XIST/Xist-Induced Epigenetic Events in Somatic Cells

by

Nancy P. Thorogood
B.Sc., University of Waterloo, 2002

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April, 2010

Abstract

Mammalian dosage compensation of X-linked genes is achieved between XX females and XY males by silencing one of the two X chromosomes. It is the expression of a functional non-coding RNA transcript, XIST that is responsible for the initiation of silencing during X-chromosome inactivation. *XIST* expression and subsequent in *cis* localization to the future inactive X chromosome initiates a cascade of epigenetic events that leads to the formation of facultative heterochromatin. The exact role the XIST RNA plays in establishing and maintaining the inactive state is uncertain.

It is hypothesized that the XIST RNA sets up a repressive nuclear compartment and transcriptionally silences through the recruitment of factors required for setting up the heterochromatic state of the inactive X. In this thesis I address XIST/Xist's recruitment of factors with two separate approaches: 1) I ask whether Xist, in the absence of silencing is able to recruit epigenetic marks in a somatic cell hybrid system, 2) I evaluate a system whereby the XIST RNA is tagged and can be isolated to identify novel RNA-protein interactions.

To assess epigenetic features that may be directly recruited by the expression of the XIST/Xist RNA, I have analyzed mouse/human somatic cell hybrids where *XIST/Xist* expression and silencing are disconnected. Loss of active chromatin marks, H3 acetylation and H3 lysine 4 methylation, was not observed with XIST/Xist expression; nor was there a gain of DNA methylation or the silencing of the Cot-1 fraction. Therefore, these marks of heterochromatin are not solely dependent upon *XIST/Xist* expression in a somatic cell.

The isolation of the XIST RNA with its interacting partners would allow us to better understand the mechanism by which XIST acts to silence the inactive X. I have developed a MS2 stem loop-tagged XIST RNA and integrated it into an inducible XIST system. The MS2 tagging system is a valid system for pulling out the XIST RNA and identifying interacting proteins since the tagged RNA was able to interact with whatever factor(s) are required to silence the Cot-1 fraction and form a localized RNA signal.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

MEF – mouse embryonic fibroblasts
HEK – human embryonic kidney
HMTase – Histone methyl-transferase
RT-PCR – reverse transcriptase polymerase chain reaction
qPCR – quantitative polymerase chain reaction
RT-qPCR – reverse transcriptase and quantitative polymerase chain reaction
Dox – Doxycycline
FRT – Flp recombinase target
CMV – cytomegalovirus
ES cells – embryonic stem cells
FACS – Fluorescence-activated cell sorting
Xi – inactive X
Xa – active X
XIST – human Xi-specific transcript - Gene names are italicized and gene products are not italicized
Xist – mouse Xi-specific transcript
PAC – P1 artificial chromosome
MacroH2A – macro histone H2A
MCB – macro-chromatin body
AURKB – Aurora B kinase
RNA polymerase II – RNA PolII
ChIP – chromatin immunoprecipitation
FISH – fluorescent in situ hybridization
PRC – polycomb repressive complex
K – lysine residue
R – arginine residue
Me – methylated
Ac – acetylated
Ub – ubiquitinated
H3 – histone H3
H4 – histone H4
H2A – histone H2A
H3K27me3 – histone H3 trimethylated at lysine residue 27
TS cells – trophoblast stem cells
CD – chromo domain
HP1 – heterochromatin protein 1
DNMT – DNA methyl transferase
SAF-A – scaffold attachment factor A
hnRNP – heterogenous nuclear ribonucleoproteins
EGFP – enhanced green fluorescent protein
5-aza – 5-azacytidine

Acknowledgements

I would like to thank all the past and present members of the Brown Lab for the many helpful discussions, advice, and camaraderie. I would especially like to thank Dr. Carolyn Brown for her guidance and patience. Carolyn has been a great teacher, supervisor and mentor.
To my committee members, Dr. Tom Grigliatti, Dr. Michael Kobor, and Dr. Wendy Robinson, thanks for your time and effort in attending my committee meetings and your analysis of this thesis.
And finally, I appreciate all the love and support received from my family – even if you didn't know what I was talking about.

I would like to acknowledge some specific contributions to this work:

- Jennifer Chow for the creation of the mouse cell line containing the PAC integration of human XIST
- Jennifer Chow and Sarah Baldry for the creation of the inducible XIST HT1080 cell lines
- Marissa Jitratkosol for assisting in the generation of the tagged XIST RNA
- Sarah Baldry for transfecting the tagged XIST into the HT1080 cell lines
- Jakub Minks for assisting with the FACS analysis

To my grandparents,

Nancy and Ashley Collie

Wilhelmina and Patrick Joesph Thorogood

Chapter 1: Introduction

1.1 Clinical importance of X-chromosome inactivation

Male and female mammalian cells differ in sex chromosome composition between XY males and XX females. This difference requires dosage compensation of X-linked genes between the sexes and is achieved by a process known as X-chromosome inactivation (18). X-chromosome inactivation involves the silencing of one of the two female X chromosomes, which is chosen at random during early embryonic development, and is maintained throughout the lifetime of the cell. The inactivation of the inactive X (Xi) is achieved by the formation of facultative heterochromatin through a highly orchestrated series of epigenetic events. Therefore during development, a cell that begins with two identical active X (Xa) chromosomes becomes a somatic cell with differentially regulated and identifiable Xa and Xi chromosomes. X-chromosome inactivation is a remarkable example of chromosome-wide epigenetic regulation and cellular memory for the maintenance a heterochromatic state.

The approximately 155 megabase human X chromosome contains more than one thousand genes (19). Of clinical importance is the influence that the X chromosome has on disease manifestation, either directly through X-linked gene mutations or indirectly by the inactivation of X-linked genes. Based on the number of genes contained on the X chromosome, there are a disproportionately large number of disorders that are linked to X-linked genes (19). Hundreds of genetic disorders are associated with X-linked genes (Online Mendelian Inheritance in Man: “X-linked disease” 948 search results (20)) and X inactivation can either protect against or contribute to the severity of disease.

X-chromosome inactivation can influence disease even when the silencing of X-linked genes occurs normally during development. Males manifest X-linked recessive disorders at a relatively high frequency as a result of possessing a single X chromosome (21,22). In females, although X inactivation results in a cell only expressing one X chromosome, the random choice of which X to inactivate creates two separate cell populations (21,22). Females are considered mosaics because some cells express the maternally inherited X, while other cells express the paternally inherited X (21,22). The cellular mosaicism that results from X inactivation is thought to protect

females from X-linked recessive disorders because females express both copies of X-linked genes (21,22).

X inactivation normally occurs by random choice, however, non-random or skewed X inactivation can also occur and influence disease. Skewing results from chance factors in about 10% of all skewed females or more commonly skewing results from cellular selection (21). When skewing occurs by chance, the expression of an X-linked mutation may be affected, conversely when skewing occurs by selection, the X-linked mutation may influence which X chromosome is to be inactivated (21). Cellular selection can result in skewing when there is a selective advantage for one of the parental X chromosomes to remain active or inactive (21,22). If there is an X-linked disease mutation and skewing occurs that favors the expression of the wild-type allele from the Xa, the disease phenotype may never develop (21,22). Conversely, if skewing favors the expression of the mutant allele from the Xa, there is an increased risk of developing the disease phenotype (21,22). Therefore, skewing of X inactivation can influence the manifestation of X-linked disorders in females.

In summary, sex-specific differences in disease can sometimes be attributed to X-linked gene mutations and the process of X inactivation. The single X chromosome in males has a profound effect on the incidence of X-linked disease compared to XX females. Although X inactivation is a fundamental process that occurs in female cells, the pattern of inactivation could also contribute to disease manifestation. When a disease presents itself with sex-specific differences, it is important to consider the effect that mutations in X-linked genes and patterns of X-chromosome inactivation may have on the manifestation of the disease.

1.2 X-chromosome inactivation as a model system

Epigenetic processes play a critical role in the regulation of gene expression, genome stability, genomic imprinting and X inactivation (reviewed in (23-25)). Current research is focused on the understanding of epigenetic mechanisms during embryonic development and disease progression (reviewed in (23-25)). The process of X-chromosome inactivation provides a model system where epigenetic regulation can be studied outside of a disease context to address basic mechanistic questions.

The epigenetic differences that distinguish the heterochromatic Xi from the Xa have been studied in cells at different states of development in both mouse and humans. Similarities have been observed with other heritable gene silencing events in the genome (26,27). What makes X inactivation unique is that its initiation is prompted by the expression of a large X-linked non-coding RNA, Xi Specific Transcripts (XIST) that is approximately 17 kb in size and is the only gene expressed solely from the Xi (28). While exactly how the XIST RNA establishes and maintains the inactive state is uncertain, studying X inactivation will contribute to our understanding of how a non-coding RNA may directly recruit proteins to initiate chromatin changes during development.

1.2 XIST/Xist

The *XIST/Xist* gene codes for a large, alternatively spliced, and polyadenylated transcript that lacks a conserved open reading frame (29,30). The XIST/Xist transcript is thought to function as a non-coding RNA that is involved in X-chromosome inactivation since it is only found in the nucleus, is expressed exclusively from the Xi in somatic cells, and its in *cis* coating of the future Xi precedes gene silencing and chromatin changes. The mechanism by which one of the two X chromosomes is chosen, at random, to express *XIST/Xist* and initiate inactivation is unknown. However, it is this fateful expression that leads to the characteristic Xi that can be visualized in female somatic cells.

Knock-out studies in mouse embryos and mouse embryonic stem (ES) cells demonstrate that *Xist* is required for X-chromosome inactivation (31-33). When *Xist* expression is abolished from one of the two X chromosomes in female mouse ES cells and the cells are subsequently differentiated, X inactivation becomes non-random and the wild-type X chromosome that retains the *Xist* gene is inactivated (32). Ectopic *Xist* expression can also induce inactivation (31). Therefore, *Xist* expression is both required and sufficient to induce inactivation.

A sequence comparison analysis of the *XIST* gene across species shows that there are conserved regions which may help to elucidate the manner in which XIST functions in X-chromosome inactivation. The *XIST* gene has little primary sequence homology

between species (34,35). Generally, long non-coding RNAs show low sequence conservation and display rapid sequence evolution (36). The lack of DNA sequence conservation does not mean there is a lack of similar RNA structure and function. It is the small regions of similarity within the *XIST* gene that hint at important secondary structures required for the initiation of X inactivation. There are six sets of tandem repeat regions throughout the gene and a predicted hairpin structure in exon 4 of the XIST/*Xist* RNA that display conservation in eutheria (34,35,37).

The most promising functional region for silencing is located at the 5' end of the XIST gene, within exon 1, and is comprised of highly conserved tandem repeats that are commonly referred to as the A-repeats. Wutz *et al* use a collection of *Xist* mutations to map functional regions experimentally within the RNA (38). The silencing ability of several *Xist* mutations from the single X chromosome in male mouse ES cells are analyzed by assessing cell lethality (as a result of inactivating the only X chromosome in male cells) and determining localization with RNA FISH (38). It is shown that the A-repeats are required for the Xist RNA to silence genes (38). Although only one region is required for silencing, the localization of the RNA depends on several regions that have little or no homology to each other and are dispersed throughout the transcript (38). Interestingly, the 3' region of the RNA is not required for silencing or localization but appears to have a specific role in recruitment of macroH2A1 (38).

Experiments to date have shown that the A-repeat region is required for gene silencing, that multiple regions are required for localization, and that the 3' region is required for macroH2A1 recruitment. There could be additional functional regions of Xist that have not been recognized. With an increasing number of potential proteins involved in the process of X-chromosome inactivation, there will be a need to determine if additional regions of Xist are crucial for protein recruitment and function at the Xi.

1.3 Features of the Xi

The expression of *XIST* from the future Xi and its in *cis* localization first establishes the difference between the two female X chromosomes. Once the process of inactivation is complete, the chromatin of the Xi can be identified in a somatic cell by decreased active histone modifications, such as H3 and H4 acetylation (4,39,40), as

well as increased inactive modifications (11,13,16,41-43), including the association of a histone variant, macro histone H2A (macroH2A) (5,44), and DNA methylation at the CpG islands of gene promoters (7). Furthermore, the Xi is associated with X-linked gene silencing, late replication in S-phase (45), and a transcriptionally silent domain, known as the Cot-1 hole, which silences the Cot-1 repetitive fraction and excludes RNA Polymerase II (RNA PolII) (12,46).

The Xi chromatin can be physically, temporally and spatially separated from its active counterpart. These changes occur early during development, thus much of our knowledge regarding the initial events is derived from studies utilizing mouse embryos or differentiating mouse ES cells. Figure 1.1 is a schematic representation of the X inactivation process that outlines when the heterochromatic events are estimated to occur. Earlier events in the timeline of X-chromosome inactivation are proposed to contribute to the establishment of the Xi compared to later events, which are important in the maintenance of the Xi. I will discuss each distinguishing event as it pertains to X inactivation.

1.3.1 Expression and localization of XIST

X-chromosome inactivation begins with the increased expression and *in cis* localization of the Xist RNA. *Xist* is expressed at a low level from both X chromosomes in females prior to X inactivation and can be visualized with RNA FISH as a dot-like signal in undifferentiated ES cells. During the process of X-chromosome inactivation, the expression of *Xist* is upregulated from the future Xi and coats the chromosome, prior to gene silencing and the establishment of a heterochromatic state (1-3). The early expression of *Xist* in the cascade of events that occur during X inactivation suggests a role for the Xist RNA in the recruitment of factors required for silencing. The localization of the Xist RNA can occur in somatic cells; however, Xist-mediated silencing is dependent on a developmental context through the requirement of temporally restricted factors (31,47)

The XIST RNA localizes in the nucleus with a specific accumulation that forms an observable single focus and corresponds to the Barr body within a female cell (30,48). XIST spreads over the area of the Xi chromosome (as seen when hybridized with X

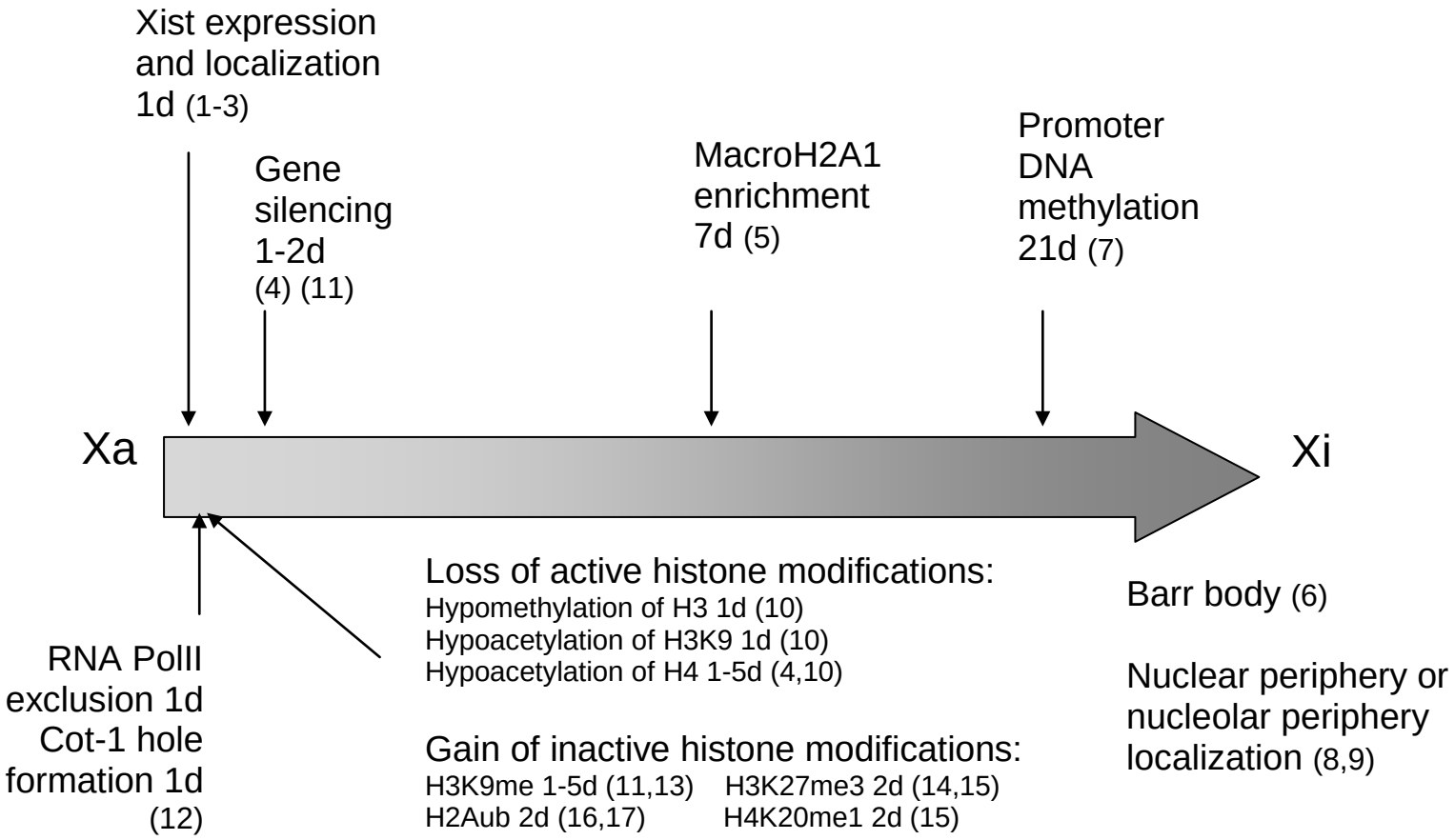

Figure 1-1: A schematic representation of the process of X-chromosome inactivation in mouse ES cells

chromosome paints) and coats the chromosome from which it is transcribed (48). Localization of the XIST transcript is a crucial step in the process of inactivation since it appears to be necessary for XIST-dependent genic silencing. There are examples in the literature of localization occurring without silencing, however, the reverse has not been documented when XIST/Xist is expressed through development (38,49). Conversely, silencing can be maintained in the absence of a localized XIST/Xist RNA in post-differentiated cells (49). The initiation of silencing is dependent on XIST localization but the maintenance of silencing is not dependent on XIST localization.

There is also a localization quality that is specific to the X chromatin itself, which confers efficient XIST/Xist RNA localization. Studies utilizing X:autosome translocations show that the RNA has limited ability to coat the autosomes (50-52). Although the exact manner in which the XIST RNA localizes in *cis* to the Xi remains unknown, it is postulated that protein partners play a role in the localization process.

During mitosis, the association of the XIST RNA to the Xi is disrupted and the RNA can be seen throughout the cytoplasm (48,53). Recently, Hall et al reason that the loss of localization during mitosis suggests that events occurring during this time affect XIST binding (53). XIST's interaction with the chromosome is influenced by the mitotic expression of Aurora B kinase (AURKB) (53). Analysis of the Xist RNA sequence show that multiple sequences dispersed throughout the RNA act cooperatively to localize Xist to the X chromosome (38). A model for XIST binding to the Xi via multiple anchor points propose that a strict localization is maintained by the binding of several factors to the RNA that cooperatively ensure proper localization (38,53).

1.2.2 Cot-1 holes and RNA PolII exclusion

The visualization of the Xi within the nucleus enables researchers to determine if epigenetic marks are enriched or depleted within the Xi chromatin territory. The expression of *Xist* immediately leads to the formation of a silent nuclear compartment (Refer to Figure 1-1). Hall *et al* were the first to identify the lack of Cot-1 RNA within the XIST territory of a female fibroblast cell line (54). The Cot-1 fraction is enriched for repetitive sequences (55) and can be hybridized to RNA within the nucleus to identify heteronuclear RNA transcription (54). A general diffuse hybridization pattern can be seen in the nucleus with a lack of Cot-1 RNA corresponding to the nucleolus and the Xi.

This lack of staining at the Xi is termed the Cot-1 hole. Originally it was thought that the Cot-1 hole represented X-linked gene silencing, however, with further analysis it appears that the Cot-1 hole represents a silent nuclear compartment that is initiated by *Xist* expression through the silencing of the Cot-1 repetitive fraction and the exclusion of transcription machinery (12,46).

To add to this picture of a silent nuclear compartment, there is also the exclusion of RNA PolII and transcription factors (TAF10 and TBP) from the Xi (46,56). RNA FISH and immunofluorescence are used to study the dynamics of both Cot-1 and RNA PolII exclusion during mouse female ES cell differentiation (12). RNA PolII exclusion, at the presumptive Xi is the earliest known event following Xist RNA coating, whereby, after one day of differentiation, 72% of the Xist RNA domains have an accompanying RNA PolII exclusion (12). The kinetics of the Cot-1 exclusion, at the future Xi, follows RNA PolII exclusion closely, whereby 63% of RNA PolII exclusions also have Cot-1 exclusion after one day of differentiation (12). It is interesting that the exclusion of RNA PolII precedes gene silencing (gene silencing being first detected around day two (4,11)), and the exclusion of both RNA PolII and Cot-1 can occur without gene silencing (12). The rapid exclusion of RNA PolII and the appearance of the Cot-1 hole following *Xist* expression, even prior to gene silencing, supports the idea that *Xist* expression leads to the formation of a silent nuclear compartment and that the majority of the domain painted by the Xist RNA is highly repetitive DNA (12,46).

1.2.3 Changes in histone modification

A number of changes in histone modifications also take place early following *Xist* expression. The Xi can be identified based on the lack of euchromatic, active modifications and the enrichment of heterochromatic, inactive modifications in somatic cells. Specifically, using RNA FISH and chromatin immunoprecipitation (ChIP) to analyze the Xi, researchers show that there is a depletion of H3 methylation (at K4, R17 and K36), H3 and H4 acetylation, and an enrichment of H3K9me2, H3K27me3, H4K20me1, and H2AK119ub1 (4,11,13-16,39-43,57-59).The specific timing of these events during the process of X inactivation is deciphered using differentiating mouse female ES cells and early mouse development. As the changes in histone modifications

occur early in the timeline of X inactivation, they are suggested to play a role in the establishment of the Xi.

The loss of the active histone marks appear to be the earliest chromatin change occurring after *Xist* expression (10,11,56). Although the loss of active marks appears to begin prior to the gain of inactive marks, the loss and gain of histone modifications is a gradual process. Therefore, the loss of active histone modifications and the gain of inactive modifications occur within an overlapping timeframe of one to five days of differentiation and 8-cell to 32-cell stage of early mouse embryonic development (4,10,11,43,56). The hypomethylation of histone H3 residues (K4, R17, and K36) and hypoacetylation at H3K9 occur with similar kinetics, approximately one day following differentiation (10). The hypoacetylation of H4 (residues K5, K8 and K12) occurs within this same time (one to five days), although, this process appears to have slightly slower kinetics (10) and perhaps is initiated following the changes on H3 (4,11) .

Closely overlapping the loss of active modifications is the enrichment of inactive marks on the Xi. Beginning at one day following differentiation H3K9 methylation appears and by two days following differentiation, H3K27me3, H4K20me1 and H2AK119ub enrichment could be seen at the Xi (14-16,42).

The exact days during differentiation in which these marks are seen varies slightly in the literature and depends on the cell cycle (i.e. interphase versus metaphase) and the differentiation method (retinoic acid treatment versus withdrawal of leukemia inhibitory factor, LIF). Nevertheless, there is a loss of active modifications, specifically hypomethylation of H3 and the hypoacetylation of both H3 and H4 that are seen in a proportion of cells with an Xist RNA focus from one day after ES cell differentiation (essentially when the Xist RNA coats the chromosome) and a gain of inactive marks beginning one to two days following ES cell differentiation.

The Xist RNA could initiate distinct changes in histone modifications associated with the Xi, as the timing of these histone modifications immediately follow *Xist* upregulation and localization, and coincide with genic silencing. Although some of the enzymes that impart these changes are known, it is not known if the Xist RNA actively recruits these factors to initiate X inactivation or whether a more passive process of histone replacement occurs during cell division that is affected by the overall silent nature of the

chromatin. The ability of the Xist RNA to initiate Xi-associated marks in the absence of the ‘normal’ gene silencing will be discussed in a later section.

Although these modifications distinguish the Xi, the Xi chromatin is not uniform in nature with respect to the histone modifications (60). Approaches that look at chromatin modifications genome-wide are enabling researchers to develop maps of epigenomes. In order for these studies to be useful for the understanding of X-chromosome inactivation, a method in which the Xi can be compared to the Xa must be established. The human Xi has a distinct distribution of H3K9me3 and H3K27me3 whereby the modifications appear in reproducible segregated regions that do not overlap when analyzed with immunofluorescence (61). The H3K9me3 regions colocalize with HP1 and H4K20me3 regions, while the H3K27me3 regions colocalize with the XIST RNA and the histone variant macroH2A (61). Valley *et al* used a SNP-based assay to identify the Xi from the Xa and performed chromatin immunoprecipitation (ChIP) on 112 X-linked loci to determine the occupation of macroH2A versus H3K9me3 (62). The majority of the loci are either enriched for macroH2A or are enriched for H3K9me3, similar to the immunofluorescence results. ChIP is a more region-specific method than immunofluorescence and the ChIP results have some loci with both macroH2A and H3K9me3 enrichment. Therefore, there does not seem to be a uniform distribution of the differential marks over the Xi; however, the inactive epigenetic marks could still be interconnected and act redundantly to keep the chromatin of the Xi silent.

1.2.4 MacroH2A recruitment

In addition to histone modifications, the Xi can also be identified by the incorporation of the histone variant macro histone H2A. There are three macroH2A variants that are derived from two genes. MacroH2A1.1 and macroH2A1.2 are derived by alternate splicing of the *macroH2A1* gene, whereas macroH2A2 is encoded by a separate *macroH2A2* gene (63,64). All of these variants are found to be enriched on the Xi and consist of an N-terminal histone domain, with similarity to H2A, and a large C-terminal non-histone macro domain (65,66). A general nuclear diffuse staining pattern is observed with immunofluorescence in both male and female somatic cells (67); however, a female-specific macroH2A enriched region corresponds to the Barr body and is referred to as the macro-chromatin body (MCB) (65).

The enrichment of macroH2A at the Xi is temporally regulated. In undifferentiated male and female mouse ES cells, macroH2A1 accumulation is seen at the centrosomes and is dependent on intact microtubules (68). During differentiation, macroH2A1 association with the centrosome decreases and the histone variant becomes enriched at the Xi. MacroH2A cannot be recruited when the expression of *Xist* is induced in undifferentiated ES cells (38,44). The initial formation of the MCB in differentiating mouse ES cells is observed seven days following differentiation and is later in the timeline when compared to histone modifications, suggesting that macroH2A1 is not involved in the early initiation phase of X inactivation (5).

The mechanism by which macroH2A is recruited to the Xi is unknown but it is dependent on the expression of *Xist*, in particular, the 3' end of the transcript (38,44,69). A conditional deletion of *Xist* abolishes the localization of macroH2A at the Xi despite the retention of late replication and H4 hypoacetylation (69). Complementary, the ectopic expression of *Xist* from an autosome is sufficient for MCB formation and the continued expression of *Xist* is required for the maintenance of MCBs (44). MacroH2A appears to associate, in some manner, either directly or indirectly with the Xist RNA and evidence to support this association will be discussed later in this introduction.

Estimates of histone protein content from rat liver suggests that up to one in every 30 nucleosomes could contain macroH2A (65) and that the Xi has approximately 1.5 fold more macroH2A1 than the autosomes (70). Cytological investigation of the X chromosome in both metaphase and interphase cells, shows that macroH2A colocalizes with both the Xist RNA and H3K27me3 to form distinct regions on the Xi that are distinct from regions enriched for H3K9me3 (61). However, analyses with an allele-specific approach show alleles that have macroH2A colocalize with the Xist RNA and H3K27me3 and alleles that have macroH2A colocalize with H3K9me3 (62).

Similar to the core histones, post-translational modifications are found on macroH2A1 (71). While macroH2A1 can be ubiquitinated, phosphorylated, methylated, and ADP-ribosylated, little is known about the functional implications of these modifications on macroH2A1 chromatin (71,72). A form of macroH2A1 that is phosphorylated at serine 127 is excluded from the Xi, and the ubiquitin ligase CULLIN3/SPOP that ubiquitinates macroH2A1 is required for macroH2A enrichment at

the Xi (73,74).The specific exclusion of a phosphorylated macroH2A1 and the dependence on a ubiquitin ligase for the inclusion at the Xi suggests there is an unknown regulation on the localization of macroH2A1 to the Xi.

