xist* RNA levels should be roughly equivalent in both cell lines. Alternatively, if the A-repeat directly affects *Xist* RNA processing or turnover,

spliced ΔA transcripts should not be detected regardless of whether *Tsix* is expressed or not.

Does the A-repeat regulate Xist RNA splicing?

Regardless of whether the A-repeat regulates spliced Xist RNA accumulation through *Xist* or *Tsix*, the observation that spliced ΔA Xist transcripts do not accumulate warrants further examination. The absence of spliced Xist RNA in A-repeat mutant cells could result from a failure to produce spliced ΔA transcripts or because these transcripts are unstable. Unspliced Xist RNA levels appear normal in A-repeat mutant cells, suggesting that the A-repeat is required for post-transcriptional processing of Xist. Because the A-repeat binds the splicing factor ASF/SF2 and because unprocessed and aberrantly processed transcripts are detected in ΔA mutant ES cells, it is reasonable to hypothesize that the A-repeats promote Xist RNA splicing. One way to test this hypothesis would be to replace the *Xist* genomic locus with an *Xist* cDNA in ΔA male ES cells and measure spliced Xist RNA levels. If the A-repeat is necessary for Xist RNA splicing, ΔA cDNA transcripts should accumulate because they are no longer subject to splicing regulation.

Are A-repeat binding proteins important for X-inactivation?

The A-repeat appears to play several roles in X-inactivation. It is necessary for Xist-mediated silencing, is required for Xist and *Tsix* RNAs to be properly expressed in ES cells and is necessary for random X-inactivation. In addition, this element is bound by two proteins: the splicing factor ASF/SF2 and the enzymatic subunit of the histone H3 lysine 27 methyltransferase complex Ezh2. Whether either of these proteins is necessary

for Xist-dependent silencing, Xist RNA processing or random choice is unknown. We have identified a mutation within the A-repeat consensus sequence that abolishes Ezh2 binding but not ASF/SF2 binding (Zhao et al. 2008) and given the degenerate nature of the ASF/SF2 binding site, it seems likely that a mutation could be found that disrupts ASF/SF2 binding but not Ezh2 binding. Replacing the endogenous A-repeat sequence in male cells with these ASF/SF2 or Ezh2 binding deficient sequences would enable us to determine if either protein is important for proper accumulation of Xist and Tsix RNAs. To determine if ASF/SF2 or Ezh2 is necessary for random choice, we could introduce these same mutations into female ES cells and assay whether random choice is affected. These studies have the potential to identify the first protein factor involved in the designation of Xa and Xi fates during random X-inactivation. The classification of ASF/SF2 or Ezh2 as a regulator of choice would nucleate new areas of investigation.

Appendix

Generating an *Xist* homozygous mutant ES cell line

In order to gain insight into the role of *Xist* in female ES cells, I attempted to generate an *Xist* homozygous mutant cell line. Our original motivation for generating this cell line was to determine if *Xist* was required for chromosome-wide coordination of X-linked FISH signals prior to differentiation. Below, I describe two unsuccessful approaches and propose alternate strategies for future attempts.

A heterozygous *Xist* mutant ES cell line, $\Delta Xist/+$ (Csankovszki et al. 1999), was chosen as the starting ES cell line because homozygous *Xist* mutant cells could be made with only one targeting. Two constructs: the ΔA construct described in chapter 2 and a construct made in Takashi Sado's lab that inserts an EGFP cassette into the first exon of *Xist* (pE4.4x6.5DTA) (Sado et al. 2005) were linearized and electroporated into $\Delta Xist/+$ cells according to the standard lab protocol. $\Delta Xist/+$ cells contain a deletion that extends from 5 kb upstream of the *Xist* promoter through intron 3 of *Xist*. As the arms for both targeting constructs are contained within the $\Delta Xist$ deletion, homologous recombination could only occur on the wild-type chromosome. ES cell colonies resistant to 5 $\mu\text{g}/\text{ul}$ puromycin were picked after 11 to 14 days of selection and screened by Southern Blot. More than 1000 potential $\Delta Xist/\Delta A$ and 400 potential $\Delta Xist/Xist/lox$ clones were analyzed, but no correctly targeted clones were identified.

The inability to obtain targeted clones could arise because targeting is inefficient and not enough colonies were screened or because deletion of both copies of *Xist* in a female cell is lethal. It is possible that continued attempts would result in the desired targeting, however given the potential lethality of heterozygous *Xist* mutations, future

attempts should employ conditional *Xist* mutations. These mutations could be introduced into an *Xist* mutant background by homologous recombination, similar to the attempts described above. Alternatively, ES cells could be derived from embryos collected from a cross between conditional *Xist* mutant male and *Xist* mutant female mice (Csankovszki et al. 1999). Because homologous recombination at *Xist* may be inefficient, this second approach may be the faster option. For both strategies, cre-mediated recombination would allow a direct comparison between homozygous and heterozygous *Xist* mutant ES cells.

