Studies on X chromosome inactivation and the X-linked disease Rett syndrome

by

Sandra Luikenhuis

Diplom, Biology
Bonn University, Germany, 1997

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biology

at the

Massachusetts Institute of Technology

July 2004
[September 2004]

Signature Author: _____
Department of Biology
July 15, 2004

Certified by: _____
Rudolf Jaenisch
Professor of Biology
Thesis Supervisor

Accepted by: _____
Stephen P. Bell
Professor of Biology
Co-Chair, Department Graduate Committee

To P. A. Trapp

ACKNOWLEDGEMENTS

First and foremost I would like to thank my parents and my husband David Sinclair for their love and support. Without them, none of this would have been possible. I would also like to thank my mentor, P. A. Trapp, who introduced me to Biology at a young age and continuously fostered my interested with an endless supply of books, conversations, and encouragement.

I would like to thank my thesis supervisor, Rudolf Jaenisch, for giving me the opportunity to work in his lab, for his guidance, and for making it possible for me to combine a successful graduate student carrier with family life. I am grateful to the past and present members of the Jaenisch lab for their help and support during my years in graduate school. Györgyi Csankovszki introduced me to the lab and taught me the basics about working in a mammalian system. Anton Wutz gave me invaluable advice when I was setting up my first project. I would like to thank our mouse specialists, Jessie Dausman and Ruth Flannery, for their tireless work. Emi Giacometti deserves special thanks for her patience with me. I thank Laurie Boyer for encouragement and many helpful discussions. Joost Gribnau was not only a great scientific collaborator, but he and his wife Susanne were also wonderful friends. I am especially grateful to Caroline Beard for her support, for being a source of many helpful ideas, comments, and suggestions, and for reading this thesis.

I would also like to thank the members of my thesis committee, Terry-Orr Weaver and David Page, for their advice and guidance throughout my graduate carrier.

Table of Contents

Studies on X chromosome inactivation and the X-linked disease Rett syndrome

by

Sandra Luikenhuis

Submitted to the Department of Biology
on July 15, 2004 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy
in Biology

Abstract

Deletion of the *Xist* gene results in skewed X-inactivation. To distinguish primary non-random choice from post-choice selection, we analyzed X-inactivation in early embryonic development in the presence of two different *Xist* deletions. We found that *Xist* is an important choice element, and that in the absence of an intact *Xist* gene, the X chromosome will never be chosen as the active X.

To understand the molecular mechanisms that affect choice we analyzed the role of replication timing prior to X-inactivation. The X chromosomes replicated asynchronously before X-inactivation but analysis of cell-lines with skewed X-inactivation showed no preference for one of the two *Xist* alleles to replicate early, indicating that asynchronous replication timing prior to X-inactivation does not play a role in skewing of X-inactivation.

Expression of the *Xist* is negatively regulated by its antisense gene, *Tsix*. In order to determine the role of transcription in *Tsix* function, we modulated *Tsix* transcription with minimal disturbance of the genomic sequence. Loss of *Tsix* transcription lead to non-random inactivation of the targeted chromosome, whereas induction of *Tsix* expression caused the targeted chromosome always to be chosen as the active X. These results for the first time establish a function for antisense transcription in the regulation of *Xist* expression.

The X-linked disease Rett syndrome (RTT), a neurodevelopmental disorder, is caused by mutations in the *MECP2* gene. We used a mouse model to test the hypothesis that RTT is exclusively caused by neuronal MeCP2 deficiency. Expression of an *Mecp2* transgene in postmitotic neurons resulted in symptoms of severe motor dysfunction. Transgene expression in *Mecp2* mutant mice, however, rescued the RTT phenotype.

Thesis Supervisor: Rudolf Jaenisch

Title: Professor of Biology

Chapter 1

Introduction

Dosage Compensation

Dosage compensation mechanisms have evolved in diploid species to equalize the difference in gene dosage caused by the unequal number of sex chromosomes in males and females. Dosage compensation has been studied intensely in fruit flies, nematode worms, and mammals, and, interestingly, each of these species solves the problem differently (Figure 1). In the fly *Drosophila melanogaster*, dosage compensation is achieved by upregulating transcription two-fold in the XY males to equal expression in XX females (Akhtar 2003). In *Caenorhabditis elegans* hermaphrodites, transcription from the two X chromosomes is repressed by half to meet expression in XO males (see Meyer (2000) for review). Mammals, on the other hand, compensate for the difference in gene dosage between the XX female and the XY male by inactivating one of the X chromosomes present in the female. In all three mechanisms, the X chromosome adopts a distinct chromatin composition in one sex and is bound along most of its length by specific factors that regulate its expression state. Mammals, however, present a special case, because the two X chromosomes that need to be distinguished are present in the same cell.

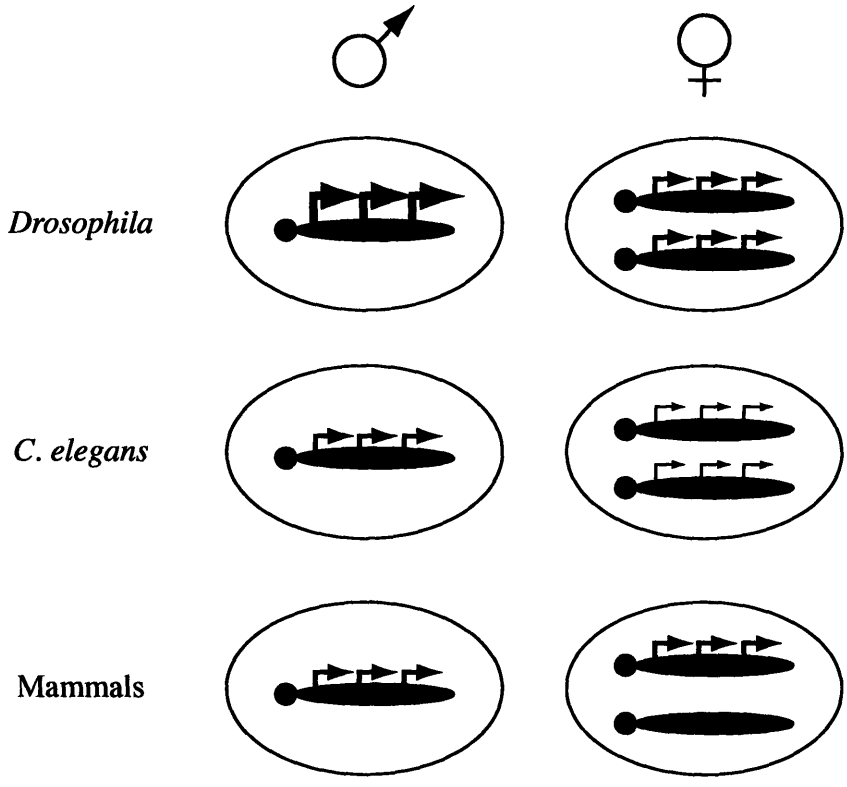


Figure 1. Mechanisms of dosage compensation. X-linked gene expression is equalized in males and females by either transcriptional activation of the single male X chromosome twofold in comparison to females (*Drosophila*), by transcription repression so that transcription from both X chromosomes of an hermaphrodite is repressed by half relative to males (*C. elegans*), or by transcriptional silencing of one of the female X chromsomes (mammals). The size of the arrows above the X chromosomes (blue) represents transcription activity. The inactive X is shown in red.

More than 40 years ago, Mary Lyon first proposed that dosage compensation in mammals is achieved by random inactivation of one of the X chromosomes in the female cell (Lyon 1961). The inactive X forms the Barr body, a condensed chromatin structure in the nucleus (Barr 1962). Based on the number Barr bodies found in X-chromosome aneuploidies (XXX individuals have two and XXXX individuals have three) a modified version of the Lyon hypothesis was formulated that stated that in a diploid nucleus only one X remains active, whereas every other X is inactivated (Lyon 1962). Studies on X-to-autosome translocations revealed that X-inactivation is initiated from one location on the X chromosome, the *X-inactivation center* (*XIC* in humans and *Xic* in mice) (Russell 1963). Silencing of the inactive X (Xi) spreads *in cis* and, once established, is stably inherited through many cell divisions. X-inactivation occurs early in embryonic development around the time when pluripotent cells differentiate to restricted lineages (Monk and Harper 1979) with the result that females are functionally mosaic for X-linked traits (Lyon 1961; Russell 1961). X-inactivation is a process that involves several steps including choice of the active X chromosome (Xa), initiation, and spreading of the inactivation along the Xi and subsequent maintenance of the silent state. Molecular mechanisms that regulate X-inactivation are best understood in the mouse, partially owing to the fact that mouse embryonic stem cells (ES) are available as a tool. ES cells undergo X-inactivation upon differentiation in culture (Paterno and McBurney 1985; Tada et al. 1993), can be genetically modified, and give rise to germ-line transmitting mice which makes it possible to study this process *in vitro* as well as in the embryo.

Xist RNA is critical for X-inactivation

A key finding in the field of X-inactivation was the discovery of *Xist*, the *Xi-specific transcript* (Borsani et al. 1991; Brockdorff et al. 1991; Brockdorff et al. 1992; Brown et al. 1992; Brown 1991). The *Xist* gene maps to the *Xic* and encodes a 17,000 nucleotide-long non-coding RNA that is spliced, polyadenylated, and associates specifically with the Xi *in cis* (Figure 2). Studies of *Xist* deletions (Csankovszki et al. 1999; Marahrens et al. 1997; Penny et al. 1996) and an inducible *Xist* cDNA transgene (Wutz and Jaenisch 2000; Wutz et al. 2002) have shown that *Xist* RNA is necessary and sufficient for chromosome-wide gene silencing. Prior to X-inactivation *Xist* RNA is expressed at low levels from all X chromosomes in female as well as male cells. Once the Xa has been chosen and X-inactivation is initiated, *Xist* RNA is stabilized (Panning et al. 1997; Sheardown et al. 1997), accumulates on the Xi *in cis*, and triggers progressive heterochromatinization (Avner and Heard 2001; Brockdorff 2002; Plath et al. 2002). *Xist* transcripts that do not associate with chromatin are inherently unstable, thus prohibiting accumulation on remote sites (Wutz et al. 2002). On the sequence level, *Xist/XIST* genes sequenced to date show a low degree of conservation (Nesterova et al. 2001) with the exception of 6 repeats, designated A through F, which may act as potential docking sites for interacting factors and/or chromatin. A recent study has shown that *Xist* RNA consists of functional domains that mediate silencing and localization (Wutz et al. 2002). The 5' A-repeat is crucial for silencing and *Xist* RNA deleted for the A-repeat spreads over the entire chromosome but does not trigger silencing. The localization of *Xist* RNA, on the other hand, is mediated by a number of functionally redundant regions that lack

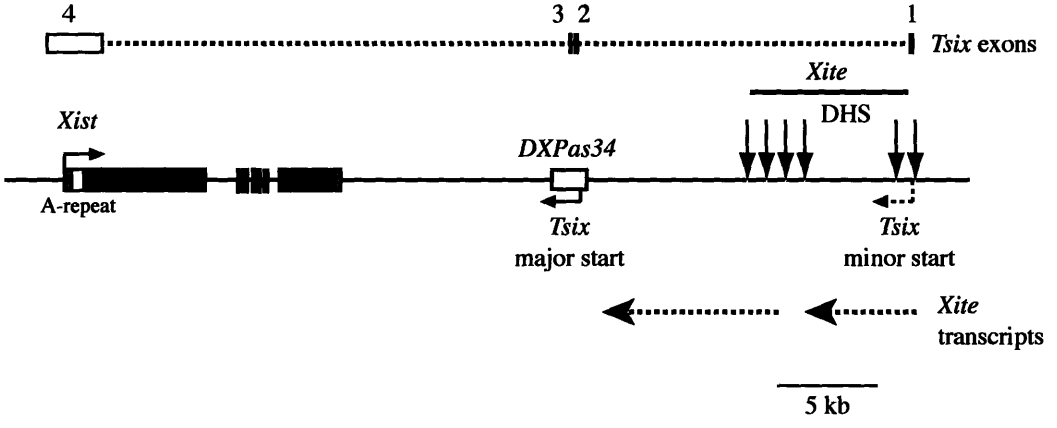


Figure 2. Structure of the mouse *Xic*. *Xist* consists of 7 exons (dark grey boxes); the *Xist* A-repeat is shown as a white box. The *Tsix* major transcriptional start site (solid arrow) overlaps with the CpG-rich region *DXPas34* (light grey box). The minor *Tsix* transcription start site is shown as a dashed arrow. The relative position of the *Tsix* exons is indicated. *Xite* intergenic transcripts initate at two sites (indicated by vertical arrows) associated with DNase hypersensitive sites (DHS).

identifiable sequence motifs and most likely contain multiple low-affinity binding sites for *trans*-acting factors that bind in a co-operative fashion.

Tsix, an antisense *Xist* transcript

Tsix, *Xist* spelled back to front, was first described in 1999, when several groups reported antisense transcription through the *Xist* locus (Debrand et al. 1999; Lee et al. 1999; Mise et al. 1999) (Figure 2). The 40 kb *Tsix* transcript initiates from a major promoter 15 kb downstream from the *Xist* 3' end within the *DXPas34* locus and extends through *Xist* into its promoter region (Lee et al. 1999). Further analysis revealed a minor promoter and, in addition, a complex pattern of alternative splicing that produces mature polyadenylated transcripts ranging from 2.7 to 4.3 kb (Sado et al. 2001). Similar to *Xist*, *Tsix* has no significant open reading frame and most likely does not encode a protein.

Tsix expression coincides with low-level *Xist* expression in undifferentiated male and female ES cells and developing embryos. Upon differentiation, *Tsix* expression is extinguished from the future Xi concurrent with *Xist* accumulation (Debrand et al. 1999; Lee et al. 1999). Expression on the Xa continues for several days, even after *Xist* expression has been extinguished (Lee et al. 1999; Sado et al. 2001). Mutations that abolish *Tsix* transcription have the opposite effect of *Xist* mutations such that the chromosome bearing the *Tsix* mutation is always chosen as the Xa (Debrand et al. 1999; Lee and Lu 1999; Luikenhuis et al. 2001; Sado et al. 2001) indicating that *Tsix* may act as a negative regulator of *Xist*.

Tsix is not the only long non-coding RNA that has been implicated in the regulation of gene expression in a *cis*-limited manner. A number of antisense RNAs have

been shown to regulate parent-specific expression of imprinted gene clusters. The *Air* RNA, for example, is required for repression of the paternally inherited *Insulin like growth factor receptor type 2* (*Igf2r*) gene in mice. *Air* also regulates genes distant to its transcriptions unit, and it has been suggested that the RNA could potentially associate in *cis* with chromatin to repress genes (Sleutels et al. 2002; Zwart et al. 2001). In the PWS/AS region, the locus responsible for Prader-Willi syndrome and Angelman syndrome, *Ube3a*, a ubiquitin-protein ligase, is paired with the antisense RNA, *Ube3a-as* (Chamberlain and Brannan 2001; Rougeulle et al. 1998; Runte et al. 2004), and imprinting the BWS (Beckwith-Wiedemann) locus may be controlled by the antisense transcript *Kcnq1-as* (Weksberg et al. 2003). In all cases, the antisense RNA has an antagonistic relationship with the sense partner, and in all cases, the antisense RNA is non-coding. The molecular mechanism of antisense regulation is still unknown, and current models include transcriptional interference, chromatin reorganization caused by antisense transcription, destabilization of the sense transcript by sense-antisense duplex formation, and RNA-independent mechanisms that rely on associated DNA sequence elements (see Ogawa and Lee (2002) for a detailed review).

Counting the X chromosomes

In mammalian ES cells and embryonic tissues, cells randomly designate one X as the Xa and inactivate any remaining Xs. The most parsimonious model proposes the presence of an autosomally encoded blocking factor that marks the future active X and is present in limited quantities such that only one X can remain active per diploid set of autosomes. Therefore *Xist* expression followed by X-inactivation would be the default

fate for all but one X. The *Xic* region that mediates blocking factor activity is referred to as the counting element (Plath et al. 2002). A test for the counting element is the ability of a sequence to induce X-inactivation in a male cell when integrated as an autosomal transgene. A number of *Xist* containing transgenes have been tested for their ability to induce counting, and the minimal region required is defined by a 35 kb transgene containing genomic *Xist* flanked by 9 kb of upstream and 3 kb of downstream sequence (Herzing et al. 1997). However, a caveat is that this transgene is present as multiple tandem copies, and no single-copy transgene has ever been shown to recapitulate *Xic* activity. By definition, deletion of the counting element leads to ectopic *Xist* expression in XY cells. The deletion of 65 kb of sequence immediately downstream of *Xist* results in an allele from which X-inactivation is initiated by default, in XX and XO (Clerc and Avner 1998) as well was in XY cells (Morey et al. 2004). Re-inserting a 37 kb sequence starting at the terminal *Xist* exons and extending 3' restores counting which narrows the location of the counting element down to a 20 kb bipartite domain within this region that excludes the *Tsix* and *Xite* loci, both of which have been shown previously not to affect counting (Morey et al. 2004) (Lee et al. 1999; Ogawa and Lee 2003).

Choosing the active X chromosome

In contrast to dosage compensation in flies and worms, where all X chromosomes in a cell are treated equally, the two X chromosomes in cells or female cells take on radically different fates. Choice is random when two equivalent X chromosomes are present in the same cell. However, a number of *cis*-acting elements have been identified that lead to skewing of the Xa choice, presumably by affecting the likelihood that the

blocking factor binds. Several lines of evidence suggest that *Xist* and its negative regulator *Tsix* play an important role in choice. In female cells heterozygous for an Xist deletion, only the wild type chromosome is inactivated (Csankovszki et al. 1999; Marahrens et al. 1998; Penny et al. 1996). At least for one of these deletions it has been shown *in vivo* that the wild type chromosome was never selected as the Xa (Marahrens et al. 1998) suggesting that the transcribed region of *Xist* or *Xist* RNA itself may act as a choice element. While loss of *Xist* function increases the chance of the mutant X to be chosen as the Xa, a chromosome with increased *Xist* expression is less likely to be chosen. Ectopic sense transcription across the *Xist* locus causes preferential inactivation of the mutated X which suggests that increased transcription of Xist has a negative influence on Xa choice (Nesterova et al. 2003; Newall et al. 2001). Loss of *Tsix* transcription leads to skewed Xa choice away from the mutated X (Lee and Lu 1999; Sado et al. 2001). In addition, *Tsix* transcription oppositely affects the abundance of *Xist* RNA *in cis* (Lee et al. 1999; Sado et al. 2001), suggesting that *Tsix* transcription acts by regulating *Xist* RNA steady-state levels.

Another genetic element implicated in choice is the X choosing element, *Xce*, which lies 3' of *Xist* and beyond *Tsix* (Simmler et al. 1993). Crossing divergent mouse strains heterozygous for the *Xce* leads to skewing of X-inactivation in the F1 female offspring such that X chromosomes bearing a stronger *Xce* allele are chosen more frequently as the Xa. Interestingly, an *Xce* of greater strength appears to be correlated with a decreased level of *Xist* expression in female somatic cells (Brockdorff et al. 1991), and before differentiation with higher levels of *Tsix* expression (Stavropoulos et al. 2001). These findings suggest that *Xce* might affect choice by acting through *Tsix* to

modulate *Xist* RNA levels *in cis*. Recently, *Xite*, a *cis*-acting element that harbors intergenic transcription start sites and DNaseI hypersensitive sites (DHSs), has been identified as a candidate locus for the *Xce* (Ogawa and Lee 2003) (Figure 2, page 12). *Xite* is located 20 - 32 kb downstream of *Xist* and shows strain-specific polymorphisms. Deleting *Xite* downregulates *Tsix in cis* and skews choice, indicating that it promotes *Tsix* persistence on the future Xa. In addition, *Xite* on the X with the greater *Xce* strength appears to be more sensitive to DNAseI and shows greater intergenic transcription (Ogawa and Lee 2003). These findings together suggest that choice is determined by a complex interplay of sense and antisense transcription prior to the onset of X-inactivation. However, the mechanism whereby this balance influences choice still remains unclear.

After choosing the Xa, the next step is to carry out the fate assigned to each X chromosome. On the presumptive Xi, *Xist* activity increases *in cis* by modulating *Xist* stability (Panning et al. 1997; Sheardown et al. 1997). *Xist* stabilization may be achieved by the developmental shut-off of *Tsix* (Debrand et al. 1999; Lee et al. 1999) or a developmentally regulated trans-acting factor. On the chosen Xa, *Tsix* continues to be expressed, and the Xa is protected from the developmentally regulated boost of *Xist* activity (Lee et al. 1999). At the end of this phase *Xist* expression on the Xa is turned off by methylation of the *Xist* promoter (Beard et al. 1995; Norris et al. 1994; Panning and Jaenisch 1996).

Imprinted X-inactivation

In contrast to random X-inactivation in embryonic tissues, X-inactivation occurs in a parent-specific (imprinted) manner in the extraembryonic tissues (Takagi and Sasaki 1975). In imprinted X-inactivation, the maternally inherited X (Xm) always remains active whereas the paternal X (Xp) expresses *Xist* and is silenced. *Xist* is not expressed in mature oocytes (Avner et al. 2000; Kay et al. 1993) but it is present in mature spermatids (McCarrey and Dilworth 1992; Richler et al. 1992; Salido et al. 1992). Recent evidence suggests that imprinted X-inactivation occurs early in the preimplantation embryo, possibly at the two- or four-cells stage and is mediated by *Xist* expression on the Xp. X-inactivation is then selectively reversed by extinguishing *Xist* expression in the inner cell mass (ICM) which gives rise to the embryo proper (Huynh and Lee 2003; Mak et al. 2004; Okamoto et al. 2004). These findings indicate that X-inactivation in the early embryo is *Xist* RNA dependent, unlike in XX somatic cells in which loss of *Xist* has little or no effect (Brown and Willard 1994; Csankovszki et al. 1999). Imprinted *Xist* expression is thought to be regulated by *Tsix*, which is imprinted oppositely of *Xist* and shows the opposite parent-of-origin-specific effect. *Xist* deletions are lethal when inherited from the father due to lack of X-inactivation in extraembryonic tissues (Marahrens et al. 1997), while *Tsix* deletions cause ectopic X-inactivation of the Xm and are lethal only when inherited from the mother (Lee 2000; Sado et al. 2001). As expected, a paternally inherited *Xist* deletion can be rescued by maternal inheritance of a *Tsix* loss-of-function allele (Sado et al. 2001). It has been suggested that the imprinted expression of *Tsix* may be directed by methylation of the CpG island, *DXPas34*, in the *Tsix* promoter (Lee 2000). Although *DXPas34* is differentially methylated on the Xa and Xi (Courtier et

al. 1995), bisulfite sequencing failed to find any methylation differences in spermatocyes and oocytes (Prissette et al. 2001). Recent evidence suggests that *DXPas34* may be bound by the CCCTC binding factor (CTCF), a protein that can act as a boundary element and regulate enhancer access (Chao et al. 2002). Differential methylation of CTCF binding sites plays a role in establishing the imprint at the *H19/Igf2* gene pair, but on the X chromosome CpG methylation has no effect on CTCF binding to *DXPas34 in vitro* (Chao et al. 2002).

Silencing the X chromosome

Studies using an inducible *Xist* cDNA transgene suggest that X-inactivation can be divided into three phases (Wutz and Jaenisch 2000), initiation, establishment, and maintenance (Figure 3A). In the initiation phase, which lasts for 1 - 2 days after onset of differentiation, high-level *Xist* expression can induce reversible, *Xist*-dependent silencing. During the establishment phase, continued coating by *Xist* is required for silencing, but *de novo* silencing cannot be initiated at this stage. After approximately 24 hours, the maintenance phase is entered and silencing becomes irreversible and *Xist*-independent.

The silent state of the Xi in the embryo proper is locked in by a cascade of chromatin modifications that are sequentially added during the process of X-inactivation (Figure 3B). The earliest detectable events are *Xist*-dependent modifications of histone H3, including methylation of histone H3 Lys-9 and Lys-27 residues by the polychrome group (PcG) protein Enx1 with its binding partner Eed, (Heard et al. 2001; Mermoud et al. 2002; Plath et al. 2003; Silva et al. 2003), hypoacetylation of H3 Lys-9, and hypomethylation of H3 Lys-4 (Heard et al. 2001). The modification of H3 is followed

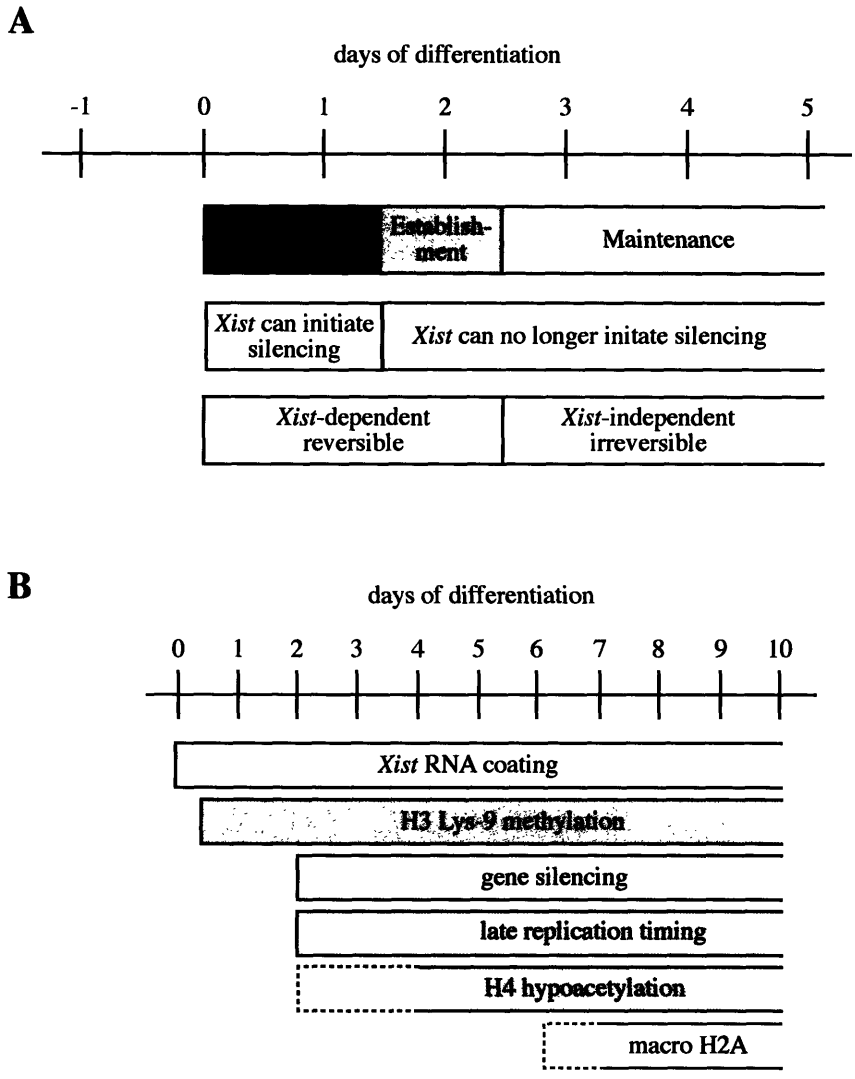


Figure 3. Silencing the Xi. **(A)** Silencing of the Xi occurs in three steps, initiation, establishment, and maintenance. Only during the initiation phase can *Xist* induce gene silencing. Both, the initiaion and establishment phase are *Xist*-dependent and reversible. The maintenance phase is characterized by irreversible, *Xist*-independent silencing. During the establishment as well as the maintenance phase *Xist* can no longer initiate silencing. **(B)** Chromatin modifications on the Xi occur sequentially. Shown is the timing of appearance of chromatin modifications that characterize the heterochromatin on the Xi. Dotted bars symbolize that conflicting data exist about the appearance of a particular modification. DNA methylation is not included in this figure (adapted from Plath et al. 2002)

by gene silencing, histone H4 hypoacetylation, and late replication of the Xi (Chaumeil et al. 2002; Keohane et al. 1996). After the irreversible maintenance phase is entered, macroH2A, a core histone variant, is recruited to the Xi (Rasmussen et al. 2001). The final step is the appearance of CpG methylation of X-linked genes about 21 days post-differentiation (Keohane et al. 1996). The mechanisms described above act redundantly, and DNA methylation appears to be the most important factor (Csankovszki et al. 1999).

Some genes on the X chromosome escape X-inactivation. Genes that escape X-inactivation do not have the chromatin modifications characteristic to the Xi. They lack DNA methylation, histone deacetylation, are early replicating, and are methylated on H3 Lys-4 (Brown and Greally 2003). *Cis*-elements may determine whether *Xist* RNA can mediate silencing in particular regions of the X chromosome but the mechanism that mediates escape from X-inactivation remains unclear.

X chromosome inactivation and human genetic disease

In humans, X-inactivation has implications for the effects seen in genetic diseases that are caused by mutation of X-linked genes. A consequence of X-inactivation is that heterozygous females are mosaics with two types of cells, expressing either the wild type or mutated X, and hence either the normal or disease gene. X-inactivation happens early in development, and the shapes and sizes of the patches expressing one or the other X depends, to a large part, on the individual tissue. Some structures are monoclonal (i.e. they arise from a single cell) and therefore only express one of the two Xs. Others are polyclonal and show various degrees of contribution by the mutant X. Typically, X-

inactivation is random, and in an individual heterozygous for an X-linked disease gene, the contribution of the two types of cell should be equal. However, skewing of X-inactivation may arise by chance due to the small numbers of cells present in the embryo at the time of X-inactivation and can lead to a wide variety of symptoms. In one survey it was found that in about 20% of women the same X was active in 80% of the cells (Naumova et al. 1996). Skewing can also be caused by cell selection, when cells with the active disease gene have a growth advantage or disadvantage, which is usually relevant for rapidly dividing cells. For example, in hypoxanthine phosphoribosyl transferase (HPRT) deficiency (Lesh-Nyhan syndrome) mutant blood cells are selected against during early stages of hematopoiesis (Hakoda et al. 1995) and red blood cells and T-cells have the X carrying the unmutated *HPRT* gene as the Xa. Primary non-random choice as has been described for *Xist* mutations in the mouse (Marahrens et al. 1998), appears to be rare in humans. Only two families with have been identified where skewed X-inactivation correlated with mutations in the *XIST* minimal promoter (Plenge et al. 1997).

A classic example of how the pattern of X-inactivation leads to a wide range of phenotypes is Rett syndrome (RTT), a X-linked dominant neurodevelopmental disorder that affects almost exclusively girls. More than 99% of RTT cases are sporadic, but the occurrence of rare familial cases lead to the discovery of the disease-causing gene encoding the methyl-CpG-binding protein 2 (MeCP2). In these cases, RTT was inherited through the maternal line and transmitted through carrier females who showed almost complete skewing of X-inactivation toward the mutated X and hence only had a slight learning disability or were symptom free (Schanen et al. 1997; Sirianni et al. 1998; Zoghbi et al. 1990). Mutations in *MECP2* do not result in growth advantages for cells

expressing the defective gene from the Xa (Shahbazian et al. 2002c), indicating that the skewed X-inactivation pattern arose most likely by chance.

Clinical features of RTT

The Austrian physician Andreas Rett first described RTT, when he observed two girls with similar abnormality in his waiting room (Rett 1966). The condition was virtually overlooked until Hagberg and colleagues reported 35 cases of RTT in European girls (Hagberg et al. 1983). It is now recognized that RTT occurs at a frequency of about 1/10,000 to 1/15,000 female children (Amir et al. 1999). Classic RTT is defined based on a number of criteria (Hagberg et al. 1985). Girls with RTT are born healthy and have apparently normal development throughout the first 6 month of life. The head size is normal at birth, but at the age of 2 to 4 months head growth begins to decelerate, ultimately resulting in microencephaly. At 6 to 12 month of age, stereotypic movements, such as midline hand wringing, clapping, and hand mouthing, replace purposeful hand use. In addition, girls show autistic features such as social withdrawal and impaired language, and they are often unable to walk and require constant care. The initial stage of regression is followed by apparent stability lasting through decades. A wide variety of atypical clinical variants have been described that do not fit the accepted diagnostic criteria. At the severe end of the spectrum lie patients who never go through a period of normal development, whereas in some mild forms, speech and normal head size are preserved (Neul and Zoghbi 2004). The pathology in RTT is subtle, and the brains of RTT patients shows no signs of neuronal degeneration or gross morphological abnormalities. The most striking feature of RTT brains is the disproportionately small

brain size, affecting the grey matter more than the white matter. Brains of patients with classic RTT weigh 14% - 34% less than healthy brains due to smaller neurons, more densely packed neurons with reduced dendritic spines and arbors (Armstrong et al. 1995).