A role for macroH2A in transcriptional repression is based on its enrichment at the heterochromatic regions (including the Xi) and its depletion from transcribed regions within the nucleus, although the exact mechanism that macroH2A functions to cause repression are unknown. There is substantial experimental evidence that links macroH2A with the inhibition of transcriptional initiation. For example, nucleosomes that contain macroH2A prevent the binding of transcription factors to their recognition sequences (75) and SWI/SNF, a chromatin remodeler involved with gene activation, preferentially binds H2A-containing nucleosomes over macroH2A-containing nucleosomes (75,76). In addition, the presence of macroH2A represses polymerase II transcription (77). Another function of macroH2A that may be distinct from X inactivation is the ability of the macroH2A1.1 variant to bind ADP-ribose and localize to sites of poly-ADP-ribose formation (78). Several macro domain-containing proteins share this binding ability and some even have enzymatic activity (78). The exact role macroH2A1.1 plays in the ADP-ribosylation pathway is just beginning to be analyzed. MacroH2A has a complex distribution within the nucleus that is influenced by several factors and several mechanisms.

There are macroH2A1 knockout mice that show normal X inactivation (79), which could be the result of macroH2A2 having a similar function to macroH2A1. Studies on double knockout mice have not yet been reported. In summary, Xist is required and sufficient for the recruitment of macroH2A to the X chromosome and several repressive mechanisms involving macroH2A likely contribute to the silent nature of the Xi.

1.2.5 DNA methylation patterns

The DNA methylation of promoters on the Xi is yet another feature that identifies the Xi and contributes to the heterochromatic state. During early mouse development, the kinetics of DNA methylation follows behind gene silencing and histone modifications (Refer to Figure 1-1) (7). Therefore it is a late event in the timeline of X inactivation and is likely involved in the maintenance of the inactive state rather than the initiation. DNA methylation in mammals occurs at the cytosine in a CpG dinucleotide. Regions with a

high density of CpG dinucleotides are called CpG islands and are found associated with the promoter region in 60-70% of all human genes (80-82).

On the Xi, the methylation of CpG islands associated with promoters contributes to transcriptional silencing of X-linked genes (83,84). While the CpG island promoters of genes on the Xi are hypermethylated compared to the Xa (7,85-88), non-island promoters have similar Xi and Xa methylation patterns (89). The promoter region of the silenced *Hprt* gene has been studied extensively and is hypermethylated on the Xi compared to the Xa (87,88). Conversely, a gene that is known to escape X inactivation, *MIC2*, remains hypomethylated on the Xi (90). More recently, a genome-wide analysis of promoters found that promoter DNA methylation is higher on the X chromosome than the autosomes (85). Specifically on the X chromosome, only the CpG island promoters of genes that undergo inactivation are hypermethylated when compared to those that escape inactivation (85).

Despite the specific hypermethylation found at promoters on the Xi, it appears that the Xi has an overall hypomethylated state compared to the Xa (86,91,92). The Xa is hypermethylated at gene bodies and hypomethylated at gene promoters (93). The hypermethylation at gene bodies on the Xa likely contributes to the overall hypomethylation seen on the Xi (93). In summary, the Xi is overall hypomethylated compared to the Xa but hypermethylation at CpG island promoter regions on the Xi occurs and is thought to contribute to the maintenance of X-linked gene silencing.

1.2.6 Other features of the Xi

In addition to physical features that highlight the differences between the chromatin content of the Xi and its active counterpart, the X chromosomes are spatially and temporally separated within the nucleus. When a human female nucleus is treated with reagents that mark nucleic acids, it is possible to detect the Xi as a DNA dense region, termed the Barr body (6). The subnuclear localization of the Xi is predominantly at the nuclear periphery or adjacent to the nucleolus (8,9,94). Although by FISH the Xa and Xi territories are indistinguishable and have similar volumes, by electron microscopy the Xi forms a unique structure that is easily distinguishable from euchromatin and constitutive heterochromatin (8). The subnuclear compartmentalization of the Xi spatially separates

it from the Xa and other euchromatin, which could contribute to the maintenance of the inactive state.

The replication timing of the Xi is also different from that of the Xa which results in the X chromosomes replicating asynchronously. The majority of the Xi replicates later in S-phase than its corresponding region on the Xa (45). However, there are regions on the Xi that replicate synchronously with the Xa and these generally correspond to genes that escape X inactivation (95). This replication pattern mirrors the general trend where later replication is associated with gene repression and earlier replication with gene expression (96,97). The change in replication timing follows the establishment of the inactive histone modifications and does not affect which origins of replication are fired (10,98,99). As the Xi and Xa use similar replication origins, it is unknown how the replication at the origins is delayed on the Xi. The late replication timing of the Xi temporally separates it from its active counterpart and may provide an additional means to maintain inactivation or be the result of the inactive state.

1.4 Model systems for the study of X inactivation

The differences between the Xa and Xi chromosomes are observed in female somatic cells. The process of X inactivation occurs early during mammalian development; therefore much of our knowledge regarding the procurement of these inactivating events requires the study of model systems. Undifferentiated female mouse ES cells have two active X chromosomes and upon *in vitro* differentiation, can initiate the process of X inactivation randomly on one of the two X chromosomes (31). For this reason, mouse ES cells are a powerful tool used to study events that occur during inactivation. The expression of different *Xist* transgenes and mutations is used to identify factors required for inactivation (31,38). Unfortunately, human ES cells are less informative since the activity of the X chromosomes in undifferentiated cells is unknown or variable (100-103). As a result, studies using human XIST focus on the expression of *XIST* transgenes in somatic human cells, cancer cells, mouse cells or in human/mouse somatic cell hybrids (49,104,105). Figure 5-1 outlines the current and potential model systems for the study of human X-chromosome inactivation.

1.4.1 Developmental window for Xist-mediated silencing

A developmental window in which X inactivation can occur has been defined with differentiating mouse ES cells. There is a male ES cell line system whereby the integration of a single copy *Xist* transgene into the X chromosome of male ES cells is able to localize and initiate late replication, H4 hypoacetylation, and silencing as a measure of cell death (as a result of the inactivation of the only X chromosome in XY male cells) (31). The timing of *Xist* expression is controlled through an inducible promoter; such that, *Xist* could be expressed sequentially during differentiation. To initiate silencing, *Xist* expression must be initiated within 48 hours of differentiation, beyond this point, the cells begin to resist Xist-mediated silencing (31). X-chromosome inactivation can be used as a model to study how epigenetic patterns are established during cellular differentiation since the initiation is intimately linked to development.

It is uncertain if there is an equivalent developmental window during human differentiation. Some aspects of X inactivation are recapitulated in a human male fibrosarcoma cell line, HT1080, with the expression of *XIST* occurring post-differentiation. A human *XIST* transgene that is expressed in the HT1080 cell line induces H4 hypoacetylation (54), late replication (54) , ubiquitination of H2A (58), recruitment of macroH2A (58), silencing of a reporter gene (106), and the formation of a Cot-1 hole (54,106). A human cell line, derived from an adult cancer, is able to initiate XIST-mediated silencing, post-differentiation, suggesting there is a cell type competency for silencing and epigenetic changes, rather than a species-specific difference.

1.4.2 Deletion of the A-repeats and lack of silencing

As mentioned previously, Wutz *et al* use the Xist-inducible male ES cell system to define functional regions of the Xist RNA that are responsible for localization, silencing, and macroH2A recruitment (see XIST/Xist section) (38). A region at the 5’ end of the gene, commonly referred to as the A-repeats, is required for silencing in the mouse ES cell system (38). In the human HT1080 model system, the A-repeats are required for silencing, as well as for the accumulation of the XIST transcript (106). Although the silencing ability is conserved, the localization requirement for the A-repeats differ between mouse and human.

Several studies in mice have used the A-repeat deletion transgene (*XistΔsx*) to assess whether features of the Xi can be recruited in the absence of silencing. If features could be recruited with the *XistΔsx* transgene, it would suggest that the A-repeats are not necessary for the feature's recruitment and that the feature does not initiate silencing alone. The results from the *XistΔsx* transgene studies are not as straightforward; the integration site and differentiation state varies among studies and some features have different degrees of recruitment, as compared to controls. Table1-1 summarizes the epigenetic marks and proteins that are recruited with the *XistΔsx* transgene. The A-repeats are required for efficient enrichment of macroH2A1, Eed, Phc1, H3K27me3, and H2A119ub (14,15,38,107,108). The exclusion of Cot-1 RNA and RNA PolII is independent of the A-repeats and silencing (12). The exclusion of these two marks occurr with the same efficiency when the *XistΔsx* transgene is induced from the endogenous *Xist* locus in male ES cells as compared to differentiating female ES cells (12).

MacroH2A is enriched at the Xi in 30% of *Xist*-expressing cells in differentiated female ES cells (38). Two lines of evidence suggest that macroH2A1 can be recruited in the absence of silencing. First, when an *Xist* cDNA transgene is integrated at the *Hprt* locus (on the X) and is expressed just outside of the developmental window, there is an X chromosome enrichment of macroH2A (26%) that is similar to female cells, despite a lack of gene silencing (38). Second, when the *XistΔsx* transgene is integrated at the Hprt locus in differentiating ES cells, an X chromosome enrichment of macroH2A occurs in 17% of Xist-coated chromosomes despite a lack of gene silencing (38). Therefore, the expression of the *XistΔsx* transgene results in a lower recruitment of macroH2A to the X chromosome. MacroH2A can be recruited to the X chromosome with the expression of Xist in the absence of silencing and the A-repeats contribute to the efficiency of macroH2A recruitment.

Interestingly, differentiation improves the efficiency in which the H3K27me3 mark is recruited to the X chromosome with the *XistΔsx* transgene. When *XistΔsx* is expressed from the endogenous *Xist* locus in undifferentiated cells, 15-25% of cells have an X chromosome enrichment of H3K27me3 (15). Following differentiation, the H3K27me3 enrichment at the X chromosome increases to 78% of cells and is similar to

Table 1-1: Epigenetic marks and enrichment at the X chromosome with the XistΔsx[a] construct in male mouse ES cells

	***Hprt* Locus Integration**				**Endogenous Locus Integration**			**Other Controls**
	Undifferentiated		**Differentiating**		**Undifferentiated**	**Differentiating**		
	XistΔsx[a]	**Xist cDNA**	**XistΔsx[a]**	**Xist cDNA**	**XistΔsx[a]**	**XistΔsx[a]**	**Xist gDNA TetOP[b]**	
MacroH2A			17% (38)	26% (38) differentiated for 2D[c] then 4d[e] Dox[f]				30% (38) differentiated ♀ ES cells
Eed	Enriched (14) 68.9% (107)	92.9% (107)						
Ezh2	Enriched (108)				9% (15)			76% (15) Clone 36[c] (chr11)
Phc1	13.7% (107)	19.3% (107)						
Cot-1 Hole						70% (12) differentiated for 1D[d] with Dox[f]		63% (12) ♀ ES cells differentiated for 1D[d]
RNA Pol II Hole						>90%[g] (12) D1[d] with Dox[f]		72% (12) ♀ ES cells differentiated for 1D[d]
H3K27m3	Enriched (14) 20%(15)	75% (15)			15-25% (15)	78% (15) differentiated for 12 D[d] with Dox[f]		~85%[g] (15) ♀ Fibroblast
H4K20m1					14% (15)			
H2AK119Ub1					Enriched (108)	~42%[g] (108) differentiated D12 with Dox[f]		69% (108) differentiated for 12D[d] with Dox[f] Clone 36[c] (chr11)
Survival **(as a measure of failure to silence)**			93% (38) differentiated for 5D[d] with Dox[f]	23% (38)		~90%[g] (38) differentiatied with Dox[f]	~20%[g] (38)	

Table footnotes:

The numbers are percentages of cells enriched for the particular mark at the chromosome that expressed Xist

a Xist∆sx – represents the Xist transgene lacking the 5' A repeats (38)

b Xist gDNA TetOP – represents the cell line whereby the endogenous Xist promoter is replaced by the inducible promoter (38)

c Clone 36 (chr11) – male ES cell line where an inducible Xist cDNA transgene is integrated in chromosome 11 (38)

d Refers to days of differentiation (D5 = day 5 of differentiation)

e Refers to days of Doxycycline induction (d)

f Refers to Doxycycline induction (Dox)

g Percentages are approximated from bar graphs when there was not an exact number found within the text

the enrichment of H3K27me3 in female fibroblast cells (15). Therefore, at the endogenous locus, the differentiation enables the *XistΔsx* transgene to recruit H3K27me3 with similar efficiency to a female fibroblast.

Overall, the A-repeats add to the recruitment efficiency of most epigenetic marks that were analyzed (Refer to Table 1-1). The expression of the *XistΔsx* RNA results in the enrichment of macroH2A1 (38), Eed (14,107), Phc1(107), H3K27me3 (14,15), and H2A119ub (108), therefore regions outside of the A-repeats could be responsible for the recruitment of proteins that impart these changes. However, the A-repeats appear to add to the recruitment in some manner not as yet defined, since the recruitment efficiency for the *XistΔsx* transgene is reduced compared to an *Xist* transgene (or endogenous) containing the A-repeats. Conversely, the ability of the Xist RNA to form a nuclear repressive compartment is not dependent on the A-repeats (12), and involves other regions of the RNA that remain to be determined.

1.4.3 Somatic cell hybrids as a model system

Somatic cell hybrids are created by the fusion of somatic cells from two different species. When human and mouse cells fuse, generally the human chromosomes become eliminated from the cells during subsequent cell divisions, at random, for an unknown reason (109,110). As a result, somatic cell hybrids can be selected with the complement of mouse chromosomes and a variety of human chromosomes. Somatic cell hybrids are useful in studies of X inactivation because they can be specifically selected to contain either the human Xi or Xa (109,110). Generally a mouse mutant is used where the cells require a human complement to survive; if the mutant gene is X-linked, the human X chromosome will be retained (109,111).

Human/mouse somatic cell hybrids are valuable in the field of X inactivation because they circumvent the need to discriminate between two identical X-linked alleles as they can be selected to only contain one X chromosome, either in its active or inactive state, and can be used to study X-linked gene expression, XIST/Xist localization and X inactivation. Human X-linked gene expression has been analyzed in human/mouse somatic cell hybrids containing either a human Xi or Xa to determine the gene's activity status (112,113). These hybrids are also used to identify human X-linked genes that escape inactivation (112,113).

To study the process of X inactivation in somatic cell hybrids, hybrids containing a human Xi are analyzed for XIST localization and silencing. Surprisingly, the human XIST RNA is not able to localize to the human Xi chromosome in the hybrids, though the X remains inactive (49). Additionally, when a hybrid containing a human Xa is demethylated to express human *XIST* and mouse *Xist*, the human XIST RNA is unable to localize and the mouse Xist localizes (49). However, in both cases, the expression of *XIST/Xist* does not cause X inactivation of either the human or mouse X chromosome (49,104).

The studies in somatic cell hybrids expressing human *XIST* show that proper localization of the human XIST transcript cannot occur and induced expression of *XIST* does not lead to X inactivation. Conversely, inactivation can be maintained on a human Xi that lacks proper XIST localization. The aberrant localization of human XIST and correct localization of mouse Xist suggests that there are species-specific factors required for localization (49). The continued maintenance of inactivation, despite a drifting XIST RNA, supports the idea that silencing is irreversible in somatic cells. While the lack of silencing with induced *XIST* expression in the somatic cell supports the idea of a developmental window for XIST/Xist mediated silencing.

1.5 XIST and potential interacting proteins

There have been several approaches undertaken to identify XIST-interacting proteins and the task remains daunting. The size of XIST/Xist and its association with the insoluble nuclear matrix hinders biochemical approaches for isolating RNA-associated proteins. An indirect strategy is to look at the changes that occur on the Xi during X-chromosome inactivation and postulate what proteins could impart such changes. These candidate proteins can then be analyzed with immunofluorescence, RNA-immunoprecipitation, and knock-out studies to determine if there is an association, whether direct or indirect, with the XIST/Xist RNA. As described in the following sections, there is a wide array of proteins with varying functions that have been identified by these methods including histones, polycomb group proteins, heterochromatin proteins, scaffold proteins, and others.

1.5.1 MacroH2A

Xist expression is both necessary and sufficient for the recruitment of macroH2A to the Xi and the formation of an MCB. A physical interaction between the Xist RNA and the histone variant is proposed to explain the Xist-dependent enrichment of macroH2A at the Xi. An association either direct, or indirect, is shown with RNA immunoprecipitation (RIP) experiments where an antibody to macroH2A immunoprecipitates the associated Xist RNA (114). Although the enrichment of macroH2A at the Xi is Xist-dependent, the exact manner in which the RNA interacts with the histone variant has not been elucidated.

As the recruitment of macroH2A is dependent on the Xist RNA, regions of both the protein and the RNA have been analyzed to identify elements that are required for Xi recruitment and possible interaction. The macroH2A histone variant is characterized by a large non-histone C-terminal region that comprises two-thirds of the protein (commonly called the macro domain) and an N-terminal histone-like region that comprises one-third of the protein and shares 65% amino acid identity to the core H2A (115). The histone-like region of macroH2A alone is capable of targeting to the Xi and formation of MCBs (116,117). Therefore, despite the similarity of the variant histone-like region to histone H2A, there are differences that are important for the specific targeting of this histone-like region to the Xi (116,117). There appears to be redundancy involving the regions required for this localization such that these Xi-targeting sequences are found dispersed throughout the histone-like region of macroH2A (116).

The non-histone macro region of macroH2A alone is not capable of forming an MCB; however, the macro region forms an MCB when the non-histone macro region is fused to H2A or H2B to create chimeric proteins (117). The fusion of core histones with the macro region causing Xi enrichment suggests that there is a region of importance within the non-histone macro region (117). Mutagenesis studies determined that the C-terminal portion of the macro region is critical for targeting to the Xi (117). A comparison of the amino acid residues in the macro region of macroH2A to bacterial and RNA viral proteins show 24-35% similarity (118).This conservation hints at a potential RNA binding ability within the non-histone macro region of macroH2A.

Complementary to studies that identify regions of the protein required for Xi enrichment, the Xist RNA has also been examined to determine if any regions are definitely required for the recruitment of macroH2A to the Xi. The 3' end of Xist is important for recruitment of macroH2A, therefore macroH2A can be recruited in the absence of the 5'A-repeats and its enrichment at the Xi is silencing independent (38).

Overall, an interaction between macroH2A and the Xist RNA is suggested by the Xist-dependent and silencing-independent enrichment of macroH2A at the X chromosome. The conservation of the macro domain suggests that macroH2A is a protein with bacterial and viral origins that has a more general purpose than its role in X inactivation since it is found in organisms that do not dosage compensate and in both males and females. Regardless of an ancestral function, a role has evolved for macroH2A in X inactivation, as it is enriched on the Xi in an Xist-dependent manner and several lines of evidence point to a possible macroH2A-Xist RNA interaction.

1.5.2 PRC2

Polycomb group proteins belonging to the PRC2 complex are enriched on the Xi. The PRC2 proteins, Eed, Ezh2, and Suz12 are transiently recruited to the Xi early during inactivation and are responsible for trimethylating H3 at K27 (14,41,119,120). Ezh2 contains a SET domain and is the histone methyl transferase (HMTase) responsible for H3K27me3, however, the activity of Ezh2 is dependent on its protein partners Eed and Suz12 (41,121,122). The Xist RNA might recruit PRC2 to set up the heterochromatic state of the Xi through the methylation of H3K27.

The kinetics of PRC2 recruitment during X-chromosome inactivation suggest an early role and a direct interaction with the Xist RNA. The transient recruitment of Eed/Ezh2 to the Xi is observed by immunofluorescence in differentiating trophoblast stem (TS) cells and ES cells, in addition to the extraembryonic tissues and the embryonic tissues of developing mice (14,41,56,123). Two separate groups show that in differentiating ES cells, the Xi enrichment of Eed/Ezh2 is the highest two to six days following differentiation and is undetectable by ten to thirteen days (14,41). It seems Eed/Ezh2 plays a critical role in the silencing of the Xi early during the process by setting up the repressive H3K27me3 mark.

Despite the repressive function of PRC2 in the initiation of X inactivation, silencing can occur without Eed recruitment and Ezh2 can be recruited in the absence of silencing (124,125). Eed knockout studies show that initiation of inactivation can occur in the absence of Eed and the H3K27me3 mark in embryonic cells, and that Eed is required for maintenance of the imprinted Xi in trophoblasts (124,126). There are some contradicting results regarding reactivation of the Xi in embryonic cells with one group finding minor reactivation, while another group showing stable silencing (41,125). Regardless, it is difficult to assess long term maintenance of the Xi in cells lacking PRC2 activity and its associated H3K27me3 since knockouts of Eed, or Ezh2, or Suz12 are all embryonic lethal. The current data suggests that PRC2 recruitment to the Xi is Xist RNA-dependent, but is not required for the initiation of X inactivation and its recruitment is not sufficient for silencing.

An indirect interaction between the Xist RNA and PRC2 is proposed on the basis of RIP experiments in ES cells as antibodies to Ezh2 and Suz12 are able to pull-down the Xist RNA (127). An in depth look at regions of Xist identified a 1.6 kb RNA within Xist, RepA, that is able to recruit Ezh2 to an autosome and its deletion prevents PRC2 enrichment and subsequent H3K27me3 in female undifferentiated ES cells (127). The interaction between this RNA and PRC2 appears to be mediated through Ezh2 which is the RNA-binding protein within PRC2 (127). Full length Xist RNA also binds PRC2, which supports the idea that the Xist transcript itself recruits chromatin modifying proteins to the X from which it is transcribed. Therefore, PRC2 and specifically Ezh2 enrichment on the Xi appears to be a result of *Xist* expression and spreading along the Xi.

1.5.3 PRC1

The PRC1 is a highly dynamic complex such that its constituents vary depending on cell type and developmental stage (107). There are about 16 known protein components of PRC1 and the functions of some of these proteins establish a link between PRC1 and X-chromosome inactivation. Cbx2, Cbx4, Cbx5, Cbx6, Cbx7, and Cbx8 all contain the highly conserved N-terminal chromodomain (CD) which is thought to play a role in the recognition of methylated lysine residues on histone tails. It is suggested that the recruitment of PRC1 to the Xi is mediated through the CD-containing

proteins of PRC1 that could potentially bind to the H3K27me3 and H3K9me2 marks of the Xi (128). Ring1b is yet another PRC1 member and is the ubiquitin ligase that monoubiquitinates H2AK119 and is transiently recruited to the Xi during differentiation (42). The recruitment of some PRC1 components occurs independently of inactive histone modifications. Ring1b and its resultant H2AK119ub1 are recruited in the absence of H3K27me3, and not all H3K27me3 enriched Xi chromosomes have a co-associated enrichment of PRC1 (107,108).

Immunofluorescence co-localization studies confirm the recruitment of several PRC1 components to the Xi and their dependence on *Xist* expression (Refer to Table 1-2). Enrichment is seen in a wide array of cells: mouse ES cells at various states during differentiation, early embryos of mice, mouse TS cells, MEFs, human HEK293, and mouse ES cells with *Xist* transgenes. The enrichment of PRC1 components varies among these cell types; for example, for those proteins analyzed there is a higher percent of cells with Xi enrichment in trophoblast cells than in all other types. Cells expressing ectopic *Xist* transgenes, even on an autosome, are able to recruit Phc1, Phc2, Ring1A and H2Aub (Refer to Table 1-2).

It can also be seen from Table 1-2 that the enrichment of different components varies during the progression of differentiation. The timing of the enrichment, during a specific period of development, could suggest a role for a particular PRC1 component in the initiation of X inactivation. For example, if a protein is enriched during the first one to three days of ES differentiation and begins to lose its Xi enrichment thereafter, the protein might be involved in the initiation phase of X inactivation, rather than its maintenance.

The kinetics of enrichment on the Xi shows that Phc1 and Ring1b could be involved during the initiation phase of X inactivation; whereas Phc2, Cbx2, and Bmi1 could be involved early in the maintenance phase (42,107). However, initiation of X-chromosome inactivation can occur in the absence of Phc1 (129), or Ring1a, or Ring1b (16). The enrichment of Phc1 and the establishment of H2AK119ub1 can occur independent of the A-repeats and therefore genic silencing, such that their recruitment is not sufficient to induce silencing (107).

Table 1-2: A literature summary of various PRC1/2 components, their respective histone modifications, and their enrichment in different cell types

	Human Embryonic Kidney (HEK293)	Mouse Embyonic Fibroblasts (MEF)	Mouse Female ES Differentiating	Mouse Male ES with an Xist Transgene	Xist Trangene Expression During Differentiation	Mouse Trophoblasts	Mouse Differentiating Trophoblasts
CBX2	IF: 45.3% (107)	IF: 13.9% (107) HA Tag: 13.1% (107)	D5[a]: 1.6% (107) D10[a]: 19.3% (107) EGFP-D3 Enriched (128)	4.1% (107) Endo cDNA[b] 24hr Dox[c]		IF: 99% (107)	
CBX4	IF (107) IF: 1-41%[d] (74)	HA Tag: 0% (107)	EGFP-D3[a]:0% (128)				
CBX6			EGFP-D3[a] Enriched (128)				
CBX7			EGFP- D3[a] Enriched (128)				
CBX8		HA Tag: 13% (107)	EGFP-D3[a] Enriched (128)				
PHC1	IF: 60.9% (107)	IF: 0.3% (107) GFP Tag: 0.6% (107)	D5: 47% (107) D10: 0% (107)	27% (107) Endo cDNA[b] 24hr Dox 19.3% (107) Hprt[e] 24hr Dox[c] 48% (108) chr11[f] 3d Dox[c] 30% (130) chr11-3d Dox		IF: 69% (107)	
PHC2	IF: 42.7% (107)	IF: 3.4% (107) GFP Tag: 27% (107)	D5[a]: 1.0% (107) D10[a]: 32.1% (107)	0% Endo cDNA[b] (107) 24hr Dox[c]	33% (108) chr11[f] D8[a] of diff[g] with Dox[c] 30% (130) chr11[f] D8[a] of diff[g] with Dox[c]	IF: 99% (107)	
PHC3		GFP Tag: 0.4% (107)					
SCMH1		GFP Tag: 25% (107)					

	HEK293	MEF	Female ES Differentiating	Male ES + Xist Transgene	Trangenes w/Differentiation	TS	Differentiating TS
BMI-1	IF: 65.2% (107) IF: 7.6% (74)	IF: 12.1% (107) GFP Tag: 12.3% (107)	D5[a]: 3.3% (107) D10[a]:6.7%(107)	3.1% (107) Endo cDNA[b] 24hr Dox[c]		IF: 99% (107)	
RNF110		GFP Tag: 9.1% (107)					
RING1A	IF (107)	GFP Tag: 3.3% (107)			60% (130) chr11[f] D3[a] of diff[g] with Dox[c]		
RNF2/ RING1B	IF: 4-38%[d] (74)	GFP Tag: 6.5% (107)	D3-4[a]: ~60%[h] (17) D13[a]: ~0%[h] (17)	56% (108) chr11[f] 3d Dox[c]			D 2: >90%[h] (17) D 12: ~0%[h] (17)
EED	IF: 1% (107)	IF: 1% (107)	D5[a]: 99%(107) D10[a]: 0%(107)	99% (107) Endo cDNA[b] 24hr dox[c] 92.9% (107) Hprt[e] 24hr Dox[c] 89% (108) chr11[f] 3d Dox[c]		IF: 99% (107)	
EZH2			D3-4[a]: ~60%[h] (17) D13[a]: ~0%[h] (17)	79% (108) chr11[f] 3d Dox[c] 96% (130) chr11[f] 3d Dox[c]	34% (130) chr11[f] D8[a] of diff[g] with Dox[c]		D2[a]: >90%[h] (17) D12[a]: ~0%[h] (17)
SUZ12				88% (108) chr11[f] 3d Dox[c] 89% (130) chr11[f] 3d Dox[c]	44%(130) chr11[f] D8[a] of diff[g] with Dox[c]		
H3K27m3	IF: 99.6% (107)	IF: 95% (107)	D5[a]: 100% (107) D10[a]: 99.5% (107) D3-4[a]: ~60%[h] (17) D13[a]: ~0%[h] (17)	97.5% (107) Endo cDNA 24hr dox 96% (108) chr11[f] 3d Dox[c] 95% (130) chr11[f] 3d Dox[c]	61% (130) chr11[f] D8[a] of diff[g] with Dox[c]	IF:99.5% (107)	D2[a]: >90%[h] (17) D12[a]: ~0%[h] (17)

	HEK293	MEF	Female ES Differentiating	Male ES + Xist Transgene	Trangenes w/Differentiation	TS	Differentiating TS
H4K20m1				82% (108) chr11[f] 3d Dox[c]			
H2AK119 Ub			D3-4[a]: ~60%[h] (17) D13[a]: ~0%[h] (17)	97% (108) chr11[f] 3d Dox[c] 90% (130) chr11[f] 3d Dox[c]	69% (108) chr11[f] D12[a] of diff[g] with Dox[c] 90% (130) chr11[f] D8[a] of diff[g] with Dox[c]		D2[a]: >90%[h] (17) D12[a]: ~0%[h] (17)

Table footnotes:

a Refers to day of differentiation (D5 = day 5 of differentiation)
b Xist cDNA transgene was inserted into the endogenous locus (Endo cDNA)
c Refers to Doxycycline induction (Dox)
d Variation in percentage because localization was found to be cell cycle-dependent
e Xist cDNA transgene was inserted into the Hprt locus (Hprt)
f Xist cDNA transgene was inserted into chromosome 11 (Chr 11)
g Differentiation (diff)
h The ~ or > symbol is used when reading percentages from bar graphs when there is not an exact number written in the text.