References

- Augui, S., Filion, G.J., Huart, S., Nora, E., Guggiari, M., Maresca, M., Stewart, A.F., and Heard, E. 2007. Sensing X chromosome pairs before X inactivation via a novel X-pairing region of the Xic. *Science* 318(5856): 1632-1636.
- Avner, P. and Heard, E. 2001. X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet* 2(1): 59-67.
- Azuara, V., Brown, K.E., Williams, R.R., Webb, N., Dillon, N., Festenstein, R., Buckle, V., Merckenschlager, M., and Fisher, A.G. 2003. Heritable gene silencing in lymphocytes delays chromatid resolution without affecting the timing of DNA replication. *Nat Cell Biol* 5(7): 668-674.
- Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., and Heard, E. 2006. Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat Cell Biol* 8(3): 293-299.
- Blewitt, M.E., Gendrel, A.V., Pang, Z., Sparrow, D.B., Whitelaw, N., Craig, J.M., Apedaile, A., Hilton, D.J., Dunwoodie, S.L., Brockdorff, N., Kay, G.F., and Whitelaw, E. 2008. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* 40(5): 663-669.
- Borsani, G., Tonlorenzi, R., Simmler, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., and et al. 1991. Characterization of a murine gene expressed from the inactive X chromosome. *Nature* 351(6324): 325-329.
- Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., and Rastan, S. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351(6324): 329-331.
- Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J., and Willard, H.F. 1992. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* 71(3): 527-542.
- Cao, R. and Zhang, Y. 2004. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* 14(2): 155-164.
- Cartegni, L., Chew, S.L., and Krainer, A.R. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 3(4): 285-298.
- Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., and Krainer, A.R. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31(13): 3568-3571.
- Cattanach, B.M. 1975. Control of chromosome inactivation. *Annu Rev Genet* 9: 1-18.
- Cattanach, B.M. and Rasberry, C. 1994. Identification of the *Mus castaneus* Xce allele. *Mouse Genome* 92: 114-115.
- Chaumeil, J., Le Baccon, P., Wutz, A., and Heard, E. 2006. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev* 20(16): 2223-2237.

- Cohen, H.R., Royce-Tolland, M.E., Worringer, K.A., and Panning, B. 2005. Chromatin modifications on the inactive X chromosome. *Prog Mol Subcell Biol* 38: 91-122.
- Csankovszki, G., Panning, B., Bates, B., Pehrson, J.R., and Jaenisch, R. 1999. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation [letter]. *Nat Genet* 22(4): 323-324.
- Ensminger, A.W. and Chess, A. 2004. Coordinated replication timing of monoallelically expressed genes along human autosomes. *Hum Mol Genet* 13(6): 651-658.
- Fazio, T.G., Huff, J.T., and Panning, B. 2008. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* 134(1): 162-174.
- Gartler, S.M., Goldstein, L., Tyler-Freer, S.E., and Hansen, R.S. 1999. The timing of XIST replication: dominance of the domain. *Hum Mol Genet* 8(6): 1085-1089.
- Goto, T. and Monk, M. 1998. Regulation of X-chromosome inactivation in development in mice and humans. *Microbiol Mol Biol Rev* 62(2): 362-378.
- Gribnau, J., Hochedlinger, K., Hata, K., Li, E., and Jaenisch, R. 2003. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* 17(6): 759-773.
- Gribnau, J., Luikenhuis, S., Hochedlinger, K., Monkhorst, K., and Jaenisch, R. 2005. X chromosome choice occurs independently of asynchronous replication timing. *J Cell Biol* 168(3): 365-373.
- Hansen, R.S., Canfield, T.K., Lamb, M.M., Gartler, S.M., and Laird, C.D. 1993. Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* 73(7): 1403-1409.
- Heard, E. 2004. Recent advances in X-chromosome inactivation. *Curr Opin Cell Biol* 16(3): 247-255.
- Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C.D., and Spector, D.L. 2001. Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* 107(6): 727-738.
- Hendzel, M.J. and Bazett-Jones, D.P. 1997. Fixation-dependent organization of core histones following DNA fluorescent in situ hybridization. *Chromosoma* 106(2): 114-123.
- Hoki, Y., Kimura, N., Kanbayashi, M., Amakawa, Y., Ohhata, T., Sasaki, H., and Sado, T. 2009. A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse. *Development* 136(1): 139-146.
- Huang, Y., Yario, T.A., and Steitz, J.A. 2004. A molecular link between SR protein dephosphorylation and mRNA export. *Proc Natl Acad Sci U S A* 101(26): 9666-9670.
- Huynh, K.D. and Lee, J.T. 2001. Imprinted X inactivation in eutherians: a model of gametic execution and zygotic relaxation. *Curr Opin Cell Biol* 13(6): 690-697.
- Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., and Wutz, A. 2004. A chromosomal memory triggered by xist regulates histone methylation in x inactivation. *PLoS Biol* 2(7): E171.