Rett syndrome is caused by a mutation in MeCP2

The identification of the genetic cause of RTT was hindered by the lack of familial RTT cases because the affected individuals rarely have children. However, the prevalence of RTT in girls and the presence of families with affected half sisters suggested an X-linked dominant inheritance pattern with male lethality (Hagberg et al. 1983). Conventional genome-wide linkage analysis was not feasible, therefore an exclusion mapping approach narrowed the region to the region distal to Xq27.3 (see Shahbazian and Zoghbi (2002) for details). Subsequently, a number of candidate genes with known expression or function in the nervous system were analyzed, but ultimately mutations in the methyl-CpG binding protein 2 (*MECP2*) were identified in RTT patients (Amir et al. 1999). In almost 80% of sporadic and 50% of familial cases mutations in *MECP2* have been identified (Shahbazian and Zoghbi 2001). Mutations in regulatory regions could account for the remaining cases.

MeCP2 and DNA methylation

The addition of methyl groups to position five of cytosine bases in the dinucleotide 5'-CG-3' (CpG) is one of the most prominent epigenetic modification of the mammalian genome. During mammalian development, a wave of demethylation occurs during embryonic cleavage, when the paternal genome is actively stripped of methylation

within hours after fertilization, and the maternal genome is passively demethylated during subsequent cell divisions (Leonhardt et al. 1992). After implantation, the embryo undergoes a wave of *de novo* methylation that establishes a new embryonic methylation pattern (Jaenisch and Bird 2003). The pattern of DNA methylation is created and maintained by a group of enzymes called DNA methyltransferases (DNMTs). Dnmt1, the first methyltransferase to be cloned (Bestor et al. 1988), functions as the major maintenance methyltransferase ensuring preservation of the DNA methylation pattern after each round of cell division (Bestor 1992; Leonhardt et al. 1992; Li et al. 1992). *De novo* methylation after implantation is mediated by Dnmt3a and Dnmt3b, both of which are highly expressed in the developing embryo (Okano et al. 1999). Mutations in the human *DNMT3B* gene have been identified as the cause of ICF (immunodeficiency, centromeric instability, facial abnormalities) syndrome (Hansen et al. 1999; Okano et al. 1999). Dnmt3L, a protein that by itself has no methyltransferase activity, colocalizes with Dnmt3a and Dnmt3b and is essential for establishing methylation imprints in the female germline (Bourc'his et al. 2001; Hata et al. 2002). The last member of the Dnmt-family, Dnmt2, is conserved in flies, but it lacks biochemical methyltransferase activity, and deletion of *Dnmt2* has no phenotype in mice (Kunert et al. 2003; Okano et al. 1998).

In non-embryonic cells, about 80% of all CpG sites in the genome are methylated. However, the promoters of housekeeping genes and tissue specific genes can contain short sequence domains with high CpG contents, so called CpG islands, which are entirely unmethylated, regardless of gene expression (Jaenisch and Bird 2003). In some cases the CpG islands become methylated during development, which leads to the long-term shut-down of the expression of the associated gene. It has been suggested that DNA

methylation acts as a system of cellular memory that is not directly involved in initiating gene silencing, but in sensing and propagating the silent state (Bird 2002). Evidence supporting this view stems from studies that show that Dnmts interact with histone deacetylases (Bachman et al. 2001; Fuks et al. 2001). Moreover, in the fungus *Neurospora crassa* and the plant *Arabidopsis thaliana* histone H3 Lys-9 methylation has been shown to be essential for DNA methylation (Jackson et al. 2002; Tamaru and Selker 2001).

The epigenetic methylation mark can cause an effect by directly preventing sequence-specific DNA-binding proteins from binding. Alternatively, the methylated CpG motif can be recognized by specific proteins that, in turn, sterically block access of other factors or transmit the signal by protein-protein interactions. MeCP1 and MeCP2 were the first activities that were identified based on their ability to bind methylated DNA (Lewis et al. 1992; Meehan et al. 1989). The characterization of the region in MeCP2 that mediates binding to methylated DNA (methyl-binding domain, MBD) (Nan et al. 1993), led to the identification of other MBD-containing proteins (MBD1-4) (Cross et al. 1997; Hendrich and Bird 1998). The unrelated protein Kaiso, which has homology to the POZ zinc finger family of DNA-binding transcription factors, does not contain the classical MBD motif but has also been shown to specifically bind methylated DNA (Daniel and Reynolds 1999; Prokhortchouk et al. 2001). Although the MBD is well conserved between members of the family of MBD-containing proteins as well as between species, MBD-proteins show very little homology (except for MBD2 and MBD3) but genetic analysis has revealed considerable functional overlap between these proteins (Guy et al. 2001; M. Tudor, pers. communication).

MeCP2 is the most extensively studied among the MBD proteins and is the only member of the family that can bind selectively to a single symmetrically methylated CpG (Lewis et al. 1992). In addition to the N-terminal MBD, MeCP2 also contains a transcriptional repression domain (TRD), which is required for transcriptional repression *in vitro* and *in vivo* (Jones et al. 1998; Kaludov and Wolffe 2000; Nan et al. 1997). MeCP2 has been shown to repress transcription by at least two different mechanisms. On the one hand, the TRD of MeCP2 interacts with the Sin3A co-repressor that exists in a complex with HDAC1 and 2, and transcriptional repression by MeCP2 can be relieved with histone deacetylases inhibitors (Jones et al. 1998; Nan et al. 1998). In addition, it also interacts with two other co-repressors, the proto-oncoprotein of the Sloan-Kettering virus (c-Ski) and the nuclear receptor co-repressor (N-CoR), which are also components of histone deacetylase complexes (Kokura et al. 2001). On the other hand, the repression of some promoters by the MeCP2-TRD is insensitive *in vivo* to histone deacetylase inhibitors (Yu et al. 2000). The TRD was found to interact with TFIIB (Kaludov and Wolffe 2000), which suggests that MeCP2 may interfere with the basal transcriptional machinery.

The *Mecp2* gene produces a variety of different transcripts by alternative codon usage as well as differential usage of polyadenylation sites (Figure 4). The gene contains an 8.5 kb 3'-untranslated region (UTR) that shows a domain-like conservation pattern between human and mouse on the sequence level as well as the predicted secondary structure (Coy et al. 1999).

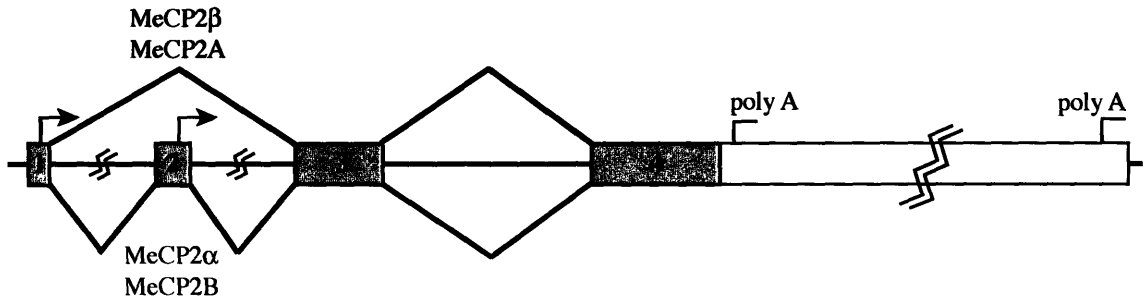


Figure 4. Structure of the Mecp2 gene. The previously described isoform (MeCP2α or MeCP2B) is encoded when all known exons are sequentially spliced (represented by red lines). The novel isoform (MeCP2β or MeCP2A) is produced when exon 1 is spliced onto exon 3, skipping exon 2 (represented by blue lines). MeCP2 exons are shown as dark grey boxes. Alternative start codons in exons 1 and 2 are shown as arrows. The 3'-UTR with its polyadenylation signals (poly A) is depicted as a light grey box).

The use of two alternative polyadenylation sites within the 3'-UTR results in two different transcripts that are 1.9 kb and 10 kb in length, respectively (Coy et al. 1999; Reichwald et al. 2000). Both transcripts have similar half lives and are found at varying levels in most tissues examined, with some tissue specificity (Coy et al. 1999; Reichwald et al. 2000). For example, the long transcript is expressed at relatively high levels in the brain, while the short transcript is barely visible in this tissue. It has been suggested that the 3'-UTR is involved in post-translational regulation of MeCP2 expression, but its function is still unknown. Recently, a new isoform of MeCP2 has been described (MeCP2B or MeCP2α) that includes exon 1 which was previously thought to be part of the 5'-UTR, but lacks exon 2 (Kriaucionis and Bird 2004; Mnatzakanian et al. 2004) (Figure 4). In the mouse brain, the new isoform appears to be more efficiently transcribed as well as translated which results 10 times higher protein compared to the previously identified isoform. However, the alternative N-terminus is located outside the functional domains of MeCP2, which suggests that the functions of these isoforms overlap considerably.

MeCP2 mutations in RTT

In human RTT, almost 70% of reported MeCP2 mutations are C-T transitions that result from spontaneous deamination of methylated cytosines at CpG-sites. About 10% of mutations are intragenic deletions or complex rearrangements that lead to frameshifts in the C-terminal region. The majority of mutations, about two-thirds, result in truncations; the remaining mutations are missense mutations, mostly in the methyl-binding domain (Lee et al. 2001). RTT is found mainly in girls, which lead to the hypothesis that boys

who carry the mutation on their single X die perinatally. However, another explanation for the low incidence of affected males is that the majority of mutations originate on the paternal X and would only be inherited by daughters. One study found that 70% of *de novo* MECP2 mutations were of paternal origin, probably due to the higher levels of methylation and greater number of mitotic divisions in the male germline (Girard et al. 2001).

It is unclear whether the different types of *MECP2* mutations account for the variability in the RTT phenotype. Genotype-phenotype correlation studies are complicated by the fact that the X-inactivation pattern affects the disease outcome in females. In the rare affected males the correlation between mutation and phenotype is clearer. For example, boys that carry the same mutation as their sisters with classic RTT suffer neonatal encephalopathy and die within 1 – 2 years of birth (see (Shahbazian and Zoghbi 2002) for review). On the other hand, *MECP2* mutations, such as late truncating mutations and some missense mutations that lead conservative amino acids exchanges, cause a much milder phenotype. Boys who carry these mutations survive, but they develop severe mental retardation and motor dysfunction. In cases, where males carry a RTT-causing *MECP2* mutation and develop symptoms of RTT, they usually also carry a karyotypic abnormality such as Klinefelter (XXY) or are somatic mosaics, and only a fraction of cells express the mutant protein (Shahbazian and Zoghbi 2002). However, the same mutation can cause distinct symptoms in different families (Cohen et al. 2002) (Klauck et al. 2002), suggesting that other factors such as genetic modifiers also play a role in determining the disease outcome.

Mouse models for RTT

Tools for studying the *in vivo* role of MeCP2 have been produced by mutating *Mecp2* in mice (Chen et al. 2001; Guy et al. 2001; Shahbazian et al. 2002a). Male mice that carry null mutations in *Mecp2* show a phenotype similar to the human condition, including delayed onset of symptoms at 4 to 6 weeks of age followed by death at around 10 weeks, small brain size due to smaller, more densely packed neurons, reduced dendritic arborization, tremors, inertia, abnormal gait and arrhythmic breathing (Chen et al. 2001; Guy et al. 2001). The deletion of *Mecp2* in developing neurons results in the same phenotype, emphasizing the neurological basis of RTT (Chen et al. 2001; Guy et al. 2001). In addition, deletion of *Mecp2* solely in postmitotic neurons results in a phenotype that is later and milder than that of the null allele suggesting MeCP2 is important for postmitotic neurons in particular (Chen et al. 2001). Heterozygous female mice, which represent the genetic model for RTT, acquire symptoms between 4 and 12 month of age. In another study, an *Mecp2* mutation removes the C-terminal third of the protein leaving the MBD and TRD intact (Shahbazian et al. 2002a). The engineered stop codon is located between a truncating mutation that causes severe encephalopathy and death in boys and a milder mutation that allows male survival with severe retardation and ataxia. In mice, this mutation leads to a milder phenotype than the null allele. Male mice show many of the features of RTT, but they survive and showed no change in brain weight (Shahbazian et al. 2002a).

MeCP2 function in RTT

In mouse cells, MeCP2 is detectable throughout the metaphase chromosome arms, but is concentrated in the pericentromeric heterochromatin, which contains highly methylated satellite DNA (Nan et al. 1996). MeCP2 is expressed at higher levels in the brain than in many other tissue (LaSalle et al. 2001; Shahbazian et al. 2002b). Within the brain, MeCP2 is present in neurons but not in glia (Akbarian et al. 2001; LaSalle et al. 2001; Shahbazian et al. 2002b). In the developing cerebral cortex of mouse, rat, human, and non-human primate embryos, the appearance of MeCP2 expression correlates with neuronal maturation, with the earlier-born neurons expressing MeCP2 before the later-born neurons (Akbarian et al. 2001; Jung et al. 2003; Shahbazian et al. 2002b). Studies using the olfactory system to monitor the expression of MeCP2 during development and regeneration suggest that MeCP2 is required for neuronal maturation and plasticity (Cohen et al. 2003). Furthermore, in an experimental system that prevents synapse formation, MeCP2 expression precedes synapse formation, but protein levels increase as synapses form indicating that MeCP2 is involved for maintaining or modulating synapses. (Cohen et al. 2003). Supporting this idea is the finding that the olfactory epithelium of RTT patients contains an excess of immature neurons that have absent or unusual dendrites, dispersed axons, and small cell bodies (Ronnett et al. 2003). Recently it has been suggested that loss of MeCP2 expression also affects neuronal viability. Female mice heterozygous for an *Mecp2* mutation showed a skewed X-inactivation pattern, favoring the expression of the wild type allele (Braunschweig et al. 2004; Young and Zoghbi 2004). Moreover, in primary neuronal cultures, neurons that express the wild type protein survive better than MeCP2-deficient cells (Young and Zoghbi 2004).

Interestingly, in heterozygous mutant mice, the level of MeCP2 expression in wild type cells is significantly decreased suggesting that *Mecp2* mutant neurons affect the development of surrounding neurons in a non-cell-autonomous manner (Braunschweig et al. 2004).

Based on the known biochemical properties of MeCP2, an obvious hypothesis is that the symptoms of RTT are due to a failure to effectively silence methylated genes. However, microarray analysis of brain RNA from wild type and mutant mice did not reveal gross transcriptional dysregulation (Tudor et al. 2002). It is possible that the absence of one methyl-binding protein in the brain may be so well compensated by other related proteins that the effects on gene expression are below detectable levels using current technology. Alternatively, these findings may indicate that MeCP2 has a function in the brain that is as yet unknown. Analysis of the olfactory proteome in mice indicates that loss of MeCP2 results in protein expression changes due to posttranslational modifications in addition to differences in mRNA levels (Matarazzo and Ronnett 2004). Recent evidence suggests that MeCP2 may, in fact, not act as a global repressor but have specific gene targets. In *Xenopus leavis*, reduction of xMeCP2 expression inhibited primary neurogenesis by increasing the expression level of the Notch/Delta pathway target xHairy2a (Stancheva et al. 2003). However, mouse *Mecp2* does not appear to affect neurogenesis, which may reflect a difference in primary neurogenesis in *Xenopus* or redundancy of methyl-CpG-binding proteins in the mouse.

The first mammalian target of MeCP2 function that has been identified is the brain-derived neurotrophic growth factor (BDNF), a secreted protein that is essential for neural plasticity, learning, and memory (Chen et al. 2003; Martinowich et al. 2003).

MeCP2 was found to be bound to methylated CpG sites near promoter III of *BDNF* in resting neurons and is released upon membrane depolarization by a combination of cytosine demethylation and MeCP2 phosphorylation which in turn results in BDNF activation. Taken together, these new findings reveal a surprising new role for MeCP2 in the induction of gene expression in the nervous system as well as in activity-dependent gene regulation.

Studies on X chromosome inactivation and Rett syndrome

In the experiments described in the following chapters we investigate several aspects of X chromosome choice including the role of *Xist*, *Tsix* and asynchronous replication timing prior to X-inactivation as well as the possibility of curing RTT in a mouse model.

In chapter 2, we approach the question how *Xist* deletions affect X chromosome choice. Two *Xist* deletions have been described that lead to skewed X-inactivation, but it is controversial whether skewing is achieved by primary non-random choice or by post-choice selection in which cells that chose to inactivate the mutated X retain two active X chromosomes and die (Marahrens et al. 1998; Penny et al. 1996). We analyzed the process of X-inactivation in the presence of a third *Xist* deletion (Csankovszki et al. 1999) by monitoring GFP expression from an X-linked transgene that is subject to X-inactivation during early embryonic development.

In order to better understand the molecular mechanisms that affect choice we analyzed the role of replication timing (chapter 2). A feature of the Xi is that it becomes late replicating in S-phase (Takagi 1974), a characteristic that it shares other mono-

allelically expressed loci such as imprinted genes, olfactory receptor genes and immunoglobulin genes (Chess et al. 1994; Kitsberg et al. 1993; Mostoslavsky et al. 2001). Recent evidence suggests that asynchronous replication timing may also be involved in determining choice. For instance, at the immunoglobulin κ-light chain locus the allele that was later chosen for recombination replicated prior to the future non-recombined allele prior (Mostoslavsky et al. 2001). We examined ES cell lines that show random or skewed X-inactivation to address the question whether X chromosomes replicate asynchronousely before the onset of X-inactivation, and how the replication timing correlates with X chromsome choice.

One aspect of X chromosome choice is the interplay between the *Xist* RNA and its negative regulator, *Tsix*. The role of *Tsix* in X-inactivation is well established, and deletion of *Tsix* results in non-random inactivation of the mutated X. However, published results at the time this work was carried out had relied on targeted deletions that not only abolished *Tsix* expression but also removed significant portions of genomic sequence such as the *DXPas34* locus that might contain crucial regulatory elements (Lee and Lu 1999; Sado et al. 2001). In chapter 3 we investigate the role of *Tsix* transcription in X-inactivation by modulating *Tsix* transcription with minimal perturbance of the genomic sequence.

In human females, the pattern of X chromosome inactivation can have a big impact on the severity of symptoms in patients with X-linked genetic diseases. The X-linked disease RTT is caused by mutations in the *MECP2* gene (Amir et al. 1999). The neurological phenotype of RTT patients, the developmental expression pattern of MeCP in the brain, the identification of MeCP2 target genes as well as studies in mice carrying a

conditional MeCP2 deletion in the brain all point to a role for MeCP2 in neuronal maturation and plasticity. We tested the hypothesis that the RTT phenotype is exclusively caused by loss of MeCP2 in postmitotic neurons by placing MeCP2 expression under the control of a neuron-specific promoter. The results of transgene expression in *Mecp2* mutant and wild type mice are described in chapter 4.

References

Akbarian, S., R.Z. Chen, J. Gribnau, T.P. Rasmussen, H. Fong, R. Jaenisch, and E.G. Jones. 2001. Expression pattern of the Rett syndrome gene MeCP2 in primate prefrontal cortex. *Neurobiol Dis* **8:** 784-791.

Akhtar, A. 2003. Dosage compensation: an intertwined world of RNA and chromatin remodelling. *Curr Opin Genet Dev* **13**: 161-9

Amir, R.E., I.B. Van den Veyver, M. Wan, C.Q. Tran, U. Francke, and H.Y. Zoghbi. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23:** 185-188.

Armstrong, D., J.K. Dunn, B. Antalffy, and R. Trivedi. 1995. Selective dendritic alterations in the cortex of Rett syndrome. *J Neuropathol Exp Neurol* **54:** 195-201.

Avner, P. and E. Heard. 2001. X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet* **2:** 59-67.

Avner, R., J. Wahrman, C. Richler, N. Ayoub, A. Friedmann, N. Laufer, and S. Mitrani-Rosenbaum. 2000. X-inactivation-specific transcript expression in mouse oocytes and zygotes. *Mol Hum Reprod* **6:** 591-594.

Bachman, K.E., M.R. Rountree, and S.B. Baylin. 2001. Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *J Biol Chem* **276:** 32282-32287.

Barr, M.L.a.C., D. H. 1962. Correlation between sex chromatin and chromosomes. *Acta Cytol.* **6:** 34-45.

Beard, C., E. Li, and R. Jaenisch. 1995. Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev* **9:** 2325-2334.

Bestor, T., A. Laudano, R. Mattaliano, and V. Ingram. 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol* **203:** 971-983.

Bestor, T.H. 1992. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *Embo J* **11:** 2611-2617.

Bird, A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev* **16:** 6-21.

Borsani, G., R. Tonlorenzi, M.C. Simmler, L. Dandolo, D. Arnaud, V. Capra, M. Grompe, A. Pizzuti, D. Muzny, C. Lawrence, and et al. 1991. Characterization of a murine gene expressed from the inactive X chromosome. *Nature* **351:** 325-329.

Bourc'his, D., G.L. Xu, C.S. Lin, B. Bollman, and T.H. Bestor. 2001. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294:** 2536-2539.

Braunschweig, D., T. Simcox, R.C. Samaco, and J.M. LaSalle. 2004. X-Chromosome inactivation ratios affect wild type MeCP2 expression within mosaic Rett syndrome and Mecp2-/+ mouse brain. *Hum Mol Genet* **13:** 1275-1286.

Brockdorff, N. 2002. X-chromosome inactivation: closing in on proteins that bind Xist RNA. *Trends Genet* **18:** 352-358.

Brockdorff, N., A. Ashworth, G.F. Kay, P. Cooper, S. Smith, V.M. McCabe, D.P. Norris, G.D. Penny, D. Patel, and S. Rastan. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351:** 329-331.

Brockdorff, N., A. Ashworth, G.F. Kay, V.M. McCabe, D.P. Norris, P.J. Cooper, S. Swift, and S. Rastan. 1992. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71:** 515-526.

Brown, C.J. and J.M. Greally. 2003. A stain upon the silence: genes escaping X-inactivation. *Trends Genet* **19:** 432-438.

Brown, C.J., B.D. Hendrich, J.L. Rupert, R.G. Lafreniere, Y. Xing, J. Lawrence, and H.F. Willard. 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.

Brown, C.J. and H.F. Willard. 1994. The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature* **368:** 154-156.

Brown, S.D. 1991. XIST and the mapping of the X chromosome inactivation centre. *Bioessays* **13:** 607-612.

Chamberlain, S.J. and C.I. Brannan. 2001. The Prader-Willi syndrome imprinting center activates the paternally expressed murine Ube3a antisense transcript but represses paternal Ube3a. *Genomics* **73:** 316-322.

Chao, W., K.D. Huynh, R.J. Spencer, L.S. Davidow, and J.T. Lee. 2002. CTCF, a candidate trans-acting factor for X-inactivation choice. *Science* **295:** 345-347.

Chaumeil, J., I. Okamoto, M. Guggiari, and E. Heard. 2002. Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells. *Cytogenet Genome Res* **99:** 75-84.

Chen, R.Z., S. Akbarian, M. Tudor, and R. Jaenisch. 2001. Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat Genet* **27:** 327-331.

Chen, W.G., Q. Chang, Y. Lin, A. Meissner, A.E. West, E.C. Griffith, R. Jaenisch, and M.E. Greenberg. 2003. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302:** 885-889.

Chess, A., I. Simon, H. Cedar, and R. Axel. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78:** 823-834.

Clerc, P. and P. Avner. 1998. Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. *Nat Genet* **19:** 249-253.

Cohen, D., G. Lazar, P. Couvert, V. Desportes, D. Lippe, P. Mazet, and D. Heron. 2002. MECP2 mutation in a boy with language disorder and schizophrenia. *Am J Psychiatry* **159:** 148-149.

Cohen, D.R., V. Matarazzo, A.M. Palmer, Y. Tu, O.H. Jeon, J. Pevsner, and G.V. Ronnet. 2003. Expression of MeCP2 in olfactory receptor neurons is

developmentally regulated and occurs before synaptogenesis. *Mol Cell Neuroscience* **22**: 417-429.

Courtier, B., E. Heard, and P. Avner. 1995. Xce haplotypes show modified methylation in a region of the active X chromosome lying 3' to Xist. *Proc Natl Acad Sci U S A* **92:** 3531-3535.

Coy, J.F., Z. Sedlacek, D. Bachner, H. Delius, and A. Poustka. 1999. A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3"-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. *Hum Mol Genet* **8:** 1253-1262.

Cross, S.H., R.R. Meehan, X. Nan, and A. Bird. 1997. A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat Genet* **16:** 256-259.

Csankovszki, G., B. Panning, B. Bates, J.R. Pehrson, and R. Jaenisch. 1999. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X-inactivation. *Nat Genet* **22:** 323-324.

Daniel, J.M. and A.B. Reynolds. 1999. The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. *Mol Cell Biol* **19:** 3614-3623.

Debrand, E., C. Chureau, D. Arnaud, P. Avner, and E. Heard. 1999. Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. *Mol Cell Biol* **19:** 8513-8525.

Fuks, F., W.A. Burgers, N. Godin, M. Kasai, and T. Kouzarides. 2001. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *Embo J* **20:** 2536-2544.

Girard, M., P. Couvert, A. Carrie, M. Tardieu, J. Chelly, C. Beldjord, and T. Bienvenu. 2001. Parental origin of de novo MECP2 mutations in Rett syndrome. *Eur J Hum Genet* **9:** 231-236.

Guy, J., B. Hendrich, M. Holmes, J.E. Martin, and A. Bird. 2001. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet* **27:** 322-326.

Hagberg, B., J. Aicardi, K. Dias, and O. Ramos. 1983. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* **14:** 471-479.

Hagberg, B., F. Goutieres, F. Hanefeld, A. Rett, and J. Wilson. 1985. Rett syndrome: criteria for inclusion and exclusion. *Brain Dev* **7:** 372-373.

Hakoda, M., Y. Hirai, M. Akiyama, H. Yamanaka, C. Terai, N. Kamatani, and S. Kashiwazaki. 1995. Selection against blood cells deficient in hypoxanthine phosphoribosyltransferase (HPRT) in Lesch-Nyhan heterozygotes occurs at the level of multipotent stem cells. *Hum Genet* **96:** 674-680.

Hansen, R.S., C. Wijmenga, P. Luo, A.M. Stanek, T.K. Canfield, C.M. Weemaes, and S.M. Gartler. 1999. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci U S A* **96:** 14412-14417.

Hata, K., M. Okano, H. Lei, and E. Li. 2002. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129:** 1983-1993.

Heard, E., C. Rougeulle, D. Arnaud, P. Avner, C.D. Allis, and D.L. Spector. 2001. Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X-inactivation. *Cell* **107:** 727-738.

Hendrich, B. and A. Bird. 1998. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* **18:** 6538-6547.

Herzing, L.B., J.T. Romer, J.M. Horn, and A. Ashworth. 1997. Xist has properties of the X-chromosome inactivation centre. *Nature* **386:** 272-275.

Huynh, K.D. and J.T. Lee. 2003. Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426:** 857-862.

Jackson, J.P., A.M. Lindroth, X. Cao, and S.E. Jacobsen. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416:** 556-560.

Jaenisch, R. and A. Bird. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33 Suppl:** 245-254.

Jones, P.L., G.J. Veenstra, P.A. Wade, D. Vermaak, S.U. Kass, N. Landsberger, J. Strouboulis, and A.P. Wolffe. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* **19:** 187-191.

Jung, B.P., D.G. Jugloff, G. Zhang, R. Logan, S. Brown, and J.H. Eubanks. 2003. The expression of methyl CpG binding factor MeCP2 correlates with cellular

differentiation in the developing rat brain and in cultured cells. *J Neurobiol* **55:** 86-96.

Kaludov, N.K. and A.P. Wolffe. 2000. MeCP2 driven transcriptional repression in vitro: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery. *Nucleic Acids Res* **28:** 1921-1928.

Kay, G.F., G.D. Penny, D. Patel, A. Ashworth, N. Brockdorff, and S. Rastan. 1993. Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* **72:** 171-182.

Keohane, A.M., P. O'Neill L, N.D. Belyaev, J.S. Lavender, and B.M. Turner. 1996. X-Inactivation and histone H4 acetylation in embryonic stem cells. *Dev Biol* **180:** 618-630.

Kitsberg, D., S. Selig, M. Brandeis, I. Simon, I. Keshet, D.J. Driscoll, R.D. Nicholls, and H. Cedar. 1993. Allele-specific replication timing of imprinted gene regions. *Nature* **364:** 459-463.

Klauck, S.M., S. Lindsay, K.S. Beyer, M. Splitt, J. Burn, and A. Poustka. 2002. A mutation hot spot for nonspecific X-linked mental retardation in the MECP2 gene causes the PPM-X syndrome. *Am J Hum Genet* **70:** 1034-1037.

Kokura, K., S.C. Kaul, R. Wadhwa, T. Nomura, M.M. Khan, T. Shinagawa, T. Yasukawa, C. Colmenares, and S. Ishii. 2001. The Ski protein family is required for MeCP2-mediated transcriptional repression. *J Biol Chem* **276:** 34115-34121.

Kriaucionis, S. and A. Bird. 2004. The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res* **32:** 1818-1823.

Kunert, N., J. Marhold, J. Stanke, D. Stach, and F. Lyko. 2003. A Dnmt2-like protein mediates DNA methylation in Drosophila. *Development* **130:** 5083-5090.

LaSalle, J.M., J. Goldstine, D. Balmer, and C.M. Greco. 2001. Quantitative localization of heterogeneous methyl-CpG-binding protein 2 (MeCP2) expression phenotypes in normal and Rett syndrome brain by laser scanning cytometry. *Hum Mol Genet* **10:** 1729-1740.

Lee, J.T. 2000. Disruption of imprinted X-inactivation by parent-of-origin effects at Tsix. *Cell* **103:** 17-27.

Lee, J.T., L.S. Davidow, and D. Warshawsky. 1999. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* **21:** 400-404.

Lee, J.T. and N. Lu. 1999. Targeted mutagenesis of Tsix leads to nonrandom X-inactivation. *Cell* **99:** 47-57.

Lee, S.S., M. Wan, and U. Francke. 2001. Spectrum of MECP2 mutations in Rett syndrome. *Brain Dev* **23 Suppl 1:** S138-143.

Leonhardt, H., A.W. Page, H.U. Weier, and T.H. Bestor. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71:** 865-873.

Lewis, J.D., R.R. Meehan, W.J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein, and A. Bird. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69:** 905-914.

Li, E., T.H. Bestor, and R. Jaenisch. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915-926.

Luikenhuis, S., A. Wutz, and R. Jaenisch. 2001. Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol Cell Biol* **21:** 8512-8520.

Lyon, M.F. 1961. Gene action in the X-chromosome in the mouse (*Mus musculus*). *Nature* **190:** 372-373.

Lyon, M.F. 1962. Sex chromatin and gene action in the mammalian X chromosome. *American Journal of Human Genetics* **14:** 135-148.

Mak, W., T.B. Nesterova, M. de Napoles, R. Appanah, S. Yamanaka, A.P. Otte, and N. Brockdorff. 2004. Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303:** 666-669.

Marahrens, Y., J. Loring, and R. Jaenisch. 1998. Role of the Xist gene in X chromosome choosing. *Cell* **92:** 657-664.

Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch. 1997. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* **11:** 156-166.

Martinowich, K., D. Hattori, H. Wu, S. Fouse, F. He, Y. Hu, G. Fan, and Y.E. Sun. 2003. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302:** 890-893.

Matarazzo, V. and G.V. Ronnett. 2004. Temporal and regional differences in the olfactory proteome as a consequence of MeCP2 deficiency. *Proc Natl Acad Sci U S A* **101:** 7763-7768.

McCarrey, J.R. and D.D. Dilworth. 1992. Expression of Xist in mouse germ cells correlates with X-chromosome inactivation. *Nat Genet* **2:** 200-203.

Meehan, R.R., J.D. Lewis, S. McKay, E.L. Kleiner, and A.P. Bird. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58:** 499-507.