In the colocalization and knock-out studies, a direct relationship between the Xist RNA and the PRC1 proteins could not be determined. Several components of PRC1 have potential RNA-binding domains. A conserved RNA binding domain is contained within Phc1, Phc2 and Phc3; and the CDs of Cbx4, Cbx7, Cbx6 and Cbx8 are able to bind RNA (107 ,128). Although the CDs bind to several RNAs in *vitro*, the association was not sequence specific to Xist. Additionally, enrichment of Cbx7 on the Xi is partially dependent on its ability to associate with RNA (128). Therefore, these proteins are candidates for direct recruitment by the Xist transcript via their RNA binding abilities.

1.5.4 HP1

Heterochromatin protein 1 (HP1) is a non-histone protein that associates with chromatin. There are three distinct isoforms of HP1 in mammals (HP1α, HP1β, HP1γ), each of which is transcribed from its own gene (reviewed in (131)). Localization within the nucleus appears to be isoform-specific and post-translational modification-specific (131). There are several proteins that are HP1-interacting partners and have a variety of roles including transcriptional regulation, chromatin-modification, DNA replication and repair, and nuclear architecture (131). HP1 is thought to propagate heterochromatization by its recruitment to methylated H3K9 via the histone methyl-transferase (HMT) (132,133), SUV39H1. There appears to be a circular recruitment since the binding of HP1 to methylated H3K9 leads to subsequent recruitment of itself and more HMT which in turn leads to methylation of H3K9. There also might be a link to DNA methylation, with HP1 and SUV39H1 recruiting Dnmt1 and Dnmt3a (134).

All HP1 isoforms contain an N-terminal CD and a C-terminal chromoshadow domain that are separated by a hinge region. The CD facilitates the binding of HP1 to methylated H3 at K9 and confers its gene-silencing function (132,135) whereas, the chromoshadow domain enables HP1 to dimerize and interact with other proteins. Interestingly, the amino acids within the hinge domain of HP1α have RNA-binding activity (136).

The role HP1 plays in the formation of heterochromatin can easily be connected to X inactivation, where H3K9 is methylated following *Xist* expression and HP1 binds to methylated H3K9. In human female nuclei there is an enrichment of all HP1 isoforms (although to varying degrees) corresponding to the Barr body. In contrast, no such

enrichment is observed in mouse cells (13,137). Despite the obvious connection between X inactivation and HP1 recruitment via H3K9 methylation, there is a lack of research investigating its possible interaction with the XIST RNA. HP1α requires both the CD and the hinge domain to achieve proper association with heterochromatin in the nucleus (132). Since the CD is responsible for the methyl-binding ability and the hinge domain is responsible for the RNA-binding ability of the protein and the CD alone cannot confer correct localization, it appears that an RNA component is crucial for HP1α action at pericentromeric heterochromatin (132). By analogy, it is possible that H3K9me on the X alone is not sufficient for HP1 recruitment and that the XIST RNA itself could play a role.

1.5.5 SAF-A

A nuclear structure, named the nuclear matrix or scaffold, is suggested to be a non-chromatin structure that plays a role in DNA replication and gene expression and is involved in higher order organization of the genome (reviewed in (138,139)). There are several protein components of the matrix. The heterogeneous nuclear ribonucleoproteins (hnRNP) are a major constituent of the matrix (140,141). One such hnRNP, the scaffold attachment factor A (SAF-A), is a multifunctional protein with both RNA- and DNA-binding domains (142-144). SAF-A is a candidate protein that could connect X-chromosome inactivation to the nuclear scaffold through an interaction between SAF-A and the XIST RNA.

Since the XIST RNA has a discrete localization within the nucleus and remains after the removal of chromatin a link to the nuclear scaffold is proposed (48,145). hnRNPs interact with XIST RNA (146) and more recently, SAF-A is enriched in the Xi territory (147). The Xi enrichment of SAF-A is dependent on the RNA-binding domain of SAF-A suggesting a possible interaction with XIST(147). Further experiments are required to elucidate the exact role SAF-A plays in X inactivation, whether it is a structural or functional one.

1.5.6 Additional proteins with XIST-interacting potential

There are several other proteins that are enriched on the Xi. I have discussed, in detail, those for which the evidence suggests a potential interaction with the XIST/Xist

RNA. Proteins might be enriched on the Xi as a result of heterochromatization, rather than Xist-specific recruitment, or have an unknown functional link to X inactivation and need further investigation to determine if there is an interaction with the XIST/Xist RNA, a topic to be explored in this thesis. Histone H1, HMGI/Y, BRCA1, and ATRX have all been shown to localize to the Xi by immunofluorescence (137,148). The list of potential XIST/Xist-interacting proteins is growing and with improved techniques for identifying RNA-protein associations, the information obtained, will add to our knowledge of how a non-coding RNA is able to initiate chromosome wide inactivation.

1.6 Thesis objective

It is commonly hypothesized that the XIST/Xist RNA recruits factors required for setting up the heterochromatic state of the inactive X. In my thesis I address XIST/Xist's recruitment of factors with two separate approaches. First, I ask whether Xist, in the absence of silencing is able to recruit epigenetic marks in a somatic cell. Second, I evaluate a system whereby the XIST RNA is tagged and can be affinity-purified to identify novel RNA-protein interactions.

Chapter 2: Materials and methods

2.1 Cell culture

The active X hybrid cell lines were grown in minimal essential media (MEM) supplemented with 7.5% fetal calf serum, penicillin/streptomycin, L-glutamine at 37°C and 5% CO_2. The inactive X hybrid cell lines were grown similarly but at 39°C. BMSL2 was grown in Dulbecco's modified Eagles medium (DMEM) with 10% fetal calf serum, penicillin/streptomycin (Invitrogen), L-glutamine (Invitrogen) at 37°C and 5% CO_2. GM04626 was grown in MEM supplemented with 20% fetal calf serum, penicillin/streptomycin, L-glutamate and non-essential amino acids at 37°C and 5% CO_2.

2.2 RNA isolation

RNA was isolated using acid guanidinium thiocyanate/phenol-chloroform extraction technique or by using the Trizol reagent (Invitrogen). To prevent genomic DNA contamination, RNA was treated with an RNase-free DNase I (Roche) for 1h at 37°C. All RNA isolations were tested for contamination by PCR of a negative RT, where there is no reverse transcriptase enzyme added to the reaction mixture.

2.3 RT-PCR

RNA was used to make cDNA for gene expression analysis. Each RT reaction consisted of 5ug of RNA, 1X first strand buffer (Invitrogen), 0.01M Dithiothreitol (DTT), 0.0625mM dNTPs (Invitrogen), 1µL random hexamers, 2µL (1 unit (U)) RNase inhibitor (Amersham), and 1µL (1U) M-MLV (Moloney Murine Leukemia Virus) reverse transcriptase (Invitrogen). The volume of the reactions were brought up to 20µL with DEPC-treated water, kept at room temperature for 5 min, incubated for two hours at 42°C, and heated at 95°C for 5 min before storage at -20°C.

2.4 Quantitative RT-PCR

Quantitative PCR was performed with cDNA as a template that was generated from the RT method described above or with genomic DNA isolated following a ChIP experiment. The dye, SYBR Green (Sigma) was used to quantitate levels of cDNA in samples and amounts of DNA from ChIP. The qPCR reactions were performed similar to a regular PCR except with the addition of 1X SYBR Green. The DNA Engine Opticon

Monitor qPCR machine (MJ Research) was used to run the qPCR reactions with the Opticon Monitor 3 software for the data analysis.

The delta-delta Ct method was used to compare the levels of RNA. The Ct values were first normalized by subtracting the C_t value obtained for the b-Actin control (ΔC_t = C_t gene of interest – C_t b-Actin) then the normalized Ct values obtained for the 'untreated' cells were subtracted from the 'treated' cells ($\Delta\Delta C_t = \Delta C_t$ treated – ΔC_t untreated) and the relative concentration was determined using ($2^{-\Delta\Delta Ct}$).

For ChIP results, the standard curve method was used to quantitate the levels of mRNA. In this method a DNA dilution series was made from input DNA and run alongside the pull-down samples. The Opticon Monitor Software created a standard curve from which the sample quantities were extrapolated.

2.5 Genomic DNA isolation

Genomic DNA was isolated using a salting-out method. Briefly, cells from a confluent T75 flask were resuspended in 2.5 mL of Tris-EDTA (TE) buffer and 1/20th of the total volume of 20% SDS and Proteinase K were added. The mixture was incubated at room temperature overnight.

The following day, 200 µL of 5M NaCl was added to the mixture and incubated at 37°C until the mixture went into a solution. 825 µL of 5M NaCl was added to the solution and it was shaken vigorously for 30 s. The solution was then centrifuged at room temperature for 15 min at 2500 rpm. The supernatant was transferred to a 15 mL tube and 75 µL of 20% SDS plus 825 µL of 5M NaCl was added. The solution was again shaken vigorously for 30 s followed by a centrifugation. The supernatant was then transferred to a 50 mL tube and 2X the volume of 100% ethanol was added and the mixture was gently rocked to precipitate the DNA. The DNA was scooped out with a pipette tip and transferred into a tube with 25 µL of water.

2.6 Methylation analysis

The isolated DNA was subject to a double restriction enzyme digest prior to methylation analysis by PCR. The double digest for methylation analysis is a four day protocol. On the first day, 40 µL of genomic DNA (concentration >100ng/uL) was

digested with 50U of *EcoR1* (Invitrogen) at 37°C overnight to digest the DNA into smaller fragments.

On the following day, 2 µL of 1mg/mL RNase was added and incubated at 37°C for 15 min. A phenol-chloroform clean up was performed after the RNase treatment and the DNA was precipitated using 2X the volume of ethanol and 1/10th the volume of 3M KOAc; overnight at -20°C. On the third day, the DNA was centrifuged at 13000 rpm for 15 min and the pellet was resuspended in 35 µL of water and was left to sit for 4 hours to allow the DNA to resuspend.

The concentration of the DNA was determined using an Ultrospec 2000 UV/Visible Spectrophotometer (Pharmacia Biotech). 2 µL of the *EcoR1* digested DNA was then digested with a methylation sensitive enzyme (20U) in a 20 µL reaction and incubated 37°C at overnight. Each sample was cut with *HpaII* (NEB), *HhaI* (NEB), and *AciI* (NEB). A mock digest was also made for each DNA sample that did not contain the methylation sentitive enzyme.

All digests were heat killed by incubating at 65°C for 15 min. Finally, all the digests were checked for complete cutting by PCR with MIC2 primers and presence of DNA with XIST primers (Refer to Table 2-1 and 2-2). The MIC2 promoter is known to be unmethylated, therefore, the methylation sensitive enzymes will cut the DNA and no PCR product will be seen.

2.7 ChIP

The ChIP protocol I used was originally adapted from an Upstate (Millipore) protocol. For one antibody pull-down I used 3-10 x 10^6 cells (1.5-5 x 10^6 cells for the antibody and 1.5-5 x 10^6 cells for the no antibody control). The cells were harvested and the cell pellets were washed twice with 1XPBS. The cells were resuspended in 1mL of cell culture media with 1% formaldehyde and incubated at 37°C for 10 min to be crosslinked. The crosslinking was stopped by the addition of 1/10th the volume of a 1.25M glycine solution and incubated for 5 min at room temperature before the cell suspensions were transferred to an ice bucket. The cells were then washed twice in ice cold 1XPBS with 1/100th protease inhibitors (Sigma). After the washes, the cells were lysed in SDS lysis buffer (1% SDS, 10mM EDTA, 10mM TRIS pH 8.1) with 1/100th

protease inhibitors for 15 min on ice. The lysing of the cells was followed by passing the solution through a 25 gauge needle five times.

The cell lysate was sonicated using the Biorupter (Diagenode) with a refrigerated water bath (Thermo Electro Corporation) 30 sec ‘on’, 30 sec ‘off’ for 15-30 min. The sonicated sample was centrifuged at 13000 rpm for 10 min at 4°C and the supernatant was transferred to a 15mL tube; 10% of supernatant was removed to be used as a 10% input. The sonicated cell supernatant was diluted 10-fold in ChIP dilution buffer with protease inhibitors (0.01%SDS, 1.1% TritonX-100,1.2mM EDTA pH 8.1, 16.7mM Tris pH 8.1, 167mM NaCl) and pre-cleared with salmon sperm DNA/agarose slurry (Upstate) rotating for 30 min at 4°C. The salmon sperm DNA/agarose slurry was removed by a brief spin at 700 rpm at 4°C for 1 min. The pre-cleared supernatant was split into three tubes: input, antibody, and no antibody. The antibody of interest was added and incubated overnight at 4°C.

The following day, 75µL of the salmon sperm DNA/agarose slurry was added to the antibody and no antibody solutions and incubated for 2 hours with rotation at 4°C. The agarose pellet (with antibody/histone complex) was centrifuged briefly with 700 rpm at 4°C for 1 min. The agarose/antibody/histone complex was brought through a series of 15 min washes performed with rotation at 4°C with 1mL of each of the following buffers: low salt immune complex wash buffer (0.1% SDS, 1% TritonX-100, 2mM EDTA pH 8.1, 20mM Tris pH 8.1, 150mM NaCl); high salt immune complex wash buffer (0.1% SDS, 1% TritonX-100, 2mM EDTA pH 8.1, 20mM Tris pH 8.1, 500mM NaCl); LiCl immune complex wash buffer (0.25M LiCl, 1% Nonidet-40, 1% deoxycholic acid, 1mM EDTA pH 8.1, 10mM Tris pH 8.1); and finally, two washes with TE (10mM Tris pH 8.1, 1mM EDTA pH 8.1).

After the washes, the sample was an agarose/antibody/histone/DNA complex and the DNA-protein-antibody complex was separated from the agarose beads by washing twice with 250µL of elution buffer (1% SDS, 0.1M $NaHCO_3$) for 15 min at room temperature.

The crosslinks were reversed by adding 20 µL of 5M NaCl and incubating for 4 hours at 65°C. The input samples also had their crosslinks reversed at this stage. The RNA was removed by treating with 1 µL of 10mg/mL RNase and incubated for 10 min at

37°C. The proteins were digested with 10 µL of 0.5M EDTA, 10 µL 1M Tris-HCl, and 2 µL of 10mg/mL proteinase K and incubated for 1 hour at 45°C. The DNA was purified with Epoch columns and the samples were used as a template for qPCR.

2.8 Fixing cells on coverslips

Cells were grown on coverslips (Gold Seal® Cover Glass, 22x22mm No.1½) in X-plates (VWR) with appropriate media and conditions for cell type. When the coverslips were 50-80% confluent, they were rinsed twice with 1X PBS and placed in a Coplin jar with 10mL of CSK buffer for 2 minutes. Then the slips were transferred to a Coplin jar with 5% Triton X (Roche) and 0.5mL RVC (ribonucleoside vandayl complex – NEB) in CSK buffer and permeabilized for 5 minutes. For fixation, the coverslips were incubated for 10 minutes in 4% PFA (16% formaldehyde in 1X PBS) and finally, rinsed in 70% ethanol and stored submerged in 70% ethanol at 4°C.

2.9 Nick translation labeling of FISH probes

A nick translation kit (Invitrogen) was used to label the probes with either dig-UTP (Roche) or biotin-UTP (Roche). A reaction mixture was made with 1µg of probe DNA, 5 µL of nucleotide solution, 5 µL nucleotide buffer, 6 µL of a labeled UTP, 5 µL of nick translation enzyme and water so that the final reaction volume was 50 µL. The reaction mixture was then incubated for 2.5 hours at 15°C and stopped with a stop reaction buffer from the kit, 5 µL was added and heated to 80°C for 10 minutes. The probe was precipitated with 2 µL salmon sperm/tRNA (make stock with 1 mL of 10mg/mL ssDNA [Sigma] and 0.01g *E.Coli* tRNA [Roche]), 0.1 volume 3M socium acetate and 2.5 volume 95% ethanol at -20°C for 1h or overnight. The probes were precipitated by spinning at 10000 rpm for 30 min and were then rinsed with 70% ethanol, air dried, resuspended in 80 µL of DEPC-treated water and stored at -20°C.

2.10 RNA FISH

In 500uL Eppendorf tubes, 12µL of the appropriate Cot-1 DNA (Invitrogen – stock 1µg/µL), 2µL salmon sperm and tRNA (make stock with 1mL of 10mg/mL ssDNA [Sigma] and 0.01g *E.Coli* tRNA [Roche]), and 5µL of the appropriate probe DNA (nick translated probe concentration - 1µg/µL), were mixed together to form the probe

mixture. The probe mixture was then dried in a speed vacuum. The coverslips were retrieved from the 4°C, 70% ethanol storage and dehydrated with 100% ethanol and then in a dried on a drying rack.

When the probe mixture was completely dried, it was resuspended in 10 μL of 100% Formamide (Sigma) and denatured at 80°C for 10min. While the probe was denatured a 4:1 mixture of RNA hybridization buffer (1mL BSA [Roche]; 1mL 20XSSC [Sigma]; 2mL autoclaved dextran sulfate)/RVC (NEB) was made. 10 μL of mixture was added to each denatured probe and 20 μL was dotted on a parafilmed glass plate. The air dried coverslips were placed cell side down onto the probe and sealed with another piece of parafilm. The coverslips were hybridized with probes overnight at 37°C.

The following day, the coverslips were rotated in Coplin jars with 5mL formamide; 5 mL 4X SSC for 15 min at 37°C followed by a rinse in 2X SSC for 15 min at 37°C. Then the coverslips were washed in 1X SSC for 15 min at room temperature with rotation followed by a 1 min rinse in 4X SSC in preparatin for the secondary antibody. (Note: if a directly labeled probe was used, after washes, the coverslips were stained with DAPI and mounted as described below.)

During the final wash, a 500 μL aliquot of 4X SSC/1% BSA (Roche) was thawed and the secondary antibody (FITC anti-dig or avidin conjugated to an Alexa) was diluted appropriately and 40 μL was dotted on a parafilmed glass plate. The coverslips were then transferred from the 4X SSC in Coplin jar onto the secondary antibody on the glass plate. The coverslips were sealed and incubated at 37°C for 1 hour or at 4°C overnight.

Following the incubation, the coverslips were washed in 4X SSC at room temperature in a dark box for 10 min with rotation followed by a wash in 4X SSC/0.1% Triton X for 10 min. An additional rinse in 4X SSC was used to remove the triton and the coverslips were transferred into DAPI (4'-6-Diamidino-2-phenylindole) for 20 sec and washed twice with 1X PBS. Finally, the coverslips were mounted onto glass slides with a dual antifade and mounting media, Vectashield (Vector Labs) and stored at -20°C.

2.11 RNA FISH combined with immunofluorescence

When RNA FISH was combined with immunofluorescence, a directly labeled probe was used and the FISH was performed first as described above. The coverslips are

kept in the dark upon completion of the FISH procedure. After the wash in 1X SSC at room temperature the coverslips were transferred to 1X PBS and rotated for 10 min at room temperature. The primary antibody was diluted into 1X PBS/1% BSA and dotted onto a parafilmed glass plate. The coverslips were placed onto the primary antibody and incubated for 1h at 37°C. The slides then went through a series of washes: 1X PBS for 10min, 1XPBS + 0.1% triton for 10 min, and 1XPBS for 10 min.

The secondary antibody was diluted into 1X PBS/1% BSA and dotted onto a parafilmed glass plate. The coverslips were transferred from the 1XPBS onto the secondary antibody and incubated for 1h at 37°C. Then the coverslips were washed three consecutive times in 1X PBS for 10 min each at room temperature, stained with DAPI for 20 sec followed by 2 washes in 1X PBS before the coverslips were mounted onto glass slides with Vectashield (Vector Labs) and can be stored at -20°C.

2.12 Fluorescent microscope imaging

All images were captured using a Leica inverted microscope (DMI 6000B) with a Retiga 4000R (Q-Imaging) camera and using the OpenLabs imaging software. A volumetric image was captured by taking several images at various focal planes through the cell, commonly called a Z-stack. The Z-stack was then subjected to volume deconvolution to decrease the blurring/noise/distortion obtained from the microscope itself and then compressed into one 3D image using the 3D rendering option in the OpenLabs imaging program.

The localization of the XIST/Xist signals in the hybrid cell lines were visually determined. The percent of Cot-1 holes associated with the Xist signals were qualitatively identified and counted by visually determining whether there was a decrease in the Cot-1 RNA signal corresponding to the XIST/Xist RNA signal. The Cot-1 hole was quantified by plotting the pixel intensity for each signal across a given region in the NIH, ImageJ software.

2.13 Cloning scheme for tagging XIST with the MS2 loops

The expression vector containing eight MS2 loops (pcDNA5/FRT/TO+8x MS2#6) and the expression vector containing the XIST cDNA transgene, VI-34 (pcDNA5/FRT/TO VI-34), were streaked on LB plates and grown overnight at 37°C,

under the selection of ampicillin. The following day, colonies were picked and grown in 2mL of TB with ampicillin overnight. Isolation of the plasmid DNA was performed; the eight MS2 loops were amplified by pFRT5A and BGHrev2 and checked for the expected 561 base pair (bp) product size.

Both the MS2 insert and the VI-34 plasmid were digested with *ApaI* overnight at 25°C, then digested with *ClaI* overnight at 37°C and cleaned with phenol/chloroform. The VI-34 vector and the MS2 insert were ligated with T4 DNA ligase (New England BioLabs) for eight hours at 16°C. The ligated samples were heat inactivated at 65°C for 20 minutes then drop dialyzed before electroporation. The ligated samples were electroporated into ElectroMAX^{TM} Stbl4^{TM} competent cells (Invitrogen) with a Bio-Rad Gene Pulser®. The electroporated ligations were then plated on LB with ampicillin and colonies were picked and grown in 2mL of TB media overnight at 37°C.

100µL of the culture was removed, spun down and the pellet was resuspended in 20µL of sterile water and boiled for five minutes. The boiled samples were directly used as a template in a PCR reaction that confirmed the eight MS2 loop insert was ligated at the 3' end of XIST with primers: qXISTtg3 and BGHrev2. If the MS2 loops inserted appropriately, then the product was 663 bp in length; however, if the MS2 insert was absent and there was a re-ligation of the vector, the product was 181 bp in length. The DNA was isolated from the positive cultures for transfection into the HT1080 HH1 TetTR cells.

The inducible expression vector containing the XIST cDNA with eight MS2 stem loops fused to the 3'end (35-8-10-16) and the pOG44 plasmid (expressing Flp recombinase) were co-transfected using Lipofectamine 2000 (Invitrogen) into HT1080 HH1 TetTR 2-3-0.5+3#4 (3q FRT integration) and HT1080 HH1 TetTR 2-3-2.0d (1p FRT integration). The cells were grown with hygromycin (150µg/mL) to select for positively integrated clones. There were seven different single-cell clones of the 3q integration and only one was positive when screened with RNA FISH for XIST expression after Dox induction, HT1080 HH1 TetTR 2-3-0.5+3#4 XIST+loops #3. There were four different single-cell clones of the 1p integration and only one was FISH positive, HT1080 HH1 TetTR 2-3-2.0d XIST+loops #10.

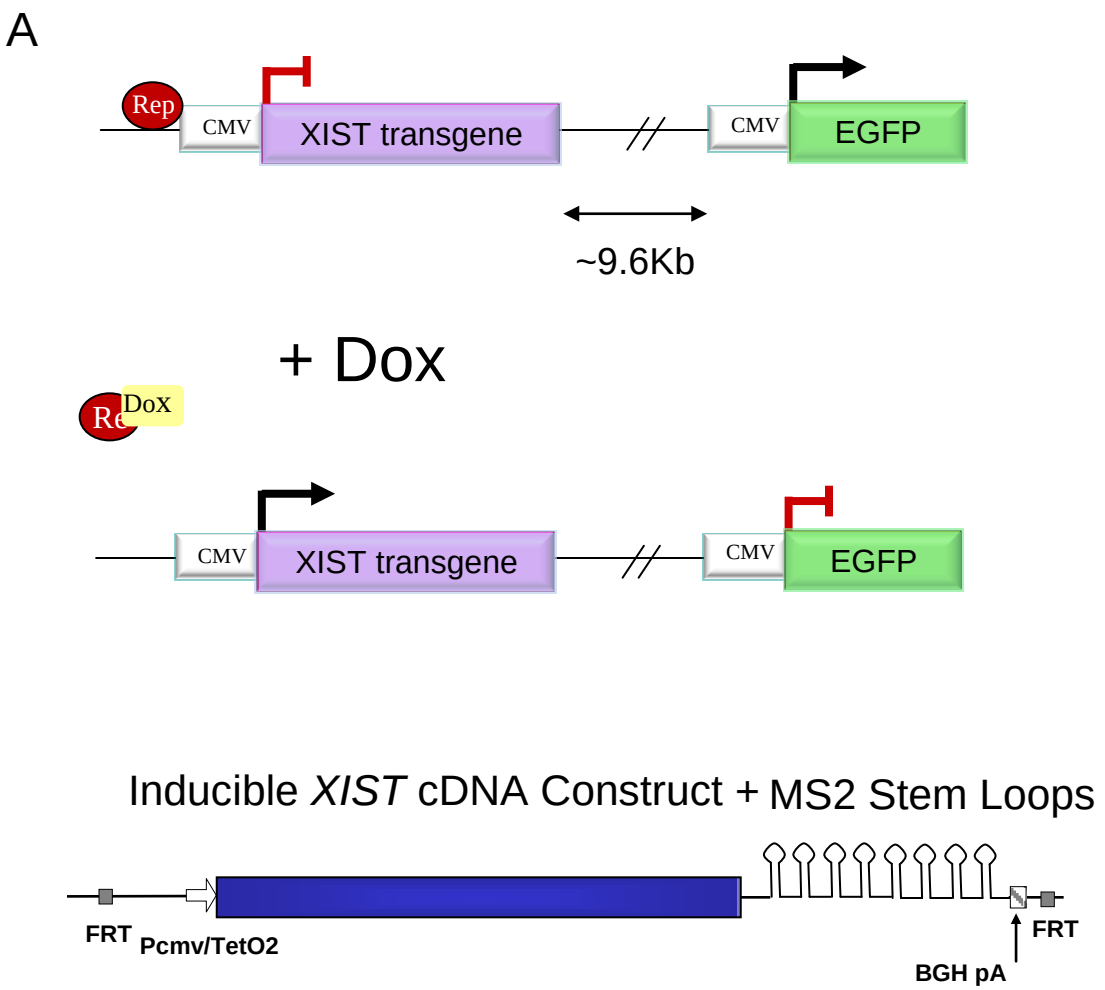

Figure 2-1: Schematic drawings of the doxycycline-inducible *XIST* transgenes in the HT1080 cell line

A) A schematic of the doxycyline-inducible *XIST* transgene in the HT1080 cell line (3q integration) with the linked EGFP reporter gene approximately 9.6Kb downstream of the *XIST* transgene. B) A schematic of the inducible XIST+loops construct showing the fusion of eight MS2 RNA stemp loops to the 5' end of the XIST cDNA transgene.