- Lee, J.T. 2002. Homozygous Tsix mutant mice reveal a sex-ratio distortion and revert to random X-inactivation. *Nat Genet* 32(1): 195-200.
- Lee, J.T., Davidow, L.S., and Warshawsky, D. 1999. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* 21(4): 400-404.
- Lee, J.T. and Lu, N. 1999. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* 99(1): 47-57.
- Lucchesi, J.C., Kelly, W.G., and Panning, B. 2005. Chromatin remodeling in dosage compensation. *Annu Rev Genet* 39: 615-651.
- Luikenhuis, S., Wutz, A., and Jaenisch, R. 2001. Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol Cell Biol* 21(24): 8512-8520.
- Lyon, M.F. 1961. Gene action in the X-chromosome of the mouse (*Mus musculus* L). *Nature* 190: 372-373.
- Marahrens, Y., Loring, J., and Jaenisch, R. 1998. Role of the Xist gene in X chromosome choosing. *Cell* 92(5): 657-664.
- Marahrens, Y., Panning, B., Dausman, J., Strauss, W., and Jaenisch, R. 1997. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* 11(2): 156-166.
- Martin, G.R., Epstein, C.J., Travis, B., Tucker, G., Yatziv, S., Martin, D.W., Jr., Clift, S., and Cohen, S. 1978. X-chromosome inactivation during differentiation of female teratocarcinoma stem cells in vitro. *Nature* 271(5643): 329-333.
- Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 288(5): 911-940.
- McMahon, A. and Monk, M. 1983. X-chromosome activity in female mouse embryos heterozygous for P_{gk}-1 and Searle's translocation, T(X; 16) 16H. *Genet Res* 41(1): 69-83.
- Meyer, B.J. 2000. Sex in the worm counting and compensating X-chromosome dose. *Trends Genet* 16(6): 247-253.
- Mlynarczyk-Evans, S., Royce-Tolland, M., Alexander, M.K., Andersen, A.A., Kalantry, S., Gribnau, J., and Panning, B. 2006. X chromosomes alternate between two states prior to random X-inactivation. *PLoS Biol* 4(6): e159.
- Morey, C., Arnaud, D., Avner, P., and Clerc, P. 2001. Tsix-mediated repression of Xist accumulation is not sufficient for normal random X inactivation. *Hum Mol Genet* 10(13): 1403-1411.
- Navarro, P., Pichard, S., Ciaudo, C., Avner, P., and Rougeulle, C. 2005. Tsix transcription across the Xist gene alters chromatin conformation without affecting Xist transcription: implications for X-chromosome inactivation. *Genes Dev* 19(12): 1474-1484.
- Norris, D.P., Patel, D., Kay, G.F., Penny, G.D., Brockdorff, N., Sheardown, S.A., and Rastan, S. 1994. Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell* 77(1): 41-51.
- Ogawa, Y. and Lee, J.T. 2003. Xite, X-Inactivation Intergenic Transcription Elements that Regulate the Probability of Choice. *Mol Cell* 11(3): 731-743.