Mermoud, J.E., B. Popova, A.H. Peters, T. Jenuwein, and N. Brockdorff. 2002. Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation. *Curr Biol* **12:** 247-251.

Meyer, B.J. 2000. Sex in the wormcounting and compensating X-chromosome dose. *Trends Genet* **16:** 247-253.

Mise, N., Y. Goto, N. Nakajima, and N. Takagi. 1999. Molecular cloning of antisense transcripts of the mouse Xist gene. *Biochem Biophys Res Commun* **258:** 537-541.

Mnatzakanian, G.N., H. Lohi, I. Munteanu, S.E. Alfred, T. Yamada, P.J. MacLeod, J.R. Jones, S.W. Scherer, N.C. Schanen, M.J. Friez, J.B. Vincent, and B.A. Minassian. 2004. A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat Genet* **36:** 339-341.

Monk, M. and M.I. Harper. 1979. Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. *Nature* **281:** 311-313.

Morey, C., P. Navarro, E. Debrand, P. Avner, C. Rougeulle, and P. Clerc. 2004. The region 3' to Xist mediates X chromosome counting and H3 Lys-4 dimethylation within the Xist gene. *Embo J* **23:** 594-604.

Mostoslavsky, R., N. Singh, T. Tenzen, M. Goldmit, C. Gabay, S. Elizur, P. Qi, B.E. Reubinoff, A. Chess, H. Cedar, and Y. Bergman. 2001. Asynchronous replication and allelic exclusion in the immune system. *Nature* **414:** 221-225.

Nan, X., F.J. Campoy, and A. Bird. 1997. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88:** 471-481.

Nan, X., R.R. Meehan, and A. Bird. 1993. Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res* **21:** 4886-4892.

Nan, X., H.H. Ng, C.A. Johnson, C.D. Laherty, B.M. Turner, R.N. Eisenman, and A. Bird. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393:** 386-389.

Nan, X., P. Tate, E. Li, and A. Bird. 1996. DNA methylation specifies chromosomal localization of MeCP2. *Mol Cell Biol* **16:** 414-421.

Naumova, A.K., R.M. Plenge, L.M. Bird, M. Leppert, K. Morgan, H.F. Willard, and C. Sapienza. 1996. Heritability of X chromosome--inactivation phenotype in a large family. *Am J Hum Genet* **58:** 1111-1119.

Nesterova, T.B., C.M. Johnston, R. Appanah, A.E. Newall, J. Godwin, M. Alexiou, and N. Brockdorff. 2003. Skewing X chromosome choice by modulating sense transcription across the Xist locus. *Genes Dev* **17:** 2177-2190.

Nesterova, T.B., S.Y. Slobodyanyuk, E.A. Elisaphenko, A.I. Shevchenko, C. Johnston, M.E. Pavlova, I.B. Rogozin, N.N. Kolesnikov, N. Brockdorff, and S.M. Zakian. 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 Res* **11:** 833-849.

Neul, J.L. and H.Y. Zoghbi. 2004. Rett syndrome: a prototypical neurodevelopmental disorder. *Neuroscientist* **10:** 118-128.

Newall, A.E., S. Duthie, E. Formstone, T. Nesterova, M. Alexiou, C. Johnston, M.L. Caparros, and N. Brockdorff. 2001. Primary non-random X-inactivation associated with disruption of Xist promoter regulation. *Hum Mol Genet* **10:** 581-589.

Norris, D.P., D. Patel, G.F. Kay, G.D. Penny, N. Brockdorff, S.A. Sheardown, and S. Rastan. 1994. Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell* **77:** 41-51.

Ogawa, Y. and J.T. Lee. 2002. Antisense regulation in X-inactivation and autosomal imprinting. *Cytogenet Genome Res* **99:** 59-65.

Ogawa, Y. and J.T. Lee. 2003. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol Cell* **11:** 731-743.

Okamoto, I., A.P. Otte, C.D. Allis, D. Reinberg, and E. Heard. 2004. Epigenetic dynamics of imprinted X-inactivation during early mouse development. *Science* **303:** 644-649.

Okano, M., D.W. Bell, D.A. Haber, and E. Li. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99:** 247-257.

Okano, M., S. Xie, and E. Li. 1998. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* **26:** 2536-2540.

Panning, B., J. Dausman, and R. Jaenisch. 1997. X chromosome inactivation is mediated by Xist RNA stabilization. *Cell* **90:** 907-916.

Panning, B. and R. Jaenisch. 1996. DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev* **10:** 1991-2002.

Paterno, G.D. and M.W. McBurney. 1985. X chromosome inactivation during induced differentiation of a female mouse embryonal carcinoma cell line. *J Cell Sci* **75:** 149-163.

Penny, G.D., G.F. Kay, S.A. Sheardown, S. Rastan, and N. Brockdorff. 1996. Requirement for Xist in X chromosome inactivation. *Nature* **379:** 131-137.

Plath, K., J. Fang, S.K. Mlynarczyk-Evans, R. Cao, K.A. Worringer, H. Wang, C.C. de la Cruz, A.P. Otte, B. Panning, and Y. Zhang. 2003. Role of histone H3 lysine 27 methylation in X-inactivation. *Science* **300:** 131-135.

Plath, K., S. Mlynarczyk-Evans, D.A. Nusinow, and B. Panning. 2002. Xist RNA and the mechanism of X chromosome inactivation. *Annu Rev Genet* **36:** 233-278.

Plenge, R.M., B.D. Hendrich, C. Schwartz, J.F. Arena, A. Naumova, C. Sapienza, R.M. Winter, and H.F. Willard. 1997. A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. *Nat Genet* **17:** 353-356.

Prissette, M., O. El-Maarri, D. Arnaud, J. Walter, and P. Avner. 2001. Methylation profiles of DXPas34 during the onset of X-inactivation. *Hum Mol Genet* **10:** 31-38.

Prokhortchouk, A., B. Hendrich, H. Jorgensen, A. Ruzov, M. Wilm, G. Georgiev, A. Bird, and E. Prokhortchouk. 2001. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev* **15:** 1613-1618.

Rasmussen, T.P., A.P. Wutz, J.R. Pehrson, and R.R. Jaenisch. 2001. Expression of Xist RNA is sufficient to initiate macrochromatin body formation. *Chromosoma* **110:** 411-420.

Reichwald, K., J. Thiesen, T. Wiehe, J. Weitzel, W.A. Poustka, A. Rosenthal, M. Platzer, W.H. Stratling, and P. Kioschis. 2000. Comparative sequence analysis of the MECP2-locus in human and mouse reveals new transcribed regions. *Mamm Genome* **11:** 182-190.

Rett, A. 1966. On a unusual brain atrophy syndrome in hyperammonemia in childhood. *Wien Med Wochenschr* **116:** 723-726.

Richler, C., H. Soreq, and J. Wahrman. 1992. X-inactivation in mammalian testis is correlated with inactive X-specific transcription. *Nat Genet* **2:** 192-195.

Ronnett, G.V., D. Leopold, X. Cai, K.C. Hoffbuhr, L. Moses, E.P. Hoffman, and S. Naidu. 2003. Olfactory biopsies demonstrate a defect in neuronal development in Rett's syndrome. *Ann Neurol* **54:** 206-218.

Rougeulle, C., C. Cardoso, M. Fontes, L. Colleaux, and M. Lalande. 1998. An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat Genet* **19:** 15-16.

Runte, M., P.M. Kroisel, G. Gillessen-Kaesbach, R. Varon, D. Horn, M.Y. Cohen, J. Wagstaff, B. Horsthemke, and K. Buiting. 2004. SNURF-SNRPN and UBE3A transcript levels in patients with Angelman syndrome. *Hum Genet* **114:** 553-561.

Russell, L.B. 1961. Genetics of mammalian sex chromosomes. *SCIENCE* **133:** 1795-1803.

Russell, L.B. 1963. Mammalian X-chromosome action: inactivation limited in spread and region of origin. *Science* **140:** 976-978.

Sado, T., Z. Wang, H. Sasaki, and E. Li. 2001. Regulation of imprinted X-chromosome inactivation in mice by Tsix. *Development* **128:** 1275-1286.

Salido, E.C., P.H. Yen, T.K. Mohandas, and L.J. Shapiro. 1992. Expression of the X-inactivation-associated gene XIST during spermatogenesis. *Nat Genet* **2:** 196-199.

Schanen, N.C., E.J. Dahle, F. Capozzoli, V.A. Holm, H.Y. Zoghbi, and U. Francke. 1997. A new Rett syndrome family consistent with X-linked inheritance expands the X chromosome exclusion map. *Am J Hum Genet* **61:** 634-641.

Shahbazian, M., J. Young, L. Yuva-Paylor, C. Spencer, B. Antalffy, J. Noebels, D. Armstrong, R. Paylor, and H. Zoghbi. 2002a. Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron* **35:** 243-254.

Shahbazian, M.D., B. Antalffy, D.L. Armstrong, and H.Y. Zoghbi. 2002b. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum Mol Genet* **11:** 115-124.

Shahbazian, M.D., Y. Sun, and H.Y. Zoghbi. 2002c. Balanced X chromosome inactivation patterns in the Rett syndrome brain. *Am J Med Genet* **111:** 164-168.

Shahbazian, M.D. and H.Y. Zoghbi. 2001. Molecular genetics of Rett syndrome and clinical spectrum of MECP2 mutations. *Curr Opin Neurol* **14:** 171-176.

Shahbazian, M.D. and H.Y. Zoghbi. 2002. Rett syndrome and MeCP2: linking epigenetics and neuronal function. *Am J Hum Genet* **71:** 1259-1272.

Sheardown, S.A., S.M. Duthie, C.M. Johnston, A.E. Newall, E.J. Formstone, R.M. Arkell, T.B. Nesterova, G.C. Alghisi, S. Rastan, and N. Brockdorff. 1997. Stabilization of Xist RNA mediates initiation of X chromosome inactivation. *Cell* **91:** 99-107.

Silva, J., W. Mak, I. Zvetkova, R. Appanah, T.B. Nesterova, Z. Webster, A.H. Peters, T. Jenuwein, A.P. Otte, and N. Brockdorff. 2003. Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell* **4:** 481-495.

Simmler, M.C., B.M. Cattanach, C. Rasberry, C. Rougeulle, and P. Avner. 1993. Mapping the murine Xce locus with (CA)n repeats. *Mamm Genome* **4:** 523-530.

Sirianni, N., S. Naidu, J. Pereira, R.F. Pillotto, and E.P. Hoffman. 1998. Rett syndrome: confirmation of X-linked dominant inheritance, and localization of the gene to Xq28. *Am J Hum Genet* **63:** 1552-1558.

Sleutels, F., R. Zwart, and D.P. Barlow. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415:** 810-813.

Stancheva, I., A.L. Collins, I.B. Van den Veyver, H. Zoghbi, and R.R. Meehan. 2003. A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos. *Mol Cell* **12:** 425-435.

Stavropoulos, N., N. Lu, and J.T. Lee. 2001. A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. *Proc Natl Acad Sci U S A* **98:** 10232-10237.

Tada, T., M. Tada, and N. Takagi. 1993. X chromosome retains the memory of its parental origin in murine embryonic stem cells. *Development* **119:** 813-821.

Takagi, N. 1974. Differentiation of X chromosomes in early female mouse embryos. *Exp Cell Res* **86:** 127-135.

Takagi, N. and M. Sasaki. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256:** 640-642.

Tamaru, H. and E.U. Selker. 2001. A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. *Nature* **414:** 277-283.

Tudor, M., S. Akbarian, R.Z. Chen, and R. Jaenisch. 2002. Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. *Proc Natl Acad Sci U S A* **99:** 15536-15541.

Weksberg, R., A.C. Smith, J. Squire, and P. Sadowski. 2003. Beckwith-Wiedemann syndrome demonstrates a role for epigenetic control of normal development. *Hum Mol Genet* **12 Spec No 1:** R61-68.

Wutz, A. and R. Jaenisch. 2000. A shift from reversible to irreversible X-inactivation is triggered during ES cell differentiation. *Mol Cell* **5:** 695-705.

Wutz, A., T.P. Rasmussen, and R. Jaenisch. 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* **30:** 167-174.

Young, J.I. and H.Y. Zoghbi. 2004. X-chromosome inactivation patterns are unbalanced and affect the phenotypic outcome in a mouse model of rett syndrome. *Am J Hum Genet* **74:** 511-520.

Yu, F., J. Thiesen, and W.H. Stratling. 2000. Histone deacetylase-independent transcriptional repression by methyl-CpG-binding protein 2. *Nucleic Acids Res* **28:** 2201-2206.

Zoghbi, H.Y., A.K. Percy, R.J. Schultz, and C. Fill. 1990. Patterns of X chromosome inactivation in the Rett syndrome. *Brain Dev* **12:** 131-135.

Zwart, R., F. Sleutels, A. Wutz, A.H. Schinkel, and D.P. Barlow. 2001. Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. *Genes Dev* **15:** 2361-2366.

Chapter 2

X chromosome choice occurs independently of asynchronous replication timing

Joost Gribnau*[†], Sandra Luikenhuis*[‡§], Konrad Hochedlinger[§], Kim Monkhorst[†], and Rudolf Jaenisch[‡§]

Data presented in this chapter have been submitted to *The Journal of Cell Biology* for publication.

Respective contributions: Joost Gribnau generated Figures 2, 3, and 4. Konrad Hochedlinger was involved in designing the replication experiments. Kim Monkhorst generated some of the FISH data.

* These authors contributed equally to this work

[‡]Massachusetts Institute of Technology,

[§] Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142

[†] Department of Cell Biology, Erasmus MC, 3015 GE Rotterdam, The Netherlands

Abstract

In mammals, dosage compensation is achieved by X chromosome inactivation in female cells. *Xist* is required and sufficient for X-inactivation, and *Xist* gene deletions result in complete skewed X-inactivation. In this study, we analyzed skewing of X-inactivation in mice with an *Xist* deletion encompassing 5 kb of promoter sequence through exon 3. We found that this mutation results in primary non-random X-inactivation, in which the wild type X chromosome is always chosen for inactivation. To understand the molecular mechanisms that affect choice, we analyzed the role of replication timing in X chromosome choice. We found that the two *Xist* alleles and most likely both entire X chromosomes replicate asynchronously prior to the start of X-inactivation. However, analysis of replication timing in cell lines with skewed X-inactivation showed no preference for one of the two *Xist* alleles to replicate early in S-phase before the onset of X-inactivation, indicating that asynchronous replication timing does not play a role in skewing of X-inactivation.

Introduction

In mammalian cells, dosage compensation of X-linked genes is achieved by X chromosome inactivation in female cells. X-inactivation is a complex process that is regulated by integrating a number of mechanisms. Genetic studies have revealed that X chromosome silencing is initiated from one location on the X chromosome, the X-inactivation center (*Xic*) (Rastan 1983). The *Xist* gene, which lies within the *Xic,* encodes a large untranslated RNA that is required and sufficient for X-inactivation (Penny et al.

1996). At the onset of X-inactivation, one of the X chromosomes is chosen to remain active, *Xist* expression is stabilized on the future inactive X (Xi), and *Xist* RNA accumulates *in cis* (Brockdorff et al. 1991; Brown et al. 1991). The *Xist* RNA associates with the X chromosome and recruits different chromatin modifying complexes, eventually rendering the chromosome transcriptionally inactive (Silva et al. 2003).

The number of X chromosomes that will be inactivated is determined relative to the ploidy of the cell (Rastan 1983; Rastan and Robertson 1985). X-inactivation is initiated when the number of X chromosomes exceeds one in a diploid nucleus. Each cell then makes the epigenetic choice to keep one X chromosome active (Xa) and to inactivate all supernumerary X chromosomes. These observations have led to a model that suggests the presence of a blocking factor in limited quantities that can protect only one X chromosome from inactivation per diploid set of chromosomes (Lyon 1996).

X chromosome choice is random in the embryonic lineage of the mouse. In contrast, X-inactivation in metatherian mammals such as kangaroos (Cooper et al. 1971) and in the extraembryonic tissues of some eutherian mammals, including mice (Takagi and Sasaki 1975), is imprinted and the paternal X chromosome always undergoes inactivation.

Deletion of the *Xist* gene also leads to non-random X-inactivation in embryonic tissues in mice (Penny et al. 1996; Marahrens et al. 1997; Csankovszki et al. 1999). However, it is controversial whether this non-random X-activation is achieved by primary non-random X-inactivation or by random inactivation followed by post-choice selection (also referred to as secondary non-random X-inactivation). Deletion of the *Xist* promoter and exon1 ($Xist^{\Delta prom\text{-}1}$) resulted in post-choice selection in which cells that

choose to inactivate the mutated X retain two active X chromosomes and die (Penny et al. 1996). In contrast, a deletion extending from part of exon 1 to exon 5 ($Xist^{\Delta 1\text{-}5}$) leaves the *Xist* promoter intact and results in primary non-random X-inactivation where the wild type allele is always chosen to be inactivated (Marahrens et al. 1997; Nesterova et al. 2003).

Considerable progress has been made in determining the different chromatin modifications and proteins involved in the establishment and maintenance of the Xi. However, much less is known about the choice process. Genetic studies have revealed that choice is influenced by the X controlling element (*Xce*) such that X-inactivation is skewed towards the chromosome possessing the weaker *Xce* allele (Simmler et al. 1993). At least four *Xce* alleles exist in mice: Xce^a, Xce^b, Xce^c, and Xce^d, with Xce^a being the weakest and Xce^d being the strongest allele. Differential replication of the *Xist* alleles prior to X-inactivation could be another mechanism to influence choice, and possibly blocking factor binding. A characteristic of imprinted as well as random X-inactivation is that the Xi replicates late in S-phase (Takagi 1974). Asynchronous replication timing is not unique to the X chromosome in somatic cells, but is also found at mono-allelically expressed autosomal gene loci, including imprinted genes, olfactory receptor genes and immunoglobulin gene loci (Kitsberg et al. 1993; Chess et al. 1994; Mostoslavsky et al. 2001). Recently, asynchronous replication timing of the immunoglobulin k-light chain locus was found to correlate with VJC rearrangement, with the functional recombined allele being replicated prior to the unrecombined allele (Mostoslavsky et al. 2001). At this locus asynchronous replication timing was present before one allele was chosen for recombination and may therefore play a role in determining choice (Mostoslavsky et al.

2001). In the case of X-inactivation, however, it is unknown whether asynchronous replication timing is present prior to X-inactivation, and whether it has an effect on determining choice.

In this study, we analyzed the nature of non-random X-inactivation in mice carrying different *Xist* deletions by monitoring X-linked GFP expression during early embryonic development. We then examined the role of replication timing in X chromosome choice by analyzing replication timing of the X chromosome before the onset of X-inactivation in embryonic stem (ES) cell lines that show skewed X-inactivation.

Materials and methods

Analysis of X-inactivation in embryos

The appropriate genotypes were obtained by crossing $Xist^{1lox/+}$, $Xist^{\Delta 1\text{-}5/+}$ or wild type females to males homozygous for the X-linked GFP (Hadjantonakis et al. 1998). Pregnant females were sacrificed at the appropriate times and embryos collected. The embryo was separated into embryonic and extraembryonic tissues. Extraembryonic tissues were used for PCR genotyping by incubating them in 20μl 1 x PCR buffer (GIBCO) supplemented with 2 mM $MgCl_2$ and 1 mg/ml proteinase K for 1 h at 50°C, followed by 10 min at 95°C. Standard PCR was carried out by using 9 μl of the above lysate in a 20 μl reaction (30 cycles, with an annealing temperature of 52°C). For the $Xist^{1lox}$ allele the primers Xint 3R (5'-CAC TGG CAA GGT GAA TAG CA-3'), XpromL (5'-TTT CTG GTC TTT GAG GGC AC-3') and 5' Lox R (5-ACC CTT GCC TTT TCC

ATT TT–3') were used which gave a 427 bp band for the wild type allele a 513 bp band for the 1 lox allele. For the *Xist*$^{\Delta 1\text{-}5}$ allele we used the primers Xist KO F (5'-AAC TGA GTG GGT GTT CAG GG-3'), Xist KO R (5'-ACC ACA AAT CAA GGC GAA TC-3') and PGK-Pr1 (5'-GGG AAC TTC CTG ACT AGG GG-3'), which gave a 200 bp band for the wild type allele and a 260 bp band for theΔ1-5 allele.

Embryonic tissues were washed twice in HEPES followed by trypsinization for 5 min and dissociation by pipetting. The trypsinization reaction was stopped by the addition of a small volume of DME/10% FCS. Cells were then pelleted and resuspended in PBS/2% FBS supplemented with a final concentration of 1 mg/ml propidium iodine and analyzed by FACS.

Cell culture

Polymorphic *Mus musculus/Mus castaneus* F1-2-1 ES cells were grown on mouse embryonic fibroblasts (MEFs) in DMEM (GIBCO), 15% fetal calf serum (FCS, Hyclone), and 1000 U LIF/ml (Marahrens et al. 1997). ES cells were differentiated for 5 days in ES media without LIF and MEFs in the presence of 100 nM all-trans-retinoic acid on gelatinized coverslips and the medium was changed every day.

DNA fluorescence in situ hybridization

DNA FISH was performed as in Selig et al. (Selig et al. 1992) with minor modifications. Briefly medium of exponentially growing cells was supplemented with 10 mM BrdU and incubated for 45 min. Cells were trypsinized, washed with HEPES and resuspended in 0.75 M KCl. After trypsinization ES cells were incubated for 10 min on

ice, all other cell types for 10 min at 37°C. Cells were fixed for 10 min in ice cold methanol:acetic acid solution (3:1 ratio), washed three times with methanol:acetic acid, and stored at 4°C or spotted onto polylysine coated slides.

Slides were treated with RNase (100 mg/ml, 2 x SSC) for 30 min at 37°C and washed 3 x 5 min in 2 x SSC and dehydrated in 70%, 90%, and 100% ethanol. Target sequences were denatured by applying 100 ml of 70% formamide, 10 mM phosphate buffer in 2 x SSC under a coverslip and incubated for 3 min on a hotplate (75°C). After removal of the coverslip, slides were washed in 2 x SSC (5 min; 4°C), in 70% ethanol (5 min; -20°C), and in 90% and 100% ethanol for 3 min. Meanwhile, nick translated BAC and cosmid probe sequences were dissolved in a hybridization mixture that contained 50% formamide, 2 x SSC, 50 mM phosphate buffer pH 7.0, 10 mg/ml salmon sperm DNA, 10% dextrane sulfate and 100 ng/ml mouse Cot DNA to a final concentration of 2 ng/ml. The probe mix was denatured for 5 min, prehybridized for a minimum of 45 min, and then applied onto the slide. Slides were incubated overnight in a humidified chamber at 37°C.

BAC probes covering the *Xic* have been sequenced and described before (Chureau et al. 2002). The BAC probes for *Scurfy* and *Irak1* have been described and were BAC 196K10 (Brunkow et al. 2001) and BAC 228O4 (Reichwald et al. 2000), respectively, and were acquired from Research Genetics Inc. The *α-globin* and L23mrp cosmid probes have been described before (Gribnau et al. 2003). The *Mecp2* probe was a 11 kb *Kpn*I fragment covering part of the *Mecp2* gene. All probes were digoxygenin labeled by nick translation (Roche) or biotin labeled with a random prime labeling kit

(Invitrogen), purified over G50 colomns, precipitated and resuspended in hybridization mix.

After hybridization, slides were washed in 2 x SSC (5 min; 37°C), in 50% formamide, 2 x SSC (3x10min; 37°C) and in 0.1 M Tris, 0.15 M NaCl, 0.05% Tween 20 (2 x 5 min; RT), then incubated in 2 mg/ml BSA in 0.1 M Tris, 0.15 M NaCl in a humidified chamber (30 min; RT). Detection was carried out during subsequent incubation steps with anti-digoxygenin (Boehringer), anti-sheep (FITC, Jackson Labs, only when necessary), anti-BrdU (DAKO), or anti-mouse antibodies (Rhodamine Red, Jackson Labs) in 0.1 M Tris, 0.15 M NaCl (30 min; RT). For double label DNA FISH, biotin was detected with an anti-biotin (Roche) and an anti-mouse antibody (Rhodamine Red, Jackson Labs), and BrdU was detected with an anti-BrdU (Abcam) and an anti-rat antibody (AMCA, Jackson Labs). Slides were washed twice between each detection step with 0.1 M Tris, 0.15 M NaCl, 0.05% Tween 20, mounted with Vectashield (Vector Labs), and stored at 4°C. Fluorescence was detected by epifluorescence/CCD. Between 100 and 150 cells were counted per cell line. For replication timing coordination studies, more than 20 BrdU positive cells with two SD signals were counted per cell line. Images were acquired using a Nikon E800 microscope equipped with a 100x DIC H oil immersion lens with 1.4 n/a. The camera was a Princeton Instruments, inc., model RTE/CCD 1317=k/2 with a Kodak KAF-1400 1317x1035 chip. As acquisition software we used Openlab 2.2 (Improvision).

RNA-DNA FISH analysis

RNA FISH analysis was performed as described (Panning and Jaenisch 1996). Differentiated ES cells were grown on coverslips, extracted with cytoskeletal buffer and fixed in 4% paraformaldehyde in PBS. The *Xist* probe was a cDNA sequence (Wutz and Jaenisch 2000), which was digoxygenin labeled by nick translation (Roche). After overnight hybridization slides were washed in 2 x SSC (5 min; 37°C), in 50% formamide, 2 x SSC (3 x 10min; 37°C), and fixed for 15 min in 4% paraformaldehyde/PBS at RT. Slides were washed twice with PBS, dehydrated, and denatured. DNA FISH was as described in the previous paragraph. The *tet* operator probe was a 500 bp fragment containing 7 direct *tet* repeats which were biotin labeled with a random prime labeling kit (Invitrogen). Image acquisition was performed as described above.

RT-PCR analysis

RNA was isolated with Trizol (Invitrogen), and 5 mg of RNA was DNase treated and reverse transcribed with Superscript II (Invitrogen). *Xist* RNA was amplified with primers Xist-forward (5'-TTCCCATGTTTCTCCTGCAT-3') and Xist-reverse (5'-GGAGACATGCAAAGGAAGGA-3'). These primers amplify a length polymorphism in exon 7 of *Xist* and amplification results in a 790 bp *M. musculus*-specific product and a 820 bp *M. castaneus*-specific product, which were resolved on a 1.5% agarose gel.

Results

Primary non-random X-inactivation in cells with an *Xist* deletion

At the onset of X-inactivation, one X chromosome is chosen for inactivation. Loss of *Xist* RNA expression leads to non-random choice which results in the inactivation of the wild type X chromosome (Penny et al. 1996; Marahrens et al. 1997; Csankovszki et al. 1999). Two different *Xist* deletions, covering either exon 1 through to exon 5 leaving the *Xist* promoter intact ($Xist^{\Delta 1\text{-}5}$) or deleting the *Xist* promoter and most of exon 1 ($Xist^{\Delta prom\text{-}1}$) have separated primary non-random X-inactivation from post-choice selection (Figure 1A). These results suggest that a region between exon 1 and 5 or the presence of residual transcription are required for choice.

To further define the sequence important for choice, we analyzed X-inactivation *in vivo* in embryos heterozygous for the previously characterized $Xist^{\Delta 1\text{-}5}$ allele as well as the $Xist^{1lox}$ allele, which deletes 18 kb of the *Xist* locus, including the promoter region, and extends into intron 3 (Csankovszki et al. 1999) (Figure 1A). The wild type X in these mice was marked with an X-linked GFP transgene that is subject to X-inactivation (Hadjantonakis et al. 1998). We used fluorescence activated cell sorting (FACS) to monitor X-inactivation as detected by a change in the number of cells that express GFP. In the embryo proper, random X-inactivation begins at around 5.5 days past coitum (dpc) and is complete by 7.5 dpc. Cells that are partially disomic for X-linked genes do not die rapidly, but cell death continues over days until it is complete by 10 dpc (Takagi and Abe 1990). We isolated wild type and mutant embryos between 7.0 and 7.5 dpc, removed extraembryonic tissues for genotyping, and dissociated the embryo proper for analysis. In wild type embryos, starting at 7.0 dpc, we observed that

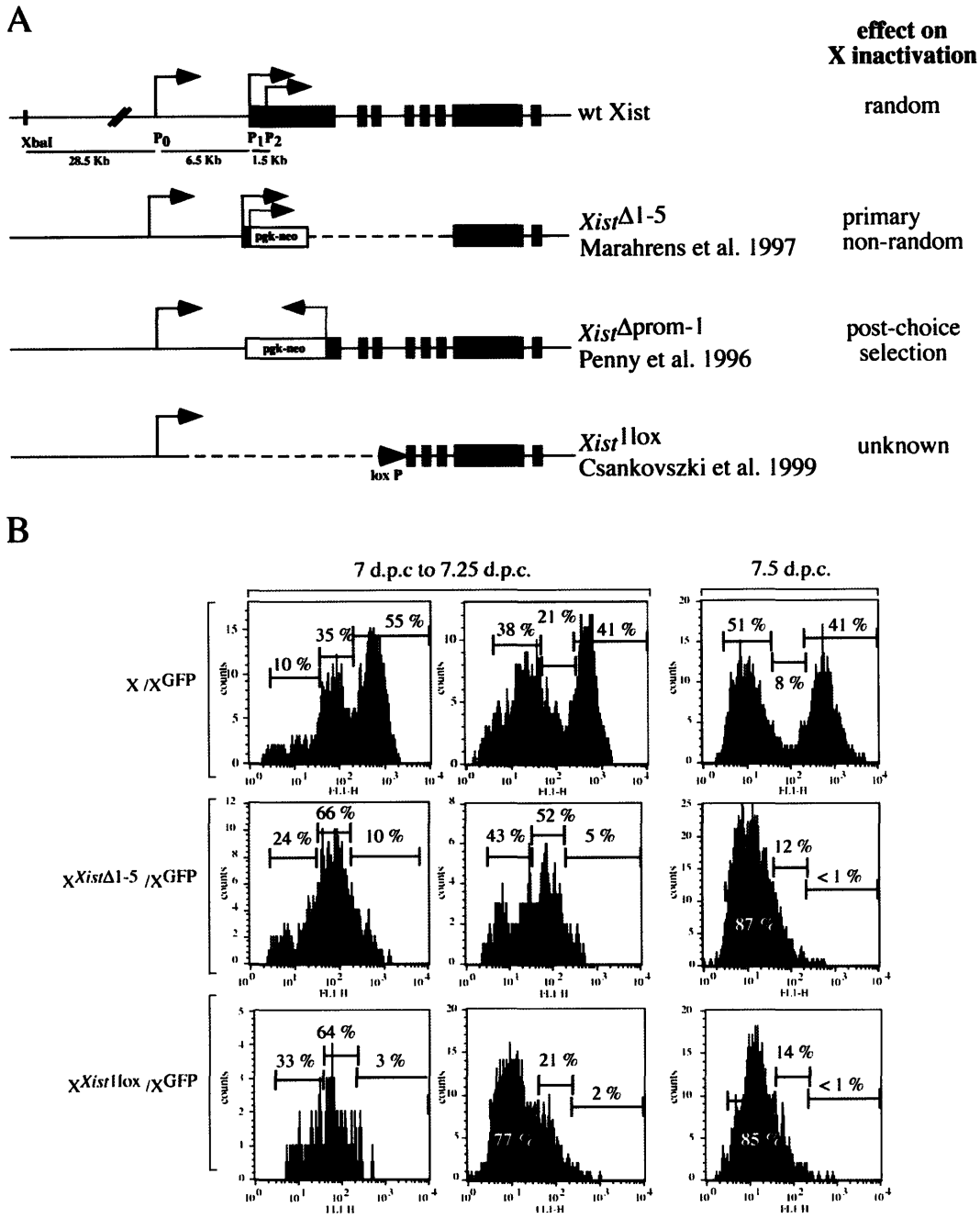


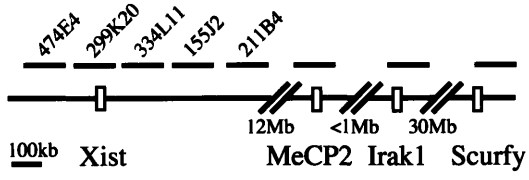
Figure 1. Primary non-random X-inactivation in the presence of different *Xist* deletions. **(A)** Maps of the different *Xist* deletions, and the effect of the deletion on X chromosome choice. **(B)** The absence of a functional *Xist* allele leads to primary non-random X chromosome choice. Top row: FACS analysis of cells isolated from wild type embryos heterozygous for X-linked GFP at 7, 7.25 and 7.5 dpc. Middle row: FACS analysis of cells isolated from $X^{Xist\Delta 1\text{-}5}/X^{GFP}$ embryos at 7, 7.25 and 7.5 dpc. Bottom row: FACS analysis with cells isolated from $X^{Xist\,1lox}/X^{GFP}$ embryos at 7, 7.25 and 7.5 dpc.

around 41% of cells stayed GFP positive. The remaining cells gradually lost GFP fluorescence consistent with random X-inactivation that was complete by 7.5 dpc (Figure 1B, top panel). Embryos heterozygous for either *Xist* deletion invariably chose the wild type chromosome for inactivation, which resulted in the loss of GFP expression in all cells over time (Figure 1B, middle and bottom panels). Heterozygous $Xist^{\Delta 1\text{-}5}$ embryos have been shown previously to undergo primary non-random X-inactivation (Marahrens et al. 1997). In agreement with this finding, we observed that during X-inactivation all cells synchronously became GFP negative, which suggests that in every cell the wild type chromosome was chosen to be inactivated (Figure 1B, middle panel). In $X^{Xist1lox}X^{GFP}$ embryos the dynamics of X-inactivation was indistinguishable from the $Xist^{\Delta 1\text{-}5}$ deletion (Figure 1B, bottom row). These results indicate that both *Xist* deletions cause primary non-random X-inactivation.