Table 2-1: List of primers used for expression analysis

Gene	Product Size	Name	Sequence	Conditions
mActin	cDNA:123-130 gDNA: 220	9F 5R	ctg gct cct agc acc atg aag atc tgc tga tcc aca tct gct gg	(95-30s, 59-30s, 72-30s) x 40
(149)Chic1	F-R: 308 F-NR: 277	F R NR	gag tgc cct tcc taa taa gtg g ctg gaa ctc ata ctg tag acc agg tga act caa aga gat ctg tgg c	(94-1, 54-1, 72-2) x 30
mPdha1	203	1 2	gag cta aag gcg gat cag ctg tat ctt cgt cct gtt agc tct gca aga	(95-30s, 59-30s, 72-30s) x 40
mPgk1	238	rA rB	aag cgc acg tct gcc gcg ctg ttc t gtt ggc tcc att gtc caa gca gaa t	(95-30s, 59-30s, 72-30s) x 40
mPhka	223	1 2	aca cct gca gtt gga tgc tac ttc ggc cat tcc aac tga act cgc att	(94-1, 54-1, 72-2) x 30
Ube	196	1a 1b	agc tgt gct gca acg atg aa gtc ttg agg ttg ctg ggt a	(95-30s, 59-30s, 72-30s) x 40
mZfx	504	1 2	cag ttg tca tcc agg atg tc tcg ttg tcc ata gtc agt cc	(95-30s, 59-30s, 72-30s) x 40
(149)Abcb7	cDNA 290	F R	aag cat tcg gca gtt ctg acc tct agt atc aac atc ctt taa ccc	(94-1, 54-1, 72-2) x 30
mXist	578	mx23b mix20	act gcc agc agc cta tac ag gtt gat cct cgg gtc att ta	(94-1, 56-1, 72-2) x 30
mXist	140	mx23b s2	act gcc agc agc cta tac ag gca aag cag caa gcc cac aa	(94-1, 62-1, 72-2) x 30
IDS	300 235	1 3	taa ggga gct gac tga tct tg gct ata cgg aga atc atc g	(94-1, 54-1, 72-2) x 30
XIST	180	5' 3'	ttg ggt cct cta tcc atc tag gta g gaa gtc tca agg ctt gag tta gaa g	(94-1, 54-1, 72-2) x 30
XIST	265	C39-1 C39-2	cct ata ctg ctt aaa tgc gc cct aag att atg cac gct aa	(94-1, 54-1, 72-2) x 30
SLC16A2	155	F R	ttc aag gca tta acc tca ag ggg ctc acc ata tca ata act	(94-1, 54-1, 72-2) x 30
PGK1	395	r1 r2	tcg gct ccc tcg ttg acc ga agc tgg gtt ggc aca ggc tt	(94-1, 54-1, 72-2) x 30
PLS3	210	1 2	tga cct tgt gaa gag tgg c acc tgc gaa tca tgc acc t	(94-1, 54-1, 72-2) x 30
TIMP1	180	C1A C1B	aga tcc agc gcc cag aga ga ccc tga tga cga ggt cgg aa	(94-1, 56-1, 72-2) x 30
PDHA1	200	5 6	cct gtg cgt ccg aga cgc gtt cac cat cct gtc ctt g	(94-1, 54-1, 72-2) x 30

Table 2-2: List of primers used for the methylation analysis

Gene	Product Size	Name	Sequence	Conditions	Enzymes
MIC2	373	5'A 5'B	aga ggt gcg tcc gat tct t cgc cgc aga tgg ac aat tt	(94-1, 52-1, 72-2) x 40 1.0mM $MgCl_2$, 4% DMSO	3 *HpaII*, 4 *HhaI*, 5 *Aci1*
(39)POLA	348	M1 M2	ctg ggg aaa acg atc caa cc ctg aaa gcc aat cag cgg c	95-1.5(95-1, 60-45s, 72-45s)x30 1M betaine	5 *HpaII*
TIMP1	294	5A 5B	ccc ttg ggt tct gca ctg a cca agc tga gta gac agg c	95-5 (94-1, 58-1, 72-2) x 30 1.0mM $MgCl_2$	2 *HpaII*
XIST	555	AT2 29r	atg ctc tct ccg ccc tca Atc agc agg tat ccg ata cc	(94-1, 54-1, 72-2) x 35	3 *AciI*
(39)SLC16A2	129	M1 M2	ctg gcc cgg ctc ctg gc gct ttg ttt gcg cca acc tg	(94-1, 62-1, 72-2) x 31 2M betaine	8 *HpaII*
(39)PGK1	145	M1 M2	acg cgg ctg ctc tgg gc tta ggg gcg gag cag gaa g	(94-1, 62-1, 72-2) x 31 2M betaine	3 *HpaII*
G6PD1	425	M1 M2	cac tac gcg gag ctg cac ctg aag cac aac aaa cag cgt	(94-1, 54-1, 72-2) x 35 1M betaine, 0.5mM $MgCl_2$	7 *HpaII*
mXist	392	meth 6 meth 7	ttc tcg agc cag tta cgc ca cca ttg cta cac acc aga ac	(94-1, 58-1, 72-2) x 30	4 *HhaI*
mSlc16a2	321	m1 m2	cct gaa ctg tgt tct gcg t agc ctg gaa ctt aga cac c	(94-1, 54-1, 72-1**)** x 35 1M betaine, 2mM $MgCl_2$	6 *HpaII*
mPgk1	181	m1 m2	ctt gag ggc agc agt acg gaa ccg gca ttc tgc acg ctt caa	(94-1, 54-1, 72-2) x 35	2 *HpaII*
mZfx	238	m1 m2	ctc gtg cgg att tta cag c agg aaa atg cgg aag ggt ag	(94-1, 54-1, 72-2) x 30 1M betaine	4 *HpaI*
mPhka1	400	m1 m2	cgt tca gtc cca gtc tct cag aag acc ccg tct cca ctc a	(94-1, 54-1, 72-2) x 30 1M betaine	6 *HpaII*
mPola	257	m1 m2	cat gcg tcc tac gga ttg tt gaa agc caa tca gcg gcc t	(94-1, 54-1, 72-2) x 30 1M betaine	3 *HpaII*

Table 2-3: List of primers used for the ChIP experiments

Gene	Product Size	Name	Sequence	Conditions
XIST	217	AT1 PM	gaa cca acc aaa tca cag aga ata aag ggt gtt ggg gga c	(94-1, 58-1, 72-2) x 30
ELK1	300	5A 5B	gca cag ctc tgt agg gaa agc tca cct gtg tgt agc g	(94-1, 54-1, 72-2) x 35
mXist	392	meth 6 meth 6	ttc tcg agc cag tta cgc ca cca ttg cta cac acc aga ac	(94-1, 58-1, 72-2) x 30
mPgk1	181	m1 m2	ctt gag ggc agc agt acg gaa ccg gca ttc tgc acg ctt caa	(94-1, 54-1, 72-2) x 35
mPola	257	m1 m2	cat gcg tcc tac gga ttg tt gaa agc caa tca gcg gcc t	(94-1, 54-1, 72-2) x 35 1M Betaine
mG6pd1	214	1 2	gcc cat gag gac tag acc tt aca tcc act gtg ggc agc ta	(94-1, 54-1, 72-2) x 35 1M Betaine

Table 2-4: List of primers used for tagging the XIST RNA with the MS2 stem loops

Use	Product Size	Name	Sequence	Conditions
Amplify stem loops from pcDNA5/FRT/TO	561	pFRT5A BGH rev2	gcg gcc gca tcg att ccg gat ctc tag cgt tta aac tag aag gca cag tcg agg ct	(94-1, 54-1, 72-2) x 30
Amplify from XIST and across the stem loops	663 (stem loops ligated) 181 (religation)	tg3 BGH rev2	cca cca gaa agt aat ctt aag cca t tag aag gca cag tcg agg ct	(94-1, 54-1, 72-2) x 30

Chapter 3: Active chromatin is retained on the X chromosome following re-activation of *XIST/Xist* expression in human/mouse somatic cell hybrids

3.1 Introduction

The process of X-chromosome inactivation begins with the expression and in *cis* localization of the XIST RNA to the future Xi. This in turn initiates gene silencing and the concurrent establishment of a heterochromatic state. Exactly how the XIST RNA localizes and the manner in which this localization causes gene silencing and changes to the chromatin is unknown. Is there a direct interaction between the RNA and the surrounding DNA or are there protein intermediates that assist in localization? Is it the XIST RNA itself that recruits proteins involved in setting up an inactive state or are some repressive chromatin marks a result of the transcriptionally inactive chromatin? In the absence of silencing could Xist localization result in changes to the chromatin?

When a human/mouse somatic cell hybrid that contains a human Xa in a mouse background is demethylated with 5-azacytidine (5-aza), it results in stable human *XIST* expression and transient mouse *Xist* expression (49,104). The human XIST RNA does not localize to the human X chromosome and the mouse Xist RNA localizes to the mouse X chromosome in the somatic cell hybrids (49). Similarly, a somatic cell hybrid retaining an Xi also has this de-localized XIST RNA, therefore demonstrating that the human XIST RNA cannot localize to a human X chromosome in a mouse background (49,104).

On the other hand, in mouse ES cells expressing a human YAC, the human XIST is able to localize to the mouse chromosome from which it is expressed. This localization suggests that there may be localization proteins that are DNA-specific, such that they may only recognize mouse DNA in a mouse background. To further explore this concept, I have analyzed the XIST localization in a mouse cell line expressing *XIST* from a mouse chromosome to resolve the localization ability of the XIST RNA.

In the human/mouse somatic cells described above, both human and mouse X chromosomes remain active, despite XIST/Xist expression. As silencing of genes closest to the *Xist* locus occurs first in early development (149,150), I revisited the question of silencing to determine if X-linked genes in close proximity to the Xist locus are silenced by the re-activated expression of the Xist RNA.

There are alternative systems in which *Xist* expression is disconnected from silencing. The analysis of a large series of deletion constructs of an inducible *Xist* cDNA in mouse ES cells identified a transgene lacking the conserved 5'A-repeats (*Xist*Δ*sx*) that was able to produce localized Xist RNA but was not able to induce transcriptional silencing (38). The inducible expression of this *Xist*Δ*sx* transgene in mouse ES cells provides another model system that separates the activity of silencing from localization, yet allows progression through the developmental stage where inactivation normally occurs (38).

Studies that utilize the *Xist*Δ*sx* construct show that despite a lack of Xist-induced silencing, a variety of downstream repressive chromatin marks are recruited following Xist expression (Refer to table 1-2) (12,14,15,38,107,108). The *Xist*Δ*sx* construct is induced from a strong viral promoter at either the endogenous locus or the X-linked *Hprt* locus. In the somatic cell hybrids, the transcription of *XIST/Xist* is induced by DNA demethylation of the endogenous promoter in its native context. These somatic cell hybrids, in which silencing and *Xist* expression are uncoupled, provide an alternative and complementary system to ascertain whether expression of *Xist* is able to induce specific epigenetic events distinct from gene silencing.

In this chapter, I describe the analysis of both active and inactive Xist-associated epigenetic events in the somatic cell hybrids to determine if *Xist* expression and localization can modulate X chromatin in the absence of silencing in a somatic cell.

3.2 Results

3.2.1 XIST/Xist are induced by demethylation but show species difference in localization

The mouse/human hybrid cell line, AHA-11aB1, retains only the human Xa, and a subclone, AHA-A52b, stably expressed human *XIST* after three rounds of demethylation with 5-aza (104). Further rounds of demethylation have yielded a new cell line, AHA-4C1 that stably expresses both human *XIST* and mouse *Xist*. In the AHA-4C1 cells, FISH analysis demonstrated that the human XIST RNA was not localized within the nucleus while mouse Xist RNA was localized (Figure 3-1 A). 76% of cells had mouse Xist localization and a diffuse human XIST signal (n=95), similar to numbers reported by Clemson *et al* when they analyzed a hybrid that transiently expressed mouse *Xist* (49).

To further assess the ability of the human XIST RNA to localize in a mouse cell, the mouse parent cell used to generate the hybrids was transfected with a phage artificial chromosome (PAC) that contained the *XIST* gene and over 100 kb of flanking sequences (151). FISH of ectopic *XIST* expression from these cells (Figure 3-1 B) showed that the XIST signal appeared to be generated from a specific site within the nucleus but drifted from the location, compared to both the mouse and human control cell lines which demonstrated a tightly localized XIST/Xist signal (Figure 3-1 C and D). The drift observed for the PAC that was integrated into a mouse chromosome (Figure 3-1 B) was less diffuse than the human XIST signal from the human X in the mouse background, AHA-4C1 (Figure 3-1 A). Thus, a hierarchy of localization is observed with very limited localization of human XIST to the human X in a mouse background and better localization when the XIST locus in integrated into a mouse chromosome but not comparable to human or mouse somatic XIST localization.

3.2.2 Absence of X-linked gene silencing upon induction of XIST/Xist

I directly assessed the transcriptional status of human and mouse X-linked genes in the AHA-4C1 cells by RT-PCR, but focused in on the *Xist* region (Figure 3-2). Previous studies have shown that both the human and mouse X chromosomes remained active upon induction of *XIST/Xist* expression (49,104). To test whether any silencing may occur closer to the *Xist* gene than previously examined, or after stable long-term *Xist* expression, I analyzed the expression of 14 genes, including two genes *Chic1* and *Abcb7* which were within 1Mb of *Xist*. Both mouse and human X chromosomes are shown schematically in Figure 3-2 with the placement of genes examined shown approximately to scale. Expression analysis with RT- PCR demonstrated that the expression of *XIST/Xist* did not cause the transcriptional silencing of any of the 14 X-linked genes examined (Figure 3-2 A). A female human cell line, GM7350, and a female mouse cell line, BMSL2, were used as controls to show the normal transcription status of the genes tested. All genes, except *XIST/Xist,* were expressed in the hybrid cell line, AHA-11aB1, prior to *XIST/Xist* induction and continued to be expressed in the AHA-4C1 hybrid with the reactivated *XIST/Xist* loci.

To assess small changes in gene expression, I performed quantitative RT-PCR on three mouse and one human X-linked gene (Figure 3-2 B). An additional cell line,

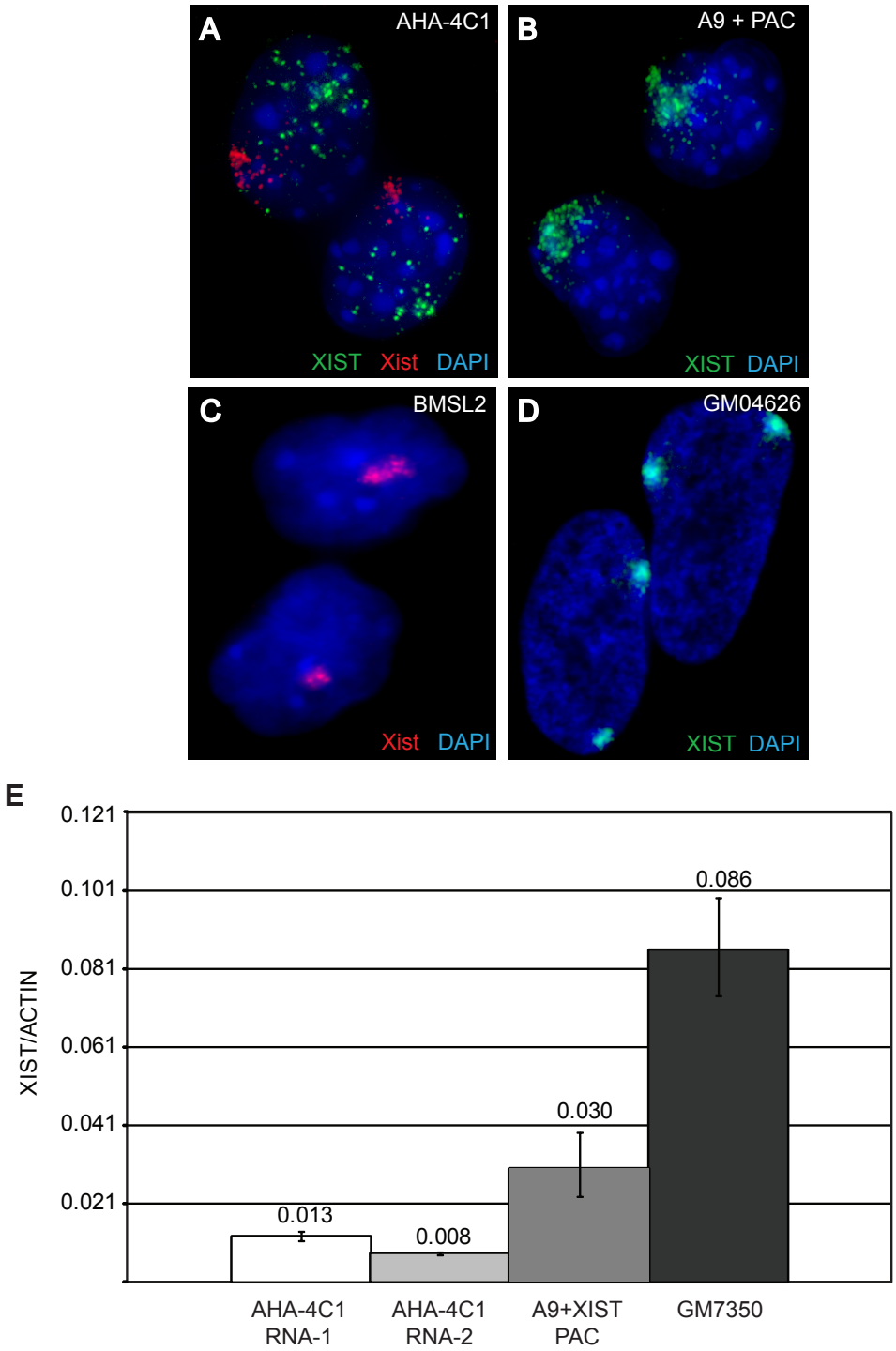

Figure 3-1: Analysis of XIST/Xist localization with dual-colour RNA FISH **(A-D)** and expression with RTqPCR **(E)**. **(A)** In the XIST+/Xist+ (AHA-4C1) somatic cell hybrid where treatment with demethylating agents resulted in XIST/Xist expression from the X chromosome, the human XIST RNA (green) signal is dispersed, while the mouse Xist RNA (red) signal is localized. **(B)** In the A9 mouse cell expressing XIST from a human PAC integrated into the mouse genome the human XIST RNA (green) signal is more disperse within the nucleus than seen for female mouse fibroblasts (BMSL2 cell line, shown in panel **C)** or the human female fibroblast line GM04626 (shown in panel **D**, karyotype: 47,XXX). **E)** qPCR for XIST expression relative to either mouse or human ACTIN. The two separate RNA samples for AHA-4C1 were isolated on separate dates. The error bars represent standard deviation of the triplicate qPCR error.

AHA-A52b, was included as a control for the 5-aza treatment to determine whether any expression changes were a result of the demethylation process itself or *XIST/Xist* expression. AHA-A52b reactivated only the human *XIST* gene upon demethylation (49), therefore any changes in the expression status of mouse X-linked genes in this control cell line, should reflect the effect of the demethylation treatment and not *Xist* expression. The relative expression ratio was determined by dividing the expression level (Ct value) of the gene of interest by *Actin* levels, and then calculating the fold change in relative expression of the reactivated cell lines versus the parental hybrid. Three out of the four genes analyzed showed a change in expression; specifically, mouse *Pgk1*, *Uba1*, and human *PGK1* had a decreased expression following the demethylation treatment since the decrease was seen in both the demethylated hybrid cell lines (AHA-A52b and AHA-4C1). The decrease in expression did not appear to be a result of the *XIST/Xist* expression since two mouse genes had this decrease in the intermediate cell line, AHA-A52b that only expressed human *XIST* and not mouse *Xist.* Hence, the expression levels of *Pgk1*, *Uba1*, and *PGK1* decreased as a result of the demethylation treatment, rather than *Xist* expression.

Both the X chromosomes in these hybrid cells remained active despite XIST/Xist expression. These results validate the somatic cell hybrid system as one in which XIST/Xist expression and the silencing of X-linked genes has been disconnected, thereby providing a model system in which X-chromosome inactivation marks can be studied with respect to *XIST/Xist* expression in the absence of silencing.

3.2.3 Cot-1 hybridization shows ‘domains’ not diminished by XIST without localization

Cot-1 DNA can be hybridized to RNA within the nucleus to identify heteronuclear RNA transcription (54). Lack of Cot-1 staining has been observed at the location of the Xi and appropriately termed a ‘Cot-1 hole’. It has been suggested that this Cot-1 hole reflects X chromosome silencing (54) and the formation of a repressive nuclear compartment (12). To determine the ability of the XIST/Xist RNA to silence the Cot-1 fraction and form a repressive nuclear compartment, I examined the presence of a Cot-1 hole in the hybrid cells using dual-RNA FISH with an *XIST/Xist* probe and the corresponding human or mouse Cot-1 probe (Figure 3-3). The commercially

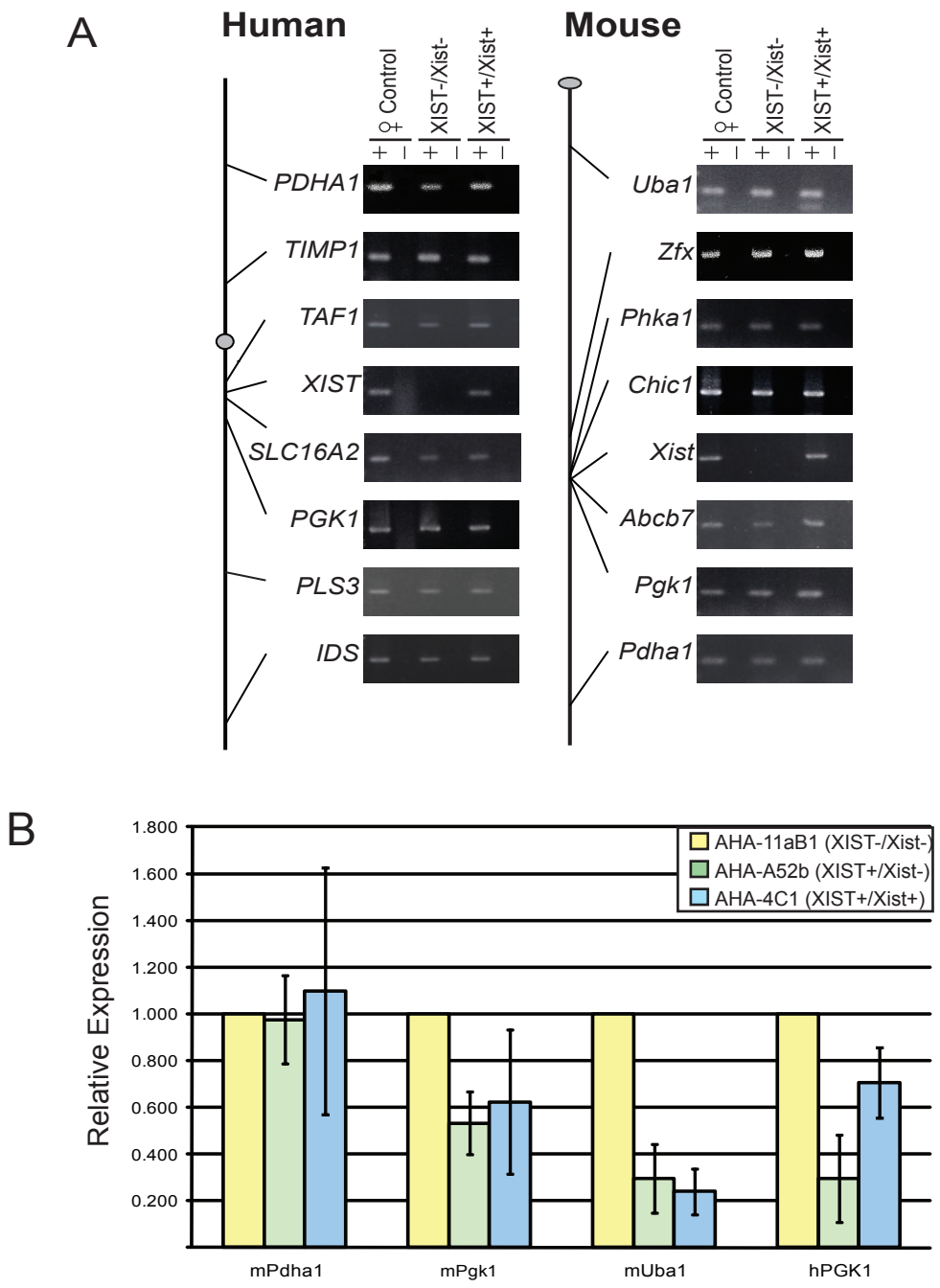

Figure 3-2: X-linked gene expression in somatic cell hybrids before and after expression of XIST/Xist. **(A)** Schematic diagram of both the human and mouse X chromosomes showing the location of the genes analyzed is at the left of gel images of RT-PCR products of cDNA from the parental hybrid not expressing AHA-11aB1 (XIST-/Xist-) and the derivative hybrid that expresses AHA-4C1 (XIST+/Xist+). Control female cDNA was from GM7350 (human) and BMSL2 (mouse) cell lines. **(B)** Quantitative RT-PCR of four X-linked genes. An intermediate hybrid AHA-A52b (XIST+/Xist-) was analyzed for expression along with the hybrid AHA-4C1 (XIST+/Xist+). The gene expression was normalized to Actin and expressed as a fold-change relative to the parent hybrid AHA-11aB1 (XIST-/Xist-). The error bars represent the standard deviation of three separate RNA isolations.

available Cot-1 DNA that was used to make the probe was species-specific; mouse Cot-1 (mCot-1) showed no hybridization to human RNA under our hybridization conditions and vice-versa.

In the hybrid cell line that expresses *XIST/Xist,* AHA-4C1, I found that only 7% of the cells contained a mCot-1 hole corresponding to the localized Xist signal (n=83), compared to 59% (n=80) in the mouse control cell line (Figure 3-3 B and C). To determine the signal intensity of mCot-1 RNA at the site of Xist localization, I used line scans (ImageJ, NIH). The line scans confirmed that the increased signal intensity for Xist RNA did not always correspond with a decrease in signal intensity for mCot-1 RNA (Figure 3-3 D). Furthermore, not only were there fewer mCot-1 holes in the hybrid cells, but in those that were identified, the lack of mCot-1 signal was also less obvious compared to a mouse control. The drastic reduction in the formation of Cot-1 holes showed that the expression of *Xist* in a somatic cell does not consistently promote the appearance of a Cot-1 hole.

The demethylated cells that expressed *XIST/Xist,* AHA-4C1, also showed several mCot-1 bright regions that varied in size and number from cell to cell (Figure 3-3 C and D, white arrows). Since these cells were demethylated to induce *XIST/Xist* expression, I analyzed an Xi hybrid cell (t11-az-10) as a control since it expressed *XIST* before the demethylation treatment (152) and I also observed the mCot-1 bright regions. The presence of Cot-1 foci in both these demethylated cell lines suggests that their appearance is a result of the demethylation treatment and not *Xist* induction.

The dispersed nature of the human XIST RNA signal prevented the assessment of hCot-1 hybridization associated with the XIST RNA. However, the pattern of the hCot-1 hybridization in these cells was of interest. One definitive hCot-1 RNA domain per cell was identified within the hybrid cells which corresponded to the one human chromosome (Figure 3-3 E and F). To ascertain if other human chromosomes would form these Cot-1 domains, I analyzed hybrid cells with more than one human chromosome. In cell lines with more human chromosomes, there was an equivalent increase in the number of hCot-1 RNA domains (Figure 3-3 G). Since Cot-1 RNA hybridization represents heteronuclear RNA transcription, this observation suggests that in human-mouse fibroblast hybrid cells, human chromosomes form a transcriptional

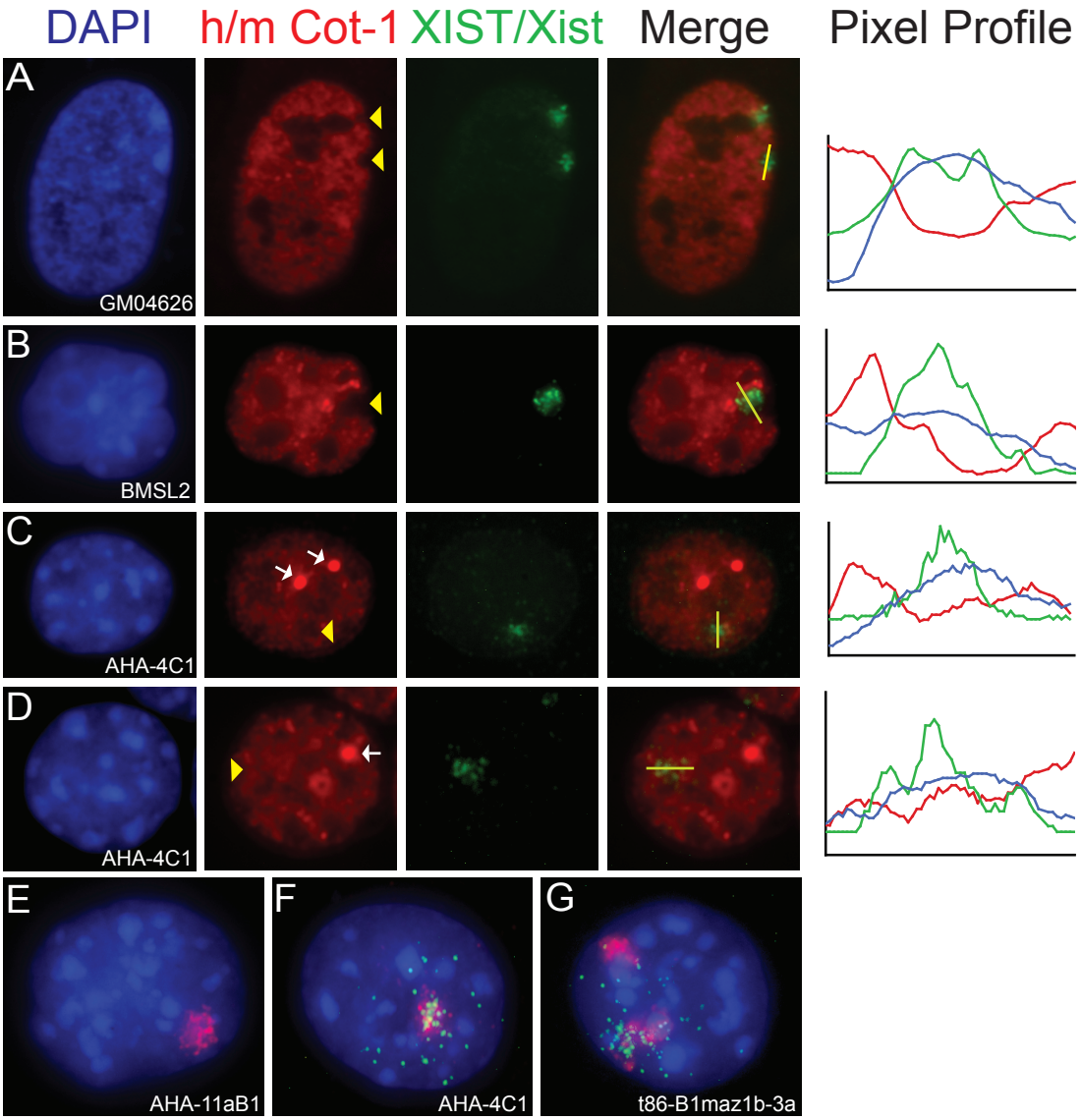

Figure 3-3: Analysis of Cot-1 hybridization in somatic cell hybrids expressing Xist/XIST by dual RNA FISH hybridization. Cot-1 RNA (mouse (m) or human(h)) signal is red, Xist/XIST RNA signal is green and DAPI is blue. (A-D) Lines were drawn in ImageJ (NIH) and pixel intensities of each signal was plotted across the yellow line and shown to the right of the images. Yellow arrow heads indicate where the Xist signal corresponds to the Cot-1 signal. (A) Human female cell line, GM04626, with two inactive X chromosomes and two hCot-1 holes corresponding to the two XIST RNA signals. (B) Mouse female cell line, BMSL2, showing one mCot-1 hole corresponding to the mouse Xist signal, this pattern was observed in 59% of cells (n = 100). (C) XIST+/Xist+ somatic cell hybrid, AHA-4C1, showing a mCot-1 hole corresponding to the mouse Xist signal, this pattern was observed in 7% of cells (n = 100). (D) AHA-4C1 without a mCot-1 hole. In panels C and D bright foci of mCot-1 staining are observed (white arrows). (E) hCot-1 RNA expression (red) in the XIST-/Xist- hybrid cell, AHA-11aB1, containing only one human chromosome (the X chromosome). (F) hCot-1 RNA expression (red) in the XIST+/Xist+ hybrid cell, AHA-4C1, containing only one human chromosome (the X chromosome), and XIST RNA (green) drifting from the hCot-1 domain. (G) An Xi-hybrid cell line, t86-B1maz1b-3a, containing two human chromosomes (X + 15). Both human chromosomes form a hCot-1 transcriptional domain, one with the XIST signal drifting away.

domain as visualized by hCot-1 hybridization.