- Ohlsson, R., Tycko, B., and Sapienza, C. 1998. Monoallelic expression: 'there can only be one'. *Trends Genet* 14(11): 435-438.
- Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. 2004. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* 303(5658): 644-649.
- Panning, B., Dausman, J., and Jaenisch, R. 1997. X chromosome inactivation is mediated by Xist RNA stabilization. *Cell* 90(5): 907-916.
- Panning, B. and Jaenisch, R. 1996. DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev* 10(16): 1991-2002.
- Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. 1996. Requirement for Xist in X chromosome inactivation. *Nature* 379(6561): 131-137.
- Percec, I., Plenge, R.M., Nadeau, J.H., Bartolomei, M.S., and Willard, H.F. 2002. Autosomal dominant mutations affecting X inactivation choice in the mouse. *Science* 296(5570): 1136-1139.
- Perocchi, F., Xu, Z., Clauder-Munster, S., and Steinmetz, L.M. 2007. Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. *Nucleic Acids Res* 35(19): e128.
- Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., de la Cruz, C.C., Otte, A.P., Panning, B., and Zhang, Y. 2003. Role of histone H3 lysine 27 methylation in X inactivation. *Science* 300(5616): 131-135.
- Plath, K., Mlynarczyk-Evans, S., Nusinow, D.A., and Panning, B. 2002. Xist RNA and the mechanism of X chromosome inactivation. *Annu Rev Genet* 36: in press.
- Rastan, S. 1983. Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. *J Embryol Exp Morphol* 78: 1-22.
- Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C.D., Reinberg, D., Avner, P., and Heard, E. 2004. Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* 24(12): 5475-5484.
- Sado, T., Hoki, Y., and Sasaki, H. 2005. Tsix silences Xist through modification of chromatin structure. *Dev Cell* 9(1): 159-165.
- . 2006. Tsix defective in splicing is competent to establish Xist silencing. *Development* 133(24): 4925-4931.
- Sado, T., Wang, Z., Sasaki, H., and Li, E. 2001. Regulation of imprinted X-chromosome inactivation in mice by Tsix. *Development* 128(8): 1275-1286.
- Savatier, P., Huang, S., Szekely, L., Wiman, K.G., and Samarut, J. 1994. Contrasting patterns of retinoblastoma protein expression in mouse embryonic stem cells and embryonic fibroblasts. *Oncogene* 9(3): 809-818.
- Selig, S., Okumura, K., Ward, D.C., and Cedar, H. 1992. Delineation of DNA replication time zones by fluorescence in situ hybridization. *Embo J* 11(3): 1217-1225.
- Sheardown, S.A., Duthie, S.M., Johnston, C.M., Newall, A.E., Formstone, E.J., Arkell, R.M., Nesterova, T.B., Alghisi, G.C., Rastan, S., and Brockdorff, N. 1997. Stabilization of Xist RNA mediates initiation of X chromosome inactivation [see comments]. *Cell* 91(1): 99-107.

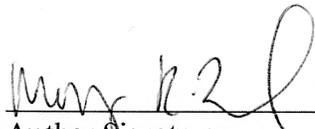
- Shibata, S. and Lee, J.T. 2003. Characterization and quantitation of differential Tsix transcripts: implications for Tsix function. *Hum Mol Genet* 12(2): 125-136.
- Shibata, S., Yokota, T., and Wutz, A. 2008. Synergy of Eed and Tsix in the repression of Xist gene and X-chromosome inactivation. *EMBO J* 27(13): 1816-1826.
- Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T.B., Webster, Z., Peters, A.H., Jenuwein, T., Otte, A.P., and Brockdorff, N. 2003. Establishment of histone h3 methylation on the inactive x chromosome requires transient recruitment of eed-enx1 polycomb group complexes. *Dev Cell* 4(4): 481-495.
- Singh, N., Ebrahimi, F.A., Gimelbrant, A.A., Ensminger, A.W., Tackett, M.R., Qi, P., Gribnau, J., and Chess, A. 2003. Coordination of the random asynchronous replication of autosomal loci. *Nat Genet* 33(3): 339-341.
- Sugawara, O., Takagi, N., and Sasaki, M. 1983. Allocyclic early replicating X chromosome in mice: genetic inactivity and shift into a late replicator in early embryogenesis. *Chromosoma* 88(2): 133-138.
- Sun, B.K., Deaton, A.M., and Lee, J.T. 2006. A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol Cell* 21(5): 617-628.
- Tada, T., Tada, M., and Takagi, N. 1993. X chromosome retains the memory of its parental origin in murine embryonic stem cells. *Development* 119: 813-821.
- Takagi, N. and Sasaki, M. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* 256(5519): 640-642.
- van Raamsdonk, C.D. and Tilghman, S.M. 2001. Optimizing the detection of nascent transcripts by RNA fluorescence in situ hybridization. *Nucleic Acids Res* 29(8): E42-42.
- Williams, B.R. and Wu, C.-t. 2004. Does random X-inactivation in mammals reflect a random choice between two X chromosomes? *Genetics* 167(3): 1525-1528.
- Wutz, A. and Gribnau, J. 2007. X inactivation Xplained. *Curr Opin Genet Dev* 17(5): 387-393.
- Wutz, A. and Jaenisch, R. 2000. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell* 5(4): 695-705.
- Wutz, A., Rasmussen, T.P., and Jaenisch, R. 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* 30(2): 167-174.
- Xiao, S.H. and Manley, J.L. 1997. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev* 11(3): 334-344.
- Xu, N., Tsai, C.L., and Lee, J.T. 2006. Transient homologous chromosome pairing marks the onset of X inactivation. *Science* 311(5764): 1149-1152.
- Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322(5902): 750-756.

Zuker, M. and Jacobson, A.B. 1995. "Well-determined" regions in RNA secondary structure prediction: analysis of small subunit ribosomal RNA. *Nucleic Acids Res* 23(14): 2791-2798.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature

25 March 2009
Date