Asynchronous replication of X-linked genes prior to X-inactivation

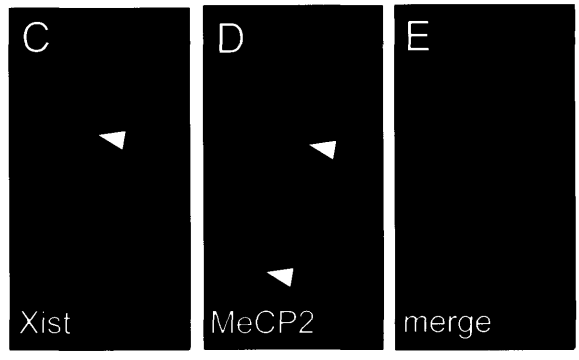
A marker associated with X-inactivation is late replication of the Xi (Takagi 1974). It is unknown, however, whether the X chromosomes also replicate asynchronously prior to X-inactivation. To determine whether X-linked gene loci replicate asynchronously before the onset of X-inactivation, we tested replication timing of several loci along the X chromosome in undifferentiated ES cells which have two transcriptionally active X chromosomes (Figure 2A). Undifferentiated ES cells were BrdU pulse labeled and methanol/acetic acid fixed. BrdU was detected in conjunction with DNA fluorescent in situ hybridization (FISH) using plasmid and BAC probes. In this analysis, three different types of BrdU positive nuclei can readily be distinguished:

A



B

	SS	SD	DD
Undifferentiated ES cells:			
Xist	22%	39%	39%
Xist (Parth.)	31%	41%	28%
BAC474E4	20%	43%	37%
BAC155J2	15%	38%	47%
BAC334L11	11%	41%	48%
BAC211B4	10%	43%	47%
Mecp2	22%	33%	44%
Irak	17%	41%	42%
Scurfy	38%	39%	23%
α globin	20%	17%	63%
L23mrp	37%	13%	50%



F

	coordinated	non-coordinated
Undifferentiated ES cells:		
Xist-Irak	77%	23%
Xist-MeCP2	83%	17%
Xist-BAC474E4	86%	14%
Xist-BAC211B4	93%	7%

Figure 2. X-linked genes replicate asynchronously before X-inactivation. **(A)** Map indicating the location of the X chromosome specific BAC and plasmid probes used for the replication timing analysis. **(B)** Replication timing analysis of different X-linked loci in undifferentiated ES cells. **(C, D and E)** Double label DNA FISH detecting *Xist* and *MeCP2*. *Xist* is shown in green **(C)**, *MeCP2* in red **(D)**, and BrdU in blue **(E)**; white triangle = replicated, gray triangle = not replicated. **(F)** The relative amount of nuclei with coordinated versus non-coordinated SD nuclei for different combinations of target sequences.

1) nuclei with two single signals (single-single, SS), which indicates that both loci have not replicated yet; 2) Nuclei with two double signals (double-double, DD), which indicates that both loci have replicated; and 3) nuclei with one single signal and one double signal (single-double, SD), which indicates that only one locus has been replicated and the other one not. For bi-allelically expressed gene loci that replicate synchronously, the relative amount of the SD nuclei is low and varies between 10 and 20%, depending on the cell type and target sequence. Asynchronously replicating gene loci display much higher relative numbers of SD nuclei in the BrdU positive population of cells that ranges from 25 to 50% (Kitsberg et al. 1993; Chess et al. 1994; Mostoslavsky et al. 2001). We used this method to analyze replication timing of the *Xic* locus, including the *Xist* gene, and found that the whole *Xic* region replicated asynchronously prior to X-inactivation. Analysis of replication timing of the *Xist* locus with a cDNA probe and probes covering either exon 1 or 7 gave similar results (data not shown). The high proportion of DD nuclei (39%) relative to SS nuclei (22%) indicates that both *Xist* alleles replicate relatively early in S-phase. Interestingly, we found that different X-linked loci across the X chromosome, including *Mecp2*, *Irak1* and *Scurfy* all exhibit asynchronous replication timing in undifferentiated ES cells (Figure 2B). In comparison, the bi-allelically expressed autosomal loci, *α-globin* and *L23mrp,* showed relatively low numbers of SD nuclei (13-17%) indicating that these loci replicated synchronously in S-phase. In mice X-inactivation is imprinted in extraembryonic tissues but random in the epiblast. To see if asynchronous replication timing in ES cells is due to imprinting, we analyzed replication timing of *Xist* in parthenogenetic ES cells, which have two maternal chromosome sets. We found that replication timing in these cells was also asynchronous (41% SD nuclei

compared to 39% in wild type ES cells) indicating that replication timing of *Xist* is not imprinted in undifferentiated ES cells (*Xist* (parth.), Figure 2B).

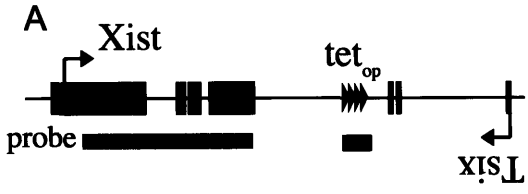
Asynchronous replication timing of non-imprinted autosomal mono-allelically expressed genes is coordinated along the chromosome (Singh et al. 2003). To test whether the same coordination is present on the X chromosome in undifferentiated ES cells, we performed double label FISH for two different loci in conjunction with BrdU staining and analyzed coordination of replication timing in nuclei with SD signals. We tested several probe combinations, including probes for *Xist* and *Mecp2*, which are separated by 12 Megabases, probes for *Xist* and *Irak1*, which are 13 Megabases apart, and probes for *Xist* and two different BACs covering the 5' and 3' end of the *Xic*. (Figure 2A). For all probe combinations, we found that in 77 - 93% of cells the SD signals were detectable on the same allele indicating coordination of replication timing along the chromosome (Figure 2C - F). Analysis of probe combinations covering a greater distance was inconclusive, because the signals were too far apart in the nucleus.

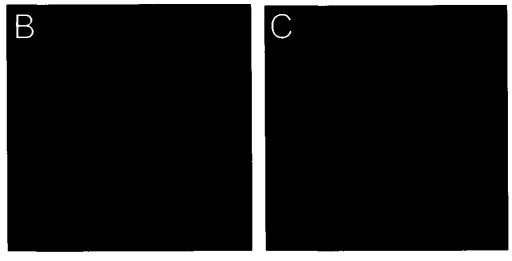
These results indicate that the X chromosomes in female cells replicate asynchronously prior to initiation of X-inactivation, and that replication timing is coordinated along the X chromosome over large distances around the Xist locus.

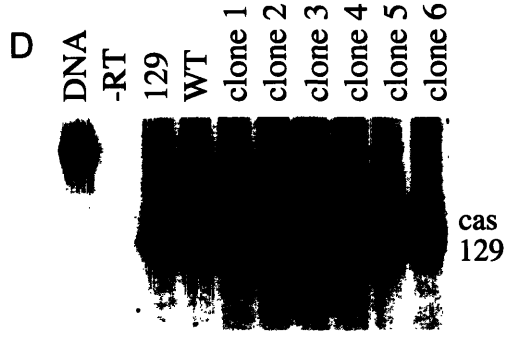
Replication timing and skewed X-inactivation

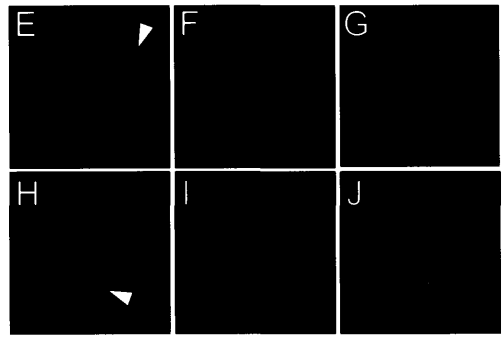
Random X-inactivation is affected by the *Xce* locus and X-inactivation is skewed in cells carrying *Xce* alleles of different strength. We assessed replication timing in F1 *Mus musculus* (129)/*Mus castaneus* (cast) ES cells, which preferentially inactivate the 129 X chromosome that carries the weaker Xce^a allele (relative to the Xce^c allele of the

cast X). To be able to distinguish the two alleles by DNA FISH analysis we introduced 56 *tet* operator repeats downstream of the 129 *Xist* locus and removed the neomycin resistance cassette (Figure 3A - C). We picked six independent subclones and induced X-inactivation by differentiating the ES cells with retinoic acid (RA). RT-PCR analysis with primers detecting a length polymorphism in the cast *Xist* gene showed the expected skewing towards inactivation of the 129 X chromosome after 5 days of differentiation (Figure 3D). Skewing of X-inactivation in the subclones was comparable to skewing found in wild type control samples indicating that the *tet* operator sequences did not affect X-inactivation. In addition, we analyzed skewing of X-inactivation in individual cells by RNA-DNA FISH in two independent subclones after 5 days of differentiation. *Xist* RNA was detected with a probe consisting of the complete cDNA sequence, followed by a brief fixation step and DNA-FISH analysis with a probe specific for the *tet* operator DNA sequences (Figure 3E - K). We determined the proportion of differentiated ES cells that contained an *Xist* signal that colocalized with the *tet* operator signal relative to the number of ES cells, in which the *Xist* signal did not colocalize with the *tet* operator signal. This experiment confirmed the RT-PCR results and showed the inactivation of the 129 X chromosome in 74 - 78% of cells. We next analyzed replication timing of the cast and 129 *Xist* alleles before ES cell differentiation in the 6 ES cell sub-clones that were used for RT-PCR analysis. Using double label DNA-FISH, we found no preference for either the 129 or cast allele to replicate first in S-phase (Figure 3L), which indicats that replication timing does not correlate with skewing of X-inactivation. Because all subclones were derived from one founder clone, replication timing of the *Xist* locus appeared to be dynamic and to switch between alleles in undifferentiated ES cells.









K

	Xist$_{cloud}$ +tet	Xist$_{cloud}$
5day differentiated ES cells:		
subcl. 1	78%	22%
subcl. 2	74%	26%

L

	129 early	Cast early
Undifferentiated *ES cells:*		
subcl. 1	55%	45%
subcl. 2	41%	59%
subcl. 3	50%	50%
subcl. 4	50%	50%
subcl. 5	60%	40%
subcl. 6	55%	45%

Figure 3. Replication timing of the *Xist* locus in ES cells with different *Xce* alleles.
(A) Schematic representation of the integration site of the *tet* operator repeats and the location of the DNA FISH probe used to distinguish the two *Xist* alleles. **(B, C)** Double label DNA FISH using and *Xist* (FITC, green) and *tet* probes (rhodamine red) shows that the *tet* signal colocalizes only with one of the two *Xist* signals. The 129 allele of the polymorphic ES cell line was targeted with the *tet* operator repeat sequence. **(D)** Individual subclones of the *tet* operator targeted ES cell line were differentiated for 5 days. Cast and 129 *Xist* transcripts were detected by RT-PCR and distinguished using a length polymorphism in exon 7.
(E - J) Combined RNA-DNA FISH on two different subclones that were differentiated for 5 days, using two probes that detect the *Xist* RNA (FITC, white triangle), or *tet* DNA (rhodamine red, gey triangle), respectively, revealed cells with an inactivated cast X **(E. F, and G)** or an inactivated 129 X **(H, I, and J)**. DAPI staining is shown in blue.
(K) Quantification of the RNA-DNA FISH experiment using two different subclones.
(L) Replication timing analysis of individual undifferentiated ES cell subclones.

In contrast, replication timing of X-linked genes is highly stable in cells that have undergone X-inactivation (Hansen et al. 1996; Xiong et al. 1998).

Replication timing and complete primary non-random X-inactivation.

We have shown that replication timing before X-inactivation does not correlate with skewing of X-inactivation. We next investigated, whether replication is correlated with complete primary non-random choice. We generated several *Xist*$^{1lox/+}$ heterozygous ES cell lines and analyzed replication timing of the wild type and mutant alleles in undifferentiated ES cells before the onset of X-inactivation using two different probes. One probe covered exon 1 and, therefore, only detected the wild type *Xist* allele; a second probe covered exon 7 and detected both the wild type and mutant alleles (Figure 4A). Double label DNA FISH in conjunction with BrdU staining showed no preference for either the wild type or mutant allele to replicate first in S-phase, which suggests that asynchronous replication timing has no causal relationship with primary non-random choice (Figure 4B - E).

Next, we tested whether the *Xist* gene is required for chromosome-wide coordination of replication timing. We analyzed coordination of replication timing of *Xist* and *Irak1* as well as *Xist* and *Mecp2* in *Xist*$^{1lox/+}$ ES cells and found coordinated SD nuclei in around 80% of the cells, which is similar to our results with wild type ES cells (Figure 4F,G and H). These results indicate that the *Xist* gene is not required for the coordination of replication timing along the X chromosome.

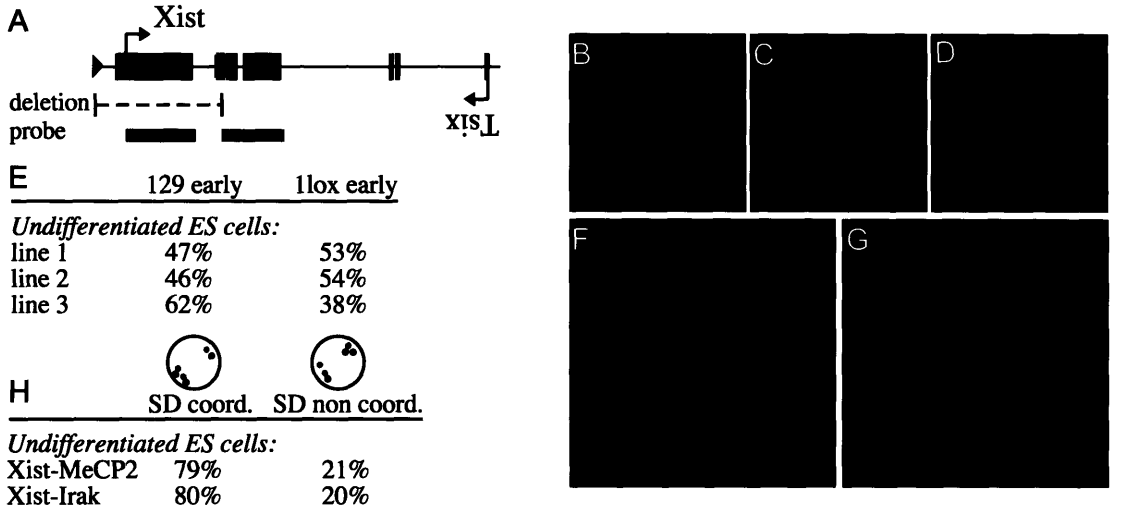


Figure 4. Replication timing of the *Xist* locus in conditional *Xist* knockout ES cells. **(A)** Map of the *Xist* gene and location of the probes used for DNA FISH analysis. **(B, C, D)** BrdU detection in combination with double label DNA FISH using an exon 1 probe (rhodamine red) and an exon 7 probe (FITC) reveals nuclei with the mutant allele being replicated before **(B)** or after **(C)** the wild type allele. **(D)** Control cells show two colocalizing signals. **(E)** Replication timing analysis with three independent $Xist^{1lox/+}$ undifferentiated ES cell lines. **(F, G)** BrdU detection and double label DNA FISH detecting the *Xist* (FITC) and *Mecp2* (rhodamine red) loci reveals cells that have coordinated replication timing **(F)** and non-coordinated replication timing **(G)**. **(H)** Relative amount of nuclei with coordinated replication timing versus non-coordinated replication timing.

Discussion

In this study, we investigated several aspects of the X-inactivation choice process. We analyzed the effect of an *Xist* deletion that spans from 5 kb upstream of the *Xist* promoter to intron 3 on choice and found that this mutation results in primary non-random X-inactivation. This result is consistent with primary non-random X-inactivation that has been previously reported for a different deletion that encompasses part of exon1 to exon 5 (Figure 1A, page 63). It therefore appears that an X chromosome can only be chosen for inactivation if *Xist* is intact. In addition, we studied the role of replication timing and found that asynchronous replication timing is present prior to X-inactivation at all tested regions, and most likely the whole X chromosome. After completion of X-inactivation, asynchronous replication timing of the two X chromosomes is stable and propagated through many cell divisions. In contrast, we found that early replication timing switches between alleles before the initiation of X-inactivation. Analysis of asynchronous replication timing in ES cell lines with severe or complete skewing of X-inactivation showed that replication timing does not correlate with choice, which indicates that replication timing does not mediate skewing of the choice process.

Primary non-random X-inactivation or post-choice selection?

Recent findings that describe the effects of targeted deletions within the *Xic* suggest that the choice process is complex and is most likely regulated by a variety of superimposed mechanisms. Targeted deletions of *Xist* coding regions and promoter sequences result in complete skewing of X-inactivation where the mutant X always remains active (Penny et al. 1996; Marahrens et al. 1997; Csankovszki et al. 1999). A

negative regulator of *Xist is Tsix*, a gene that encodes an antisense RNA to *Xist,* which is transcribed from both alleles before the onset of X-inactivation. In contrast to *Xist* mutations, targeted deletions of the *Tsix* promoter or mutations that truncate the *Tsix* transcript result in preferential inactivation of the mutated X chromosome. (Lee and Lu 1999; Luikenhuis et al. 2001; Sado et al. 2001). Similarly, the deletion of a *Tsix* regulatory element, *Xite*, which normally promotes *Tsix* persistence on the Xa, also causes non-random inactivation of the mutated X chromosome (Ogawa and Lee 2003). Finally, deletion of *Xist* upstream regions that reduce or abolish intergenic transcription resemble *Xist* deletions in that these mutated cells preferentially inactivate the wild type allele (Nesterova et al. 2003).

Non-random X-inactivation due to an *Xist* mutation can be achieved by either primary non-random X-inactivation or random choice followed by post-choice selection, Primary non-random X-inactivation leads to the inactivation of the wild type X in a heterozygous embryo and was observed in cells carrying a deletion extending from part of exon 1 through to exon 5, $Xist^{\Delta 1\text{-}5}$ (Marahrens et al. 1997) as well as in the longer deletion ($Xist^{1lox}$) (Csankovszki et al. 1999) described here. In contrast, Penny et al. (Penny et al. 1996) concluded that random X-inactivation was followed by post-choice selection in a cell line carrying a deletion of the transcriptional start site of *Xist* and part of exon 1. This is similar to previous results in cells that inherit both products of the Searle's X-to-autosome translocation (T(X;16) 16H) (McMahon and Monk 1983). In these embryos, random choice results in either the inactivation of the wild type X chromosome, which results in balanced cells, or in the inactivation of the translocation product that bears the *Xic*, which results in unbalanced cells. The cells that inactivate the

translocation product are then progressively lost from the embryo with about 75% being lost between 7 dpc and 8 dpc (McMahon and Monk 1983). Despite this massive cell loss, the embryos develop into healthy, but somewhat smaller, newborn pups (Lyon et al. 1964).

Because the two *Xist* mutations that do affect choice (Marahrens et al. 1997; Csankovszki et al. 1999) encompass all sequences deleted in the Penny et al. mutation (Figure 1A, page 63), it is difficult to define genetic elements that regulate this process. We consider the following possibilities to reconcile these differences: 1) It is possible that incomplete differentiation of the ES cells or the stability of the gene products measured by Penny et al. affected the outcome of the experiment. 2) It also cannot be excluded that the presence of the selectable marker in the antisense orientation and of exons 2 and 3 in the mutant allele somehow abrogated the role of *Xist* in choice.

Replication timing does not correlate with choice.

Using DNA FISH to analyze replication timing of different X-linked loci including *Xist* in ES cells, we found that all loci tested replicate asynchronously prior to X-inactivation, and that replication timing along the entire *Xic* is coordinated. Therefore, the possibility that replication timing of a small region within the *Xic* is not coordinated is unlikely. Our results show that prior to X-inactivation the proportion of asynchronously replicating (SD) X chromosomes is comparable to that of imprinted genes and non-imprinted mono-allelically expressed genes (Simon et al. 1999; Gribnau et al. 2003; Singh et al. 2003). Two differences distinguish asynchronous replication timing before and after X-inactivation. Asynchronous replication timing prior to X-inactivation

switches between alleles and is not affected by the *Xce* allele carried on the respective X. In contrast, in somatic cells, asynchronous replication timing of the Xa and Xi is clonal and stable through many cell divisions (Hansen et al. 1996; Xiong et al. 1998), and is influenced by the *Xce*. Another difference is the time window between the replication of the two alleles. Prior to X-inactivation, we found 40% of the nuclei with an SD signal by FISH. Earlier studies using S-phase fractionation analysis, which measures the DNA content at different stages in S-phase, have shown that this proportion of SD nuclei reflects a difference in replication timing of 1.5 to 2 hours (Simon et al. 1999; Gribnau et al. 2003; Singh et al. 2003). In contrast, after X-inactivation is completed, the time window is significantly larger and extends through almost the entire S-phase (Xiong et al. 1998). We conclude that replication timing prior to X-inactivation does not correlate with skewing of the X chromosome choice and most likely reflects distinct chromatin states of the two X chromosomes.

Asynchronous replication timing, what does it do?

The vast majority of mono-allelically expressed gene loci have been shown to replicate asynchronously in S-phase, which suggests a direct role for replication timing in the choice processes. In this study, we found that asynchronous replication timing is present throughout the X-inactivation process. However, replication timing does not correlate with skewing of X chromosome choice.

Asynchronous replication timing of X-linked genes and other non-imprinted monoallelically expressed genes is random with respect to the parental origin (Mostoslavsky et al. 2001). In contrast, asynchronous replication timing of imprinted

gene loci is parent specific (Simon et al. 1999). Interestingly, loss of imprinting caused by the erasure of methylation marks in the germline or after fertilization does not result in a loss of asynchronous replication timing of imprinted gene loci (Gribnau et al. 2003). In addition, at the imprinted *Igf2/H19* locus, a 3 Megabase inversion, which results in the establishment of a paternally imprinted *Igf/H19* locus in the female germline, does not change the replication timing characteristics of this locus (Cerrato et al. 2003). Therefore, gene expression and asynchronous replication timing appear to be separable mechanisms, which raises questions about the significance of asynchronous replication. We found that asynchronous replication timing prior to X-inactivation is random, switches between alleles, and is independent of the *Xce* and *Xist* mutations. Therefore, we cannot exclude a role for asynchronous replication upstream of choice. It is possible that asynchronous replication timing before X-inactivation reflect epigenetic differences between the two X chromosomes such as transient blocking factor binding, which may be involved in the counting process upstream of choice (Brockdorff 1998).

Asynchronous replication timing could also be the remnant of an ancient imprinting or choice mechanism and could have played a role in setting up and maintaining imprints or determining choice processes for random mono-allelically expressed genes. Over time, different epigenetic mechanisms, like DNA methylation or chromatin modifications, may have taken over the role of replication timing.

Acknowledgements

We thank Kathrin Plath for critically reading the manuscript, Barbara Panning, Andy Chess, Albrecht Sippel and Alex Emsinger for stimulating discussions. This work was conducted using the W. M. Keck biological imaging facility at the Whitehead Institute and was supported by a fellowship from the Human Frontier Science Program to J.G. and by NIH/NCI grant 5-R01-CA87869 to R.J.

References

Brockdorff, N. 1998. The role of Xist in X-inactivation. *Curr Opin Genet Dev* **8:** 328-333.

Brockdorff, N., A. Ashworth, G.F. Kay, P. Cooper, S. Smith, V.M. McCabe, D.P. Norris, G.D. Penny, D. Patel, and S. Rastan. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351:** 329-331.

Brown, C.J., A. Ballabio, J.L. Rupert, R.G. Lafreniere, M. Grompe, R. Tonlorenzi, and H.F. Willard. 1991. A gene from the region of the human X-inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349:** 38-44.

Brunkow, M.E., E.W. Jeffery, K.A. Hjerrild, B. Paeper, L.B. Clark, S.A. Yasayko, J.E. Wilkinson, D. Galas, S.F. Ziegler, and F. Ramsdell. 2001. Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nat Genet* **27:** 68-73.

Cerrato, F., W. Dean, K. Davies, K. Kagotani, K. Mitsuya, K. Okumura, A. Riccio, and W. Reik. 2003. Paternal imprints can be established on the maternal Igf2-H19 locus without altering replication timing of DNA. *Hum Mol Genet* **12:** 3123-3132.

Chess, A., I. Simon, H. Cedar, and R. Axel. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78:** 823-834.

Chureau, C., M. Prissette, A. Bourdet, V. Barbe, L. Cattolico, L. Jones, A. Eggen, P. Avner, and L. Duret. 2002. Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Res* **12:** 894-908.

Cooper, D.W., J.L. VandeBerg, G.B. Sharman, and W.E. Poole. 1971. Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X-inactivation. *Nat New Biol* **230:** 155-157.

Csankovszki, G., B. Panning, B. Bates, J.R. Pehrson, and R. Jaenisch. 1999. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X-inactivation. *Nat Genet* **22:** 323-324.

Gribnau, J., K. Hochedlinger, K. Hata, E. Li, and R. Jaenisch. 2003. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* **17:** 759-773.

Hadjantonakis, A.K., M. Gertsenstein, M. Ikawa, M. Okabe, and A. Nagy. 1998. Generating green fluorescent mice by germline transmission of green fluorescent ES cells. *Mech Dev* **76:** 79-90.

Hansen, R.S., T.K. Canfield, A.D. Fjeld, and S.M. Gartler. 1996. Role of late replication timing in the silencing of X-linked genes. *Hum Mol Genet* **5:** 1345-1353.

Kitsberg, D., S. Selig, M. Brandeis, I. Simon, I. Keshet, D.J. Driscoll, R.D. Nicholls, and H. Cedar. 1993. Allele-specific replication timing of imprinted gene regions. *Nature* **364:** 459-463.

Lee, J.T. and N. Lu. 1999. Targeted mutagenesis of Tsix leads to nonrandom X-inactivation. *Cell* **99:** 47-57.

Luikenhuis, S., A. Wutz, and R. Jaenisch. 2001. Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol Cell Biol* **21:** 8512-8520.

Lyon, M.F. 1996. X-chromosome inactivation. Pinpointing the centre. *Nature* **379:** 116-117.

Lyon, M.F., A.G. Searle, C.E. Ford, and S. Ohno. 1964. A Mouse Translocation Suppressing Sex-Linked Variegation. *Cytogenetics* **15:** 306-323.

Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch. 1997. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* **11:** 156-166.

McMahon, A. and M. Monk. 1983. X-chromosome activity in female mouse embryos heterozygous for Pgk-1 and Searle's translocation, T(X; 16) 16H. *Genet Res* **41:** 69-83.

Mostoslavsky, R., N. Singh, T. Tenzen, M. Goldmit, C. Gabay, S. Elizur, P. Qi, B.E. Reubinoff, A. Chess, H. Cedar, and Y. Bergman. 2001. Asynchronous replication and allelic exclusion in the immune system. *Nature* **414:** 221-225.

Nesterova, T.B., C.M. Johnston, R. Appanah, A.E. Newall, J. Godwin, M. Alexiou, and N. Brockdorff. 2003. Skewing X chromosome choice by modulating sense transcription across the Xist locus. *Genes Dev* **17:** 2177-2190.

Ogawa, Y. and J.T. Lee. 2003. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol Cell* **11:** 731-743.

Panning, B. and R. Jaenisch. 1996. DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev* **10:** 1991-2002.

Penny, G.D., G.F. Kay, S.A. Sheardown, S. Rastan, and N. Brockdorff. 1996. Requirement for Xist in X chromosome inactivation. *Nature* **379:** 131-137.

Rastan, S. 1983. Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. *J Embryol Exp Morphol* **78:** 1-22.

Rastan, S. and E.J. Robertson. 1985. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J Embryol Exp Morphol* **90:** 379-388.

Reichwald, K., J. Thiesen, T. Wiehe, J. Weitzel, W.A. Poustka, A. Rosenthal, M. Platzer, W.H. Stratling, and P. Kioschis. 2000. Comparative sequence analysis of the MECP2-locus in human and mouse reveals new transcribed regions. *Mamm Genome* **11:** 182-190.

Sado, T., Z. Wang, H. Sasaki, and E. Li. 2001. Regulation of imprinted X-chromosome inactivation in mice by Tsix. *Development* **128:** 1275-1286.

Selig, S., K. Okumura, D.C. Ward, and H. Cedar. 1992. Delineation of DNA replication time zones by fluorescence in situ hybridization. *Embo J* **11:** 1217-1225.

Silva, J., W. Mak, I. Zvetkova, R. Appanah, T.B. Nesterova, Z. Webster, A.H. Peters, T. Jenuwein, A.P. Otte, and N. Brockdorff. 2003. Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev Cell* **4:** 481-495.

Simmler, M.C., B.M. Cattanach, C. Rasberry, C. Rougeulle, and P. Avner. 1993. Mapping the murine Xce locus with (CA)n repeats. *Mamm Genome* **4:** 523-530.

Simon, I., T. Tenzen, B.E. Reubinoff, D. Hillman, J.R. McCarrey, and H. Cedar. 1999. Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* **401:** 929-932.

Singh, N., Y. Bergman, H. Cedar, and A. Chess. 2003. Biallelic germline transcription at the kappa immunoglobulin locus. *J Exp Med* **197:** 743-750.

Takagi, N. 1974. Differentiation of X chromosomes in early female mouse embryos. *Exp Cell Res* **86:** 127-135.

Takagi, N. and K. Abe. 1990. Detrimental effects of two active X chromosomes on early mouse development. *Development* **109:** 189-201.

Takagi, N. and M. Sasaki. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256:** 640-642.

Wutz, A. and R. Jaenisch. 2000. A shift from reversible to irreversible X-inactivation is triggered during ES cell differentiation. *Mol Cell* **5:** 695-705.

Xiong, Z., W. Tsark, J. Singer-Sam, and A.D. Riggs. 1998. Differential replication timing of X-linked genes measured by a novel method using single-nucleotide primer extension. *Nucleic Acids Res* **26:** 684-686.

Chapter 3

Antisense transcription through the *Xist* locus mediates *Tsix* function

Sandra Luikenhuis[1], Anton Wutz[2], and Rudolf Jaenisch[1,2,]

Data presented in this chapter have been published in *Molecular and Cellular Biology* (2001) **21**: 8512-8520.

Respective contributions: Anton Wutz introduced the tetracycline inducible system to our lab and generated male and female ES cell lines carrying the *nls-rtTA* cDNA.