After I found that the hCot-1 transcriptional domain corresponded to the Xa in the AHA-4C1 hybrid cells, I wanted to examine the effect of gene silencing on the Cot-1 transcriptional domain. To this end, I hybridized hCot-1 in a hybrid cell that contained an Xi, t86-B1maz1-3a (110). In this particular cell line, the only human chromosomes contained within the mouse cell are the Xi and chromosome 15. If the Xi silenced hnRNA expression, as generally observed with the establishment of a Cot-1 hole, then only one hCot-1 domain would be observed corresponding to the human chromosome 15. However this was not the case, two hCot-1 domains were observed, one of which corresponded to the origin of the XIST signal (Figure 3-3 G). This suggests that the repetitive DNA elements contained within the hCot-1 fraction were expressed from the Xi, similar to an autosome, in these somatic cell hybrids, despite the ongoing inactivation of X-linked genes (112).

3.2.4 Retention of DNA hypomethylation at X-linked gene promoters upon induction of Xist/XIST

To analyze the methylation status of X-linked genes in the hybrid cells, I digested genomic DNA with a methylation-sensitive restriction enzyme followed by PCR with primers designed to flank three or more recognition sites. If any site within the amplicon was unmethylated, the enzyme would cut the template and there would not be a resultant PCR product. However, if the sites were methylated, the enzyme would not be able to cut the template and a band would be observed. Methylation analysis of 11 X-linked genes showed that all the genes were hypomethylated in the cell line prior to *XIST/Xist* induction, AHA-11aB1, and continued to be hypomethylated in the hybrid expressing *XIST/Xist,* AHA-4C1 (Figure 3-4). Both the *XIST/Xist* genes lacked a change in promoter methylation that corresponded to their expression state. In the hybrid cell that did not express *XIST/Xist,* AHA-11aB1, both *XIST/Xist* gene promoters were methylated and in the hybrid that did express *XIST/Xist,* AHA-4C1, both promoters were not methylated. Therefore, the DNA methylation state at the X-linked gene promoters is concordant with the X chromosome remaining active. *XIST/Xist* expression alone did not initiate DNA methylation changes at the promoters of X-linked genes in the somatic cell hybrids.

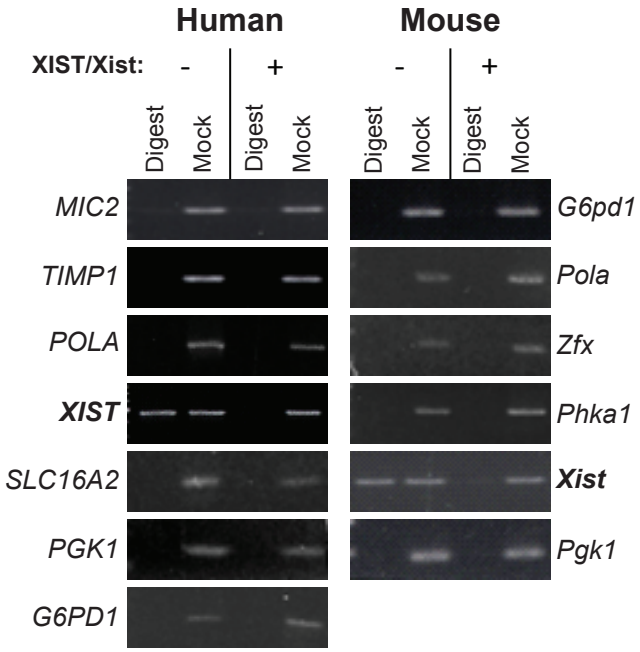

Figure 3-4: Analysis of DNA methylation at the promoters of X-linked genes in somatic cell hybrids, AHA-11aB1 (XIST-/Xist-) and AHA-4C1 (XIST+/Xist+). Genomic DNA was cut with a methylation-sensitive restriction enzyme prior to amplification by PCR with primers flanking the enzyme cut sites (lanes labeled Digest). Mock digested samples were used as a positive control (lanes labeled Mock). If a band occurs in the Digest lane, it means that the site/sites within the amplicon are methylated (e.g. XIST/Xist). If no band was observed after PCR then there was no intact template to amplify as the internal restrictin enzyme sites were not methylated.

3.2.5 Retention of active histone modifications at X-linked gene promoters

To further assess whether active chromatin can be maintained in the presence of XIST/Xist, I examined the promoter regions of X-linked genes for the active chromatin modifications H3K4 dimethylation (H3K4me2) and H3 acetylation (H3Ac) using chromatin immunoprecipitation (ChIP) followed by quantitative PCR. The only genes shown to change expression as a result of the demethylation treatment were the *XIST/Xist* genes which were not expressed in the AHA-11aB1 hybrid and were expressed in the AHA-4C1 hybrid (Figure 3-2). After expression of *XIST/Xist* in the AHA-4C1 hybrid cell line, a gain of H3K4me2 and H3Ac was observed at both *XIST/Xist* promoter regions (Figure 3-5). Since human XIST did not localize to the human X chromosome while mouse Xist was localized, I focused on mouse genes as these were the most likely to be affected by *Xist* activation. None of the mouse X-linked genes, or the human *ELK1* gene, showed any significant decrease of these active marks as a result of *XIST/Xist* expression (Figure 5). Therefore, these active state chromatin marks are retained at the promoter regions of X-linked genes despite expression and localization of the Xist RNA.

3.3 Discussion

I have described the analysis of a somatic cell hybrid that stably expresses *XIST/Xist* from the endogenous loci on active human and mouse X chromosomes in an attempt to determine which epigenetic features associated with the Xi directly result from *XIST/Xist* expression rather than developmentally-associated silencing. My results support the view of a developmental window of opportunity for Xist-dependent silencing (31) and further show that the inability to silence also applies to genes located in close proximity to the *Xist* locus. I found that localized *Xist* expression without silencing did not initiate the efficient formation of a repressive nuclear compartment or the loss of active chromatin marks, which are thought to be among the earliest events in the timeline of X-chromosome inactivation (Refer to Figure 1-1). I suggest that these events require additional factors beyond *Xist* expression with proper localization. My results also propose that a developmental context may contribute to the establishment of a repressive nuclear compartment by the absence of Cot-1 and that the lack of Cot-1

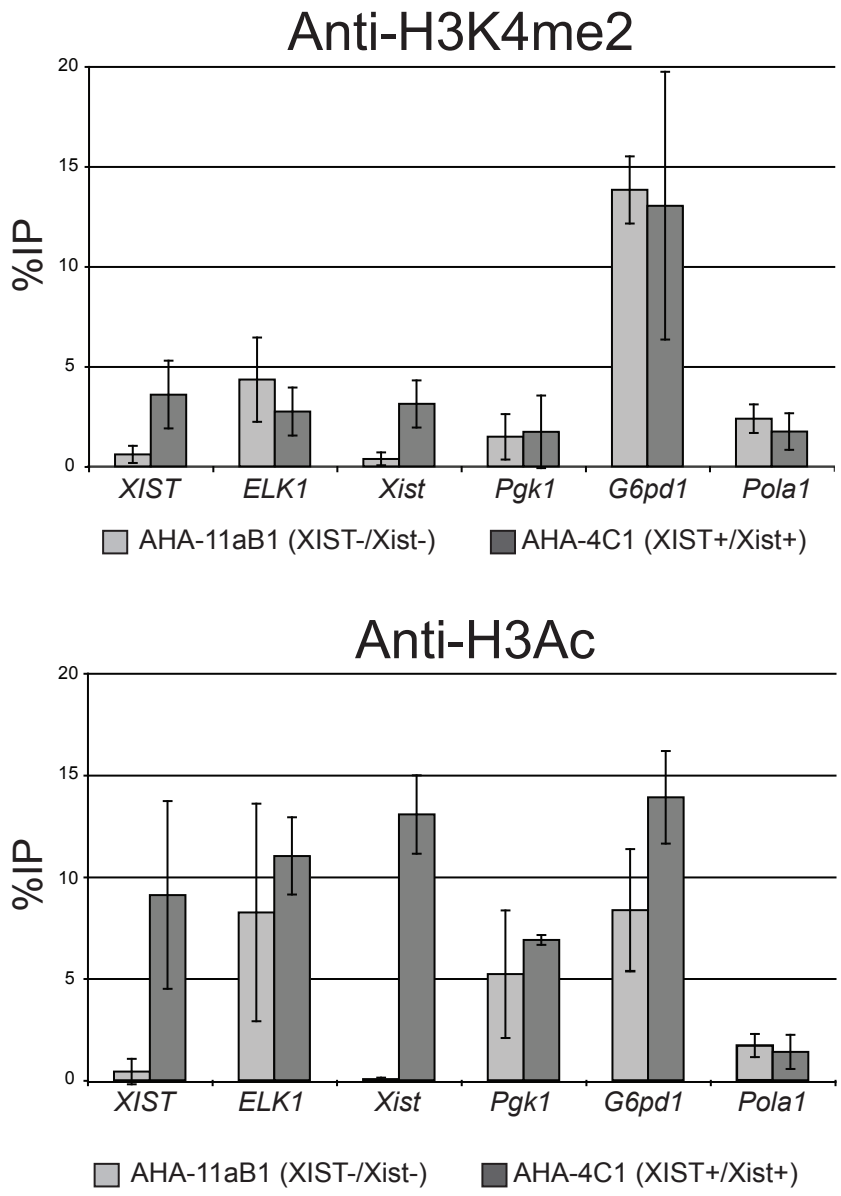

Figure 3-5: Active histone modifications are retained at the promoters of X-linked genes after expression of *XIST/Xist*. H3K4me2 (upper) and H3Ac (lower) marks were assessed by ChIP and followed by quantitative PCR at the promoter region for the two human and four mouse genes listed. The %IP is relative to input and error bars indicate error between two to three replicate immunoprecipitations.

does not represent X-linked gene silencing.

There are a variety of localization states that have been observed with XIST/Xist transgenes and interspecies expression in the literature. I have shown that a hybrid cell line, AHA-4C1, which stably expressed *XIST/Xist,* displayed a delocalized human XIST RNA signal within the nucleus, despite localization of mouse Xist. Other studies that have looked at *XIST* expression from a human X in rodent cells have similarly shown a delocalized XIST signal. These include the precursors to these Xa-containing hybrid cells (49,104), human/hamster Xa hybrid cells (105), and Xi-containing hybrid cells (49). Together this suggests that human XIST RNA is not able to localize to the human X chromosome in a mouse somatic background. Since mouse Xist was able to localize in the somatic cell hybrids, it has been suggested that the lack of proper XIST localization could be the result of some missing species-specific localization factor(s) (49).

I also looked at the localization of the human XIST RNA, when it was expressed from a mouse chromosome in a mouse background (Figure 3-1 B). The XIST RNA was more localized when it was expressed from a mouse chromosome in a mouse somatic cell compared to the XIST RNA when it was expressed from a human chromosome in the same mouse background (Figure 3-1 A); however, not as precisely localized as when expressed in a human background (Figure 3-1 C). The hypothesis that XIST/Xist coats the chromosome via way-stations (reviewed in (153)) at the DNA level suggests an unknown link between the XIST/Xist RNA and its association with the DNA.

Interestingly, when a YAC containing the *XIST* gene and over 300 kb of human DNA was integrated into a mouse ES cell line, the human XIST RNA localized in *cis* to a mouse autosome (154,155). The localization occurred in undifferentiated ES cells, differentiated EB cells that had expressed *XIST* throughout differentiation, and embryonic tissues of chimeric mice. The lack of proper XIST localization in somatic cells combined with the localization in ES cells suggest a developmental context may be required for proper localization of human XIST RNA to mouse DNA in a mouse background. However, copy number variability, integration site, and expression levels of XIST may also influence localization, thereby making it difficult to determine if there is an additional developmental influence on the ability to localize XIST.

The localization of the mouse Xist RNA in the somatic cell hybrids supports the species-specific requirement, since *Xist* is expressed in a mouse background. While the exact mechanism of Xist localization in *cis* remains unknown, the somatic cell hybrids do show that localization is not dependent on a developmental cue. The localization of the Xist RNA is consistent with previous studies that have observed XIST/Xist localization outside a developmental context (2,31,54,106). If there are proteins that allow Xist to localize to DNA, then they must be expressed in these somatic cells.

There is evidence that prior to implantation in mice, a gradient of silencing exists whereby genes closer to the *Xic* are silenced first (149,150); however, no such gradient was shown in the *XIST/Xist*-expressing hybrid cells. The ongoing transcription of X-linked genes, even those in close proximity to the *Xist* locus, confirmed that localization and silencing are disconnected in the hybrid system and enabled the analyses of epigenetic events associated with *Xist* expression alone in somatic cell hybrids. The lack of silencing in the hybrid cells, supports the concept that there is a window of opportunity during development when *Xist*-mediated silencing can occur (31) and further show that this developmentally restricted silencing also holds true for those genes located close to the *Xist* locus.

To determine if Xist-induced epigenetic marks are dependent on silencing during ES cell differentiation, several studies have used an *Xist* transcript lacking the 5' A-repeats, *Xist*Δ*sx*, that can localize properly without inducing transcriptional silencing (12,15,38,107,108). The XistΔsx studies have shown that during differentiation, silencing is not required for the enrichment of H3K27me3 and H4K20me1(15); the recruitment of macroH2A1(38); and the exclusion of RNA PolII (12) or Cot-1 RNA (12). However, it is important to note that compared to controls, the degree to which these events occurred with a silencing impaired Xist RNA was varied (See Table 1-1). If domains of the RNA, other than the 5' A-repeat, recruit epigenetic marks, it could be possible for the Xist RNA lacking the A-repeat to recruit epigenetic events in the absence of silencing; although, efficiency of recruitment appears dependent on the developmental window and the A-repeats.

Cot-1 is a fraction of genomic DNA obtained by the extraction, shearing, denaturation and reannealing of DNA under conditions that enrich for repetitive DNA

sequences. When Cot-1 is used as a probe for RNA FISH, a Cot-1 hole is associated with the Xi. This likely represents a repressive nuclear compartment that predominantly associates with repetitive non-genic silencing (12,46). I have shown that human chromosomes, within a mouse nucleus, form transcriptional domains. When I analyzed human Cot-1 expression with RNA FISH in the hybrid cells containing an Xa (AHA-11aB1 and AHA-4C1), both the cell lines had a single region of hCot-1 expression. The observation of a human Cot-1 domain suggested that in these Xa-containing hybrid cells, human XIST was not able to silence hCot-1 RNA, in addition to X-linked genes. When there were additional human autosomes within the nucleus, there was the equivalent number of these hCot-1 transcription domains. Therefore a Cot-1 domain, rather than a Cot-1 hole, was observed corresponding to the X chromosome, similar to an autosome, when the XIST RNA was not able to form a repressive compartment.

I also observed an hCot-1 transcriptional domain associated with the Xi in a hybrid cell, where the X-linked genes have been shown to be inactive (113). Therefore, the Cot-1 RNA was not representative of genic silencing, since an Xi had an associated Cot-1 domain (Figure 3G). In the Xi-containing somatic cell hybrids, gene silencing was maintained while repetitive silencing was not; thus the Xi was represented by a Cot-1 domain region and not a Cot-1 hole. The idea that the Cot-1 hole is not representative of genic silencing was also supported by a previously published study that observed the kinetics of Cot-1 hole formation during mouse ES cell differentiation (12). The silencing of the Cot-1 RNA was shown to occur prior to genic silencing and even a silencing-deficient *Xist* transgene, *Xist*Δ*sx*, was able to induce Cot-1 holes (12).

A Cot-1 hole can occur without genic silencing and conversely, my results have shown that genic silencing was maintained without a Cot-1 hole. This suggests that silencing of the Cot-1 fraction and genic silencing can be mutually exclusive. My results support the idea that the Cot-1 hole reflects a repressive nuclear compartment that predominantly associates with repetitive non-genic silencing (12,46).

In the Xa hybrid cells, where mouse Xist was able to localize but not induce gene silencing (AHA-4C1), I observed a substantially lower number of Cot-1 holes (7%), as compared to a control cell line (59%). The observation of very few Cot-1 holes contrasts with the previously mentioned study that saw the *Xist*Δ*sx* transgene, expressed during

ES differentiation, form a Cot-1 hole in the absence of gene silencing (12). Therefore, when *Xist* was activated post-differentiation in these hybrid cells, Xist localizes, gene silencing was not observed and repetitive silencing was not efficiently initiated. In contrast, when an established Xi was introduced, the XIST RNA showed aberrant localization and gene silencing was maintained while repetitive silencing was not. This suggested that a silent nuclear compartment was not able to be maintained without a localized XIST RNA signal, or alternatively, that a developmental context was required for the initiation of the repressive compartment.

While it was previously shown that *XIST* expression is not required for maintenance of gene silencing (156), it would be of interest to determine if repeat silencing was also maintained in the absence of XIST. My findings have shown that the somatic cell hybrids were not efficient at forming Cot-1 holes following *Xist* expression and Chaumeil *et al* have shown that formation of Cot-1 holes occurred without silencing in differentiating ES cells (12). Taken together, these results support the idea that Xist-induced silencing, and now the formation of a Cot-1 hole, is dependent, to some extent, on a developmental context.

Overall, my results support the view that the Cot-1 hole, associated with the Xi, is not a direct indicator of gene silencing. The somatic cell hybrids were not able to efficiently form a Cot-1 hole and therefore are unable to silence Cot-1 RNA, in addition to X-linked genes. Consistent with the formation of a Cot-1 hole, human chromosomes in a mouse nucleus can form a transcriptional domain. Furthermore, it appears that a developmental context is required for effective Cot-1 hole formation.

A cascade of epigenetic events leads to the characteristic Xi chromatin following *Xist* expression. X-linked gene silencing immediately follows Xist accumulation and histone modifications are amongst the earliest known events, with the loss of active chromatin modifications appearing to precede the gain of inactive marks (157).Therefore, I looked at active marks at the promoters of X-linked genes following *Xist* reactivation in the hybrid cells. I found that the promoters I examined retained both H3 acetylation and H3K4me2 despite *Xist/XIST* expression (AHA-4C1). It was of interest that both the human XIST and mouse Xist promoters acquired these active marks following demethylation induced expression. Therefore, in the somatic cell hybrids, Xist

localization was not able to induce the earliest known chromatin modifications outside of a developmental context.

Methylation at X-linked gene promoters appears to play a role in the maintenance of the inactive state as it occurs after initial inactivation (7). I looked at the promoters of X-linked genes in the hybrid cells where both human and mouse X chromosomes were determined to be active. All the gene promoters that I examined were hypomethylated in the cell line expressing *Xist/XIST*. Therefore, the demethylation treatment and the resultant *Xist/XIST* expression did not cause these X-linked genes to become hypermethylated. The methylation state of these genes corresponded to the expression status, such that, all genes examined remained active and hypomethylated at their promoters with *Xist/XIST* expression.

In summary, studying the somatic cell hybrid system has shown that stable *Xist* expression from its native locus leads to proper localization outside a developmental context. Human XIST was not able to properly localize whether expressed from a mouse or human chromosome, whereas mouse Xist localized, therefore suggesting that factors required for localization are not conserved in mouse somatic cells. X-linked genes remained active despite localized *Xist* expression and proximity to the *Xist* locus. The localization of the Xist RNA and lack of silencing in the hybrid cell, allowed me to demonstrate that the ability to silence repeat elements and form a Cot-1 hole was hampered. In addition, the X chromosome remained identifiable as an Xa, retaining DNA hypomethylation and both H3K4me2 and H3Ac at promoter regions. My results support the view that a developmental context is required for Xist-mediated genic silencing and the formation of a repressive nuclear compartment; additionally that the Cot-1 hole associated with the Xi is not an indicator of gene silencing.

Chapter 4: The validation of a tagged XIST RNA system

4.1 Introduction

X-chromosome inactivation is initiated by the XIST RNA which results in a cascade of events leading to the characteristic Xi chromatin (reviewed in (158)). It is thought that XIST recruits proteins involved in setting up the Xi chromatin; however, to date these specific proteins remain elusive. To determine the mechanism by which the expression of a non-coding RNA leads to chromosome-wide transcriptional silencing, it is critical to identify the XIST-interacting proteins.

Several of the methods that are used to identify RNA-interacting proteins are immunoprecipitation-based. In immunoprecipitation-based methods, there must be a candidate protein with experimental evidence that suggests an interaction with the RNA of interest, and an antibody against the candidate protein to immunoprecipitate the protein and its interacting RNAs. The antibodies to macroH2A, Ezh2, and Suz12 have each immunoprecipitated the XIST/Xist RNA and suggest either a direct or indirect association as a result of formaldehyde fixing prior to the immunoprecipitation that crosslinks molecules within close proximity (114,127).

With the identification of novel protein-XIST interactions, a clearer picture would emerge regarding the mechanism by which XIST acts to silence the X chromosome. It would be advantageous to isolate the XIST RNA itself and analyze what proteins are bound. To do this, the RNA must be tagged, so that the tag can be used to isolate the RNA and its associated proteins. There are several methods that have been used for the purpose of tagging an RNA including chemical tagging, hybridization of affinity-tagged oligonucleotides, incorporation of an artificially selected RNA motif or of a known protein-binding RNA sequence during transcription (159). Unfortunately, no single method is perfect and there are downfalls to each tagging method.

Chemical tagging with modified ribonucleotides might cause detrimental structural changes to the RNA. With the hybridization of affinity-tagged oligonucleotides method, a region of the RNA of interest is selected that is accessible. Accessibility may pose a problem for highly organized RNAs. Additionally, this method must be performed under denaturing conditions or with a competitor oligonucleotide to release the RNA and its associated proteins under native conditions. The incorporation of an RNA sequence from a known protein-RNA interaction has been used under native conditions. Since the

binding affinity for the interaction is strong, the efficiency of the purification/elution could be challenging and requires a separation scheme.

To isolate the XIST RNA with its associated proteins, I exploited an RNA-protein interaction found in the MS2 bacteriophage. MS2 is an RNA bacteriophage of *Escherichia coli* that contains a capsid protein which, in addition to its structural role, also represses translation by binding to an RNA stem loop in the replicase gene (reviewed in (160)). This RNA tagging and coat protein affinity method takes advantage of the well characterized binding of the MS2 coat protein to a specific RNA stem loop sequence. The interaction has been optimized to reduce capsid formation and increase affinity of the protein for the stem loop (160).

The MS2 tagging system has been used as a method for the *in vivo* labeling of RNA transcripts where the RNA of interest contains the stem loops (a range from 3-24 MS2 stem loops have been used) and the coat protein is fused to a fluorescent reporter protein (i.e. GFP) for visualization (161-164). Knowledge has been gained in the transcriptional dynamics and movement of RNA molecules within human cell lines (161-165) . Additionally, potential RNA-protein interactions of a non-coding RNA, NRON, were also identified using the MS2 phage tagging system (166), as were the protein components of the human spliceosome (167). The NRON RNA was tagged with three MS2 loops at the 3' end of an exon and a MS2/maltose fusion binding-protein was used to purify interacting proteins from whole cell protein extracts and for identification by mass spectrometry (166).

In order to isolate the XIST RNA and identify its interacting partners, the MS2-tagged XIST RNA would have to be expressed in a human system that allows for silencing. The human male HT1080 fibrosarcoma cell line is one such cell line that has been published in the literature (106). I utilized an inducible XIST system in the HT1080 cells where any transgene can be integrated into a single reproducible site within the genome. A previous study that used this model system demonstrated that induced expression of an *XIST* cDNA transgene led to localization and some degree of silencing in the HT1080 cell line (106). Therefore, it is possible to integrate the MS2-tagged XIST into the same site within the HT1080 genome and compare its localization and silencing ability with a non-tagged XIST RNA.

To determine novel XIST-protein interactions, I fused a series of eight MS2-stem loops to the 3' end of a full length XIST cDNA construct, and integrated this into the human HT1080 fibrosarcoma cell line. I set out to characterize the expression of *XIST* and some inactive marks to determine if the hybrid RNA was behaving similar to its full-length counterpart. With an affinity tagged XIST RNA functioning similar to XIST without the tag, it would be possible to isolate the RNA and determine what proteins are interacting with XIST. These proteins would provide evidence for the manner in which XIST initiates transcriptional silencing, formation of a silent compartment, and chromatin remodeling.

4.2 Results

4.2.1 Generation of the MS2-tagged XIST transgene

To identify XIST-interacting proteins, I generated a tagged XIST RNA. The RNA was tagged by the fusion of eight MS2 stem loops to the 3' end of an XIST cDNA transgene under the control of a doxycycline-inducible CMV promoter. The MS2 loops were amplified with primers spanning the loops; the resulting MS2 loops insert and the XIST cDNA expression vector were digested and ligated together. The ligations were electroporated into competent cells and screened for the appropriate integration of the eight MS2 loops at the 3' end of the XIST cDNA transgene with PCR primers located at the 3' end of the transgene and spanning the MS2 loops insert.

The DNA was isolated from positive cultures containing the XIST cDNA with eight MS2 loops fused to the 3'end and was co-transfected with a plasmid expressing the Flp recombinase protein into the male human fibrosarcoma cells, HT1080. Two separate HT1080 cell lines were used; each contained one FRT site with a known location which allows for retargeting of transgenes to the same chromosomal location. In one clone the FRT site was located in the p arm of chromosome 1 (1p integration), and in the second clone the FRT site was located in the q arm of chromosome 3 (3q integration). Therefore the MS2-tagged XIST transgene was integrated into the p arm of chromosome 1 (1p integration) in one cell line and into the q arm of chromosome 3 (3q integration) in the second cell line. The use of these specific HT1080 cells allowed for the reproducible integration of different transgenes and enabled me to analyze the

functioning of the XIST RNA versus the MS2-tagged XIST RNA in the same integration site (either 1p or 3q) under the control of the same inducible promoter.

4.2.2 Expression of the MS2-tagged XIST transgene

To examine the expression of the MS2-tagged transgene, the HT1080 transfected clones were grown in selective hygromycin media and were induced with doxycycline (Dox) to express *XIST* and screened for expression using RNA FISH with an XIST probe. Seven single cell clones with the MS2-tagged transgene at the 1p integration were successfully grown under hygromycin selection, however, only one of these were shown to express XIST by RNA FISH and qPCR. Similarly, only one of four single cell clones with the MS2-tagged transgene at the 3q integration, grown under hygromycin selection, was shown to express XIST. Therefore, in total, two clones containing the MS2-tagged XIST were found to have upregulated XIST RNA with Dox induction (Figure 4-1). One clone was integrated at 3q and contained a downstream EGFP reporter gene and the other was integrated at 1p.