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts

[2]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts

Abstract

Expression of the *Xist* gene, a key player in mammalian X-inactivation, has been proposed to be controlled by the antisense *Tsix* transcript. Targeted deletion of the *Tsix* promoter encompassing the *DPXas34* locus leads to nonrandom inactivation of the mutant X, but it remains unresolved whether this phenotype is caused by loss of *Tsix* transcription or by deletion of a crucial DNA element. In this study, we determined the role of *Tsix* transcription in random X-inactivation using mouse embryonic stem cells as a model system. Two approaches were chosen to modulate *Tsix* transcription with minimal disturbance of genomic sequences. First, *Tsix* transcription was functionally inhibited by introducing a transcriptional stop signal into the transcribed region of *Tsix*. In the second approach, an inducible system for *Tsix* expression was created. We found that the truncation of the *Tsix* transcript led to complete nonrandom inactivation of the targeted X chromosome. Induction of *Tsix* transcription during ES cell differentiation, on the other hand, caused the targeted chromosome always to be chosen as the active chromosome. These results for the first time establish a function for antisense transcription in the regulation of X-inactivation.

Introduction

Dosage compensation in mammals is achieved by a unique mechanism that leads to transcriptional silencing of almost all genes present on one of the two X chromosomes in female cells, a process known as X chromosome inactivation (X-inactivation) (Lyon 1961). In undifferentiated cells, all X chromosomes are active and X-inactivation is

initiated at the onset of differentiation both *in vivo* and *in vitro* (Monk and Harper 1979). The number of X chromosomes that are to be inactivated is determined relative to the ploidy of the cell and X-inactivation is initiated only when the number of X chromosomes exceeds one in a diploid nucleus. Each cell then makes the epigenetic choice to keep one X chromosome active and to inactivate all supernumerary X chromosomes (n-1" rule, with "n" being the total number of X chromosomes in the cell, and "n-1" the number of X chromosomes that are being inactivated). One model suggests the presence of a blocking factor in limited quantities that can protect only one X chromosome from inactivation per diploid set of chromosomes (Avner and Heard 2001). X chromosome choice is random in the embryonic lineage of the mouse and other mammals, but is influenced by a number epigenetic and genetic factors. In metatherian mammals such as kangaroos (Cooper et al. 1971), and in the extraembryonic tissues of some eutherian mammals, including mice (Takagi and Sasaki 1975), the paternal X chromosome is imprinted and always undergoes inactivation. In addition, choice in embryonic tissues of mice is influenced by the 'X-controlling element' (*Xce*) such that X-inactivation is skewed towards the chromosome possessing the weaker Xce allele (Migeon 1994). It has been shown that at least four alleles exist in mice: Xce^a, Xce^b, Xce^c, and Xce^d, with Xce^a being the weakest and Xce^d being the strongest allele.

X-autosomal translocations and transgenic studies in mice have defined a master switch required for X-inactivation, the *Xic* (X-inactivation center), a multifunctional locus that carries elements for choosing and silencing (Rastan and Brown 1990; Lee et al. 1996; Herzing et al. 1997; Lee et al. 1999b). One of these elements is the *Xist* gene which produces a spliced and polyadenylated untranslated RNA that coats the inactive X

chromosome (Brockdorff et al. 1992; Brown et al. 1992; Clemson et al. 1996; Memili et al. 2001) . Targeted deletions have demonstrated that *Xist* is necessary for both the initiation of X-inactivation and the silencing of X-linked genes (Penny et al. 1996; Marahrens et al. 1997) as well as for X chromosome choice (Marahrens et al. 1998). *Xist* RNA shows a distinctive expression pattern during differentiation as analyzed by fluorescent *in situ* hybridization (FISH). Prior to the onset of X-inactivation, *Xist* is expressed in an unstable form from every X chromosome in male and female cells. Upon differentiation, *Xist* transcription is downregulated in male cells, and on the future active X in female cells. On the future inactive X, the *Xist* transcript is stabilized and accumulates along the chromosome *in cis* (Panning et al. 1997; Sheardown et al. 1997).

Recently, a new *Xic* element, *Tsix*, has been identified as a potential negative regulator of *Xist* expression. *Tsix* is an antisense gene to *Xist* whose major promoter is located 15 kb downstream of the *Xist* gene. (Lee et al. 1999a). The promoter overlaps with the previously defined *DXPas34* locus (Lee et al. 1999a), a 2.4 kb CpG island that is hypermethylated on the active X chromosome (as opposed to other CpG islands on the active X which remain hypomethylated) (Courtier et al. 1995). The primary *Tsix* transcript completely overlaps *Xist* and has been shown to cover at least 40 kb (Lee et al. 1999a). Recently, *Tsix* has been shown to undergo complex processing, including alternative splicing and polyadenylation (Sado et al. 2001). According to this study the *Tsix* gene contains 4 exons, one of which (exon 4) overlaps with the highly conserved *Xist* A-repeat. However, the functional significance and cellular localization of the spliced transcript still remains to be characterized.

Like *Xist*, *Tsix* shows a characteristic expression pattern during X-inactivation. In undifferentiated ES cells, *Tsix* is transcribed at low levels from all X chromosomes. Following differentiation, *Tsix* transcription is shut off in adult tissues. Interestingly, at the onset of X-inactivation *Tsix* expression becomes confined to one of the two Xs and appears to specifically mark the future active X which lead to the hypothesis that *Tsix* may regulate *Xist* expression during the initiation of X-inactivation (Debrand et al. 1999; Lee et al. 1999a; Mise et al. 1999).

Targeted deletions that encompass the *Tsix* promoter as well as the CpG island were shown to lead to skewed *Xist* expression and nonrandom inactivation of the mutated X chromosome in ES cells. Introduction of the targeted alleles into mice demonstrated the importance of this locus for imprinted X-inactivation in extraembryonic tissues (Lee 2000; Sado et al. 2001). *Tsix* expression in early blastocysts (3.5 days past coitum) is imprinted oppositely of *Xist* such that it is transcribed only from the maternal allele. Expression from both alleles is observed in late blastocysts at the time when the imprint is erased and both Xs are primed for random X-inactivation. Paternal inheritance of the deletion has no effect, but maternal inheritance leads to ectopic expression of maternal *Xist* in males, and biallelic expression from both X chromosomes in female extraembryonic tissues. The majority of the embryos that inherit the deletion from the mother die during development due to placental defects. Lee proposed the presence of an imprinting center (IC) within the CpG-rich domain overlapping the *Tsix* promoter which responds to activating maternal and repressive paternal signaling to facilitate imprinted *Tsix* expression (Lee 2000). Analysis of the *DXPas34* locus and the *Tsix* promoter by bisulphite sequencing revealed that methylation does not represent the imprinting mark

and is also unlikely to play a role in the regulation of *Tsix* transcription from this site (Prissette et al. 2001). However, it cannot be excluded that the imprinting is associated with a methylation-independent mechanisms. Taken together, *in vitro* and *in vivo* data suggest that the *Tsix* locus and/or transcript influence X chromosome choice and imprinting of the X chromosome.

A number of genes have been identified whose imprinted expression is associated with the transcription of non-coding RNAs. For example at the *H19/Igf2* locus, *H19* encodes an untranslated RNA whose expression is imprinted oppositely to *Igf2* (Moore et al. 1997). In many instances these non-coding RNAs are transcribed in the antisense direction, as is it the case for the *Igf2r* gene and its antisense transcript *Air* (Wutz et al. 1997). Antisense transcripts have also been identified at the *UBE3A* locus (Rougeulle et al. 1998), and the *Gnas* locus (Li et al. 2000). However, no evidence for the functionality of the transcripts exists in any of these cases. In fact, the only non-coding transcript that has been analyzed for its functionality, *H19*, has been shown to play no role in the regulation of *Igf2* (Jones et al. 1998; Schmidt et al. 1999).

In the case of *Tsix*, the role of antisense transcription also still remains to be established. Published results relied on targeted deletions that not only abolished *Tsix* transcription but also eliminated significant portions of genomic sequence such as portions of the *DXPas34* locus that might harbor crucial regulatory elements (Lee and Lu 1999; Lee 2000; Sado et al. 2001).

Here, we determined the role of *Tsix* transcription in random X-inactivation using embryonic stem (ES cells) as a model system for random X-inactivation (Rastan and Robertson 1985; Lee et al. 1996; Penny et al. 1996; Lee and Lu 1999). Two approaches

were chosen to modulate *Tsix* transcription with minimal perturbation of genomic sequences. In the first approach, *Tsix* transcription was functionally inhibited by introducing a transcriptional stop signal into the transcribed region of *Tsix*. In the second approach, we created an inducible system for *Tsix* expression. We found that the truncation of the *Tsix* transcript recapitulates the phenotype of a targeted *Tsix* deletion and lead to complete nonrandom inactivation of the targeted X chromosome. Induction of *Tsix* transcription during ES cell differentiation, on the other hand, caused the targeted chromosome always to be chosen as the active chromosome. These results for the first time establish a function for antisense transcription in the regulation of X-inactivation.

Materials and Methods

Plasmid construction

The 5 kb *Sac*II/*Sal*I fragment of fragment 2 (cosmid MB4-14A, accession number U41395) cloned into pBluescript (Stratagene), pSacSal, was used for all targeting constructs. This fragment contains a unique *Hind*III site located 10959 bp away from the end of the *Xist* cDNA (Brockdorff et al. 1992) (GenBank Accession Number L04961) as insertion site. For the stop construct, a 0.65 kb *Xho*I/*Hind*III fragment from pPGKneotpAlox2 (Soriano 1999) containing a triple polyA signal (tpA) flanked by a loxP site was ligated into the *Sal*I/*Hind*III digested vector pSAßgeo (Soriano 1999) containing a splice acceptor (SA). The SAtpA1lox cassette from this vector was then released by *Xho*I digestion and ligated into a *Xho*I digested pSacSal derivative lacking a

*Xho*I site in the polylinker and harboring a *Xho*I adapter at unique *Hind*III site, yielding vector pSS-SAtpA1lox. In order to create a unique restriction site for linearization of the finished construct, an *Sfi*I adapter was ligated into the unique *Sac*II site of this vector. A HygTK selection cassette flanked by 2 loxP sites was isolated from vector pBS246CMVHygTK (gift from Chris Wilson and Peggy Lee) that lacked the *Sfi*I site 5' of the selection cassette by *Not*I/*Sfi*I digestion and inserted in the correct orientation into a unique *Sma*I site of pSS-SAtpA1lox by blunt end ligation, yielding plasmid pSS-Stop. For electroporation, pSS-Stop was linearized by *Sfi*I digestion. For the inducible construct, a *Cla*I adapter was introduced into the unique *Sal*I site of pSacSal, yielding pSSC. This step facilitated linearization of the final construct. A 0.8 kb *Xho*I/*Hind*III fragment from pTETOP-H/X (Wutz and Jaenisch 2000) was cloned into the unique *Sfi*I site pBS246CMVHygTK-*Sfi*I by blunt-end ligation, creating pTetopHygTK. The *Not*I/*Xba*I fragment containing the tetop and the HygTK selectable marker flanked by loxP sites was then inserted into the unique *Hind*III site of pSSC by blunt-end ligation, yielding pIndTsix. For electroporation, pIndTsix, was linearized with *Cla*I.

Cell culture and generation of transgenic ES cells

Undifferentiated ES cells were grown on mouse embryonic fibroblasts in DMEM (GIBCO), 15% calf serum (FCS, Hyclone), and 1000 U LIF/ml. Differentiation of ES cells was induced by in medium containing 10% FCS without LIF and with 40 ng/ml all-*trans*-retinoic acid (Sigma). Doxycycline (Sigma) was added to the medium at a final concentration of 40 ng/ml.

Transgenic ES cells were generated by transfecting cells with 30 μg of linearized plasmid by electroporation using a Bio-Rad Genepulser set at 25 μF and 400V. After selection with 140 μg/ml hygromycin B (Roche) for 6 - 9 days, colonies were screened by Southern analysis:

For the both the stop as well as the inducible construct, *Kpn*I digested genomic DNA was probed with the 1 kb probe. A 4.8 kb fragment identified the targeted allele (wild type at 16 kb). To check the integrity of the recombination event of the stop construct, blots were stripped in 0.04 M NaOH and reprobed with the 3' internal probe. 50% of male and 6% of female cells were targeted correctly. The integrity of the recombination event for the inducible construct was confirmed on *Sma*I digested DNA with the 3' internal probe. The correct targeting event was detected in 30% of male cells and 3% of female cells.

Cre-mediated recombination was achieved by electroporating targeted cells with 25 μg of supercoiled pOG231 and selecting for the absence of the TK gene in medium containing 2 μM gancyclovir. Resistant clones were analyzed by Southern blotting. For the stop construct, *Eco*RI digested DNA was probed with the 5' internal probe. The wild type allele yielded a 1.7 kb band, whereas the 1 lox and the 2 lox (stop) alleles were detected by a 1.1 kb and a 1.8 kb band, respectively. For the inducible construct, *Xba*I digested DNA was probed the 5' internal probe. The loop-out was confirmed by a 4.4 kb band, compared to the wild type and the targeted band which gave a 3.8 kb and 7.7 kb band respectively.

For this study, two independent clones were chosen for each construct, which behaved similar in all experiments.

Southern probes

The 1 kb probe was a 986 bp PCR product specific for a region 2207 – 3193 bp downstream of the insertion site (relative to the direction of the *Xist* gene) and was created using the primers indF (5'-AAGGCAGGGATTTTAGCGAT-3') and IndR (5'-TGCAGCCATTCTTTTCTGTG-3'). This probe was specific for a region outside the right arm of the targeting vector. The 3'-internal probe is a 876 bp PCR product specific for a region 1291 – 2167 bp downstream of the insertion site (right arm of both constructs) using primers HS3intF (5'-AATCCGCATCAAAACCAAAG-3') and HS3intR (5'-GCAGAGCAGAGGTGATGAAA-3'). The 5'-internal probe was a 257 bp PCR product specific for a region 1000 – 743 bp upstream of the insertion site (left arm of the construct) created using primers 5s and 5as (see below).

RT PCR and primers

RNA was prepared using RNAzol B reagent (Teltest) and treated with RNase-free DNase I (Gibco) according to the manufacturers instructions. Strand-specific RT PCR was carried out essentially as described (Lee et al. 1999a) using 1 μg of total RNA per reaction. Control reactions that were carried out in the absence of reverse transcriptase were included in all cases. The following primers are described relative to the first bp of the *Xist* cDNA: 1s (272-294); 1 as (755-775); 2s (10162-10183); 2sE7 (11011-11030); 2as (11279-11308). The following primers are located relative to the end of the *Xist* cDNA: 3s (1131-1150); 3as (1553-1572); 4s (5118-5757); 4as (6142-6161); 5s (10033-10056); 5as (10197-10216); 6s (12250-12270); 6as (12714-12734). Primers 2s and 2as

were used to amplify a product specific for the spliced *Xist* transcript (referred to as primer pair 2a). Primers 2sE7 and 2as were used to amplify the *Tsix* transcript.

Allele specific detection of *Xist* and *Tsix*

The *Nco*I polymorphism in *Xist* exon 1 has been described previously (Lee and Lu 1999). First strand synthesis with primers specific for *Xist* (1as) or *Tsix* (1s), followed by PCR and detection was as described (Lee and Lu 1999). PCR was carried out for 28 cycles. The *Pst*I polymorphism is located in *Xist* exon 7. First-strand synthesis with primers specific for *Xist* (2as) or *Tsix* (2sE7) was carried out at 50°C. cDNA was amplified using *Xist*-specific primer pair 2a or *Tsix*-specific primer pair 2b for 30 cycles (linear range). PCR products were purified through QIAquick columns (QIAGEN) and digested with *Pst*I for 3 hours. Fragments specific for the 129 and cast alleles were resolved on 2% agarose gels and detected with ethidium bromide.

FISH

Cells were harvested by trypsinization and allowed to adhere on poly-L-lysine coated microscope slides (Sigma) for 2 min before fixation in 4% formaldehyde/5% acetic acid/saline for 18 min. FISH was performed as previously described (Wijgerde et al. 1995) with the following modification. After initial washes in 2 x SSC at 37°C, unhybridized RNA was digested with 25 μg/ml RNaseA in 0.5 M NaCl, 10 mM Tris pH 7.5, 0.1% Tween-20 under a coverslip for 45 min at 37°C.

Strand-specific probes were generated by riboprobe synthesis (Promega) using T3 or T7 polymerase in the presence of biotin-16-UTP (Roche). Antibody detection included

three amplification steps using mouse-anti-biotin antibodies (Roche), followed by rhodamine-conjugated donkey-anti-mouse and goat-anti-horse antibodies (Jackson-labs). We used pBgl5K containing a 5 kb *Xist* cDNA fragment from exons 4 - 7 in pBluescript (Stratagene) as a template for *Tsix* and *Xist* specific probes. A template for a probe specific for *Tsix* exon 4, pTsixE7, was generated by cloning a –1615 - +1883 bp *BamH*I fragment (relative to start of *Xist* cDNA) into pBluescript (Stratagene).

Results

Generation of ES cells expressing a truncated *Tsix* allele

In order to assess the functionality of the *Tsix* transcript, we created a truncated *Tsix* allele by introducing a transcriptional stop signal (triple polyadenylation site, tpA) (Soriano 1999) into a *Hind*III site 4 kb downstream of the major *Tsix* transcriptional start site (Figure 1A). The presence of a strong splice acceptor (SA) (Friedrich and Soriano 1991) ensured the trapping of any transcripts that originated from all upstream promoters. In addition, the construct contained a HygroTK selectable marker and three loxP sites. After introduction of the construct into male and female ES cells by homologous recombination, the loxP sites facilitated the generation of two alleles after transient transfection with Cre-recombinase and selection for the loss of the TK gene: 1) A 2 lox allele (stop allele) was created by deleting the selectable marker that was transcribed from the CMV promoter. 2) A 1 lox allele was generated by removal of both, the selectable marker as well as the tpA/SA element.

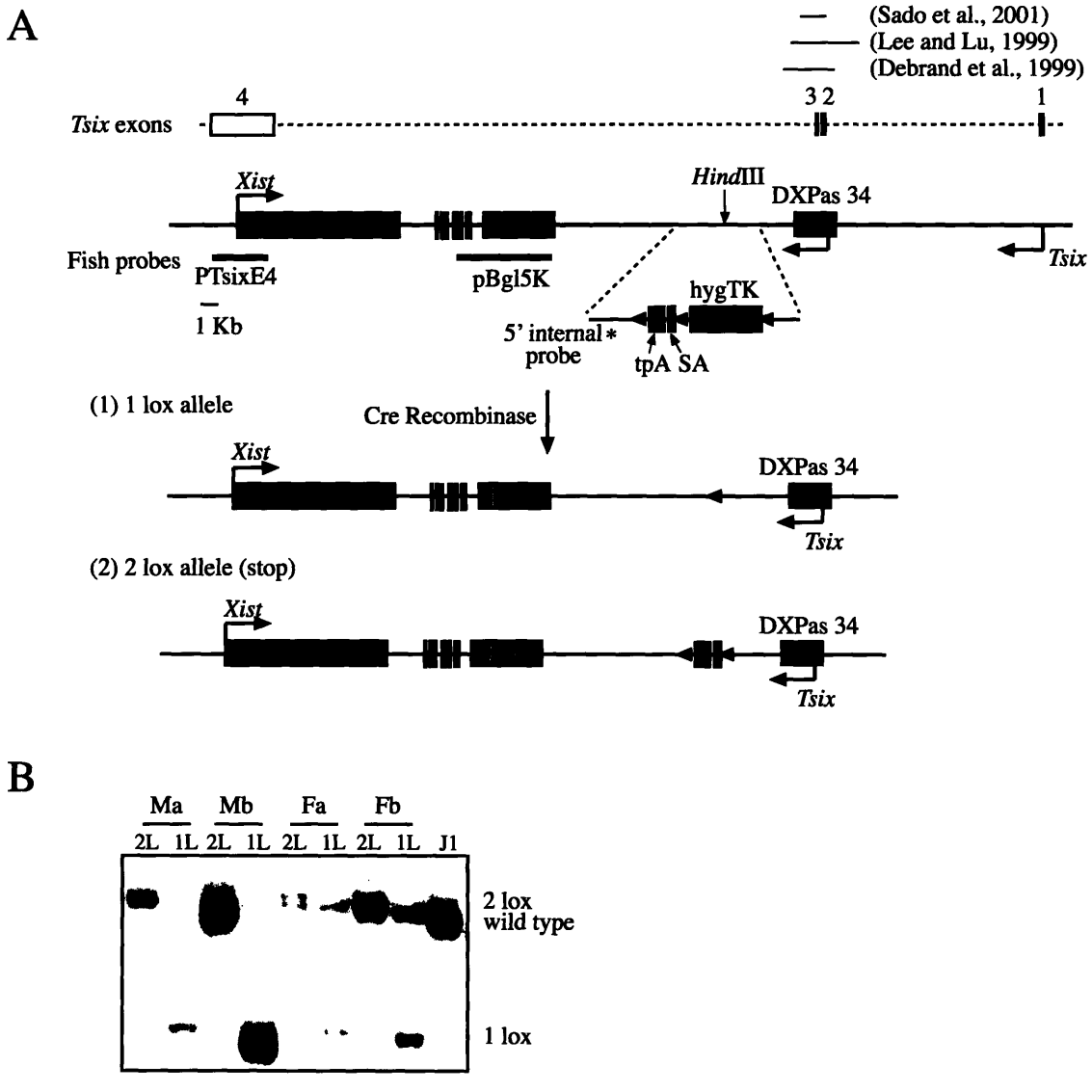


Figure 1. **(A)** Scheme for the generation of a truncated *Tsix* (2 lox) and the control 1 lox alleles (see text for details). The *Tsix* intron/exon structure, the two putative *Tsix* promoters, and regions deleted in previous studies (Debrand et al. 1999; Lee and Lu 1999; Sado et al. 2001) are indicated. **(B)** Southern analysis of targeted male and female ES cell lines after Cre-mediated recombination. DNA was digested with *EcoR*I and probed with the 5'-internal probe (indicated in Figure 1A) in order to detect wild type (1.7 kb), 1 lox (1 kb) and 2 lox (1.8 kb) alleles.

The 1 lox allele served as a control to ensure that effects observed in the presence of the stop allele did not simply arise from physical disturbance of the locus.

We targeted the X chromosome in the male ES cell line J1 (Li et al. 1992) and the 129 allele in the female cell line F1-2.1 (Panning et al. 1997) that carries a 129 and a *Mus castaneus* X chromosome. Two independent clones from each cell line were analyzed, and subclones that carried either the 1 lox (male Ma1L, Mb1L and female Fa1L, Fb1L) or the 2 lox (stop) allele (male Ma2L, Mb2L and female Fa2L, Fb2L) were generated by transient transfection with Cre-recombinase (Figure 1B).

We determined whether the presence of the 2 lox (stop) allele led to the expression of a truncated *Tsix* transcript. This was particularly important because transcriptional readthrough has been seen even in the presence of an ectopic stop signal (Maxwell et al. 1989). Total RNA was prepared from undifferentiated male ES cells that were wild type or carried either the 1 lox or the 2 lox (stop) allele, and strand-specific RT-PCR (30-cycles) was carried out at six sites across the *Xist/Tsix* locus (Figure 2A). In wild type cells as well as in the 1 lox cell line the *Tsix* transcript was detectable with all tested primers (Figure 2A, left panel), which indicated that the presence of the one loxP site did not inhibit *Tsix* transcription. In contrast, in the 2 lox (stop) cell line, *Tsix* RNA could be detected only with primers 5 and 6 which encompass the insertion site. No *Tsix* signal was detected at the other four positions tested. Primer 6 corresponded to a region upstream of the insertion site, whereas primer 5 was located 927 nucleotides downstream of it, indicating a transcriptional readthrough of at least 1 kb. These results were confirmed by RNA FISH using a strand-specific probe for a *Tsix* region downstream of the stop signal (probe pBgl5K, Figure 1A): male and female 1 lox cells showed a wild

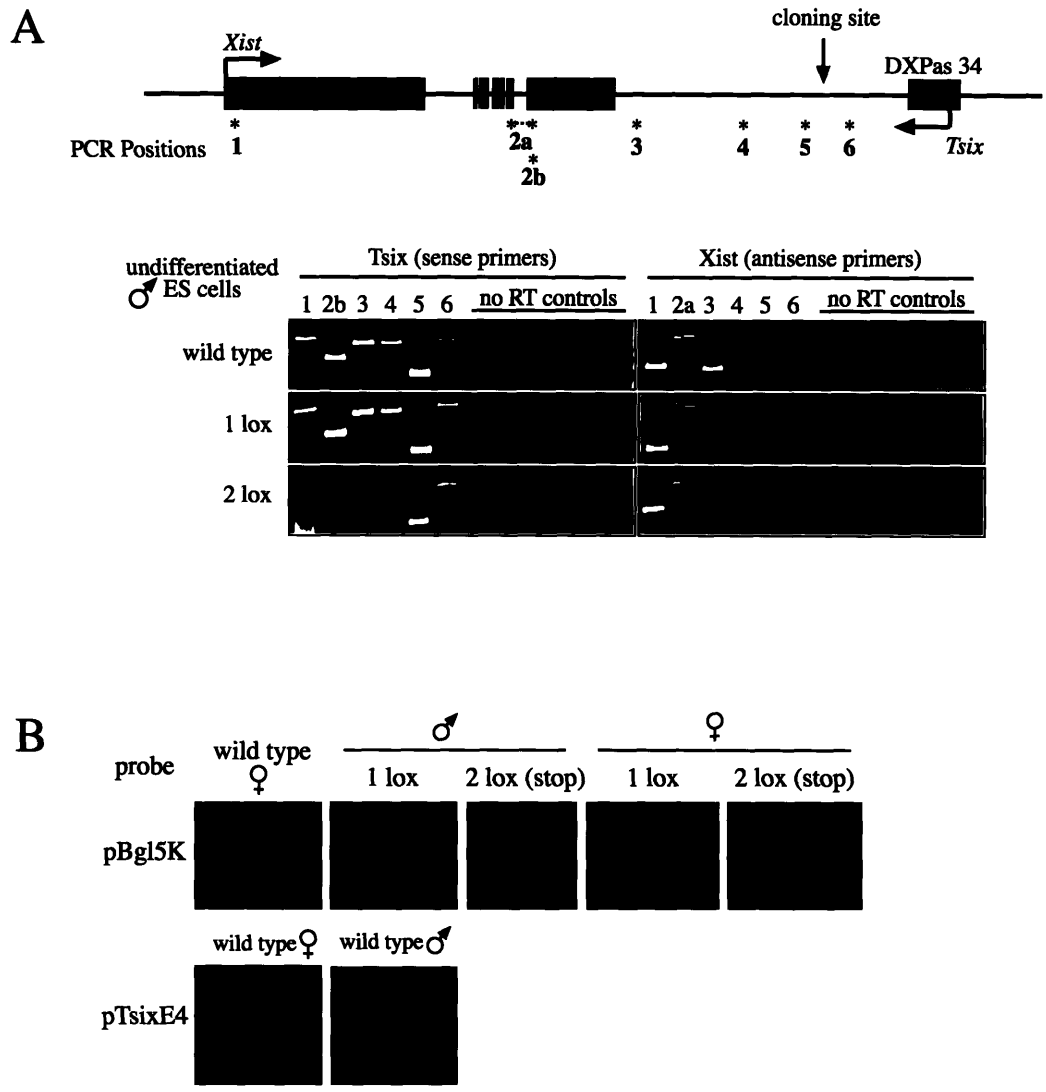


Figure 2. Presence of the 2 lox (stop) allele leads to expression of a truncated *Tsix* transcript while leaving *Xist* expression unaffected. The 1 lox allele has no effect on *Tsix* transcription. **(A)** Strand-specific RT-PCR for *Tsix* and *Xist* was performed on wild type and targeted undifferentiated male ES cells. The location of primer pairs is indicated as "PCR Positions". Primer pair 2a spans intron 6 and was used to amplify a product specific for the spliced *Xist* transcript. Primer pair 2b is located in exon 7 and was used to amplify the *Tsix* transcript. "No RT controls" refer to reactions carried out in the absence of reverse transcriptase. **(B)** Strand-specific RNA FISH for *Tsix* in undifferentiated ES cells. Male and female wild type and targeted ES cells were probed with riboprobes specific for either the *Tsix* region overlapping *Xist* exons 4 - 7 (probe pBgl5K) or *Tsix* exon 4 (probe pTsixE4). The location of both probes is indicated in Figure 1A.

type pattern of *Tsix* expression, whereas no *Tsix* signal was detectable in male 2 lox (stop) cells, and only one signal in female 2 lox (stop) cells (Figure 2B).

Xist transcription was detectable in all three cell lines with primers 1 and 2 by RT PCR (Figure 2A, right panel). The signal obtained with primer 3 was variable which may have resulted from sporadic transcriptional readthrough since this primer corresponded to an area 1kb downstream of the end of the published *Xist* cDNA sequence (Brockdorff et al. 1992) (GenBank Accession Number L04961).

We conclude that the presence of the 2 lox (stop) allele lead to the transcription of a truncated *Tsix* allele without affecting *Xist* expression in undifferentiated cells. This effect was reversed by Cre-mediated deletion of the stop cassette.

Tsix transcription is a negative regulator of *Xist* expression

Recent work has shown that deletion of the major *Tsix* transcriptional start site, including portions of the surrounding CpG island, leads to primary non-random inactivation of the mutant X in differentiated female ES cells as well in the epiblast lineages of the embryo proper (Lee and Lu 1999; Lee 2000; Sado et al. 2001). However, it remained unresolved whether *Tsix* action was mediated by a DNA element that had been deleted in the mutant locus, or whether the observed phenotype was caused by loss of the antisense transcript and/or transcription through the region. We reasoned that if the phenotype was caused by loss of *Tsix* transcription rather then by deletion of a crucial DNA element, then the 2 lox (stop) allele would lead to skewed X-inactivation in differentiated female ES cells. In order to assess this hypothesis, DNA and RNA were isolated from undifferentiated ES cells as well as ES cells that were induced to

differentiate with all-*trans*-retinoic acid in the absence of LIF (Hogan 1994) for 6 days. The presence of two X chromosomes of different parental origin in the female cells lines as confirmed by Southern analysis provided a means of analyzing allele-specific expression of *Xist* by strand-specific RT-PCR. Two restriction fragment-length polymorphisms in exon 1 (*Nco*I polymorphism) (Lee and Lu 1999) and exon 7 (*Pst*I polymorphism), respectively, were analyzed (Figure 4A, page 103). In wild type male cells, *Xist* was expressed in undifferentiated cells and undetectable in differentiated male cells (Figure 4A, compare lanes 1 and 2), in agreement with *Xist* expression being silenced in male cells upon differentiation. Control female brain cells showed expression from both *Xist* alleles (Figure 4A, lane 3), because either X in these cells can be stochastically chosen for X-inactivation. However, *Xist* expression from the 129 allele carrying the lower strength *Xce* (Xce^a) relative to that of *M. castaneus* (Xce^c) (Avner et al. 1998) was more abundant. A similar pattern of *Xist* expression was observed in differentiated 1 lox cells (Figure 4A, lane 5) and differentiated wild type ES cells (data not shown) indicating that random X-inactivation was not affected by the presence of the loxP site. In contrast, cells heterozygous for the 2 lox (stop) allele showed complete skewing of *Xist* expression toward the targeted 129 allele, which does not express *Tsix* (Figure 4A, lane 7).

The presence of the 2 lox (stop) allele also resulted in increased *Xist* expression from the targeted allele in undifferentiated cells (Figure 4A, lane 6) similar to what had been observed in cells heterozygous for the *Tsix* promoter deletion (Lee and Lu 1999). In contrast, in undifferentiated 1 lox cells, like in wild type ES cells, *Xist* was expressed from both alleles (Figure 4A, lane 4 and data not shown). To examine whether the

skewing of *Xist* expression observed in undifferentiated ES cells was caused by a high proportion of cells in the population that had entered differentiation, *Xist* expression was examined by RNA FISH using a strand-specific *Xist* probe. An *Xist* cloud characteristically seen only in differentiated cells was detectable only in 6.5% of 2 lox (stop) cells (n = 184) and 8.6% of 1 lox cells (n = 209) suggesting that this effect was specific for undifferentiated cells. It cannot be excluded, however, that the PCR assays were biased towards the more abundant *Xist* allele. Taken together these results indicate that loss of *Tsix* transcription led to skewed *Xist* expression, and the X chromosome lacking *Tsix* transcription was always chosen for inactivation.

To assess the effect of loss of *Tsix* transcription on *Xist* expression in the absence of a second X chromosome, male cells were differentiated for 2 or 3 days in all-*trans*-retinoic acid in the absence of LIF and *Xist* expression was analyzed by strand-specific RNA FISH (Figure 4B, page 103). In wild type and 1 lox cells *Xist* accumulation was never detected (n > 300). In contrast, ectopic high level *Xist* expression was found in 9.6% (n = 271) of 2 lox (stop) cells on day 2 of differentiation and 9.5 % (n = 232) on day 3.

These results indicate that *Tsix* transcription acts as a negative regulator of *Xist* expression, resulting in skewed *Xist* expression toward the chromosome lacking *Tsix* transcription in female cells and stochastic upregulation of *Xist* expression in male cells upon differentiation.

Generation of an inducible *Tsix* allele

To generate an experimental system that allows regulated expression of *Tsix* RNA in response to doxycycline addition to the medium, we inserted a tetracycline-responsive promoter (Gossen et al. 1995) into the *Hind*III site analogous to the insertion site of the stop construct (tetop, Figure 3A). In addition, the construct contained a HygroTK selectable marker flanked by loxP sites. Using homologous recombination, the construct was introduced into a male (J1) ES cell line and into the 129 *Xist* allele of a female (F1-2.1) ES cell line carrying the *nls-rtTA* cDNA encoding a doxycycline-controlled transcriptional activator at the ubiquitously expressed *ROSA26* locus (Wutz and Jaenisch 2000). Targeted male and female cell lines were transfected with Cre-recombinase and selected for loss of the TK gene. Male lines M.ind1a and M.ind1b, and female lines F.ind1a and F.ind1b containing the inducible promoter were identified by Southern blotting (Figure 3B).

Tsix expression was assessed in cells that were grown in the presence of absence of doxycyline. Strand-specific RT-PCR and RNA FISH analysis revealed that *Tsix* was expressed at similar levels from both X chromosomes in uninduced and induced female cells prior to differentiation (Figure 4C, lanes 1 and 2, Figure 4D, middle panel). Doxycycline treatment of undifferentiated male cells did not result in an increased *Tsix* signal as detected by strand-specific RNA FISH (Figure 4D, left panel), suggesting that it might be impossible to induce *Tsix* expression above the physiological level in undifferentiated cells. Alternatively it is possible that elevated *Tsix* expression levels were not detectable due to the instability of the transcript. In contrast, when cells were differentiated in the presence, but not in the absence of doxycycline for 6 days,

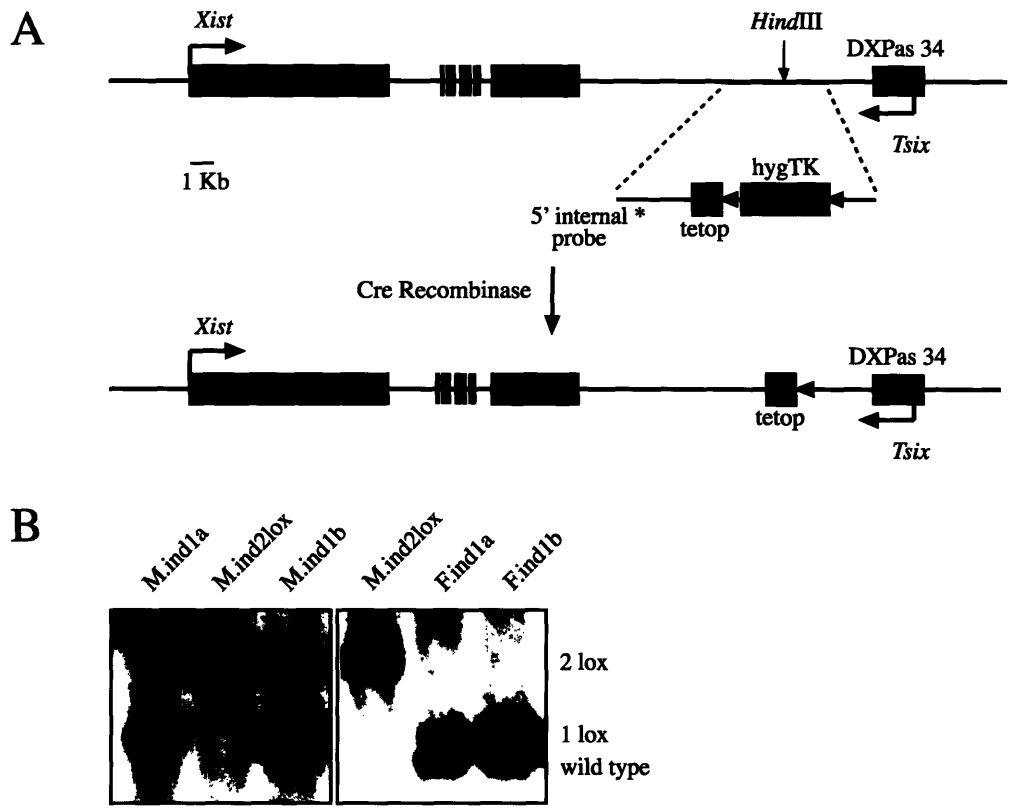


Figure 3. **(A)** Strategy for the generation of an inducible *Tsix* allele (see text for details). **(B)** Southern analysis of targeted male and female cell lines after Cre-mediated recombination. *Xba*I digested DNA was probed with the 5'-internal probe (indicated in Figure 1A). The wild type allele gave a 3.8 kb band compared to the 1 lox (loop out, 4.4 kb) and 2 lox (no loop-out, 7.7 kb) alleles.

Tsix was expressed from the targeted 129 allele at a level that was comparable to that in undifferentiated cells (Figure 4C, lanes 3 and 4, Figure 4D, right panel). These results indicate that the presence of doxycycline does not lead to detectable overexpression of *Tsix* in undifferentiated cell. However, when cells were differentiated in the presence of doxycyline, *Tsix* expression was clearly induced.

Induction of *Tsix* transcription prevents the inactivation of an X chromosome *in cis*

We have shown that the X chromosome lacking *Tsix* transcription was always chosen for X-inactivation. If the role of *Tsix* transcription was to block the accumulation of *Xist*, then overexpression of *Tsix* should protect the X chromosome *in cis* from inactivation. To test this hypothesis, we compared *Xist* expression patterns from cells that were grown in the presence or absence of doxycycline. In both induced and uninduced undifferentiated cells *Xist* was expressed from both alleles (Figure 4E, lanes 1 and 2). This result was in agreement with the finding that *Tsix* overexpression from the 129 allele was not detectable in undifferentiated cells (Figure 4C and D).

In contrast, biased X-inactivation was observed when cells were differentiated in the presence of doxycycline for 6 days (Figure 4E, compare lanes 3 and 4). Ectopic *Tsix* transcription from the induced 129 allele led to almost exclusive upregulation of *Xist* expression from the non-induced cast allele (> 90% as quantified by phosphorimaging). Background expression from the 129 allele was probably attributable to cells that had initiated differentiation and undergone random X-inactivation the addition of doxycycline. In agreement with this, an *Xist* cloud was detectable in 6 – 10% of cells before induction of differentiation by strand-specific RNA FISH.

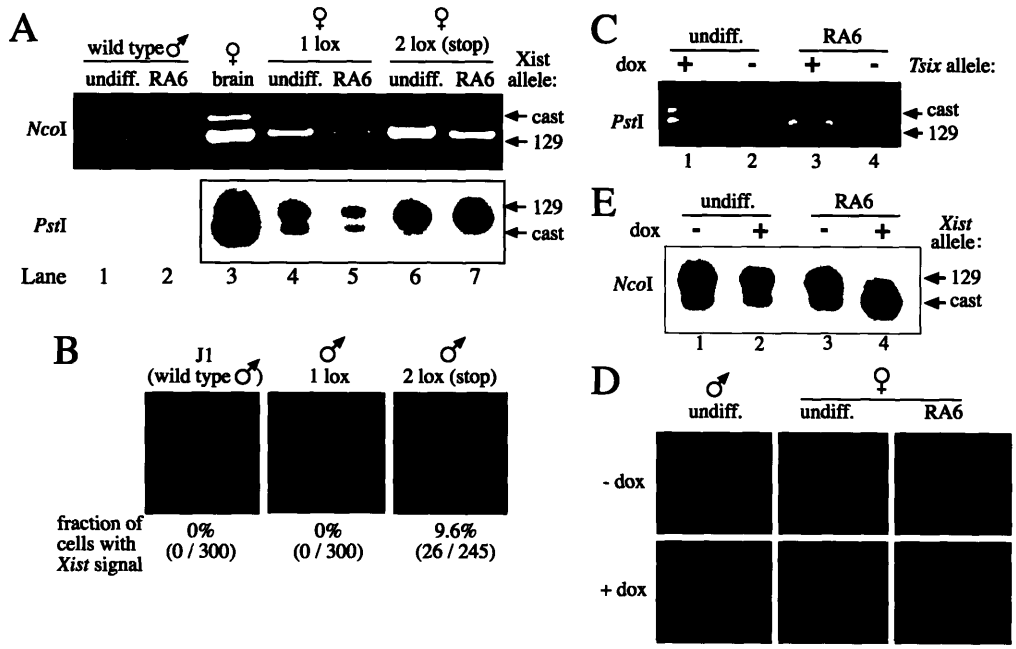


Figure 4. *Tsix* transcription is a negative regulator of *Xist* expression. (**A**) Allele specific RT-PCR for *Xist* expression was performed on control female brain cells, undifferentiated ES cells (undiff.) and ES cells that were induced for differentiated with all-*trans*-retinoic acid for 6 days (RA6). *Pst*I indicates a polymorphism in *Xist* exon 7 which was detected after *Pst*I digestion of a PCR product amplified with primer pair 2a (28 cycles). First strand synthesis was performed using primer 2as. *Nco*I indicates a polymorphism in *Xist* exon 1 that has been described previously (Lee and Lu 1999). Amplification was carried out with the primer pair 1s/as (28 cycles) after first strand synthesis with primer 1as. No bands were detected in control reactions performed in the absence of reverse transcriptase (not shown). The polymorphic bands are indicated: "129" corresponds to the targeted 129 derived X chromosome, whereas "cast" corresponds to the non-targeted *M. castaneus* X. (**B**) Wild type and targeted male ES cells were induced to differentiate with all-*trans*-retinoic acid for 2 days. Strand-specific RNA FISH for *Xist* was performed using a riboprobe specific for *Xist* exons 4-7 (probe pBgl5K, Figure 1A). The proportion of cells expressing ectopic *Xist* signals is indicated below each panel. (**C**) Allele-specific expression of *Tsix* in the presence (+dox) and absence (-dox) of the inducing agent doxycycline in female ES cells. Cells were either undifferentiated (undiff.) or were differentiated for 6 days in all-*trans*-retinoic acid (RA6). To detect the *Pst*I polymorphism, strand-specific RT-PCR for *Tsix* was carried out using primer 2sE7 for first strand synthesis, followed by amplification with the primer pair 2b (30 cycles) and digestion of the PCR product with *Pst*I. "129" corresponds to the *Tsix* allele derived from the 129 X chromosome that harbors the inducible promoter; "cast" refers to the non-targeted *M. castaneus*-derived X. (**D**) *Tsix* signals in induced cells are indistinguishable from the signals in uninduced and wild type cells. Cells were either undifferentiated (undiff.) or differentiated with all-*trans*-retinoic acid for 6 days (RA6) and grown in the presence (+ dox) or absence (- dox) of doxycycline. Strand-specific RNA FISH was performed using a riboprobe specific for the *Tsix* region overlapping *Xist* exons 4-7 (probe pBgl5K, Figure 1A). (**E**) Allele-specific expression of *Xist* in female ES cells. The X chromosome that constitutively expresses *Tsix* during differentiation is always chosen to remain active. *Xist* expression is skewed almost completely toward the non-targeted *M. castaneus* allele (cast) when *Tsix* was constitutively expressed from the 129 allele (129) [compare lanes 3 (-dox) and 4 (+dox)]. In undifferentiated cells induction of *Tsix* expression had no effect on *Xist* expression [compare lanes 1 (-dox) and 2 (+dox)]. Strand-specific RT-PCR was performed as described above.

The presence of both X chromosomes in these cells was confirmed by Southern blotting (data not shown). These results demonstrate that an X chromosome that ectopically expresses *Tsix* RNA is always chosen to remain active.

Discussion

In this study, we have generated two *Tsix* alleles in order to assess the role of *Tsix* transcription in the regulation of X chromosome choice during X-inactivation. As a model system, we used ES cells, which are known to undergo random X-inactivation when induced to differentiate. In the first approach, a transcriptional stop signal was introduced into the transcribed region of *Tsix* flanked by a strong splice acceptor. The splice acceptor ensured trapping of all transcripts arising from upstream promoters, including the major promoter initially identified by Lee (Lee et al. 1999a) and the minor promoter recently characterized by Sado et al. (Sado et al. 2001). In the second approach, we created a doxycycline inducible allele which enabled us to ectopically activate *Tsix* expression in a regulated manner. The objective was to functionally modify *Tsix* expression while leaving potentially important regulatory DNA elements intact. This approach differs from previously published studies in which targeted deletions of *Tsix* not only abolished *Tsix* transcription but, in addition, deleted major parts of the *DXPas34* locus (Debrand et al. 1999; Lee and Lu 1999; Sado et al. 2001), a CpG island that has been shown to be differentially methylated on the active and inactive X chromosomes (Courtier et al. 1995).

The results described here clearly establish for the first time the importance of active transcription for *Tsix* function. Loss of *Tsix* transcription lead to nonrandom inactivation of the X chromosome with the mutant X being always inactivated (Figure 4A, page 103), whereas constitutive expression of *Tsix* causes the mutant X chromosome to remain active (Figure 4E, page 103). This unequivocally demonstrates that *Tsix* transcription through the *Xist* locus is crucial for preventing X chromosome inactivation and confirms the role of *Tsix* as a negative regulator of *Xist*.

In agreement with this, we found that loss of *Tsix* transcription led to ectopic *Xist* expression in a subset of differentiating male cells (Figure 4B, page 103), an effect that was specific for differentiating cells since *Xist* accumulation is detectable only in about 1% (n = 367) of undifferentiated cells. These results differ from previous studies that found ectopic *Xist* expression to be restricted to cells of extra-embryonic tissues (Lee and Lu 1999; Lee 2000; Sado et al. 2001). In these tissues X-inactivation is normally imprinted such that expression from the maternal *Xist* allele is silenced. Deletion of the *Tsix* promoter/*DXPas34* locus erased this imprint leading to ectopic *Xist* expression from the maternal allele in addition to upregulation of *Xist* expression from the paternal allele. However, it is possible that ectopic *Xist* upregulation still occurred in the epiblast lineage of these embryos. Cells that undergo aberrant dosage compensation are not viable and are progressively lost from the population. Since our findings suggest that only a small fraction of cells of the epiblast lineage are affected, loss of these cells may not have an obvious effect on the overall cell population. Interestingly, Sado et al. (2001) did observe *Xist* accumulation in a subset of mutant male ES cells grown under non-differentiating

conditions consistent with these cells representing a subpopulation that had already entered differentiation and initiated stochastic upregulation of *Xist* expression.

In addition, the phenotype described here differs from that of a 65 kb knockout which included the *Tsix* promoter (Clerc and Avner 1998). This knockout also resulted in nonrandom X-inactivation of the deleted X in female ES cells. However, in cells carrying only one intact X chromosome (X0), ectopic *Xist* upregulation was observed in virtually all cells. In contrast, we see ectopic expression only in a subset of cells. An explanation for this difference may be provided by a model proposed by Lee and Lu (2000). This model suggests that *Xist* expression is regulated by a complex interplay of positive and negative factors. A blocking factor, present in limited amount, may mark one X chromosome per diploid set of autosomes as the future active X. Our work indicates that this blocking factor most likely acts by regulating *Tsix* transcription and may perhaps consist of a complex of autosomal and X-linked factors that bind to the *Tsix* promoter. In addition, Lee and Lu suggest the presence of competence factors that are required to initiate *Xist* upregulation on the future inactive X, once the active X is chosen. These competence factors are thought to be expressed only in cells that need to undergo dosage compensation, i.e. in cells that carry more than one X per diploid set of autosomes. Therefore loss *Tsix* transcription as the target of the blocking factor would not be sufficient to induce upregulation of *Xist* expression in male cells. However, the observation that stochastic ectopic *Xist* expression is still observed in some cells may indicate that competence factors are not be strictly limited to cells that carry more than one X chromosome. Male cells might simply be "less competent" to upregulate *Xist* expression in the absence of the blocking factor because they are not poised for X-

inactivation and might therefore provide a less permissive environment for *Xist* upregulation and stabilization.

Recent findings suggest that *Tsix* consists of four exons and is subject to complex alternative splicing (Sado et al. 2001). We found that the localization of these spliced forms most likely overlaps with the localization of the unspliced transcript as indicated by FISH using a strand-specific probe specific for the common exon 4 (probe pTsixE4, Figure 2B, page 96). In our study, both the transcriptional stop signal as well as the inducible promoter were inserted downstream of exons 1 through 3, leaving exon 4 as the only exon whose expression was modified. In view of our results two types of models for the mechanism of *Tsix* action can be proposed. 1) *Tsix* function could be mediated by a functional RNA. According to this model exon 4 would be the only exon that is necessary and sufficient for *Tsix* function. *Tsix* RNA may interact directly with *Xist* DNA or local chromatin to interfere with *Xist* transcription. Alternatively, the formation of an *Xist/Tsix* RNA duplex may trigger RNA degradation or mask functional domains in *Xist* RNA. Interestingly exon 4 overlaps with the *Xist* transcription unit and the *Xist* A-repeat, a conserved region that is crucial for *Xist* function (Allaman-Pillet et al. 2000; Wutz et al. 2002). 2) It is possible that transcriptional activity of the locus per se, rather than the RNA itself, is important for function so that the processed form of *Tsix* would play only a minor role in the regulation of *Xist* expression. Antisense transcription might, for example, be involved in regulating the chromatin domain around the *Xist* locus. This mechanism has been shown to be crucial for the expression of the human ß-globin gene (Gribnau et al. 2000) where intergenic transcription was found to be required for the remodeling of the chromatin structure to allow expression of this locus in a

developmentally regulated manner. Similarly, *Tsix* transcription may directly modulate the chromatin structure through the *Xist* to allow access of developmentally regulated factors. Alternatively, transcriptional interference between *Tsix* and *Xist* might inhibit efficient transcription of the *Xist* gene, thereby preventing *Xist* accumulation.

It will be interesting to see, by what mechanism *Tsix* transcription regulates *Xist* expression. In addition, upstream factors that control *Tsix* expression and therefore constitute the potential blocking factor still remain to be identified

Acknowledgments

We thank Nicki Watson for help with microscopy; Györgyi Csankovszki, Joost Gribnau, Caroline Beard, and Ted Rasmussen for discussions and critical reading of the manuscript. This work was conducted using the W.M. Keck Foundation biological imaging facility at the Whitehead Institute and was supported by NIH grants CA44339 and CA87869 to RJ.

References

Allaman-Pillet, N., A. Djemai, C. Bonny, and D.F. Schorderet. 2000. The 5' repeat elements of the mouse Xist gene inhibit the transcription of X-linked genes. *Gene Expr* **9:** 93-101.

Avner, P. and E. Heard. 2001. X chromosome inactivation: counting, choice and initiation. *Nat Rev Genet* **2:** 59-67.

Avner, P., M. Prissette, D. Arnaud, B. Courtier, C. Cecchi, and E. Heard. 1998. Molecular correlates of the murine Xce locus. *Genet Res* **72:** 217-224.

Brockdorff, N., A. Ashworth, G.F. Kay, V.M. McCabe, D.P. Norris, P.J. Cooper, S. Swift, and S. Rastan. 1992. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71:** 515-526.

Brown, C.J., B.D. Hendrich, J.L. Rupert, R.G. Lafreniere, Y. Xing, J. Lawrence, and H.F. Willard. 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.

Clemson, C.M., J.A. McNeil, H.F. Willard, and J.B. Lawrence. 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.

Clerc, P. and P. Avner. 1998. Role of the region 3' to Xist exon 6 in the counting process of X- chromosome inactivation. *Nat Genet* **19:** 249-253.

Cooper, D.W., J.L. VandeBerg, G.B. Sharman, and W.E. Poole. 1971. Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X-inactivation. *Nat New Biol* **230:** 155-157.

Courtier, B., E. Heard, and P. Avner. 1995. Xce haplotypes show modified methylation in a region of the active X chromosome lying 3' to Xist. *Proc Natl Acad Sci U S A* **92:** 3531-3535.

Debrand, E., C. Chureau, D. Arnaud, P. Avner, and E. Heard. 1999. Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. *Mol Cell Biol* **19:** 8513-8525.

Friedrich, G. and P. Soriano. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5:** 1513-1523.

Gossen, M., S. Freundlieb, G. Bender, G. Muller, W. Hillen, and H. Bujard. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268:** 1766-1769.

Gribnau, J., K. Diderich, S. Pruzina, R. Calzolari, and P. Fraser. 2000. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. *Mol Cell* **5:** 377-386.

Herzing, L.B., J.T. Romer, J.M. Horn, and A. Ashworth. 1997. Xist has properties of the X chromosome inactivation centre. *Nature* **386:** 272-275.

Hogan, B., Beddington, R., Costantini, F., and Lacy, F. 1994. *Manipulating the mouse embryo: A laboratory manual.* Cold Spring Habor Laboratory Press, Cold Spring Habor, NY.

Jones, B.K., J.M. Levorse, and S.M. Tilghman. 1998. Igf2 imprinting does not require its own DNA methylation or H19 RNA. *Genes Dev* **12:** 2200-2207.

Lee, J.T., W.M. Strauss, J.A. Dausman, and R. Jaenisch. 1996. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* **86:** 83-94.

Lee, J.T. and N. Lu. 1999. Targeted mutagenesis of Tsix leads to nonrandom X-inactivation. *Cell* **99:** 47-57.

Lee, J.T., L.S. Davidow, and D. Warshawsky. 1999a. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* **21:** 400-404.

Lee, J.T., N. Lu, and Y. Han. 1999b. Genetic analysis of the mouse X-inactivation center defines an 80-kb multifunction domain. *Proc Natl Acad Sci U S A* **96:** 3836-3841.

Lee, J.T. 2000. Disruption of imprinted X-inactivation by parent-of-origin effects at Tsix. *Cell* **103:** 17-27.

Li, E., T.H. Bestor, and R. Jaenisch. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915-926.

Li, T., T.H. Vu, Z.L. Zeng, B.T. Nguyen, B.E. Hayward, D.T. Bonthron, J.F. Hu, and A.R. Hoffman. 2000. Tissue-specific expression of antisense and sense transcripts at the imprinted Gnas locus. *Genomics* **69:** 295-304.

Lyon, M.F. 1961. Gene action in the X chromosome of the mouse (Mus musculus L.). *Nature* **190:** 372-373.

Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch. 1997. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* **11:** 156-166.

Marahrens, Y., J. Loring, and R. Jaenisch. 1998. Role of the Xist gene in X chromosome choosing. *Cell* **92:** 657-664.

Maxwell, I.H., G.S. Harrison, W.M. Wood, and F. Maxwell. 1989. A DNA cassette containing a trimerized SV40 polyadenylation signal which efficiently blocks spurious plasmid-initiated transcription. *Biotechniques* **7:** 276-280.

Memili, E., Y.K. Hong, D.H. Kim, S.D. Ontiveros, and W.M. Strauss. 2001. Murine Xist RNA isoforms are different at their 3' ends: a role for differential polyadenylation. *Gene* **266:** 131-137.

Migeon, B.R. 1994. X chromosome inactivation: molecular mechanisms and genetic consequences. *Trends Genet* **10:** 230-235.

Mise, N., Y. Goto, N. Nakajima, and N. Takagi. 1999. Molecular cloning of antisense transcripts of the mouse Xist gene. *Biochem Biophys Res Commun* **258:** 537-541.

Monk, M. and M.I. Harper. 1979. Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. *Nature* **281:** 311-313.

Moore, T., M. Constancia, M. Zubair, B. Bailleul, R. Feil, H. Sasaki, and W. Reik. 1997. Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse Igf2. *Proc Natl Acad Sci U S A* **94:** 12509-12514.

Panning, B., J. Dausman, and R. Jaenisch. 1997. X chromosome inactivation is mediated by Xist RNA stabilization. *Cell* **90:** 907-916.

Penny, G.D., G.F. Kay, S.A. Sheardown, S. Rastan, and N. Brockdorff. 1996. Requirement for Xist in X chromosome inactivation. *Nature* **379:** 131-137.

Prissette, M., O. El-Maarri, D. Arnaud, J. Walter, and P. Avner. 2001. Methylation profiles of DXPas34 during the onset of X-inactivation. *Hum Mol Genet* **10:** 31-38.

Rastan, S. and E.J. Robertson. 1985. X chromosome deletions in embryo-derived (EK) cell lines associated with lack of X chromosome inactivation. *J Embryol Exp Morphol* **90:** 379-388.

Rastan, S. and S.D. Brown. 1990. The search for the mouse X chromosome inactivation centre. *Genet Res* **56:** 99-106.

Rougeulle, C., C. Cardoso, M. Fontes, L. Colleaux, and M. Lalande. 1998. An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat Genet* **19:** 15-16.

Sado, T., Z. Wang, H. Sasaki, and E. Li. 2001. Regulation of imprinted X chromosome inactivation in mice by Tsix. *Development* **128:** 1275-1286.

Schmidt, J.V., J.M. Levorse, and S.M. Tilghman. 1999. Enhancer competition between H19 and Igf2 does not mediate their imprinting. *Proc Natl Acad Sci U S A* **96:** 9733-9738.

Sheardown, S.A., S.M. Duthie, C.M. Johnston, A.E. Newall, E.J. Formstone, R.M. Arkell, T.B. Nesterova, G.C. Alghisi, S. Rastan, and N. Brockdorff. 1997. Stabilization of Xist RNA mediates initiation of X chromosome inactivation. *Cell* **91:** 99-107.

Soriano, P. 1999. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21:** 70-71.

Takagi, N. and M. Sasaki. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256:** 640-642.

Wijgerde, M., F. Grosveld, and P. Fraser. 1995. Transcription complex stability and chromatin dynamics in vivo. *Nature* **377:** 209-213.

Wutz, A., O.W. Smrzka, N. Schweifer, K. Schellander, E.F. Wagner, and D.P. Barlow. 1997. Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature* **389:** 745-749.

Wutz, A. and R. Jaenisch. 2000. A shift from reversible to irreversible X-inactivation is triggered during ES cell differentiation. *Mol Cell* **5:** 695-705.

Wutz, A., T.P. Rasmussen, and R. Jaenisch. 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* **30:** 167-174.

Chapter 4

Expression of MeCP2 in postmitotic neurons rescues Rett Syndrome in mice

Sandra Luikenhuis[*†], Emanuela Giacometti[*], Caroline F. Beard[*], Rudolf Jaenisch[*†]

Data presented in this chapter have been published in *Proc Natl Acad Sci USA* (2004) **101**:6033-6038.

Respective contriburions: Emanuela Giacometti assisted in targeting the *Tau-Mecp2* construct into ES cells and generated Figures 2B and 2C - R. Caroline Beard was involved in the experimental design.

[†]Massachusetts Institute of Technology,

[*] Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142

Abstract

Mutations in *MECP2* are the cause of Rett syndrome (RTT) in humans, a neurodevelopmental disorder that affects mainly girls. MeCP2 is a protein that binds CpG dinucleotides and is thought to act as a global transcriptional repressor. It is highly expressed in neurons, but not in glia, of the postnatal brain. The timing of MeCP2 activation correlates with the maturation of the central nervous system and recent reports suggest that MeCP2 may be involved in the formation of synaptic contacts and may function in activity-dependent neuronal gene expression. Deletion or targeted mutation of *Mecp2* in mice leads to a RTT-like phenotype. Selective mutation of *Mecp2* in postnatal neurons leads to a similar although delayed phenotype, suggesting MeCP2 plays a role in postmitotic neurons. Here we test the hypothesis that the symptoms of RTT are exclusively caused by a neuronal MeCP2 deficiency by placing *Mecp2* expression under the control of a neuron-specific promoter. Expression of the *Mecp2* transgene in postmitotic neurons resulted in symptoms of severe motor dysfunction. Transgene expression in *Mecp2* mutant mice, however, rescued the RTT phenotype.

Introduction

Rett syndrome (RTT), a neurodevelopmental disorder, is a leading cause of mental retardation in females with an estimated prevalence of 1 in 10,000 to 15,000 female births. RTT patients develop normally until 6 - 18 month of age, when they start to show symptoms including respiratory irregularities, progressive loss of motor skills, stereotypic hand movements, seizures, and features of autism. Examination of the brain reveals profound microencephaly due, at least in part, to smaller, more densely packed

neurons. Other abnormalities include a reduction in dendritic arborization (Kriaucionis and Bird 2003; Zoghbi 2003). In about 80% of cases, RTT is associated with mutations in the X-linked *MECP2* gene that is subject to inactivation when located on the inactive X-chromosome (Amir et al. 1999). Therefore, heterozygous mutant females are mosaic for MeCP2 deficiency and show a wide range of phenotypes. Males, however, show a more severe phenotype usually involving encephalopathy, motor abnormalities and respiratory dysfunction. They rarely live beyond 2 years (Kriaucionis and Bird 2003).

Mecp2 encodes a protein that binds specifically to methylated CpG dinucleotides and recruits chromatin remodeling complexes that contain the transcriptional repressor Sin3A and histone deacetylases 1 and 2 (Nan et al. 1998). In mouse, the protein localizes to highly methylated pericentromeric heterochromatin (Shahbazian et al. 2002). Though MeCP2 is found in most tissues and cell types, highest expression levels are detected in the brain, where it is primarily present in neurons but not in glia (Coy et al. 1999; Akbarian et al. 2001; Shahbazian et al. 2002). The timing of *Mecp2* expression correlates with the maturation of the central nervous system (CNS) (Shahbazian et al. 2002; Mullaney et al. 2004), and recent reports suggest that MeCP2 may be involved in the formation of synaptic contacts (Cohen et al. 2003). Although biochemical evidence suggests that MeCP2 acts as a global silencer, transcriptional profiling has failed to detect global changes in gene expression (Tudor et al. 2002). A candidate approach has identified *BDNF*, a gene involved in neuronal survival, development, and plasticity, as a target for MeCP2 (Chen et al. 2003). These findings are consistent with MeCP2 playing a role in the maintenance and modulation of neuronal maturity. In particular, MeCP2 may function as a key regulator of activity-dependent neuronal gene expression.

Complete or partial deletion of *Mecp2* in mice leads to a neurological phenotype that is similar but less severe than human RTT (Chen et al. 2001; Guy et al. 2001). Heterozygous females remain healthy into adulthood. In contrast, *Mecp2* mutant males appear normal and healthy at birth but begin to show a phenotype that resembles the human condition at 3 to 8 weeks of age, and die at 6 to 10 weeks of age. Mutant brains show a reduction in brain weight and neuronal cell size but no obvious structural defects or signs of neurodegeneration. Conditional mutation of *Mecp2* in the neural progenitor cells at embryonic day 12 results in a phenotype identical to that of the null mutation (Chen et al. 2001). Mutation of *Mecp2* in the postnatal neurons of restricted regions in the brain leads to a similar although delayed neuronal phenotype, suggesting MeCP2 plays a role in postmitotic neurons (Chen et al. 2001). Here we test the hypothesis that the phenotype is exclusively caused by a neuronal MeCP2 deficiency by placing *Mecp2* expression under the control of a neuron-specific promoter. Overexpression of the *Mecp2* transgene in postmitotic neurons proved to be detrimental and lead to symptoms of severe motor dysfunction. Transgene expression in *Mecp2* mutant mice, however, resulted in a rescue of the RTT phenotype.