Prior to Dox induction, a very small pinpoint XIST + loops RNA signal could be seen in some cells with RNA FISH and upon induction with Dox the XIST + loops RNA became a localized focus within the nucleus similar to the full length XIST cDNA construct (Figure 4-1 A). The same nuclear pattern of expression of the XIST + loops construct and the XIST cDNA was seen which suggested that the XIST + loops RNA was functioning to localize appropriately.

To assess the level of *XIST* gene expression in the XIST + loops clones, I performed RT-qPCR. The expression of *XIST* in both uninduced and Dox-induced cells is shown side by side for the two XIST + loops clones in addition to the full length XIST cDNA clone (Figure 4-1 B). The expression levels are expressed relative to *ACTIN* and compared to that of a female lymphoblast cell line, GM7350. The expression of the XIST RNA from both the XIST + loops integrations was less than that in the XIST cDNA cell line and the control female cell line. The expression of the XIST + loops transgene at the 1p integration was 34% of the female lymphoblast level and at the 3q integration was 46% of the female level. The expression level of the XIST cDNA transgene at 3q was 69% of the female level (Figure 4-1 B). Dox-induced expression of the XIST cDNA

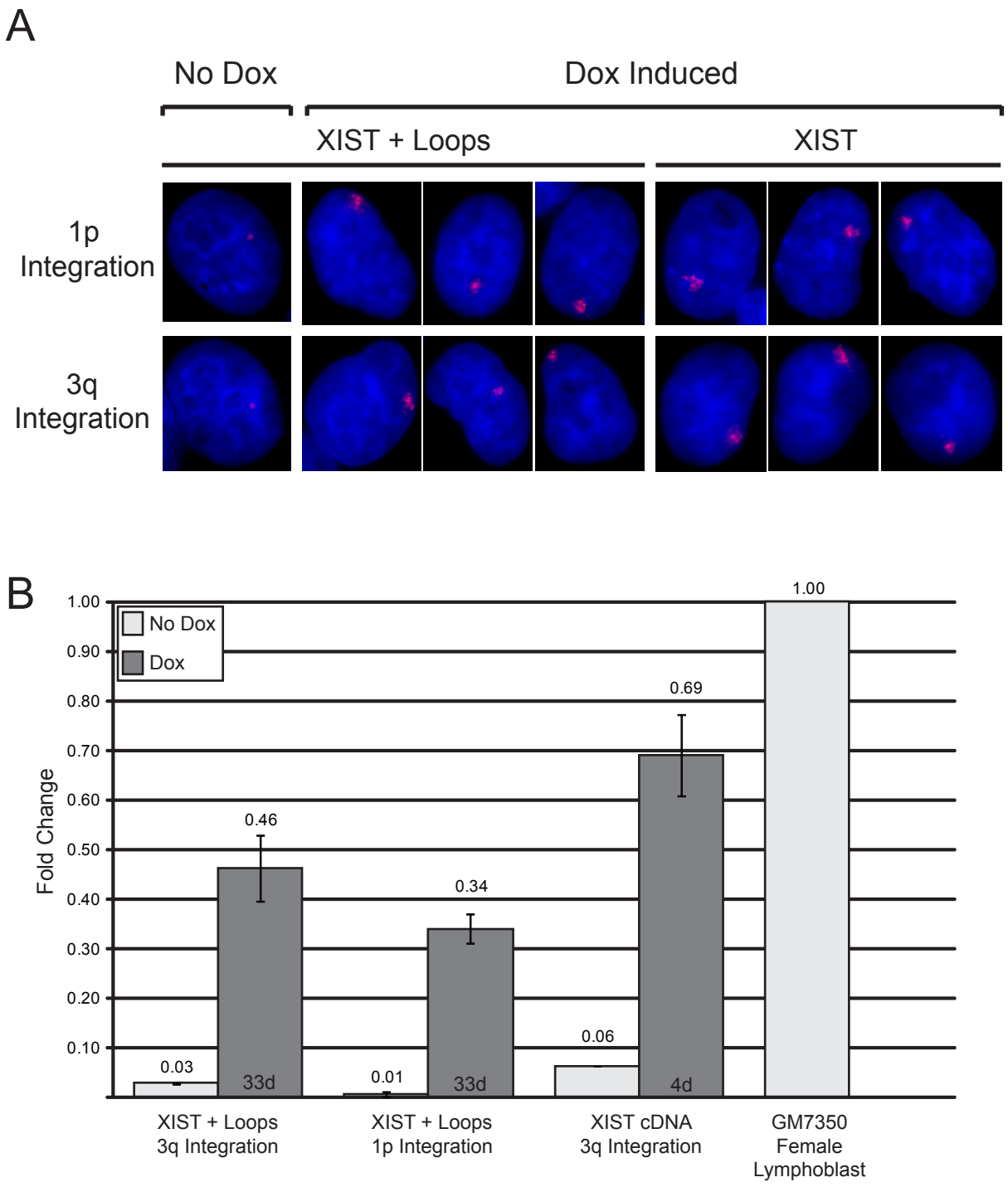

Figure 4-1: The expression of the tagged *XIST* transgene in the HT1080 cells. **A)** RNA FISH with XIST to show expression of the *XIST+Loops* transgene with No Dox and with 42 days in Dox. **B)** XIST expression compared to GM7350 relative to *ACTIN* using qPCR and analyzed with the ddCt method. The data represents a duplicate qPCR performed with the same RNA therefore the error bars show the standard error between the two values.

and the XIST + loops in the HT1080 cell lines were less than the expression of a control female lymphoblast.

4.2.3 Partial *EGFP* repression by the MS2-tagged transgene

The 3q integration site in the HT1080 cell line harbours a downstream *EGFP* reporter gene. Therefore, to examine in *cis* silencing of the *EGFP* promoter I analyzed *EGFP* expression in uninduced and Dox-induced cells using RT-qPCR and fluorescence-activated cell sorting (FACS). Figure 4-2 shows that both methods demonstrated a loss in *EGFP* expression compared to the no Dox control. The RT-qPCR data of the *EGFP* expression is shown in Figure 4-2 A, compared to the appropriate No Dox control and relative to *ACTIN* expression. The amount of *EGFP* expression for the XIST + loops is 38% reduced compared to no Dox and the full length XIST cDNA where expression is 82% reduced. I confirmed the qPCR results with the FACS data shown in Figure 4-2 B; where the numbers are displayed relative to the no Dox control having 100% *EGFP* expression. After one day of Dox induction, the expression of *EGFP* decreased approximately 10% and further decreased 45% at five days. After five days, the *EGFP* expression maintained an average reduction of 40% over the 19 days that were analyzed. The histogram plot from the FACS analysis showed that there was a single population of cells that reduced EGFP expression and not a population that silenced EGFP and a population that did not silence EGFP. I was not able to visualize both *EGFP* expression and *XIST* expression in individual cells because the RNA FISH protocol results in a loss of EGFP signal. Previously, the XIST cDNA transgene was expressed at the same 3q integration site as the XIST + loops clone and the EGFP decreased by 80% after four days of induction (106). These numbers complement the qPCR data and verify that the XIST + loops RNA silenced the *EGFP* reporter gene to a lesser extent than the XIST RNA without the loops.

4.2.4 Cot-1 hole analysis

As previously described, RNA hybridization of a Cot-1 probe can identify heteronuclear RNA transcription and the lack of Cot-1 hybridization, termed the Cot-1 hole, has been observed to correspond with the Xi in female cells. Therefore, to further examine the effects of the induced XIST + loops RNA, I performed dual-RNA FISH with

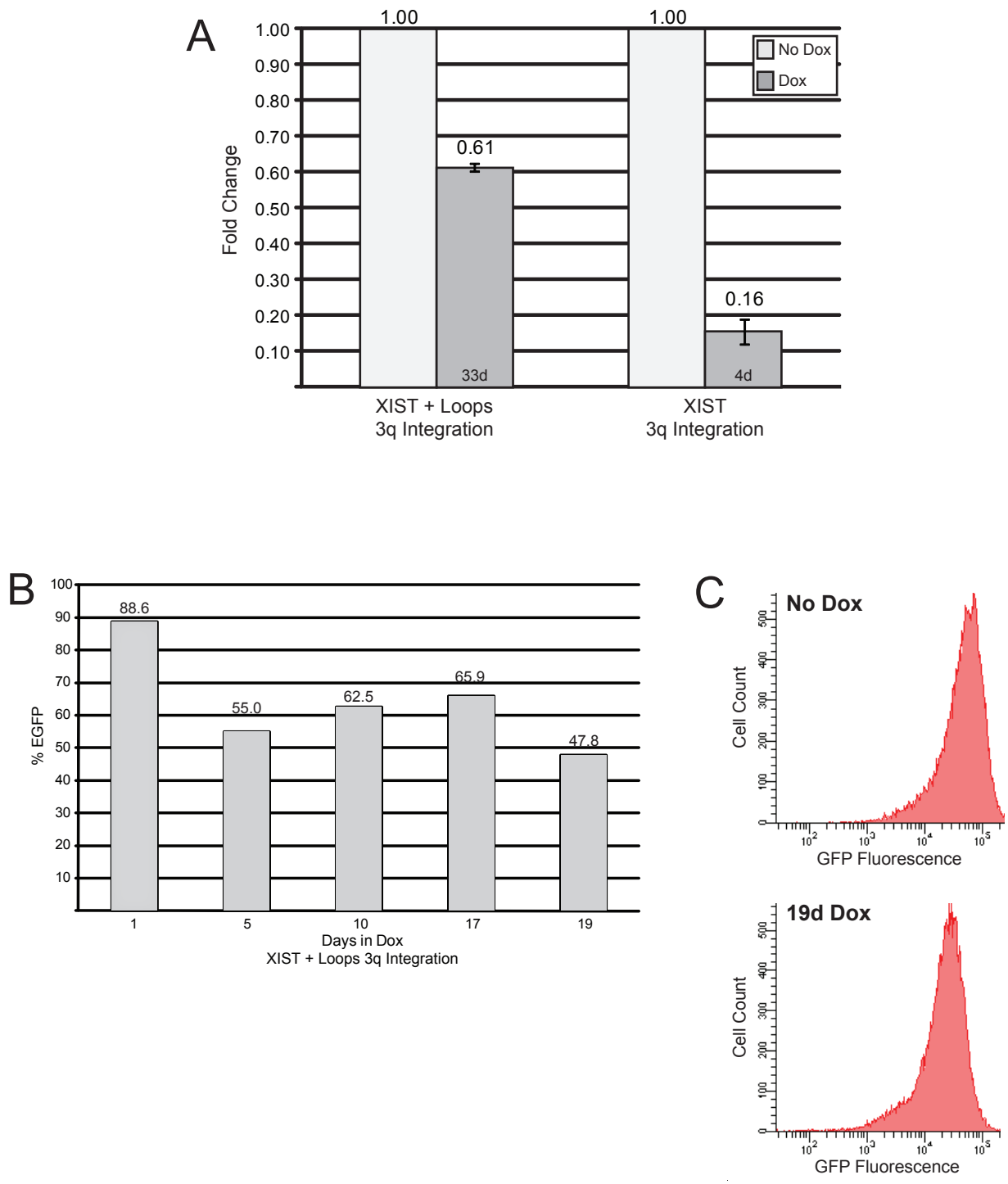

Figure 4-2: *EGFP* silencing with the expression of the tagged XIST RNA. **A)** *EGFP* expression compared to no Dox control and relative to *ACTIN* determined with qPCR and analyzed with the ddCt method. **B & C)** Results of FACS experiment with the XIST+Loops 3q integration. B) The data is represented as %EGFP relative to the no Dox control set to 100%. C) The data is represented in the FACS plot that shows No Dox versus 19d Dox.

Cot-1 and XIST probes to determine the percent of cells that had a visible Cot-1hole. In the XIST + loops cell line with the 3q integration, I observed Cot-1 holes in 35% (n=112), and in the XIST + loops cell line with the 1p integration, I observed Cot-1 holes in 30% (n=183) of cells analyzed. These numbers were similar to the numbers calculated for the XIST cDNA transgene lacking the MS2 loops, where the 3q integration had Cot-1 holes in 32% (n=123) and in the 1p integration 26% (n=161) of cells analyzed had Cot-1 holes. To compare to a control female cell line, the GM04626 had 73% of XIST signals giving Cot-1 holes. Figure 4-3 shows examples of representative Cot-1 holes (yellow arrow heads) with pixel intensity shown across the XIST focus (yellow line).

All the HT1080 integrations had fewer Cot-1 holes corresponding with XIST RNA signals when compared to a control female. The XIST cDNA transgene and the XIST + loops transgene gave a similar number of Cot-1 holes in both integration sites, suggesting that the MS2 loops did not affect the ability of the XIST RNA to form Cot-1 holes and thereby silence repetitive elements.

4.2.5 Lack of H3K27me3 and H4K20me1 enrichment

Two of the epigenetic marks appearing on the Xi following XIST RNA coating are trimethylation of histone H3 at lysine 27 and monomethylation of histone H4 at lysine 20. These marks can be observed as a specific enrichment corresponding to the XIST focus within a cell; with the use of RNA-FISH to locate the XIST RNA combined with immunofluorescence to highlight the inactive marks. The 3q integration with the full length XIST cDNA transgene had been previously examined for the enrichment of chromatin marks and no enrichment was observed for H3K27me3, H4K20me1 and macroH2A. To determine if integration site and therefore the local chromatin neighbourhood affected recruitment of H3K27me3 and H4K20me1, I analyzed the 1p integration with the XIST cDNA transgene and the XIST + loops transgene. The enrichment analysis was performed using a 3D rendered Z-stack comprised of several images taken through the nuclei to ensure that any enrichment would be observed.

Figure 4-4 shows the results of the H3K27me3 analysis. In the control female cell line that has two Xi, the H3K27me3 enrichment could be seen using immunofluorescence alone (Figure 4-4, white arrows shown in the first 3 columns)

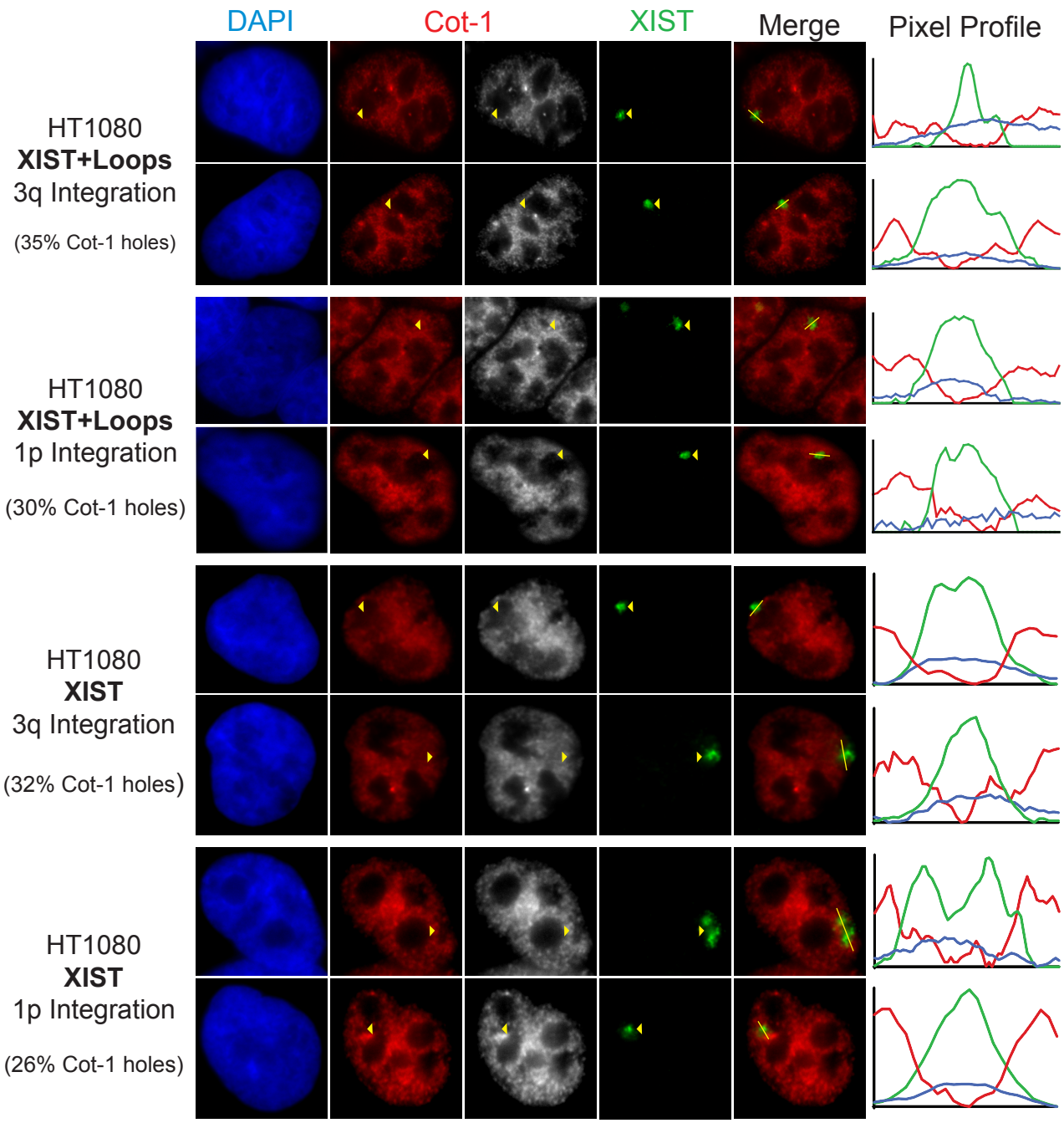

Figure 4-3: Analysis of Cot-1 hybidizaion in the HT1080 1p and 3q integrations by dual RNA FISH hybridization. The Cot-1 RNA signal is red, XIST RNA signal is green and DAPI is blue. The top four panels show the XIST+Loops transgene expression and the bottom four panels show the XIST cDNA transgene. Lines were drawn in ImageJ (NIH) and pixel intensities of each signal was plotted across the yellow line and shown to the right of the images. Yellow arrow heads indicate where the Xist signal corresponds to reductions in the Cot-1 signal.

and again when combined with RNA-FISH (Figure 4-4, arrows shown in the last 4 columns). The XIST + loops and the XIST cDNA alone at the 1p integration did not show any enrichment for H3K27me3 at the site of the localized XIST RNA (Figure 4-4, Rows 3-6). This lack of enrichment was also observed for the H4K20me1 mark (Figure 4-5). Therefore, the H3K27me3 and H4K20me1 marks were not able to be recruited by this Dox-inducible XIST system in the HT1080 male fibrosarcoma cell line at the two site (1p and 3q integration sites), regardless of whether or not the XIST cDNA had the MS2 Loops fused to its 3'end.

4.3 Discussion

I have described a method for tagging and expressing the XIST RNA with the potential to isolate novel XIST-interacting proteins that are involved in the localization of the RNA and the silencing of the surrounding chromatin. The MS2-tagged XIST RNA localized similar to the XIST RNA without a tag suggesting that the tag did not interfere with the factors required for localization. I have shown that, although reduced compared to its untagged counterpart, the MS2-tagged RNA was able to silence a reporter gene and initiated the formation of a silent nuclear compartment. Therefore, the induced expression of an MS2-tagged XIST RNA in the HT1080 system represents a valuable tool for the isolation of a functioning XIST RNA and its associated proteins.

The XIST + loops clones were up-regulated with Dox induction and the nuclear localization looked similar to XIST cDNA alone. This result suggested that the MS2 stem loops fused to the 3' end of the XIST cDNA did not have an effect on the ability of the XIST + loops RNA to localize. Although the mechanism behind XIST RNA localization is unknown, the potential protein, DNA, or RNA components required for localization are present in the HT1080 cell line and the XIST + loops RNA functions efficiently to localize.

A quantitative analysis of expression levels between the clones using qPCR showed that all the transgene-expressing HT1080 clones, regardless of whether the loops were attached, expressed XIST to a lesser extent relative to a female lymphoblast cell line. This result could suggest that the cytomegalovirus (CMV) promoter driving the transgene was not as efficient at expressing XIST compared to the endogenous promoter in a female cell line or that the differences in expression levels were a result of

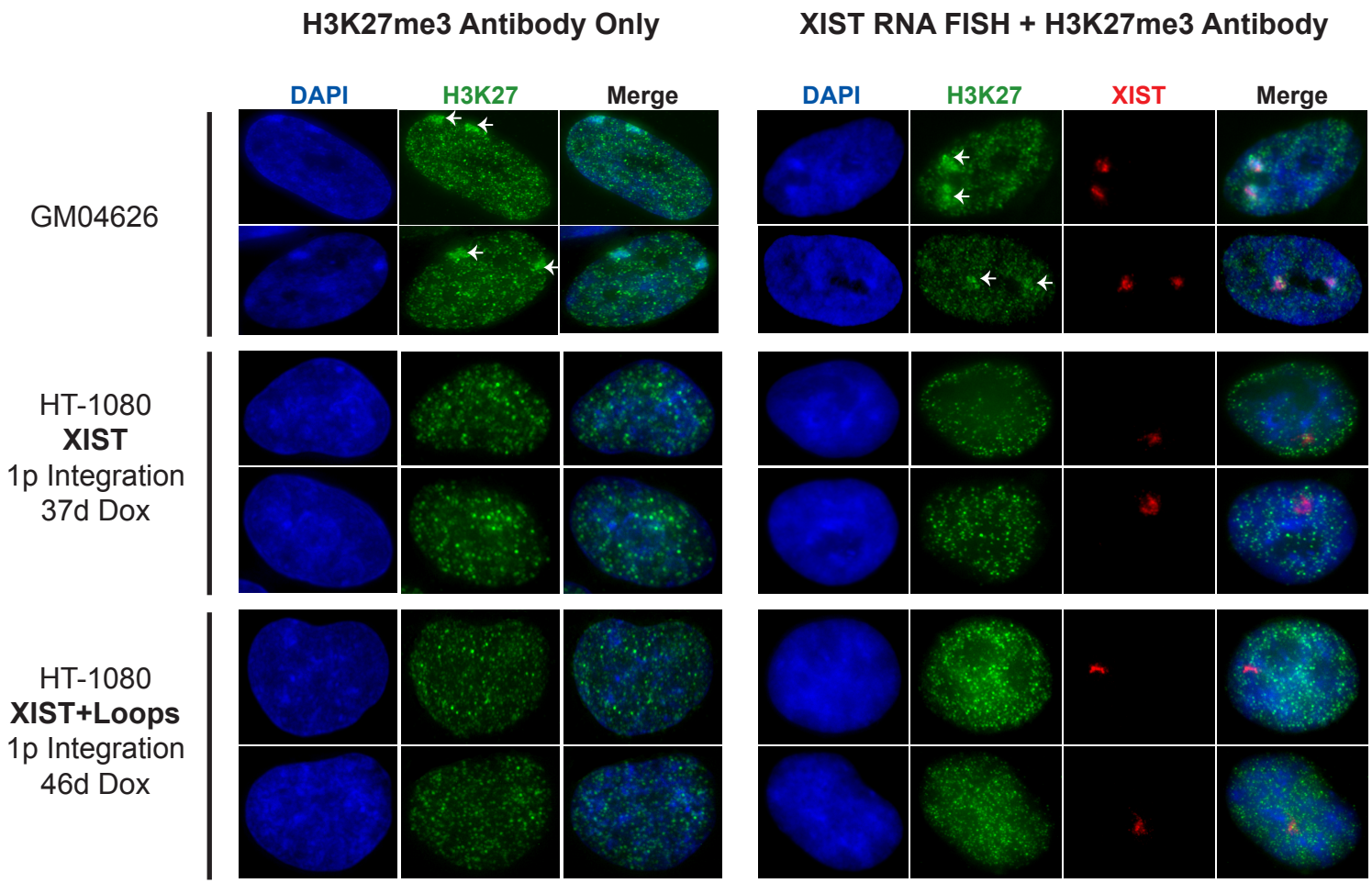

Figure 4-4 : H3K27me3 enrichment analysis in the HT1080 1p integration. All cell lines were subject to two separate procedures; the first three columns show the resulting images from an immunostaining with the H3K27me3 antibody alone; and the last four columns show the images from an RNA FISH with XIST and combined immunostaining with the H3K27me3 antibody. The first two rows show the control female cell line with two inactive X chromosomes corresponding to the two XIST signals and the two H3K27me3 foci (shown with white arrows). The middle two rows show the HT1080 male fibrosarcoma with the XIST cDNA transgene integrated in 1p and the last two rows show the HT1080 with the XIST+Loops transgene integrated at the same 1p location. There is no enrichment of H3K27m3e or evident Barr bodies observed at the localized XIST signal in either HT1080 clone.

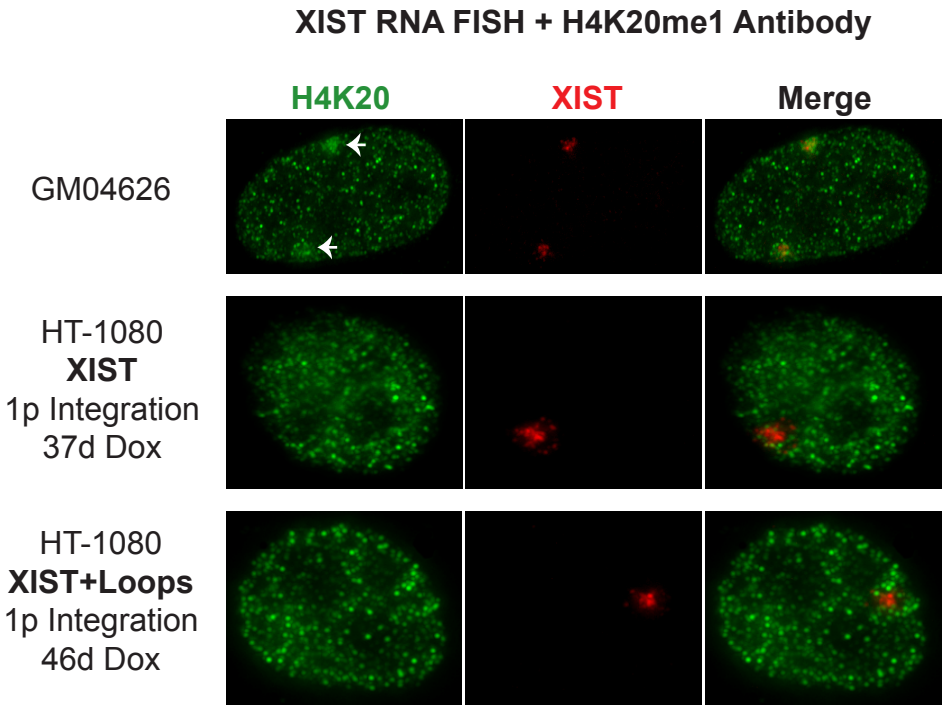

Figure 4-5: H4K20me1 enrichment analysis in the HT1080 1p Integration. The images shown are from an RNA FISH with XIST and combined immunostaining with the H3K20me3 antibody. The first row shows the control female cell line with two inactive X chromosomes corresponding to the two XIST signals and the two H4K20me1 foci (shown with white arrows). The second row shows the HT1080 male fibrosarcoma with the XIST cDNA transgene integrated in 1p and the last row shows the HT1080 with the XIST+Loops transgene integrated at the same 1p location. There is no enrichment of H4K20me1 observed at the localized XIST signal in either HT1080 clone.

normal variability observed with XIST expression. Additionally, it was found that the XIST + loops clones expressed XIST to a lesser extent than the XIST cDNA alone, which could imply that the loops interfered with the transcription or the stability of the XIST + loops RNA. The MS2 loops fused to the 3' end of the *XIST* cDNA are composed of a 21-nucleotide stem loop repeated eight times in succession. It could be possible that the nature of these loops have initiated transcript degradation.

The ability of the XIST + loops RNA to silence *EGFP* was reduced compared to the *XIST* cDNA transgene without the loops. It seemed unlikely that this reduction reflected normal silencing variability since the difference was dramatic and seen with both FACS and qPCR analysis. The *XIST + loops* clone appeared to level out the extent of its silencing to 60% of the no Dox control, therefore, the RNA silenced the *EGFP* expression by 40%. It is possible that the MS2 loops altered the secondary structure of the XIST RNA to an extent that the RNA hybrid was unable to interact efficiently with components required for silencing, or that the level of *XIST + loops* expression was significantly lower so that it could effect the RNA's silencing ability. It has been shown previously that XIST RNA transcript levels do not affect localization (54,151). However, there is no published data to suggest that a minimal amount of XIST expression is required for efficient silencing.