Materials and Methods

Gene targeting construct

To introduce the *Mecp2* coding sequence as an in-frame fusion into exon 1 of the *tau* locus, we first cloned a 3.8 Kb *KpnI/EcoRI* fragment from pHV, which contains 14 kb of *tau* genomic sequence (kindly provided by K. Tucker) into pBluescript (Stratagene) generating pTau-KR with a unique *NcoI* cloning site. Next we eliminated a unique *SpeI*

site in the polylinker of pTau-KR by cutting the vector with *Spe*I, treating it with the DNA polymerase I Klenow fragment (New England Biolabs), and religating it. We created pTau-KR-linker by introducing suitable restriction sites to allow the in-frame fusion. We inserted an adapter that destroyed the *Nco*I cloning site while introducing a new *Nco*I site that was shifted by 2 bp, plus a *Spe*I and an *EcoR*V site. The primers used were TAUadapt-F (5'-TTT GGT CAT GAT GCC ATG GAC TAG TCG ATA TCT CAT GAG ATT A-3') and TAU-link-R (5'-TAA TCT CAT GAG ATA TCG ACT AGT CCA TGG CAT CAT GAC CAA A-3'). The 1455 nucleotide-long coding sequence of *Mecp2* was amplified by PCR from IMAGE clone 1446972 (AF158181) and confirmed by sequencing. The PCR primers introduced a modified Kozak sequence including an *Nco*I site (ATTCC**ATG**G was changed to CCACC**ATG**G) and restriction sites that facilitated cloning. Primer RI-MeCP2-F (5'-CGGAATTCCGCCACCATGGTAGCTGGGATGTTAGGG-3') added an *EcoR*I site 5' of the sequence, and primer Xba-MeCP2-R (5'-GCTCTAGAGCTCAGCTAACTCTCTCGGTCACG-3') added an *Xba*I site to the 3' end. To provide the construct with an SV40 late polyadenylation signal (SV40pA), we next cloned the *EcoR*I/*Xba*I MeCP2 fragment into the *EcoR*I/*Spe*I digested vector pZ12-I-PL2 (Ariad). MeCP2-SV40pA was released as an *Nco*I/*Spe*I fragment and cloned into pTAU-KR-linker creating pTAU-MeCP2pA. A neo^R-resistance selectable marker from pPGKNR was isolated using *EcoR*I and *Sal*I, and overhangs were filled in using Klenow. This blunt-ended fragment was cloned into the unique *EcoR*V site of pTAU-MeCP2pA. The targeting vector pTAU-MeCP2pAneo was confirmed by sequencing.

Generation of mice

The targeting vector was linearized with *Sac*II and electroporated into the V6.5 (129 x C57BL/6) F1 embryonic stem (ES) cell line. We picked 96 neomycin-resistant clones of which 63 were analyzed by Southern blots as described (Sambrook 1989). The 5'-external probe consisted of a 528 bp PCR fragment amplified from pHV and was located 1.5 kb upstream of the *Nco*I insertion site in exon 1. The primers were Tau5'-F (5'-GAG CTG CTG CCA TCT TCA C-3') and Tau5'-R (5'-TTT GAT GTG TGC CCT ACA GAA-3'). The 3'-external probe consisted of a 600 bp *BamH*I/*Eco*RI genomic fragment (Tucker et al. 2001) that was located 6.1 kb downstream of the insertion site. An internal probe was used to test for additional non-homologous insertions and consisted of a 200 bp *Pst*I fragment from the neo^R ORF (Tucker et al. 2001). Eleven clones were targeted correctly, which corresponds to a targeting efficiency of 17%. Two clones were used to generate chimeras by injection into (DBA/2 x C57BL/6) F1 blastocysts as described (Hogan 1994). Chimeras were mated to C57BL/6 females, and offspring were analyzed for germline transmission. The heterozygous knock-in strain (*Tau-Mecp2* ki/+) was maintained on a mixed background that was predominantly C57BL/6 but also contained some 129 contributed by the original ES cells used for targeting. To obtain rescued males, *Tau-Mecp2* ki/+ heterozygous males were mated to heterozygous $Mecp^{1lox/+}$ females (Chen et al. 2001) that had been backcrossed to C57BL/6. After germline transmission was confirmed, animals were routinely genotyped by PCR. For the *tau* locus the primer set Tau138 (5'-CTG GCA GAT CTT CCC GTC TA-3'), Tau1078 (5'-TGC CTG ACA GAG TCC AGA TG-3') and Neo1323 (5'-AGG GGA TCC GTC CTG TAA GT-3') gave a 941 bp band for the wild type allele and a 796

bp band for the ki allele. The *Mecp2* allele was determined using primers Nsi-5 (5'-CAC CAC AGA AGT ACT ATG ATC-3'), 2lox-3 (5'-CTA GGT AAG AGC TCT TGT TGA-3') and Nsi-3 (5'-ATG CTG ACA AGC TTT CTT CTA-3'), which generated a 180 bp wild type band and 300 bp band for the 1lox allele.

Immunoblot analysis

Organs were routinely harvested and snap frozen in liquid N_2. Tissues were homogenized with a Polytron homogenizer (Biospec Products Inc.) in a lysis buffer containing 125 mM Tris, 1% SDS (pH 6.8) supplemented with a proteinase inhibitor cocktail (Roche). Protein concentrations were determined with a BCA protein assay kit (Pierce, Rockford, IL). Sample buffer containing bromophenol blue was added to final concentrations of 12.5% glycerol and 0.25% ß-mercaptoethanol. 40μg of protein (unless indicated otherwise) was loaded on 7.5% Tris-HCL acrylamide gels (Ready Gels, BIO-RAD), probed with an anti-MeCP2 rabbit polyclonal antibody (Upstate Biotech) or an anti-GAPDH rabbit polyclonal (Abcam) and visualized using the Amersham ECL system.

Immunohistochemistry

Brains were harvested from 8 weeks old animals, weighed, immersion-fixed for 20 h in 10% phosphate buffered formalin and cryoprotected for 24 - 36 h in 30% sucrose and embedded in OCT. A series of 15 μm sagittal sections were slide mounted and stored at -20°C. For immunohistochemistry slides were blocked in 5% normal goat serum and incubated with rabbit antibody raised against the C-terminal peptide of MeCP2

(Chen et al. 2003) at a dilution of 1:500 and a mouse anti-neuron-specific nuclear protein (NeuN, Chemicon) at a dilution of 1:200. The MeCP2 antibody was detected with a FITC conjugated goat anti-rabbit IgG and the NeuN antibody with a Rhodamine conjugated donkey anti-mouse IgG (both from Jackson Imunoresearch), both at a dilution of 1:500. Sections were finally rinsed in PBS/DAPI (1:10000) and mounted in Vectashield (Vector laboratories). Images were taken with Leica fluorescence microscope under 20x magnification.

RNA samples and RNase Protection assay

Embryo heads and adult brains were harvested, snap frozen in liquid N_2 and subsequently extracted with RNA-Bee (TEL-Test, INC.). Embryos were isolated from *Tau-Mecp2* ki/+ females mated with *Tau-Mecp2* ki/+ males. Embryos of the correct genotype were identified by PCR genotyping using DNA extracted from embryos (10.5 to 15.5 dpc), yolk sac (9.75 dpc) or tail tips. For the RNase protection assay the probe template was generated by PCR from genomic DNA obtained from a transgenic mouse with primers TauExon1-27F (5'-GCCAGGAGTTTGACACAATG-3') and MeCP2-282R (5'-CATACATAGGTCCCCGGTCA-3'). The PCR product was cloned into the vector pCR2.1-TOPO (TOPO TA Cloning Kit, Invitrogen) and subcloned as a *PstI* Fragment into the transcription vector pSP72 (Promega). In vitro transcription was carried out with the Maxiscript T7 Kit (Ambion,). UTP$\propto P^{32}$ (3,000 Ci/mmol, 10 mCi/ml, Amersham Pharmacia) was diluted 1:10 with unlabeled nucleotide and the probes were gel-purified. 2000 cpm of probe and 10 μg of total RNA was used in the assay according to the

manufacturer's specification (Ambion RPA III kit). The samples were resolved on a 6% acrylamide (1:30 bisacrylamide) and quantified by phosphor imaging.

Spontaneous activity measurements

Motor activity was measured using an infrared beam activated movement-monitoring chamber (Opto-Varimax-Mini-A, Columbus Instruments). For each experiment, a mouse was placed in the chamber at least 3 h before recordings started. Movement was monitored during the normal 12 h dark cycle (7 pm to 7 am).

Results

Expression of MeCP2 from the *tau* locus

In order to express MeCP2 specifically in post-mitotic neurons, we designed a targeting construct that places *Mecp2* expression under the control of the promoter of the microtubule-binding-protein, Tau. Tau protein is strongly expressed in neurons (Binder et al. 1985) and the endogenous *tau* locus has been used previously to drive neuron-specific expression of EGFP (Tucker et al. 2001). Homozygous animals mutant for *tau* have been shown to be phenotypically indistinguishable from wild type littermates (Harada et al. 1994). A cDNA containing the *Mecp2* coding sequence including a modified Kozak sequence (Kozak 1986) was placed into exon 1 of the *tau* gene in-frame with the endogenous initiation codon thereby creating a fusion protein that contains the first 31 amino acids of Tau (Tau-MeCP2, Figure 1A). The targeting construct was introduced into V6.5 embryonic stem cells by electroporation,

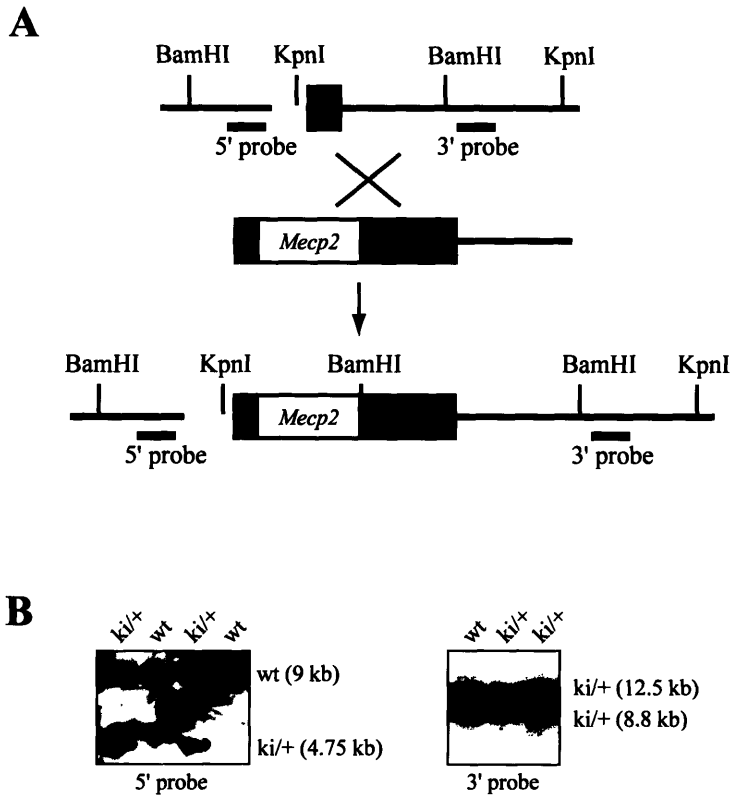


Figure 1. Targeting the *Mecp2* cDNA to the *tau* locus. **(A)** Targeting strategy to insert the *Mecp2* cDNA and the neomycin resistance marker (*NEO*^{R}) into exon 1 of the *tau* locus. The upstream and downstream targeting arms are shown in light and dark grey, respectively. The locations of 5'- and 3'-external probes used for Southern blot analysis are indicated. **(B)** Southern blot analysis of targeted ES cells clones (ki/+) after digestion with *BamHI* (left panel) and *KpnI* (right panel). When hybridized with the 5'-external probe, wild type clones display a 9 kb band. The correct targeting event resulted in a band-shift to 4.75 kb for the targeted allele. Hybridization with the 3'-external probe results in a 8.8 kb wild band and a 12.5 kb band for the targeted allele.

and 63 neomycin-resistant clones were analyzed. Eleven clones were targeted correctly, and two of these (Figure 1B) were used to generate germ line-transmitting chimeras.

To examine and compare the expression pattern of Tau-MeCP2 and endogenous MeCP2, we determined their protein levels in various adult mouse tissues by immunoblot analysis (Figure 2A, B). We found that Tau-MeCP2 expression in the brain was 2-4-fold higher than endogenous MeCP2 (Figure 2A, lanes 4 to 9). Endogenous MeCP2 was expressed highly in lung and spleen, and moderately in kidney and heart and liver (Fig 2B, wt). Tau-MeCP2 protein expression was high in lung and kidney, and was low in heart (Figure 2B, R). Very low expression in liver and spleen was detectable after long exposure times.

MeCP2 localizes to highly methylated centromeric heterochromatin (Nan et al. 1996), and immunohistochemical analysis revealed the typical punctate nuclear MeCP2 staining pattern in wild type brains (Figure 2C). This pattern was not detectable in *Mecp2* mutant animals. Instead a faint diffuse nuclear staining was observed (Figure 2G), which probably arose from the detection of the truncated MeCP2 protein (Chen et al. 2001). The deletion of exon 3 encompasses most of the methyl-CpG-binding domain (MBD) but the C-terminal epitope including the nuclear localization sequence (NLS) remains largely intact (Chen et al. 2001). We found that in *Mecp2* mutant mice heterozygous for the *Tau-Mecp2* transgene (rescued animals) MeCP2 distribution was indistinguishable from wild type MeCP2 (Figure 2K), suggesting correct localization of the fusion protein. This was to be expected, because the MBD and the NLS have both been shown to be necessary and possibly sufficient for the specific localization of MeCP2 (Nan et al. 1996). The axonal localization signal of Tau is located in the 3'-untranslated region (UTR) and therefore not

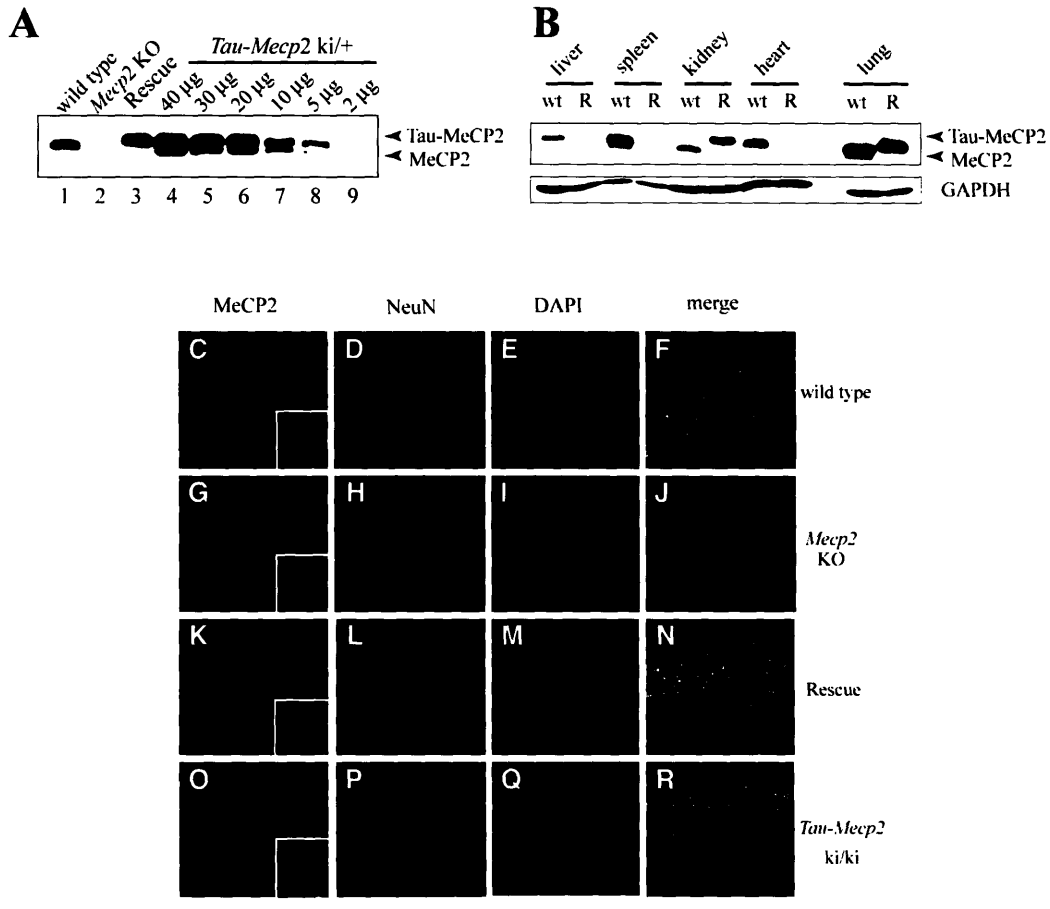


Figure 2. Expression of *Mecp2* from the *tau* locus. (**A**) Immunoblot analysis of protein extracted from whole brain samples. As controls, 40 μg of protein were loaded from wild type and MeCP2 mutant (MeCP2 KO) animals, and MeCP2 mutant animals heterozygous for the *Tau-Mecp2* transgene (Rescue). The protein extract from a MeCP2 wild type animal heterozygous for the transgene (Tau-MeCP2 ki/+) was loaded as serial dilutions as indicated. The Tau-MeCP2 fusion protein contains 31 amino acis of the Tau protein which results in a band shift. (**B**) Endogenous MeCP2 (wild type animal, wt) is highly expressed in the lung and spleen, and less in the liver, kidney and heart. Tau-MeCP2 expression (rescued animal, R) is high in lung and kidney. Low level expression is also detectable in the heart. In the liver and spleen, Tau-MeCP2 is detectable only after long exposure times. (**C - R**) Tau-MeCP2 expression is neuron-specific in the brain and localizes to heterochromatic foci. Double labled immuno-fluorescence of MeCP2 (green) and neuron-specific nuclear protein (NeuN, red) from wild type (**C - F**), MeCP2 KO (**G - J**), rescued (**K - N**) and Tau-MeCP2 homozygous (Tau-MeCP2 ki/ki, **O - R**) hippocampi of adult animals. Punctate MeCP2 staining is detectable in wild type animals (**C**) as well as animals carrying the *Tau-Mecp2* transgene (**K, O**). Inserts in panels **C, G, K, O** show enlargements of a small number of cells to illustrate MeCP2 staining. Endogenous MeCP2 as well as Tau-MeCP2 expression overlaps with NeuN immunoreactivity (**F, N, R**). Only weak diffuse MeCP2 staining is present in MeCP2 null cells (**G**). Nuclear DAPI stain is shown in blue (**E, I, M, Q**).

part of the Tau-MeCP2 mRNA (Aronov et al. 2001). In addition we found that all MeCP2-positive cells were double labeled with an antibody against the neuron-specific marker protein NeuN (Figure2E, M, Q) (Mullen et al. 1992), indicating that endogenous MeCP2 as well as Tau-MeCP2 were detectable only in neuronal cells (Figure 2F, N, R).

Timing of Tau-MeCP2 expression

In order to quantify the relative expression of *Mecp2* and *Tau-Mecp2* during embryonic development we carried out an RNase protection assay (RPA) on embryos harvested between 9.75 and 15.5 days past coitum (dpc). The probe was designed to span the *Tau-Mecp2* junction including 82 bp of *tau* exon 1 and 164 bp of the *Mecp2* cDNA. This allowed the detection of the *Tau-Mecp2* RNA as well as the endogenous *Mecp2* and *tau* transcripts (Figure 3A). *Mecp2* RNA was detectable as early as 9.75 dpc with levels steadily increasing until they reached adult levels at 15.5 dpc (Figure 3B, compare lane 5 with lanes 6 and 9). *Tau-Mecp2* expression followed closely that of endogenous *tau* with RNA being present starting at 10.5 dpc and, like *Mecp2*, reaching maximum expression levels at 15.5 dpc (Figure 3B, compare lane 5 with lanes 6 and 9). The timing of *Tau-Mecp2* expression therefore appeared to closely mimic that of *Mecp2*.

Quantification of expression levels revealed that in 15.5 dpc embryonic heads as well as in adult brains the amount of *Mecp2* RNA appeared to be similar to *Tau-Mecp2* RNA. On the protein level, however, Tau-MeCP2 was 2 - 4 times more abundant (Figure 2A), suggesting a difference in either translation efficiency of the transcripts or protein stability.

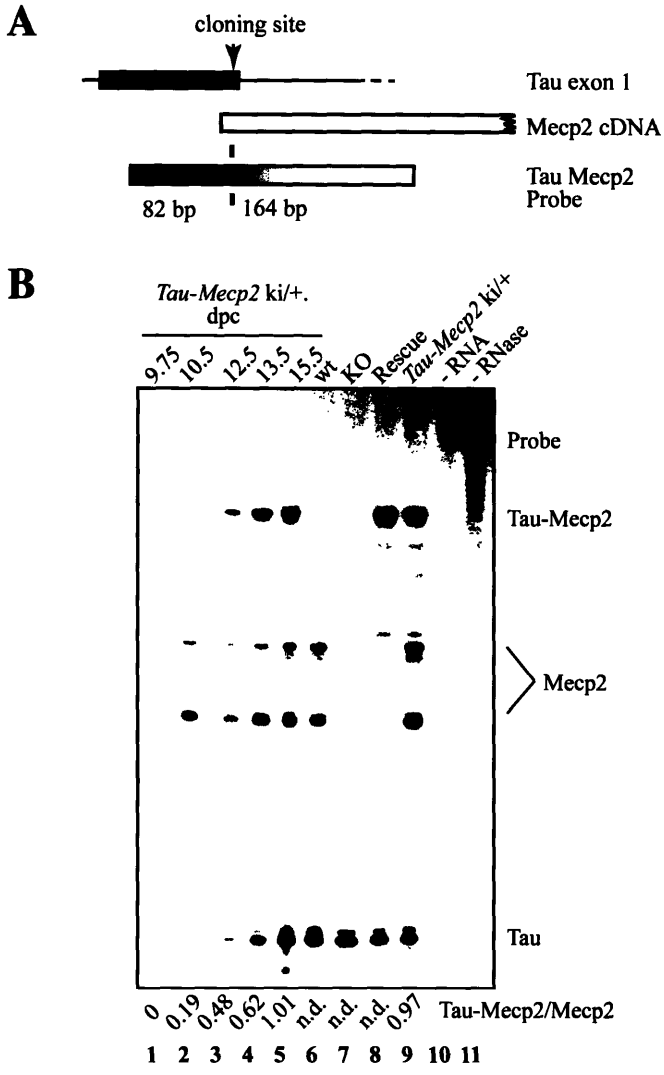


Figure 3. Timing of *Tau-Mecp2* expression. **(A)** Probe design for the RNase protection assay. The probe overlaps the *Tau-Mecp2* junctions of the transgene and consists of 82 bp of *tau* sequence and 164 bp of *Mecp2* cDNA sequence (26 bp from exon 2 and 138 bp of exon 3). **(B)** RNase protection assay. RNA was extracted from embryos (9.75 dpc) and embryonic heads (10.5 - 11.5 dpc) of *Mecp2* wild type animals heterozygous for the *Tau-Mecp2* transgene (lanes 1 - 5), or from brains of adult control animals (lanes 6 - 9). Lane 10 shows the no sample control, and lane 11 is 15% of loading of the no RNase control. Wild type *Mecp2* RNA is detectable as two bands that represent different splice products (lanes 1 - 6 and 9). No *Mecp2* signal is detectable in animals carrying the mutant *Mecp2* allele that lacks exon 3 (lanes 7 and 8). The ratio of *Tau-Mecp2* RNA expression to endogenous *Mecp2* RNA as quantitated by phosphorimaging is given below the gel.

Overexpression of MeCP2 is detrimental

Wild type animals heterozygous for the transgene were fertile and healthy with no obvious phenotypic abnormalities. In contrast, wild type as well as *Mecp2* mutant animals homozygous for the transgene suffered from profound motor dysfunction including side-to-side swaying, tremors and gait ataxia. There was no reduction in brain weight in animals at 2 - 4 month of age (data not shown). Animals were of normal weight at birth but by weaning age pups were severely runted and up to 60% smaller than wild type littermates (Figure 4B). The condition improved when animals were fostered in small groups, indicating that the failure to thrive was largely due to their inability to compete with littermates for food. Once the animals reached weaning age, the phenotype appeared to be stable but they remained small and do not mate. By 9 month ataxia and tremors appeared to have intensified. The animals were emaciated, had a disheveled look, and developed additional problems such as cataracts and lesions. The lesions were probably caused by excessive stereotypic scratching. However, no premature death was observed. Preliminary data suggest that *Mecp2*-mutant animals homozygous for the transgene were slightly less affected compared to *Mecp2* wild type animals homozygous for the transgene as judged by weight and appearance (data not shown). The phenotype was specific and not due to the lack of Tau protein because animals homozygous for other transgenes targeted to exon 1 of the *tau* locus were phenotypically normal (Tucker et al. 2001, and unpublished data).

Immunohistochemical analysis of brain sections revealed a punctate MeCP2 staining pattern (Figure 2M, page 131), suggesting that the level of Tau-MeCP2 overexpression had no effect on Tau-MeCP2 localization in adult mice.

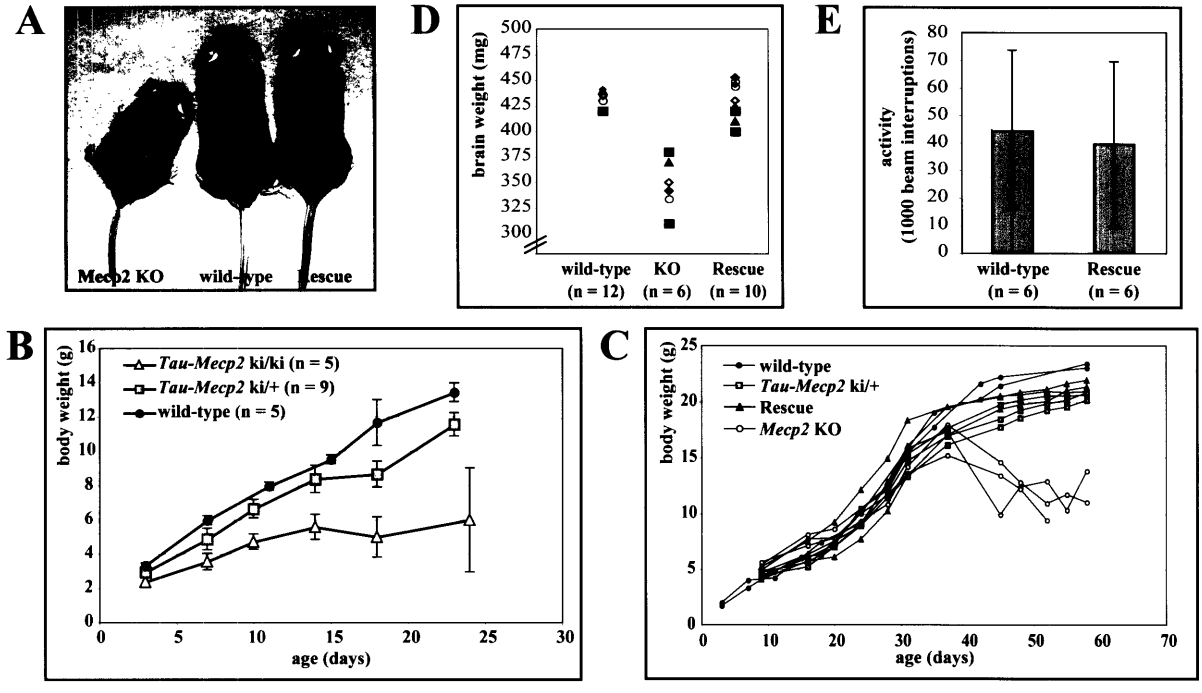


Figure 4. Expression of MeCP2 in postmitotic neurons rescues the RTT phenotype, but overexpression of MeCP2 is detrimental. **(A)** Rescued animals are indistinguishable from their wild type littlermates. Animals are shown at 8 weeks of age. **(B)** *Mecp2* wild type animals heterozygous for *Tau-Mecp2* (*Tau-Mecp2* ki/+) show normal physical development. *Mecp2* overexpressing animals (*Tau-Mecp2* ki/ki) have a normal birth weight but progressively lose weight so that animals are up to 60% smaller than wild type littermates at weaning age. **(C)** Neuronal expression of MeCP2 rescues the weight-loss phenotype. *Mecp2* null mutants start to lose weight at about 5 weeks of age. The weight of rescued as well as *Tau-Mecp2* ki/+ animals is comparable to wild type littermates. **(D)** Rescued animals have a brain weight comparable to wild type littermates. In contrast, *Mecp2* mutant animals show a 15 - 18% recuction in brain weight. **(E)** Spontaneous activity of rescued animals is comparable with the activity of wild type littermates as measured by an infrared beam activated movement detector.

Previous reports detected diffuse nuclear MeCP2 staining at 10 dpc which became increasingly punctate through 16.5 dpc (Shahbazian et al. 2002b). Our results suggest that MeCP2 dosage is critical, and that overexpression of MeCP2 4 - 6-fold above wild type level is detrimental to the health of the animal.

Expression of MeCP2 in postmitotic neurons rescues the RTT phenotype

MeCP2 mutant mice die on average around 10 weeks of age (Chen et al. 2001). In contrast, mutant animals heterozygous for the transgene were healthy and fertile and, in various crosses, passed on all alleles at the expected Mendelian ratios (data not shown). They were phenotypically indistinguishable from their wild type littermates and displayed no RTT-like symptoms (Figure 4A). A distinct feature of *Mecp2* mutant animals are weight abnormalities, and in agreement with previous reports (Guy et al. 2001), *Mecp2* mutant animals were severely underweight from four to five weeks (Figure 4C). In contrast, the weight of rescued animals was indistinguishable from that of wild type littermates throughout postnatal development (Figure 4C).

Human RTT as well as the *Mecp2* mutant mouse phenotype, are characterized by a decrease in head growth and neuronal cell size as well as increased cell packing density throughout the brain (Jellinger et al. 1988; Bauman et al. 1995; Chen et al. 2001; Hagberg et al. 2001). As published previously, we found a 14 - 18 % reduction in brain weight in 8 to 13 week old *Mecp2* mutant animals. In contrast, rescued animals showed no difference in brain weight compared to wild type littermates, even when brains were harvested as late as 5 months of age (Figure 4D). Hypoactivity is another characteristic of *Mecp2* mutant mice (Chen et al. 2001; Guy et al. 2001; Shahbazian et al. 2002a). We tested

exploratory response as well as total nocturnal activity by placing animals in cages equipped with an infrared beam movement detector. The age of the mice tested ranged from 4 to 6.5 months, long after the onset of RTT symptoms in *Mecp2* mutant mice, and in all cases after *Mecp2* mutant littermates had died. In both behavioral tests, rescued animals were indistinguishable from wild type littermates (Figure 4E).

In summary, our data indicate that rescued animals do not display any of the common RTT phenotypes that have been described for *Mecp2* mutant animals. Therefore expression of *Mecp2* in postmitotic neurons is sufficient to alleviate the RTT phenotype, even at slightly elevated protein levels.

Discussion

Mutation of the methyl-binding protein MeCP2 leads to RTT in humans, and increasing evidence suggests that this disorder is primarily caused by a defect in neuronal maintenance and maturation. In a mouse model, loss of MeCP2 expression leads to a RTT-like phenotype including tremors, heavy breathing, hypoactivity and smaller brain size associated with smaller, more densely packed neurons. After the onset of symptoms at 4 - 6 weeks of age and progressive physical deterioration, the animals typically die around 10-12 weeks of age (Chen et al. 2001; Guy et al. 2001). Deletion of *Mecp2* exon 3, encoding most of the methyl-binding domain in neural progenitor cells, produces a phenotype that is indistinguishable from the germline mutation. In contrast, loss of MeCP2 function in postmitotic neurons of the postnatal brain led to a delayed onset of symptoms by up to 3 months (Chen et al. 2001). MeCP2 is primarily expressed in mature

neuronal populations (Coy et al. 1999; Akbarian et al. 2001; Shahbazian et al. 2002b) with an expression pattern that follows the maturation of the CNS (Shahbazian et al. 2002b). Recent studies suggest that MeCP2 may be important for the maintenance and modulation of synapses (Cohen et al. 2003). In addition, neurons of RTT individuals appear to show a reduction in dendritic complexity (Armstrong et al. 1995). Also, recent experiments in Xenopus revealed a specific function of MeCP2 in early neural development (Stancheva et al. 2003).

Here we provide functional evidence for the requirement of MeCP2 in postmitotic neurons by placing the *Mecp2* cDNA under the control of the endogenous promoter of the microtubule-binding-protein, Tau. By introducing the cDNA into exon 1 of the *tau* gene in-frame with the endogenous start codon we created a fusion protein that contained the first 31 amino acids of Tau (Tau-MeCP2). Tau-MeCP2 was expressed at a level 2 - 4-fold higher than endogenous MeCP2, and the onset of expression correlated closely with endogenous *tau* expression being first detectable at 10.5 dpc. In agreement with previous reports, endogenous *Mecp2* RNA was detectable at 9.75 dpc. Immunohistochemical analysis of brain sections revealed that Tau-MeCP2 localized like endogenous MeCP2 to heterochromatic foci of postmitotic neurons in adult mice.