The Cot-1 hole data supported previous work which showed that ectopic *XIST* expression in the HT1080 cell line can silence the Cot-1 RNA fraction (54,106). I validated this observation by cell counting to estimate the percentage of cells that formed a Cot-1 hole following *XIST* induction. The *XIST* transgene and the *XIST + loops* transgene integrated at 1p and 3q gave similar numbers with 26-35% of XIST RNA signals having an associated Cot-1 hole. This suggested that the MS2 loops did not affect the silencing of the Cot-1 fraction.

Curiously, the extent of *EGFP* silencing and the repetitive element silencing differed, with the later being just as efficient with the MS2-tagged XIST RNA. It is possible, that if different regions of the XIST RNA are involved with silencing genes versus repetitive elements; the MS2 loops could interfere with a component only affecting gene silencing and not repetitive silencing. If an element in the 5' region of the XIST RNA is required for silencing *EGFP* and an element in the 3' region is required for repetitive silencing

and the loops are found only to interfere with the 5' silencing element, the result could be less efficient gene silencing and maintenance of repetitive silencing. When a mouse *Xist* transgene lacking the 5' A-repeats is expressed in differentiating male mouse ES cells, genic silencing is abolished (as determined by lack of cell lethality and transcription of three X-linked genes) and Cot-1 hole formation is efficiently initiated (12). In the present study, the XIST + loops RNA silenced a reporter gene to a lesser extent yet silenced the Cot-1 fraction to the same extent, when compared with an XIST RNA without the MS2 loops.

It is important to note that when *XIST* was expressed in the HT1080 cell line, the degree of Cot-1 repression was lower than observed for a control female cell line. An appropriate developmental context might be required for efficient formation of a repressive compartment as assessed by a Cot-1 hole. It is also possible that the X chromosome with *XIST* expression forms the most efficient Cot-1 hole and in the case of the HT1080 *XIST* transgenes, the integrations are on autosomes.

Finally I looked at the acquisition of inactive chromatin marks in the HT1080 cell line that contained the 1p integration site. There was no recruitment found for the inactive marks H3K27me3 or H4K20me1 with the *XIST* or the *XIST + loops* transgene. This is the second integration site in this HT1080 inducible system that has been unable to accumulate inactive marks to the XIST RNA focus as normally seen in a female nucleus. Previously, the 3q integration site did not recruit H3K27me3, H4K20me1 or macroH2A1 with a localized XIST RNA (106). Ectopic *XIST* expression in HT1080 cells induced H4 hypoacetylation, late replication timing, H2Aub, and macroH2A at the site of XIST RNA accumulation (54,58). Therefore, it does seem probable that epigenetic modifications are recruited in our HT1080 system. The lack of recruitment at the 1p and 3q integrations might be the result of the surrounding chromatin. There are additional transgene integration sites that have been mapped and require analysis to determine if chromatin marks can be recruited in this specific HT1080 system.

To summarize, the XIST+ loops RNA localized similarly to XIST RNA without the loops, the expression of the XIST + loops RNA was less than the expression of the XIST RNA alone and the ability of the XIST RNA to silence *EGFP* was decreased by the addition of the MS2 loops. However, the ability to silence the Cot-1 fraction was not

inhibited by the addition of the MS2 loops as shown with the similar formation of a Cot-1 hole. Despite the decreased expression and reporter gene silencing, the loops remain a valid system for tagging the RNA and identifying interacting proteins as the tagged RNA localized efficiently and maintained a certain degree of *EGFP* and repetitive element silencing, suggesting that the MS2-tagged XIST must be interacting with the elusive localization components and to some extent the silencing factors.

Chapter 5: Overall discussion and future directions

The expression of the XIST/Xist RNA is a requirement for the initiation of X-chromosome inactivation. Exactly how this large non-coding RNA functions to coat an entire chromosome and set up the heterochromatic state is unknown. Information regarding the Xi chromatin, specifically what epigenetic marks are recruited or excluded and the kinetics of their involvement has suggested a model whereby XIST/Xist is a multi-functional RNA. The action of the XIST/Xist RNA ultimately results in the transcriptional silencing of the X chromosome through the formation of a repressive nuclear compartment and the recruitment of proteins to set up and maintain the inactive state.

5.1 Lack of Xist-dependent epigenetic events in hybrid cells

In mammalian females, *Xist* expression and localization occurs early during development and is followed by X-linked gene silencing and concurrent epigenetic changes. Therefore, the occurrence of the epigenetic changes could be the result of *Xist* expression or the subsequent associated genic silencing. To determine if any of the events are caused directly by *Xist* expression, I analyzed a somatic cell hybrid system where *Xist* expression and localization occurred in the absence of developmental silencing.

In the human/mouse somatic cell hybrids, the mouse Xist RNA was able to localize outside of a developmental context but this localization did not result in the transcriptional silencing of X-linked genes, particularly those within close proximity of the *Xist* locus. I analyzed epigenetic events and based on the results shown in Chapter 3, Xist localization in a somatic cell was not sufficient to induce the four epigenetic changes I analyzed. Therefore, it is likely that the loss of active epigenetic marks associated with X inactivation are dependent on a developmental state or genic silencing.

The somatic cell hybrids support the developmental window concept (31) where a developmentally competent cell is required for Xist-mediated silencing, adding to this Xist-induced epigenetic events. There have also been studies that used a developmental context to identify Xi-associated epigenetic marks in the absence of silencing. Mouse ES cells with an inducible *Xist* transgene (*Xist*Δ*sx*) were able to localize the transgenic Xist RNA but unable to silence. When Xist localized alone,

without concurrent silencing, there was a recruitment of epigenetic marks in a developmental context (Refer to Table 1-2 for a summary). The occurrence of epigenetic marks without silencing, which followed *Xist* expression in a developmental context, suggests that the role the Xist RNA plays in inactivation extends beyond genic silencing and involves the appearance of specific epigenetic events. These include the establishment of a repressive nuclear compartment, the recruitment of macroH2A, polycomb group proteins and H3K27me3, H4K20me1, H2AK119ub. However, the manner in which Xist causes the appearance of these events is unknown.

To complement the question of what epigenetic events Xist induces in the absence of silencing; it would be interesting to identify the molecular candidates that are required for silencing. To address what is required is not a simple task. Knock-out experiments have shown that none of the chromatin components or modifications appear to be essential for silencing. However, these marks may be redundant and therefore removing a single component is not sufficient. Additionally, several marks have more than one protein that are responsible for its maintenance and for others eliminating the writer protein is lethal. Consequently, it is difficult to assess their requirement in X inactivation with knock out experiments.

5.2 Identifying XIST and associated proteins

To understand how the XIST RNA functions to inactivate a chromosome it is imperative to isolate XIST-interacting proteins. In this thesis I describe a method for tagging the XIST RNA that involves the 3' fusion of several MS2 stem loops, which interacts with an MS2 coat protein and can be utilized to isolate the RNA and its associated proteins. The tagged RNA must be expressed in a human cell line that has competency for XIST-mediated silencing. The HT1080 cell line seemed an appropriate system to integrate the tagged XIST transgene because it has reproducible integration sites where a transgene can be integrated under the control of an inducible promoter. Additionally, in the HT1080 cell line, an XIST transgene has shown reporter gene silencing and some degree of epigenetic changes (106).

The results in Chapter 4 showed that the localization of the tagged XIST RNA was indistinguishable from the localization of the XIST RNA and similarly, the ability to form Cot-1 holes was indistinguishable. However, the tagged RNA silenced the *EGFP*

reporter gene to a lesser extent than the native XIST RNA. The tag did not affect the ability of the RNA to localize or form a Cot-1 hole, but for an unknown reason, the tag decreased the ability of the XIST RNA to silence an adjacent reporter gene. The stem loop tag could have impaired a structure within the RNA that is required for silencing or blocked assess of important silencing factors to the RNA. If this is the case, it is intriguing that the gene silencing impaired RNA was still able to effectively localize and silence the Cot-1 fraction. This result adds more evidence to the idea that the XIST RNA has many roles and different regions of the RNA could be responsible for such roles.

The lack of H3K27me3 and H4K20me1 enrichment at the XIST foci for both the tagged RNA and the XIST RNA was surprising, given that the HT1080 cell line was previously shown to be capable of recruiting repressive chromatin marks (54,58). The lack of enrichment could be the result of the integration site, such that the local chromatin is restrictive to modifications. There are several additional FRT integration sites and future experiments analyzing for the enrichment of repressive chromatin marks could address this concern.

Additionally, the MS2-tagged XIST transgene was also integrated into a human HEK293 cell line. The HEK293 cell line contained two inactive X chromosomes and two separately localzied XIST RNA signals, therefore any XIST transgene integrated into the genome resulted in a third signal. Expression of an XIST cDNA transgene in this HEK293 cell line resulted in the recruitment of macroH2A1 and H3K27me3 (106). The HEK293 cell line expressing the tagged XIST RNA will be analyzed in the future for the recruitment of macroH2A1 and H3K27me3 to determine if the MS2-tagged RNA is capable of recruiting these inactive epigenetic marks.

Follow-up experiments from this thesis also include isolating the tagged XIST RNA to analyze the associated proteins. Despite the inefficient silencing of the *EGFP* reporter and the lack of recruitment of inactive marks, the tagged RNA was able to localize and form a Cot-1 hole. If the XIST RNA requires protein partners for these tasks, then they should be precipitated with the tagged RNA. As proteins are identified that interact with XIST, the studies of functional regions of the RNA will continue in the anticipation of identifying the regions of XIST that are responsible for recruiting protein components and setting up the heterochromatic state.

5.3 The Cot-1 hole

Several research groups have utilized the lack of Cot-1 hybridization to analyze the XIST territory, here referred to as the Cot-1 hole (12,46,54,106,149). The analysis of the XIST territory has suggested that it is a silent nuclear compartment, lacking transcriptional components and predominately comprised of inactive repetitive DNA with X-linked genes located at the periphery of this compartment (12,46). It is difficult to assess what the Cot-1 hole represents and how XIST functions to form the compartment since the exact constituents of the Cot-1 fraction are not known. However, it is known that the Cot-1 fraction by its nature is enriched for the repetitive fraction of the genome. The Cot-1 RNA could represent repeats within non-coding regions (UTRs and introns) of precursor mRNAs that are expressed and could additionally represent expressed intergenic repeats. Since gene silencing and Cot-1 silencing has been separated in my Xi-containing hybrid cell line and in the mouse ES cells expressing a silencing-deficient Xist transgene, it suggests that Xist functions to silence this Cot-1 fraction.

The Cot-1 analysis on both the somatic cell hybrids and the HT1080 illustrated that the Cot-1 hole should not be used as an indicator of gene silencing. My results in Chapter 3 presented a Cot-1 domain associated with an Xi in a mouse somatic cell. Additionally in Chapter 4, the appearance of a Cot-1 hole to the same extent with an XIST RNA that silenced a reporter gene and a tagged XIST RNA that did not efficiently silence the same gene. I have also shown an impaired ability to form a Cot-1 hole in both the somatic cell hybrids and the HT1080 cell line, as compared to a female cell that silenced an X chromosome within a normal developmental context (Figure 5-1).

The efficient formation of a Cot-1 hole with the XistΔsx transgene expressed in differentiating mouse ES (12) cells suggests that the Xist-initiated formation of a Cot-1 hole does not require X-linked gene silencing. A gene silencing impaired Xist RNA was still able to silence the Cot-1 fraction efficiently compared to my results that showed impaired or inefficient formation of a Cot-1 hole in somatic cells with somatic cell hybrids and HT1080 respectively. During ES cell development, genic silencing is not required for efficient Cot-1 hole formation, in contrast to somatic cell hybrids where gene silencing does not occur and no Cot-1 hole is formed. Likewise, in the HT1080 cell line,

where gene silencing of *EGFP* occurs, there is not an efficient formation of a Cot-1 hole. To summarize, Cot-1 hole formation is less efficient (regardless of genic repression) in the somatic cell lines then it is when *XIST/Xist* is expressed in a normal developmental context. This could suggest that a normal developmental context is required for Xist to efficiently silence the Cot-1 fraction (Figure 5-1). To assess the requirement for a developmental state, it would be beneficial to use the inducible expression of an Xist transgene during ES cell development and identify if the Cot-1 hole formation occurs only during the *Xist*-responsive state.

In the Xi-containing hybrid cell line, the silencing of the Cot-1 fraction was not maintained, despite the maintenance of X-linked gene silencing where the XIST signal was de-localized. This observation suggests the Cot-1 hole is XIST-dependent, such that the continued expression of *XIST* was required for the maintenance of Cot-1 silencing. However, the mouse cell background could have prevented the proper functioning of human XIST and thereby the silencing of the human Cot-1 fraction. I think it would be of interest to determine if the Cot-1 fraction is silenced in a somatic cell where *XIST*/*Xist* is conditionally deleted. Additionally, similar to the mapping of the A-repeats and their requirement for gene silencing, it would be interesting to determine the regions of XIST/Xist required for the formation of a repressive nuclear compartment.

Recently, an analogy has been suggested in the literature between the formation of facultative heterochromatin in X-chromosome inactivation and the formation of constitutive heterochromatin in yeast (168,169). In yeast it is thought that the transcription of repeat elements initiates the formation of heterochromatin through the RNAi pathway (170). This idea could be furthered with the formation of the Cot-1 hole occurring early following *Xist* expression during development and representing the silencing of repeat elements thereby providing a scaffold for chromosome-wide inactivation. However, there has been controversy in the literature with contrasting findings on the role of RNAi in Dicer-deficient ES cells (reviewed in(168)). Additionally, the transcription of the Cot-1 fraction is detected by RNA FISH only in the nucleus while the Dicer protein has a cytoplasmic localization.

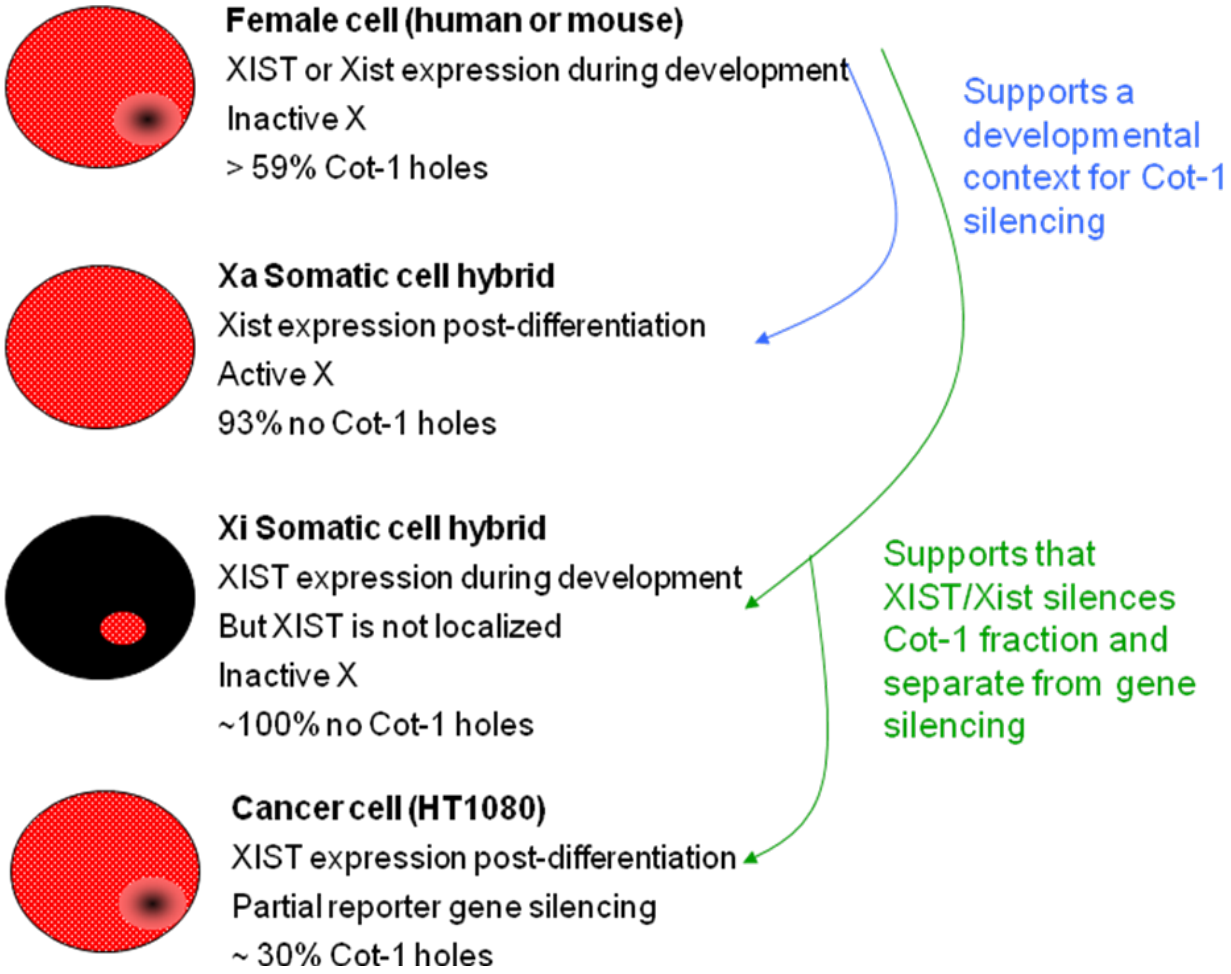

Figure 5-1: A schematic of Cot-1 results in model systems showing support for XIST/Xist silencing the Cot-1 repetitive fraction with a developmental dependence

5.4 The developmental window

Although a developmental window where X inactivation can occur has been defined with differentiating mouse ES cells (31), it remains uncertain if there is the equivalent window in human cells and if the timing is similar. While some inactive chromatin marks are initiated by *XIST* expression in certain human somatic cells, such as, HeLa cells (14), HEK293 cells (106), and the HT1080 cells (54,58,106), X inactivation has not been completely recapitulated in these systems. What properties of these cell lines enable them to respond to the XIST RNA remains unknown. Do they express silencing competency factors that other somatic cells do not? Or is it that the chromatin is responsive to change?

SATB1 is a nuclear matrix protein that has recently become a candidate for a critical factor required by a cell to be Xist-responsive. The expression of SATB1 in ES cells and its down regulation during differentiation corresponds to the developmental window (47). When a SATB1 transgene is introduced and expressed in Xist-resistant tumor cells or XY fibroblasts, the cells become Xist-responsive (47). Unfortunately, knockout studies have proven difficult since a homolog SATB2 exists and the lethality inflicted with a double knockout (47).

The data available on the gene portal system BioGPS, an online resourse with information on gene expression and protein function (http://biogps.gnf.org) shows that *SATB1* is expressed in the HT1080 cell line. Future experiments could include a quantitative analysis of *SATB1* expression in the HT1080 cell line and over expression or knock down experiments to see if changing the expression of *SATB1* has an effect on the silencing of the reporter *EGFP* gene by the XIST RNA.

5.5 The search for the ‘perfect’ model of human X inactivation

In this thesis I have looked at two model systems, the human/mouse somatic cell hybrids, and the human HT1080 fibrosarcoma cell line. The somatic cell hybrids, although useful for X-linked gene expression, were not able to recapitulate X inactivation with the expression of human *XIST*. In the HT1080 cell line, the expression of *XIST* was able to silence a closely linked reporter gene and form a repressive nuclear compartment; however, chromosome-wide gene silencing is uncertain as are all aspects of X inactivation. There is a need for a human model system which

recapitulates X inactivation, so that we may better understand the mechanism by which XIST initiates transcriptional inactivation and the formation of heterochromatin.

A variety of cell types in both mice and humans have been analyzed with *XIST/Xist* induction to determine the competency to XIST/Xist-mediated silencing (Refer to Figure 5-1). The key to an ideal human cell system depends on identifying a cell line that can recapitulate the process of X-chromosome inactivation. Some hematopoietic precursor cells have potential, as do induced pluripotent stem (iPS) cells, and perhaps eventually new human ES cell lines will be analyzed that are in the correct developmental state.

The immature hematopoietic precursor cells show potential since in an adult mouse, these cells induce Xist-mediated silencing and H3K27me3 enrichment (171). It remains to be known if these cells recapitulate the entire process of X inactivation (as in mouse ES cells). If similar cells in human hematopoiesis also support XIST-mediated silencing, immature hematopoietic precursor cells could be a human model cell system to study XIST-mediated events.

Mouse iPS cells generated from mouse embryonic fibroblasts are shown to have a reactivated Xi that lacks Xist coating and had biallelic Tsix and Pgk1 expression, in addition to a lack of H3K27me3 and H4K20me1 enrichment (172). The iPS cells also undergo some aspects of inactivation when induced to differentiate. Upon differentiation, an Xist focus appears with an exclusion of RNA PolII and an enrichment of H3K27me3 (172). With further studies on human iPS cells, we might discover that they are a suitable model for studying X inactivation.

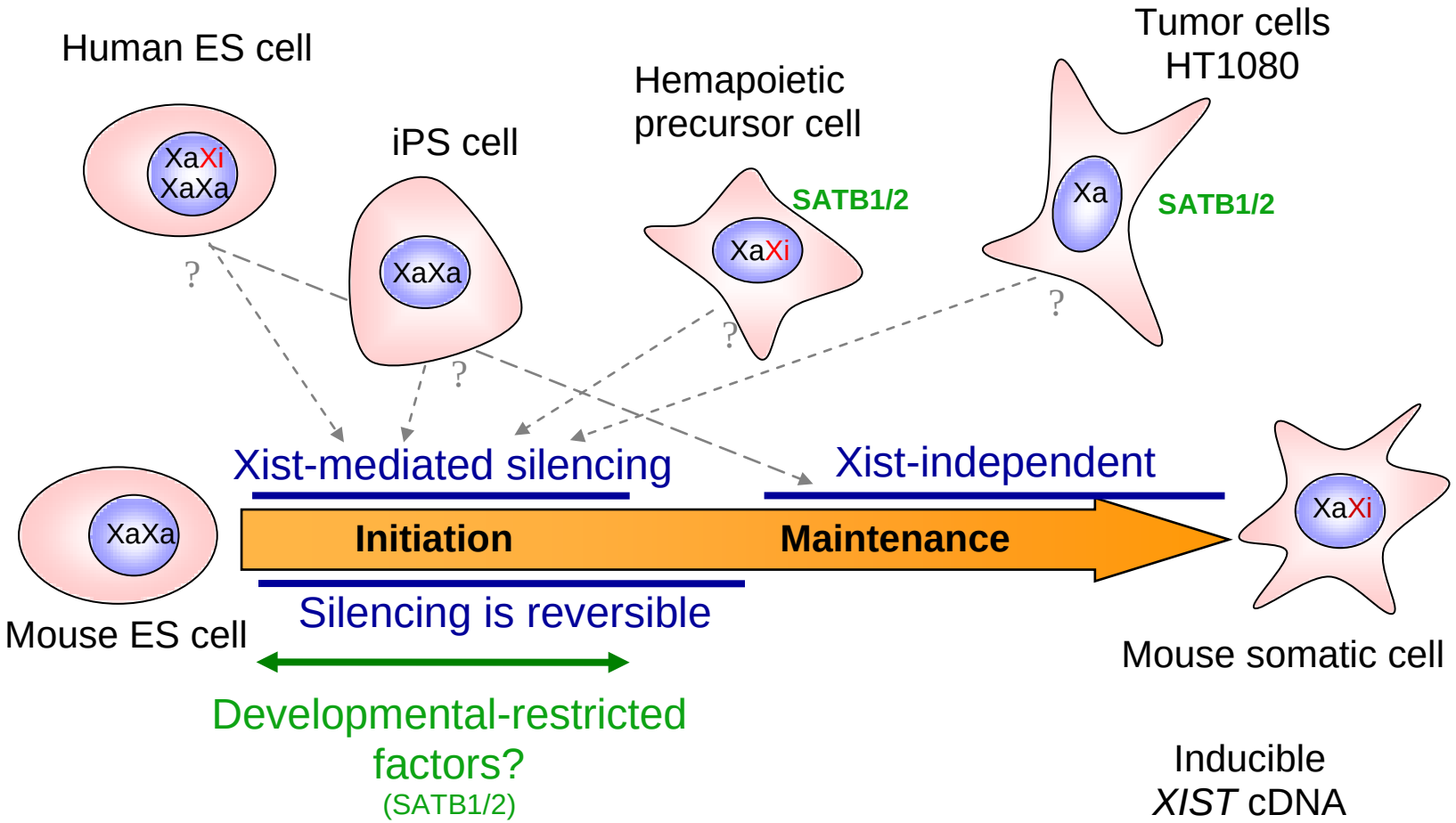

Figure 5-2: Current and prospective model systems for the study of human X-chromosome inactivation

Table 5-1: XIST/Xist-mediated silencing and epigenetic events in a variety of cell systems

Undifferentiated ES Cells

Female
Mouse ESC
- Two Xa
- Two pinpoint Xist signals

Mouse ESC
Inducible Xist Transgene
- Xist-mediated silencing
- Xist-dependent and silencing is reversible
- Cell lethality – inactivating the single X in male ES cells
- Late replication
- H4 hypoacetylation
- H3K27me3
- H4K20me2
- H2Aub
- No macroH2A

Human
ESC
- Variability of inactivation status of X chromosomes

Differentiating ES Cells

Female
Mouse ESC
Differentiating
- Choice of one Xi, one Xa
- One Xist focus at the Xi
- Xist-mediated silencing
- RNA-PolII exclusion
- Cot-1 hole
- H3K27me3
- H3K9me2
- H4K20me1
- H2A112ub
- Late replication
- DNA methylation
- MacroH2A recruitment occurs after silencing is established

Mouse ESC
Differentiating
Inducible Xist Transgene
- Xist-mediated silencing immediately following differentiation

After 48 hrs
- No Xist-mediated silencing

After 72 hrs
- Xist-independent and silencing is irreversible
- H3K27me3
- H2Aub
- MacroH2A

Post-Differentiated Cells

Female Mouse
Immature Hematopoietic Precursors
- One Xi, one Xa
- One Xist focus at the Xi

Immature Mouse
Hematopoietic Precursors
Inducible Xist Transgene
- Xist-mediated silencing
- Pgk1 and Hprt repression
- H3K27me3

Mouse Hematopoietic Stem Cells
Inducible Xist
- No Xist-mediated silencing

Somatic Cells

Mature Blood Cells
Inducible Xist
- No Xist-mediated silencing

MEF
- No Xist-mediated silencing
- MacroH2A recruitment

HEK293
- H3K27me3
- H4K20me1
- MacroH2A

HeLa
- Transient H3K27me3

HT1080
XIST Transgene
- H4 hypoAc
- Late replication
- H2Aub
- MacroH2A
- Cot-1 hole
- XIST-mediated silencing
- Silence EGFP reporter

Somatic Cell Hybrids
- No Xist-mediated silencing
- Localized mouse Xist
- De-localized human XIST

Retention of active epigenetic mark

References

1. Panning, B., Dausman, J. and Jaenisch, R. (1997) X chromosome inactivation is mediated by RNA stabilization. *Cell*, **90**, 907-916.
2. Sheardown, S.A., Duthie, S.M., Johnston, C.M., Newall, A.E.T., Formstone, E.J., Arkell, R.M., Nesterova, T.B., Alghisi, G.-C., Rastan, S. and Brockdorff, N. (1997) Stabilisation of Xist RNA mediates initiation of X chromosome inactivation. *Cell*, **91**, 99-107.
3. Kay, G.F., Penny, G.D., Patel, D., Ashworth, A., Brockdorff, N. and Rastan, S. (1993) Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell*, **72**, 171-182.
4. Keohane, A.M., O'Neill, L.P., Belyaev, N.D., Lavender, J.S. and Turner, B.M. (1996) X-inactivation and histone H4 acetylation in embryonic stem cells. *Dev. Biol.*, **180**, 618-630.
5. Mermoud, J.E., Costanzi, C., Pehrson, J.R. and Brockdorff, N. (1999) Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J. Cell Biol.*, **147**, 1399-1408.
6. Barr, M.L. and Bertram, E.G. (1949) A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*, **163**, 676-677.
7. Lock, L.F., Takagi, N. and Martin, G.R. (1987) Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell*, **48**, 39-46.
8. Rego, A., Sinclair, P.B., Tao, W., Kireev, I. and Belmont, A.S. (2008) The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. *Journal of cell science*, **121**, 1119-1127.
9. Zhang, L.F., Huynh, K.D. and Lee, J.T. (2007) Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. *Cell*, **129**, 693-706.
10. Chaumeil, J., Okamoto, I., Guggiari, M. and Heard, E. (2002) Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells. *Cytogenet Genome Res*, **99**, 75-84.
11. Heard, E., Rougelle, 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**, 727-738.
12. 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**, 2223-2237.
13. Peters, A.H., Mermoud, J.E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N. and Jenuwein, T. (2002) Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nature genetics*, **30**, 77-80.
14. 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**, 131-135.
15. 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**, E171.
16. de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M. *et al.* (2004) Polycomb Group Proteins Ring1A/B Link Ubiquitylation of Histone H2A to Heritable Gene Silencing and X Inactivation. *Developmental cell*, **7**, 663-676.
17. Fang, J., Chen, T., Chadwick, B., Li, E. and Zhang, Y. (2004) Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in Initiation of X-inactivation. *J Biol Chem*.
18. Lyon, M.F. (1961) Gene action in the X-chromosome of the mouse *(Mus musculus L.)*. *Nature*, **190**, 372-373.

19. Ross, M.T., Grafham, D.V., Coffey, A.J., Scherer, S., McLay, K., Muzny, D., Platzer, M., Howell, G.R., Burrows, C., Bird, C.P. *et al.* (2005) The DNA sequence of the human X chromosome. *Nature*, **434**, 325-337.
20. McKusick, V.A. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD).
21. Migeon, B.R. (2007) *Females are Mosaics X inactivation and sex differences in disease*. Oxford University Press, New York.
22. Migeon, B.R. (2006) The role of X inactivation and cellular mosaicism in women's health and sex-specific diseases. *Jama*, **295**, 1428-1433.
23. De Sario, A. (2009) Clinical and molecular overview of inherited disorders resulting from epigenomic dysregulation. *European journal of medical genetics*, **52**, 363-372.
24. Reik, W. (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, **447**, 425-432.
25. Minard, M.E., Jain, A.K. and Barton, M.C. (2009) Analysis of epigenetic alterations to chromatin during development. *Genesis*, **47**, 559-572.
26. Zakharova, I.S., Shevchenko, A.I. and Zakian, S.M. (2009) Monoallelic gene expression in mammals. *Chromosoma*, **118**, 279-290.
27. Nagano, T. and Fraser, P. (2009) Emerging similarities in epigenetic gene silencing by long noncoding RNAs. *Mamm Genome*, **20**, 557-562.
28. Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R. and Willard, H.F. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, **349**, 38-44.
29. Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S. and Rastan, S. (1992) The product of the mouse Xist gene is a 15kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, **71**, 515-526.
30. 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**, 527-542.
31. Wutz, A. and Jaenisch, R. (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell*, **5**, 695-705.
32. Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S. and Brockdorff, N. (1996) Requirement for Xist in X chromosome inactivation. *Nature*, **379**, 131-137.
33. 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**, 156-166.
34. Nesterova, T.B., Slobodyanyuk, S.Y., Elisaphenko, E.A., Shevchenko, A.I., Johnston, C., Pavlova, M.E., Rogozin, I.B., Kolesnikov, N.N., Brockdorff, N. and Zakian, S.M. (2001) Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. *Genome research*, **11**, 833-849.
35. Yen, Z.C., Meyer, I.M., Karalic, S. and Brown, C.J. (2007) A cross-species comparison of X-chromosome inactivation in Eutheria. *Genomics*, **90**, 453-463.
36. Pang, K.C., Frith, M.C. and Mattick, J.S. (2005) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet*.
37. Caparros, M.L., Alexiou, M., Webster, Z. and Brockdorff, N. (2002) Functional analysis of the highly conserved exon IV of XIST RNA. *Cytogenet Genome Res*, **99**, 99-105.
38. Wutz, A., Rasmussen, T.P. and Jaenisch, R. (2002) Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.*, **30**, 167-174.
39. Gilbert, S.L. and Sharp, P.A. (1999) Promoter-specific hypoacetylation of X-inactivated genes. *Proc. Natl. Acad. Sci., USA*, **96**, 13825-13830.
40. Boggs, B.A., Connors, B., Sobel, R.E., Chinault, A.C. and Allis, C.D. (1996) Reduced levels of histone H3 acetylation on the inactive X chromosome in human females. *Chromosoma*, **105**, 303-309.
41. 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. *Developmental cell*, **4**, 481-495.
42. Fang, J., Chen, T., Chadwick, B., Li, E. and Zhang, Y. (2004) Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J Biol Chem*, **279**, 52812-52815.
43. Mermoud, J.E., Popova, B., Peters, A.H., Jenuwein, T. and Brockdorff, N. (2002) Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation. *Curr Biol*, **12**, 247-251.
44. Rasmussen, T.P., Wutz, A., Pehrson, J.R. and Jaenisch, R. (2001) Expression of Xist RNA is sufficient to initiate macrochromatin body formation. *Chromosoma*, **110**, 411-420.
45. Boggs, B.A. and Chinault, A.C. (1994) Analysis of replication timing properties of human X-chromosomal loci by fluorescence in situ hybridization. *Proc. Natl. Acad. Sci.*, **91**, 6083-6087.
46. Clemson, C.M., Hall, L.L., Byron, M., McNeil, J. and Lawrence, J.B. (2006) The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences. *Proc. Natl Acad. Sci. USA*, **103**, 7688-7693.
47. Agrelo, R., Souabni, A., Novatchkova, M., Haslinger, C., Leeb, M., Komnenovic, V., Kishimoto, H., Gresh, L., Kohwi-Shigematsu, T., Kenner, L. *et al.* (2009) SATB1 defines the developmental context for gene silencing by Xist in lymphoma and embryonic cells. *Dev. Cell*, **16**, 507-516.
48. Clemson, C.M., McNeil, J.A., Willard, H.F. and Lawrence, J.B. (1996) XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.*, **132**, 259-275.
49. Clemson, C.M., Chow, J.C., Brown, C.J. and Lawrence, J.B. (1998) Stabilization and localization of Xist RNA are controlled by separate mechanisms and are not sufficient for X inactivation. *J. Cell Biol.*, **142**, 13-23.
50. Hall, L.L., Clemson, C.M., Byron, M., Wydner, K. and Lawrence, J.B. (2002) Unbalanced X;autosome translocations provide evidence for sequence specificity in the association of XIST RNA with chromatin. *Hum. Mol. Genet.*, **11**, 3157-3165.
51. Duthie, S.M., Nesterova, T.B., Formstone, E.J., Keohane, A.M., Turner, B.M., Zakian, S.M. and Brockdorff, N. (1999) Xist RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in cis. *Hum. Mol. Genet.*, **8**, 195-204.
52. Keohane, A.M., Barlow, A.L., Waters, J., Bourn, D. and Turner, B.M. (1999) H4 acetylation, XIST RNA and replication timing are coincident and define X;autosome boundaries in two abnormal X chromosomes. *Hum. Mol. Genet.*, **8**, 377-383.
53. Hall, L.L., Byron, M., Pageau, G. and Lawrence, J.B. (2009) AURKB-mediated effects on chromatin regulate binding versus release of XIST RNA to the inactive chromosome. *The Journal of cell biology*, **186**, 491-507.
54. Hall, L.L., Byron, M., Sakai, K., Carrel, L., Willard, H.F. and Lawrence, J.B. (2002) An ectopic human XIST gene can induce chromosome inactivation in postdifferentiation human HT-1080 cells. *Proc. Natl Acad. Sci. USA*, **99**, 8677-8682.
55. Britten, R.J. and Kohne, D.E. (1968) Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science*, **161**, 529-540.
56. 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**, 644-649.
57. Boggs, B.A., Cheung, P., Heard, E., Spector, D.L., Chinault, A.C. and Allis, C.D. (2002) Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. *Nature genetics*, **30**, 73-76.
58. Smith, K.P., Byron, M., Clemson, C.M. and Lawrence, J.B. (2004) Ubiquitinated proteins including uH2A on the human and mouse inactive X chromosome: enrichment in gene rich bands. *Chromosoma*, **113**, 324-335.
59. Marks, H., Chow, J.C., Denissov, S., Francoijs, K.J., Brockdorff, N., Heard, E. and Stunnenberg, H.G. (2009) High-resolution analysis of epigenetic changes associated with X inactivation. *Genome research*, **19**, 1361-1373.
60. Sidhu, S.K., Minks, J., Chang, S.C., Cotton, A.M. and Brown, C.J. (2008) X chromosome inactivation: heterogeneity of heterochromatin. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **86**, 370-379.

61. Chadwick, B.P. and Willard, H.F. (2004) Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 17450-17455.
62. Valley, C.M., Pertz, L.M., Balakumaran, B.S. and Willard, H.F. (2006) Chromosome-wide, allele-specific analysis of the histone code on the human X chromosome. *Human molecular genetics*, **15**, 2335-2347.
63. Pehrson, J., Costanzi, C. and Dharia, C. (1997) Developmental and tissue expression patterns of histone MacroH2A1 subtypes. *J. Cell. Biochem.*, **65**, 107-113.
64. Costanzi, C. and Pehrson, J.R. (2001) MACROH2A2, a new member of the MACROH2A core histone family. *J Biol Chem*, **276**, 21776-21784.
65. Costanzi, C. and Pehrson, J.R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature*, **393**, 599-601.
66. Chadwick, B.P. and Willard, H.F. (2001) Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum. Mol. Genet.*, **10**, 1101-1113.
67. Rasmussen, T.P., Huang, T., Mastrangelo, M.A., Loring, J., Panning, B. and Jaenisch, R. (1999) Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. *Nucleic acids research*, **27**, 3685-3689.
68. Rasmussen, T.P., Mastrangelo, M.A., Eden, A., Pehrson, J.R. and Jaenisch, R. (2000) Dynamic relocalization of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation. *The Journal of cell biology*, **150**, 1189-1198.
69. 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. *Nat. Genet.*, **22**, 323-324.
70. Mietton, F., Sengupta, A.K., Molla, A., Picchi, G., Barral, S., Heliot, L., Grange, T., Wutz, A. and Dimitrov, S. (2009) Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome. *Molecular and cellular biology*, **29**, 150-156.
71. Chu, F., Nusinow, D.A., Chalkley, R.J., Plath, K., Panning, B. and Burlingame, A.L. (2006) Mapping post-translational modifications of the histone variant MacroH2A1 using tandem mass spectrometry. *Mol Cell Proteomics*, **5**, 194-203.
72. Thambirajah, A.A., Li, A., Ishibashi, T. and Ausio, J. (2009) New developments in post-translational modifications and functions of histone H2A variants. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, **87**, 7-17.
73. Bernstein, E., Muratore-Schroeder, T.L., Diaz, R.L., Chow, J.C., Changolkar, L.N., Shabanowitz, J., Heard, E., Pehrson, J.R., Hunt, D.F. and Allis, C.D. (2008) A phosphorylated subpopulation of the histone variant macroH2A1 is excluded from the inactive X chromosome and enriched during mitosis. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 1533-1538.
74. Hernandez-Munoz, I., Lund, A.H., van der Stoop, P., Boutsma, E., Muijrers, I., Verhoeven, E., Nusinow, D.A., Panning, B., Marahrens, Y. and van Lohuizen, M. (2005) Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 7635-7640.
75. Angelov, D., Molla, A., Perche, P.Y., Hans, F., Cote, J., Khochbin, S., Bouvet, P. and Dimitrov, S. (2003) The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol Cell*, **11**, 1033-1041.
76. Chang, E.Y., Ferreira, H., Somers, J., Nusinow, D.A., Owen-Hughes, T. and Narlikar, G.J. (2008) MacroH2A allows ATP-dependent chromatin remodeling by SWI/SNF and ACF complexes but specifically reduces recruitment of SWI/SNF. *Biochemistry*, **47**, 13726-13732.
77. Doyen, C.M., An, W., Angelov, D., Bondarenko, V., Mietton, F., Studitsky, V.M., Hamiche, A., Roeder, R.G., Bouvet, P. and Dimitrov, S. (2006) Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Molecular and cellular biology*, **26**, 1156-1164.
78. Kraus, W.L. (2009) New functions for an ancient domain. *Nature structural & molecular biology*, **16**, 904-907.

79. Changolkar, L.N., Costanzi, C., Leu, N.A., Chen, D., McLaughlin, K.J. and Pehrson, J.R. (2007) Developmental changes in histone macroH2A1-mediated gene regulation. *Molecular and cellular biology*, **27**, 2758-2764.
80. Gardiner-Garden, M. and Frommer, M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.*, **196**, 261-282.
81. Fatemi, M., Pao, M.M., Jeong, S., Gal-Yam, E.N., Egger, G., Weisenberger, D.J. and Jones, P.A. (2005) Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level. *Nucleic acids research*, **33**, e176.
82. Illingworth, R.S. and Bird, A.P. (2009) CpG islands--'a rough guide'. *FEBS letters*, **583**, 1713-1720.
83. Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes & development*, **16**, 6-21.
84. Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics*, **33 Suppl**, 245-254.
85. Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature genetics*, **39**, 457-466.
86. Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L. and Schubeler, D. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature genetics*, **37**, 853-862.
87. Wolf, S.F., Jolly, D.J., Lunnen, K.D., Friedmann, T. and Migeon, B.R. (1984) Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. *Proc. Natl. Acad. Sci. USA*, **81**, 2806-2810.
88. Yen, P.H., Patel, P., Chinault, A.C., Mohandas, T. and Shapiro, L.J. (1984) Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human X chromosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 1759-1763.
89. Cotton, A.M., Avila, L., Penaherrera, M.S., Affleck, J.G., Robinson, W.P. and Brown, C.J. (2009) Inactive X chromosome-specific reduction in placental DNA methylation. *Human molecular genetics*, **18**, 3544-3552.
90. Goodfellow, P.J., Mondello, C., Darling, S.M., Pym, B., Little, P. and Goodfellow, P.N. (1988) Absence of methylation of a CpG-rich region at the 5' end of the *MIC2* gene on the active X, and inactive X, and the Y chromosome. *Proc. Natl. Acad. Sci. USA*, **85**, 5605-5609.
91. Viegas-Pequignot, E., Dutrillaux, B. and Thomas, G. (1988) Inactive X chromosome has the highest concentration of unmethylated Hha I sites. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 7657-7660.
92. Bernardino, J., Lombard, M., Niveleau, A. and Dutrillaux, B. (2000) Common methylation characteristics of sex chromosomes in somatic and germ cells from mouse, lemur and human. *Chromosome Res*, **8**, 513-525.
93. Hellman, A. and Chess, A. (2007) Gene body-specific methylation on the active X chromosome. *Science*, **315**, 1141-1143.
94. Rappold, G.A., Cremer, T., Hager, H.D., Davies, K.E., Müller, C.R. and Yang, T. (1984) Sex chromosome positions in human interphase nuclei as studied by in situ hybridization with chromosome specific DNA probes. *Hum Genet.*, **67**, 317-325.
95. Schempp, W. and Meer, B. (1983) Cytologic evidence for three human X-chromosomal segments escaping inactivation. *Hum. Genet.*, **63**, 171-174.
96. Gilbert, D.M. (2002) Replication timing and transcriptional control: beyond cause and effect. *Curr. Op. Cell Biol.*, **14**, 377-383.
97. Hiratani, I., Takebayashi, S., Lu, J. and Gilbert, D.M. (2009) Replication timing and transcriptional control: beyond cause and effect--part II. *Current opinion in genetics & development*, **19**, 142-149.
98. Gomez, M. and Brockdorff, N. (2004) Heterochromatin on the inactive X chromosome delays replication timing without affecting origin usage. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 6923-6928.
99. Cohen, S.M., Brylawski, B.P., Cordeiro-Stone, M. and Kaufman, D.G. (2003) Same origins of DNA replication function on the active and inactive human X chromosomes. *Journal of cellular biochemistry*, **88**, 923-931.

100. Silva, S.S., Rowntree, R.K., Mekhoubad, S. and Lee, J.T. (2008) X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4820-4825.
101. Hall, L.L., Byron, M., Butler, J., Becker, K.A., Nelson, A., Amit, M., Itskovitz-Eldor, J., Stein, J., Stein, G., Ware, C. *et al.* (2008) X-inactivation reveals epigenetic anomalies in most hESC but identifies sublines that initiate as expected. *Journal of cellular physiology*, **216**, 445-452.
102. Hoffman, L.M., Hall, L., Batten, J.L., Young, H., Pardasani, D., Baetge, E.E., Lawrence, J. and Carpenter, M.K. (2005) X-inactivation status varies in human embryonic stem cell lines. *Stem cells (Dayton, Ohio)*, **23**, 1468-1478.
103. Shen, Y., Matsuno, Y., Fouse, S.D., Rao, N., Root, S., Xu, R., Pellegrini, M., Riggs, A.D. and Fan, G. (2008) X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4709-4714.
104. Tinker, A.V. and Brown, C.J. (1998) Induction of XIST expression from the human active X chromosome in mouse/human somatic cell hybrids by DNA demethylation. *Nucleic Acids Res.*, **26**, 2935-2940.
105. Hansen, R.S., Canfield, T.K., Stanek, A.M., Keitges, E.A. and Gartler, S.M. (1998) Reactivation of *XIST* in normal fibroblasts and a somatic cell hybrid: abnormal localization of XIST RNA in hybrid cells. *Proc. Natl Acad. Sci. USA*, **95**, 5133-5138.
106. Chow, J.C., Hall, L.L., Baldry, S.E., Thorogood, N.P., Lawrence, J.B. and Brown, C.J. (2007) Inducible XIST-dependent X-chromosome inactivation in human somatic cells is reversible. *Proc. Natl Acad. Sci. USA*, **104**, 10104-10109.
107. Plath, K., Talbot, D., Hamer, K.M., Otte, A.P., Yang, T.P., Jaenisch, R. and Panning, B. (2004) Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J. Cell. Biol.*, **167**, 1025-1035.
108. Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T. and Wutz, A. (2006) Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.*, **25**, 3110-3122.
109. Brown, C.J., Powers, V.E., Monroe, D.M., Sheinin, R. and Willard, H.F. (1989) A gene on the short arm of the human X chromosome complements the murine tsA1S9 DNA synthesis mutation. *Som. Cell Mol. Genet.*, **15**, 173-178.
110. Willard, H.F., Brown, C.J., Carrel, L., Hendrich, B. and Miller, A.P. (1993) Epigenetic and chromosomal control of gene expression: molecular and genetic analysis of X chromosome inactivation. *Cold Spring Harb. Quant. Symp Biol.*, **58**, 315-322.
111. Brown, C.J. and Willard, H.F. (1989) Noninactivation of a selectable human X-linked gene that complements a murine temperature-sensitive cell cycle defect. *Am. J. Hum. Genet.*, **45**, 592-598.
112. Carrel, L., Cottle, A.A., Goglin, K.C. and Willard, H.F. (1999) A first-generation X-inactivation profile of the human X chromosome. *Proc. Natl Acad. Sci. USA*, **96**, 14440-14444.
113. Brown, C.J., Carrel, L. and Willard, H.F. (1997) Expression of genes from the human active and inactive X chromosomes. *Am. J. Hum. Genet.*, **60**, 1333-1343.
114. Gilbert, S.L., Pehrson, J.R. and Sharp, P.A. (2000) XIST RNA associates with specific regions of the inactive X chromatin. *J Biol Chem*, **275**, 36491-36494.
115. Pehrson, J.R. and Fried, V.A. (1992) MacroH2A, a core histone containing a large nonhistone region. *Science*, **257**, 1398-1400.
116. Nusinow, D.A., Sharp, J.A., Morris, A., Salas, S., Plath, K. and Panning, B. (2007) The histone domain of macroH2A1 contains several dispersed elements that are each sufficient to direct enrichment on the inactive X chromosome. *Journal of molecular biology*, **371**, 11-18.
117. Chadwick, B.P., Valley, C.M. and Willard, H.F. (2001) Histone variant macroH2A contains two distinct macrochromatin domains capable of directing macroH2A to the inactive X chromosome. *Nucl. Acids Res.*, **29**, 2699-2705.
118. Pehrson, J.R. and Fuji, R.N. (1998) Evolutionary conservation of histone macroH2A subtypes and domains. *Nucleic acids research*, **26**, 2837-2842.
119. de la Cruz, C.C., Fang, J., Plath, K., Worringer, K.A., Nusinow, D.A., Zhang, Y. and Panning, B. (2005) Developmental regulation of Suz 12 localization. *Chromosoma*, **114**, 183-192.
120. Erhardt, S., Su, I.H., Schneider, R., Barton, S., Bannister, A.J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovsky, A. and Surani, M.A. (2003) Consequences of the depletion of

zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. *Development (Cambridge, England)*, **130**, 4235-4248.
121. Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P. and Magnuson, T. (2005) The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr Biol*, **15**, 942-947.
122. Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E. and Helin, K. (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *The EMBO journal*, **23**, 4061-4071.
123. Mak, W., Baxter, J., Silva, J., Newall, A.E., Otte, A.P. and Brockdorff, N. (2002) Mitotically stable association of polycomb group proteins eed and enx1 with the inactive X chromosome in trophoblast stem cells. *Curr Biol*, **12**, 1016-1020.
124. Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J.C., Nagy, A. and Magnuson, T. (2001) Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nature genetics*, **28**, 371-375.
125. Kalantry, S. and Magnuson, T. (2006) The Polycomb Group Protein EED Is Dispensable for the Initiation of Random X-Chromosome Inactivation. *PLoS Genet*, **2**, e66.
126. Kalantry, S., Mills, K.C., Yee, D., Otte, A.P., Panning, B. and Magnuson, T. (2006) The Polycomb group protein Eed protects the inactive X-chromosome from differentiation-induced reactivation. *Nat Cell Biol*, **8**, 195-202.
127. 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**, 750-756.
128. Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E. and Allis, C.D. (2006) Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Molecular and cellular biology*, **26**, 2560-2569.
129. Takihara, Y., Tomotsune, D., Shirai, M., Katoh-Fukui, Y., Nishii, K., Motaleb, M.A., Nomura, M., Tsuchiya, R., Fujita, Y., Shibata, Y. *et al.* (1997) Targeted disruption of the mouse homologue of the Drosophila polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects. *Development (Cambridge, England)*, **124**, 3673-3682.
130. Leeb, M. and Wutz, A. (2007) Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. *The Journal of cell biology*, **178**, 219-229.
131. Lomberk, G., Wallrath, L. and Urrutia, R. (2006) The Heterochromatin Protein 1 family. *Genome biology*, **7**, 228.
132. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*, **410**, 116-120.
133. Stewart, M.D., Li, J. and Wong, J. (2005) Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Molecular and cellular biology*, **25**, 2525-2538.
134. Fuks, F., Hurd, P.J., Deplus, R. and Kouzarides, T. (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic acids research*, **31**, 2305-2312.
135. Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C. and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, **410**, 120-124.
136. Muchardt, C., Guilleme, M., Seeler, J.S., Trouche, D., Dejean, A. and Yaniv, M. (2002) Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO reports*, **3**, 975-981.
137. Chadwick, B.P. and Willard, H.F. (2003) Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome. *Hum. Mol. Genet.*, **12**, 2167-2178.
138. Nickerson, J. (2001) Experimental observations of a nuclear matrix. *Journal of cell science*, **114**, 463-474.
139. Tsutsui, K.M., Sano, K. and Tsutsui, K. (2005) Dynamic view of the nuclear matrix. *Acta medica Okayama*, **59**, 113-120.

140. Mattern, K.A., Humbel, B.M., Muijsers, A.O., de Jong, L. and van Driel, R. (1996) hnRNP proteins and B23 are the major proteins of the internal nuclear matrix of HeLa S3 cells. *Journal of cellular biochemistry*, **62**, 275-289.
141. Mattern, K.A., van Goethem, R.E., de Jong, L. and van Driel, R. (1997) Major internal nuclear matrix proteins are common to different human cell types. *Journal of cellular biochemistry*, **65**, 42-52.
142. Kiledjian, M. and Dreyfuss, G. (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *The EMBO journal*, **11**, 2655-2664.
143. Romig, H., Fackelmayer, F.O., Renz, A., Ramsperger, U. and Richter, A. (1992) Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. *The EMBO journal*, **11**, 3431-3440.
144. Fackelmayer, F.O., Dahm, K., Renz, A., Ramsperger, U. and Richter, A. (1994) Nucleic-acid-binding properties of hnRNP-U/SAF-A, a nuclear-matrix protein which binds DNA and RNA in vivo and in vitro. *European journal of biochemistry / FEBS*, **221**, 749-757.
145. Fackelmayer, F.O. (2005) A stable proteinaceous structure in the territory of inactive X chromosomes. *J Biol Chem*, **280**, 1720-1723.
146. Brown, C.J. and Baldry, S.E.L. (1996) Evidence that heteronuclear proteins interact with the XIST RNA in vitro. *Somat. Cell Mol. Genet.*, **22**, 403-417.
147. Helbig, R. and Fackelmayer, F.O. (2003) Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. *Chromosoma*, **112**, 173-182.
148. Baumann, C. and De La Fuente, R. (2009) ATRX marks the inactive X chromosome (Xi) in somatic cells and during imprinted X chromosome inactivation in trophoblast stem cells. *Chromosoma*, **118**, 209-222.
149. Huynh, K.D. and Lee, J.T. (2003) Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature*, **426**, 857-862.
150. Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Le Baccon, P., Chow, J. and Heard, E. (2009) Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc. Natl Acad. Sci. USA*, **106**, 5198-5203.
151. Chow, J.C., Hall, L.L., Lawrence, J.B. and Brown, C.J. (2002) Ectopic XIST transcripts in human somatic cells show variable expression and localization. *Cytogenet. Genome Res.*, **99**, 92-98.
152. Anderson, C.L. and Brown, C.J. (2002) Variability of X chromosome inactivation: effect on levels of TIMP1 RNA and role of DNA methylation. *Hum. Genet.*, **110**, 271-278.
153. Lyon, M.F. (2006) Do LINEs Have a Role in X-Chromosome Inactivation? *Journal of biomedicine & biotechnology*, **2006**, 59746.
154. Migeon, B.R., Kazi, E., Haisley-Royster, C., Hu, J., Reeves, R., Call, L., Lawler, A., Moore, C.S., Morrison, H. and Jeppesen, P. (1999) Human X inactivation center induces random X chromosome inactivation in male transgenic mice. *Genomics*, **59**, 113-121.
155. Heard, E., Mongelard, F., Arnaud, D., Chureau, C., Vourc'h, C. and Avner, P. (1999) Human XIST yeast artificial chromosome transgenes show partial X inactivation center function in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **96**, 6841-6846.
156. Brown, C.J. and Willard, H.F. (1994) The human X inactivation center is not required for maintenance of X inactivation. *Nature*, **368**, 154-156.
157. Heard, E. (2005) Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. *Curr. Opin. Genet. Dev.*, **15**, 482-489.
158. Heard, E. and Disteche, C.M. (2006) Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev.*, **20**, 1848-1867.
159. Walker, S.C., Scott, F.H., Srisawat, C. and Engelke, D.R. (2008) RNA affinity tags for the rapid purification and investigation of RNAs and RNA-protein complexes. *Methods in molecular biology (Clifton, N.J*, **488**, 23-40.
160. Keryer-Bibens, C., Barreau, C. and Osborne, H.B. (2008) Tethering of proteins to RNAs by bacteriophage proteins. *Biology of the cell / under the auspices of the European Cell Biology Organization*, **100**, 125-138.
161. Rook, M.S., Lu, M. and Kosik, K.S. (2000) CaMKIIalpha 3' untranslated region-directed mRNA translocation in living neurons: visualization by GFP linkage. *J Neurosci*, **20**, 6385-6393.
162. Shav-Tal, Y., Darzacq, X., Shenoy, S.M., Fusco, D., Janicki, S.M., Spector, D.L. and Singer, R.H. (2004) Dynamics of single mRNPs in nuclei of living cells. *Science*, **304**, 1797-1800.

163. Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H. *et al.* (2004) From silencing to gene expression: real-time analysis in single cells. *Cell*, **116**, 683-698.
164. Fusco, D., Accornero, N., Lavoie, B., Shenoy, S.M., Blanchard, J.M., Singer, R.H. and Bertrand, E. (2003) Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol*, **13**, 161-167.
165. Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S.M., Phair, R.D. and Singer, R.H. (2007) In vivo dynamics of RNA polymerase II transcription. *Nature structural & molecular biology*, **14**, 796-806.
166. Willingham, A.T., Orth, A.P., Batalov, S., Peters, E.C., Wen, B.G., Aza-Blanc, P., Hogenesch, J.B. and Schultz, P.G. (2005) A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science*, **309**, 1570-1573.
167. Zhou, Z., Licklider, L.J., Gygi, S.P. and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature*, **419**, 182-185.
168. Kota, S.K. (2009) RNAi in X inactivation: contrasting findings on the role of interference. *Bioessays*, **31**, 1280-1283.
169. Kanduri, C., Whitehead, J. and Mohammad, F. (2009) The long and the short of it: RNA-directed chromatin asymmetry in mammalian X-chromosome inactivation. *FEBS letters*, **583**, 857-864.
170. Grewal, S.I. RNAi-dependent formation of heterochromatin and its diverse functions. *Current opinion in genetics & development*.
171. Savarese, F., Flahndorfer, K., Jaenisch, R., Busslinger, M. and Wutz, A. (2006) Hematopoietic precursor cells transiently reestablish permissiveness for x inactivation. *Molecular and cellular biology*, **26**, 7167-7177.
172. Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R. *et al.* (2007) Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell stem cell*, **1**, 55-70.