One copy of the *Tau-Mecp2* transgene led to the complete rescue of all assessed phenotypes in *Mecp2* mutant animals. They had a normal lifespan and showed normal physical development as judged by weight. The rescued animals had normal brain weight and showed no signs of hypoactivity, tremors or any other physical symptoms typically associated with RTT, even at an advanced age of 6 month or more. We therefore conclude that expression of MeCP2 in postmitotic neurons is sufficient to alleviate the

phenotype in MeCP2 mutant mice. Our data strongly suggest that the RTT phenotype is caused by lack of MeCP2 in the brain. Nevertheless, we cannot fully exclude a contribution of MeCP2 in some peripheral tissues, because the physiological basis of some symptoms in Rett patients as well as in *Mecp2* mutant animals still remains to be determined.

It has been reported previously that there is no obvious correlation between MeCP2 protein and RNA levels in adult tissues, suggesting that *Mecp2* translation may be postranscriptionally regulated (Shahbazian et al. 2002b). The endogenous *Mecp2* sequence contains a 182 bp 5'-UTR and a highly conserved 8.5 Kb 3'-UTR with alternative polyadenylation (pA) signals. Depending on which pA signal is used, either a 1.9 kb or a 10 kb transcript is produced. In the brain, the 10 kb transcript is the predominant one (Reichwald et al. 2000). In our experiments, the *Tau-Mecp2* construct did not include either of the untranslated regions. Since both the short and the long transcript have similar half-lives (Reichwald et al. 2000) it is possible that the 3'-UTR contains regulatory elements important for translation efficiency. In fact, the *Mecp2* 3'-UTR has recently been predicted to contain potential target sequences for microRNAs (miRNAs), which confer posttranslational repression (Lewis et al. 2003).

Moderate overexpression of MeCP2 in wild type mice heterozygous for *Tau-Mecp2* had no adverse effect. In contrast, homozygosity for the transgene led to severe motor dysfunction in a wild type as well as a *Mecp2* mutant background. The result was an impaired ability to compete with littermates for food, leaving the pups severely runted. Symptoms included tremors, gait ataxia and side-to side swaying. These observations suggest that neurons can tolerate a 2 - 3-fold higher level of MeCP2 expression but that

higher expression levels result in obvious detrimental consequences. Interestingly, *Mecp2* mutant animals homogzygous for the transgene showed a slightly less severe phenotype, which emphasizes the importance of MeCP2 dosage.

We found that MeCP2 expression was dispensable through early embryonic development. In agreement with this, previous studies described that MeCP2 expression was undetectable before 9.5 dpc on the RNA or protein level, indicating that it plays no essential role in early embryonic development (Shahbazian et al. 2002b; Kantor et al. 2003). Moreover, we find that MeCP2 is not necessary for neurogenesis, despite the fact that MeCP2 appears to be expressed in some neuronal precursor cells of the rodent brain (Jung et al. 2003). The delayed phenotype of the conditional mutant observed after postnatal loss of MeCP2 function in postmitotic neurons may be attributable to the fact that the expression of CamKII-driven Cre-recombinase used in this experiment is limited to certain parts of the postnatal brain, including the forebrain, hippocampus and brainstem. It is only marginally active in the cerebellum, a region where *Mecp2* mRNA is highly expressed (Jung et al. 2003). In contrast, loss of MeCP2 function in neural progenitor cells affects essentially all neurons. It is also possible that the abatement of the phenotype of CamKII-Cre mutants is due to the presence of MeCP2 in neurons during prenatal development or the lack of Cre expression in a crucial cell type.

Our results are consistent with MeCP2 playing no essential role in the early stages of brain development. It is possible in fact, that neurons are functionally normal in the young postnatal patient, and that neural dysfunction becomes manifest only later due to prolonged MeCP2 deficiency. If correct, therapeutic strategies could be aimed at preventing postnatal dysfunctions from developing in *MECP2* mutant neurons.

Acknowledgments

We thank members of the Jaenisch lab for discussions and critical comments on the manuscript. We also thank K. Tucker for providing reagents, J. Dausman for blastocyst injections, R. Flannery for animal care, and A. Caron (MIT Center for Cancer Reasearch histology facility) for technical support. This work was conducted using the W. M. Keck biological imaging facility at the Whitehead Institute and was supported by NIH grant CA87869 to R.J. and the Rett Syndrome Research Foundation.

References

Akbarian, S., R.Z. Chen, J. Gribnau, T.P. Rasmussen, H. Fong, R. Jaenisch, and E.G. Jones. 2001. Expression pattern of the Rett syndrome gene MeCP2 in primate prefrontal cortex. *Neurobiol Dis* **8:** 784-791.

Amir, R.E., I.B. Van den Veyver, M. Wan, C.Q. Tran, U. Francke, and H.Y. Zoghbi. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl- CpG-binding protein 2. *Nat Genet* **23:** 185-188. java/Propub/genetics/ng1099_1185.fulltext java/Propub/genetics/ng1099_1185.abstract.

Armstrong, D., J.K. Dunn, B. Antalffy, and R. Trivedi. 1995. Selective dendritic alterations in the cortex of Rett syndrome. *J Neuropathol Exp Neurol* **54:** 195-201.

Aronov, S., G. Aranda, L. Behar, and I. Ginzburg. 2001. Axonal tau mRNA localization coincides with tau protein in living neuronal cells and depends on axonal targeting signal. *J Neurosci* **21:** 6577-6587.

Bauman, M.L., T.L. Kemper, and D.M. Arin. 1995. Microscopic observations of the brain in Rett syndrome. *Neuropediatrics* **26:** 105-108.

Binder, L.I., A. Frankfurter, and L.I. Rebhun. 1985. The distribution of tau in the mammalian central nervous system. *J Cell Biol* **101:** 1371-1378.

Chen, R.Z., S. Akbarian, M. Tudor, and R. Jaenisch. 2001. Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat Genet* **27:** 327-331.

Chen, W.G., Q. Chang, Y. Lin, A. Meissner, A.E. West, E.C. Griffith, R. Jaenisch, and M.E. Greenberg. 2003. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302:** 885-889.

Cohen, D.R., V. Matarazzo, A.M. Palmer, Y. Tu, O.H. Jeon, J. Pevsner, and G.V. Ronnett. 2003. Expression of MeCP2 in olfactory receptor neurons is developmentally regulated and occurs before synaptogenesis. *Mol Cell Neurosci* **22:** 417-429.

Coy, J.F., Z. Sedlacek, D. Bachner, H. Delius, and A. Poustka. 1999. A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3"-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. *Hum Mol Genet* **8:** 1253-1262.

Guy, J., B. Hendrich, M. Holmes, J.E. Martin, and A. Bird. 2001. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet* **27:** 322-326.

Hagberg, G., Y. Stenbom, and I.W. Engerstrom. 2001. Head growth in Rett syndrome. *Brain Dev* **23 Suppl 1:** S227-229.

Harada, A., K. Oguchi, S. Okabe, J. Kuno, S. Terada, T. Ohshima, R. Sato-Yoshitake, Y. Takei, T. Noda, and N. Hirokawa. 1994. Altered microtubule organization in small-calibre axons of mice lacking tau protein. *Nature* **369:** 488-491.

Hogan, B., Beddington, R., Costantini, F., & Laciy, E. 1994. *Manipulating the Mouse Embryo: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Jellinger, K., D. Armstrong, H.Y. Zoghbi, and A.K. Percy. 1988. Neuropathology of Rett syndrome. *Acta Neuropathol (Berl)* **76:** 142-158.

Jung, B.P., D.G. Jugloff, G. Zhang, R. Logan, S. Brown, and J.H. Eubanks. 2003. The expression of methyl CpG binding factor MeCP2 correlates with cellular differentiation in the developing rat brain and in cultured cells. *J Neurobiol* **55:** 86-96.

Kantor, B., K. Makedonski, R. Shemer, and A. Razin. 2003. Expression and localization of components of the histone deacetylases multiprotein repressory complexes in the mouse preimplantation embryo. *Gene Expr Patterns* **3:** 697-702.

Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44:** 283-292.

Kriaucionis, S. and A. Bird. 2003. DNA methylation and Rett syndrome. *Hum Mol Genet* **12 Spec No 2:** R221-227.

Lewis, B.P., I.H. Shih, M.W. Jones-Rhoades, D.P. Bartel, and C.B. Burge. 2003. Prediction of mammalian microRNA targets. *Cell* **115:** 787-798.

Mullaney, B.C., M.V. Johnston, and M.E. Blue. 2004. Developmental expression of methyl-CpG binding protein 2 is dynamically regulated in the rodent brain. *Neuroscience* **123:** 939-949.

Mullen, R.J., C.R. Buck, and A.M. Smith. 1992. NeuN, a neuronal specific nuclear protein in vertebrates. *Development* **116:** 201-211.

Nan, X., H.H. Ng, C.A. Johnson, C.D. Laherty, B.M. Turner, R.N. Eisenman, and A. Bird. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393:** 386-389.

Nan, X., P. Tate, E. Li, and A. Bird. 1996. DNA methylation specifies chromosomal localization of MeCP2. *Mol Cell Biol* **16:** 414-421.

Reichwald, K., J. Thiesen, T. Wiehe, J. Weitzel, W.A. Poustka, A. Rosenthal, M. Platzer, W.H. Stratling, and P. Kioschis. 2000. Comparative sequence analysis of the MECP2-locus in human and mouse reveals new transcribed regions. *Mamm Genome* **11:** 182-190.

Sambrook, J., Fritsch, E.F. & Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shahbazian, M., J. Young, L. Yuva-Paylor, C. Spencer, B. Antalffy, J. Noebels, D. Armstrong, R. Paylor, and H. Zoghbi. 2002a. Mice with truncated MeCP2

recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron* **35:** 243-254.

Shahbazian, M.D., B. Antalffy, D.L. Armstrong, and H.Y. Zoghbi. 2002b. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum Mol Genet* **11:** 115-124.

Stancheva, I., A.L. Collins, I.B. Van den Veyver, H. Zoghbi, and R.R. Meehan. 2003. A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos. *Mol Cell* **12:** 425-435.

Tucker, K.L., M. Meyer, and Y.A. Barde. 2001. Neurotrophins are required for nerve growth during development. *Nat Neurosci* **4:** 29-37.

Tudor, M., S. Akbarian, R.Z. Chen, and R. Jaenisch. 2002. Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. *Proc Natl Acad Sci U S A* **99:** 15536-15541.

Zoghbi, H.Y. 2003. Postnatal neurodevelopmental disorders: meeting at the synapse? *Science* **302:** 826-830.

Chapter 5

Perspectives

How does *Tsix* regulate *Xist*?

Much progress has been made in recent years in identifying and understanding *cis*-acting elements that control the initiation X-inactivation. We tested the hypothesis that asynchronous replication prior to X-inactivation may be one of the first hallmarks that distinguished the future Xa from the future Xi (chapter 2). However, we found that although X chromosomes in undifferentiated ES cells replicate asynchronously, their replication timing does not correlate with X chromosome choice. At the center of the X-inactivation process lies the *Xist* RNA, which is expressed at low levels from both X chromosomes prior to X-inactivation, and is stabilized on the future Xi during X-inactivation. Our work shows clearly that *Xist* is not only necessary and sufficient to induce X-inactivation but that it is also an important choice element (chapter 2). X chromosomes that do not have an intact *Xist* gene are always directly chosen as the Xa, consistent with the hypothesis that *Xist* RNA decreases the affinity of a *cis*-linked counting element for the blocking factor.

Xist is negatively regulated *in cis* by its antisense gene *Tsix* (Lee and Lu 1999; Sado et al. 2001), and throughout the process of X-inactivation, the sense and antisense

transcripts show a dynamic pattern of expression that reflects their regulatory relationship (Lee et al. 1999). The experiments described in chapter 3 establish the importance of active transcription through the *Xist* locus for *Tsix* function but how *Tsix* blocks *Xist* stabilization is one of the most compelling questions in the field. Several mechanisms of *Tsix* action are conceivable (Lee 2003). 1) *Tsix* transcription could be purely incidental and reflect the organization of a chromatin domain; 2) *Tsix* may act through transcriptional interference and provide a repressive force on *Xist* expression; 3) The *Tsix* RNA itself may be functional and hybridize to the sense transcript, promoting its degradation or blocking *Xist*'s silencing domains. *Tsix* is a spliced and polyadenylated RNA that contains 4 exons (Sado et al. 2001); however, recent work suggests that only about half of *Tsix* transcripts are spliced at the known exon-intron boundaries (Shibata and Lee 2003). Intriguingly, only exon 4 of *Tsix* overlaps with the *Xist* open reading frame and is complementary to the recently identified silencing domain of *Xist* (Wutz et al. 2002). We found that modulation *Tsix* transcription downstream of the third exon is sufficient to affect *Tsix* function (chapter 3) suggesting that either *Tsix* acts by transcription through *Xist* per se or that *Tsix* acts a functional RNA. In the latter case, exon 4 would be necessary and sufficient for *Tsix* function. An example for a functional antisense RNA is found at the imprinted *Igf2* locus where the paternal antisense transcript *Air* originates within an intron of *Igf2* and is transcribed for more than 108 kb off the opposite strand. Paternal *Air* expression is not only necessary to repress *Igf2* but also silences two non-overlapping genes, *Slcc22a2* and *Slcc22a3* (Zwart et al. 2001; Sleutels et al. 2002).

In order to test the functionality of the *Tsix* RNA, a targeting construct could be designed to express all *Tsix* exons while preventing transcription through the *Xist* locus. If *Tsix* acts as a functional RNA, X-inactivation should be unaffected. Skewing of X-inactivation toward the mutated X, on the other hand would suggest that transcription through the *Xist* gene is necessary for *Tsix* function. However, this latter outcome does not entirely exclude a contribution of the *Tsix* transcript itself. If *Tsix* arises as a functional RNA, the next step will be to identify how this transcript interferes with *Xist* function.

The role of MeCP2 in RTT

Mutations of the methyl-binding protein *MECP2* cause RTT in humans. Considerable effort has been devoted to studying the role of MeCP2 in the brain, and it is now commonly accepted that MeCP2 is involved in the maintenance of synaptic function and plasticity. However, how MeCP2 deficiency causes this devastating disease, and whether a cure of the postnatal brain is possible, still remains unclear. Our work suggests that the neurons are functionally normal at birth, and that neural dysfunction is caused by prolonged MeCP2 deficiency and becomes manifest only later (chapter 4). To understand the role of MeCP2 in postnatal development, and to assess the feasibility of postnatal intervention in RTT patients, it is important to design a system that allows brain-specific expression of MecP2 in the postnatal brain. Experiments outlined in the appendix describe efforts to use an inducible system (Gossen et al. 1995) to efficiently control *Mecp2* expression in the brain using doxcycycline. We found that the system is limited by the ability of the drug to cross the blood-brain barrier, the level of transactivator

expression in the brain, and the toxicity of the activated transcription factor in peripheral tissues. To avoid these problems, the transactivator needs to be placed under the control of a strong, neuron-specific promoter. An excellent candidate is the promoter of the CaMKIIα gene, which encodes α-subunit of a serine-thronine kinase that is predominantly expressed in neurons and involved in the regulation of a diverse set of cellular processes, including synaptic plasticity (Fox 2003). Efforts to develop such an inducible system for MeCP2 expression are currently under way in collaboration with Caroline Beard in the lab.

Clearly, a lot remains to be understood about the role of MeCP2 in the brain before the task of treating RTT can be tackled. Using a candidate approach, only one MeCP2 target gene has been identified so far, and much of the knowledge on MeCP2 function stems from *in vitro* experiments that may not adequately reflect the *in vivo* situation. The majority of *in vivo* evidence relies on deducing MeCP2 function from its developmental expression pattern. Understanding MeCP2-function is complicated by the existence of multiple MeCP2 splice-forms that arise from alternative codon usage as well as the utilization of different polyadenylation sites. The different transcripts and isoforms appear to be expressed in a tissue-specific manner but their roles remain unclear. For example, the recently identified isoform of MeCP2 that includes exon 1 rather than exon 2 (exon 1-form) appears to be 10 times more abundant in the brain than the previously identified form that includes all exons (exon 2-from) (Kriaucionis and Bird 2004; Mnatzakanian et al. 2004). Both proteins, however, contain identical functional domains, and our work has shown that the expression exon 2-form alone is able to rescue RTT in the mouse (chapter 4), which raises questions about the purpose of alternative codon

usage. The *Mecp2* gene contains a conserved 8.5 kb long 3'-UTR with alternative polyadenylation sites. The differential use of these sites results in the expression of a short 1.9 kb and long 10 kb transcript, the longer of which is the predominant transcript in the brain (Coy et al. 1999; Reichwald et al. 2000).

Recently, the *Mecp2* gene has been found to be potential target for regulation by microRNAs (miRNAs) (Lewis et al. 2003). MiRNAs are an abundant family of 21 - 22 nucleotide non-protein coding RNAs that play an important role in gene regulation by pairing to mRNAs to specifiy posttranscriptional repression (Bartel and Chen 2004). Target repression by miRNAs is be an extremely responsive and sensitive way to ensure uniform gene expression and to customize the expression level for each distinct cell type (Bartel and Chen 2004). This fine-tuning of target expression is particularly relevant for proteins that might have a narrow window of optimal expression and where several-fold more or less would have undesired effects. MeCP2 expression is tightly regulated during neuronal differentiation, and we show in chapter 4 that a 4 – 6-fold overexpression has dire consequences, suggesting that MeCP2 is one of those proteins whose expression needs to be tightly regulated. To specify repression, metazoan miRNAs appear to require only short stretches of complementarity to an mRNA (referred to as "miRNA seed"). The *Mecp2* 3'-UTR contains 3 seed-matches for the miRNAs miR-19a and miR-130 (Lewis et al. 2003). Coexpression studies using a reporter system that tested the two 5' target sites showed that MeCP2 is not regulated by miR-19a (Lewis et al. 2003). However, disruption of these target sites leads to a 3-fold increase in expression when the reporter construct was coexpressed with miR-130 (I-hung Shih, personal communication). Interestingly, we found that the Tau-MeCP2 fusion protein that does not contain the

endogenous 3' UTR is translated 2 - 4-fold more efficiently than endogenous MeCP2 (chapter 4, compare RNA levels in Figure 3 on page 128 with protein levels in Figure 2 on page 126).

RTT bears striking similarities to other neurodevelopmental disorders as recently described by Zoghbi (2003). For instance, some symptoms of autism include late onset, reduced dendritic arborization of some neurons, social withdrawal, and stereotypic behavior. Both diseases appear to be disorders of neuronal modulation or maintenance (Zoghbi 2003). Angelman syndrome, a neurodevelopmental disorder that is characterized by developmental delay, movement disorder, tremors, and hand-flapping among other symptoms, is caused by mutations in the imprinted *UBE3A* gene. Proper regulation of this genes depends on DNA methylation, and some children with *MECP2* mutations display an Angleman phenotype (Watson et al. 2001). *MECP2* mutations have also been found in other disorders including neonatal onset encephalopathy and X-linked recessive mental retardation (MRX) (Hammer et al. 2002). It is becoming apparent that these disorders may share at least some molecular pathways, which may not only lead to the identification of new MeCP2 target genes but will also help to unravel the complex pathogenesis of neurodevelopmental disorders.

References

Bartel, D.P. and C.Z. Chen. 2004. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat Rev Genet* **5:** 396-400.

Coy, J.F., Z. Sedlacek, D. Bachner, H. Delius, and A. Poustka. 1999. A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3"-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. *Hum Mol Genet* **8:** 1253-1262.

Fox, K. 2003. Synaptic plasticity: the subcellular location of CaMKII controls plasticity. *Curr Biol* **13:** R143-145.

Gossen, M., S. Freundlieb, G. Bender, G. Muller, W. Hillen, and H. Bujard. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268:** 1766-1769.

Hammer, S., N. Dorrani, J. Dragich, S. Kudo, and C. Schanen. 2002. The phenotypic consequences of MECP2 mutations extend beyond Rett syndrome. *Ment Retard Dev Disabil Res Rev* **8:** 94-98.

Kriaucionis, S. and A. Bird. 2004. The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res* **32:** 1818-1823.

Lee, J.T. 2003. Molecular links between X-inactivation and autosomal imprinting: X-inactivation as a driving force for the evolution of imprinting? *Curr Biol* **13:** R242-254.

Lee, J.T., L.S. Davidow, and D. Warshawsky. 1999. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* **21:** 400-404.

Lee, J.T. and N. Lu. 1999. Targeted mutagenesis of Tsix leads to nonrandom X-inactivation. *Cell* **99:** 47-57.

Lewis, B.P., I.H. Shih, M.W. Jones-Rhoades, D.P. Bartel, and C.B. Burge. 2003. Prediction of mammalian microRNA targets. *Cell* **115:** 787-798.

Mnatzakanian, G.N., H. Lohi, I. Munteanu, S.E. Alfred, T. Yamada, P.J. MacLeod, J.R. Jones, S.W. Scherer, N.C. Schanen, M.J. Friez, J.B. Vincent, and B.A. Minassian. 2004. A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat Genet* **36:** 339-341.

Reichwald, K., J. Thiesen, T. Wiehe, J. Weitzel, W.A. Poustka, A. Rosenthal, M. Platzer, W.H. Stratling, and P. Kioschis. 2000. Comparative sequence analysis of the MECP2-locus in human and mouse reveals new transcribed regions. *Mamm Genome* **11:** 182-190.

Sado, T., Z. Wang, H. Sasaki, and E. Li. 2001. Regulation of imprinted X-chromosome inactivation in mice by Tsix. *Development* **128:** 1275-1286.

Shibata, S. and J.T. Lee. 2003. Characterization and quantitation of differential Tsix transcripts: implications for Tsix function. *Hum Mol Genet* **12:** 125-136.

Sleutels, F., R. Zwart, and D.P. Barlow. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415:** 810-813.

Watson, P., G. Black, S. Ramsden, M. Barrow, M. Super, B. Kerr, and J. Clayton-Smith. 2001. Angelman syndrome phenotype associated with mutations in MECP2, a gene encoding a methyl CpG binding protein. *J Med Genet* **38:** 224-228.

Wutz, A., T.P. Rasmussen, and R. Jaenisch. 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* **30:** 167-174.

Zoghbi, H.Y. 2003. Postnatal neurodevelopmental disorders: meeting at the synapse? *Science* **302:** 826-830.

Zwart, R., F. Sleutels, A. Wutz, A.H. Schinkel, and D.P. Barlow. 2001. Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. *Genes Dev* **15:** 2361-2366.

Appendix

Inducible expression of MeCP2 in the brain

Respective contributions: Caroline Beard created the mouse strains carrying the *CAGGS-M2rtTA* and *R26-M2rtTA* constructs and generated Figures 1B and 3.

Introduction

The goal of these experiments was to develop a system that enables us to analyze the requirement for MeCP2 in pre- and postnatal neurons. Inducible gene expression systems allow for the regulation of a target gene placed under the control of an inducible promoter by small molecules that can be administered to animals at any postnatal age. The tetracycline inducible system uses the M2rtTA transactivator to activate transcription of the target gene by binding to a *tet* operator sequence in its promoter in the presence of the tetracycline analog doxycycline (dox) (Gossen et al. 1995). Described below is a detailed analysis of the tetracycline inducible system using a reporter system.

Results

We tested a number of different promoters to optimize expression of *M2rtTA*. Pan-neuronal promoters included the promoters of the mouse neuron-specific enolase gene (*mNSE*) (Forss-Petter et al. 1990), the rat neuron-specific enolase gene (*rNSE*) (Mucke et al. 1994), and the promoter of the gene encoding the microtubule-binding protein, Tau (*tau*) (Tucker et al. 2001). Ubiquitously expressed promoters were *ROSA26* (*R26*) (Friedrich and Soriano 1991) and *CAGGS* (Niwa et al. 1991). In order to avoid unpredictable transgenic position effects, all strains were constructed by homologous recombination mediated insertion of the transgene either into a *Pst*I site downstream of the last exon of the *Col1A1* gene (*mNSE-M2rtTA*, *rNSE-M2erTA*, *CAGGS-M2rtTA*), into the endogenous *tau* locus (as an in-fame fusion to exon 1) or into the *R26* locus. Northern analysis revealed that of the pan-neuronal promoters, only the *tau* promoter produced a detectable *M2rtTA* transcript in the brain. Low-level *Tau-M2rtTA* expression was also detectable in the kidney (Figure 1A). We then compared transcript levels of *Tau-M2rtTA* with *CAGGS*-, and *R26-M2rtTA* in brain and liver by Northern blot (Figure 1B). Wild type RNA (lanes 1 and 7) and RNA from a mouse carrying the non-expressing *rNSE-M2rtTA* construct (lanes 2 and 8) were included as negative controls. In the brain, highest expression of *M2rtTA* was observed from the *CAGGS* promoter (lane 4), followed by *R26* (lane 5). The level of *Tau-M2rtTA* expression in the brain was low (lane 3). In the liver, no *Tau-M2rtTA* transcript was detectable (lane 9), *CAGGS-M2rtTA* expression was low (lane 10), and *R26-M2rtTA* expression was comparable to the *R26-M2rtTA* expression level in the brain (lane 11).

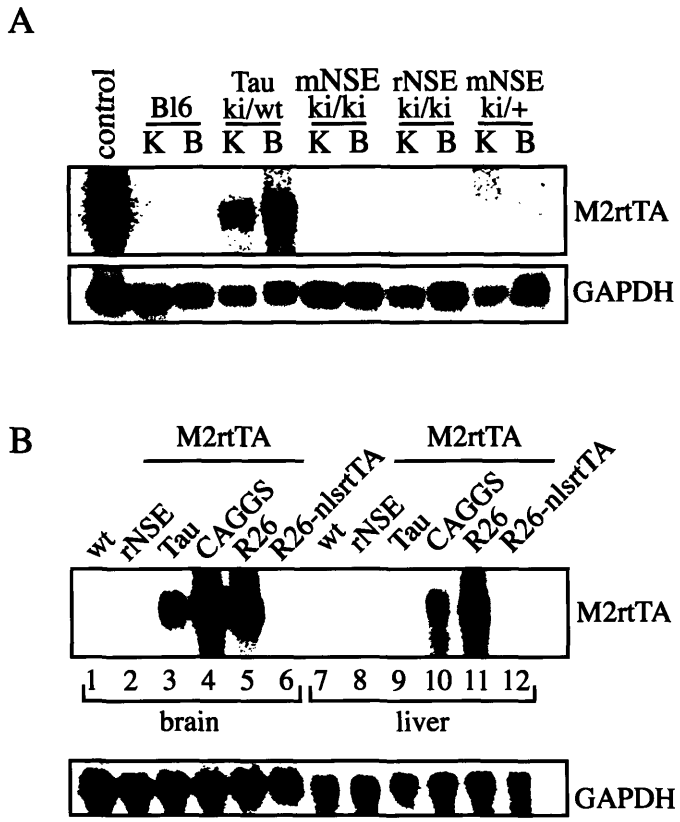


Figure 1. *M2rtTA* expression from different promoters. **(A)** Northern analysis of *M2rtTA* transcript levels driven by the Tau, mouse neuron-specific enolase (mNSE) or rat neuron-specific enolase (rNSE) promoters in kidney (K) and brain (B). C57/Bl6 wild type (Bl6) RNA is included as a negative control. Only the *Tau* promoter produces a detectable transcript. **(B)** Northern analysis of *M2rtTA* transcript levels driven by the rNSE, Tau, CAGGS or R26 promoters in brain (lanes 1 - 6) and liver (lanes 7 - 12). Transcript levels of the *nlsrtTA* transactivator driven by the R26 promoter are also assessed (lanes 6 and 12). The *CAGGS-M2rtTA* construct shows the highest level of expression in the brain, followed by *R26-M2rtTA*. Weakest expression is detected from the Tau promoter.

We tested the activity of the transactivators by carrying out reporter experiments. Transgenic animals were crossed with mice carrying a ROSA26 Cre reporter (Soriano 1999) and an inducible Cre-recombinase gene (Figure 2). In this reporter system, efficiency of dox-induced Cre expression can be monitored on a per cell basis by X-gal staining of tissue slices. In addition, recombination efficiency in tissues can be analyzed by Southern blotting after digestion with *Eco*RV, because loss of the stop-cassette results in a band-shift from 3.8 kb (2 lox allele) to 3.3 kb (1 lox allele). We tested the efficiency of various promoters as well as different types of drug delivery (Figure 3). The drug was administered for 5 days in food (0.6 g dox per 100 g soft pellets, prepared daily, lanes 1 to 9), water (0.2 g dox and 5 g sucrose per 100ml water in light protected bottles, changed every 2 days, lanes 10 to 18), or by IP injection (2 mg/day as 200 μl of 10 mg dox per 1ml of 0.0 5M Na-phosphate buffer, lanes 19 to 27). The drug caused serious side effects in mice carrying the *CAGGS-M2rtTA* construct. Animals became hunched, unkempt, and dehydrated within 24 - 48 h and were sacrificed after 3 days. Only in one experiment, tissues of *CAGGS-M2rtTA* mice were harvested after 5 days (lanes 28 to 32). Untreated control animals were included in the Southern analysis (lanes 33 to 41). All types of drug delivery achieved comparable results. Only *CAGGS-M2rtTA* and *R26-M2rtTA* induced Cre activity in the peripheral tissues tested (spleen and liver). In the brain, Cre-mediated recombination was observed only in samples from mice carrying *CAGGS-M2rtTA* (lanes 8, 17, and 26). Extending the treatment of these animals from 3 to 5 days led to increased recombination (lane 32). *Tau-M2rtTA* did not induce of Cre expression in liver or brain with any treatment.

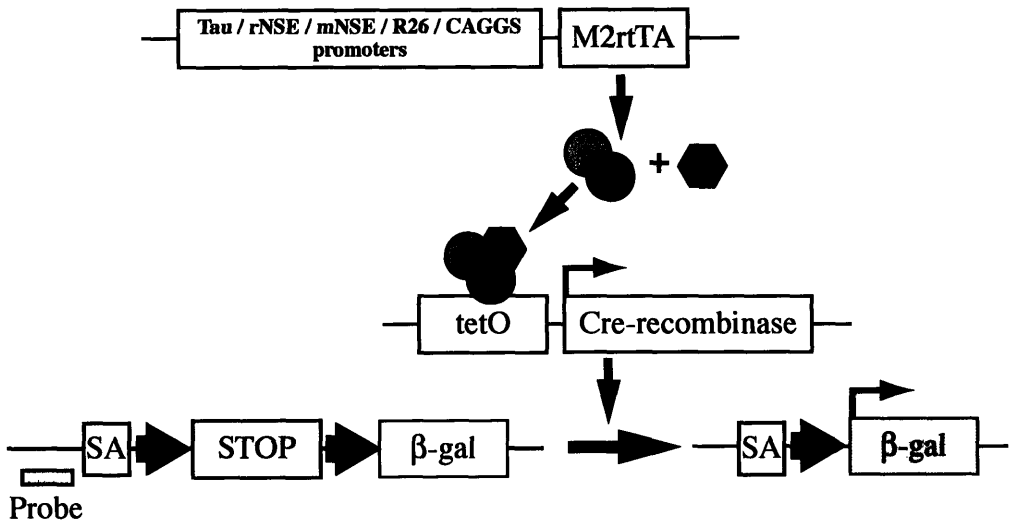


Figure 2. Schematic representation of the reporter system used to test a variety of different promoters. The gene encoding the *M2rtTA* transactivator is either integrated into the *Col1A1* locus and driven by either the rat neuron-specific enolase (rNSE), the mouse neuron-specific enolase (mNSE) or the CAGGS promoter; or into the *ROSA26* (R26) locus; or into the *tau* locus as an in-frame fusion to exon 1. The constitutively expressed transactivator is activated by doxycycline (dox) and induces the expression of Cre-recombinase. The reporter transgene at the R26 locus consists of the ß-galactosicase gene (ß-gal), flanked by a splice acceptor (SA), and a floxed (loxP) neomycin selection marker and transcriptional stop signal (STOP). Expression of Cre-recombinase leads to looping-out of the STOP cassette and expression of ß-gal.

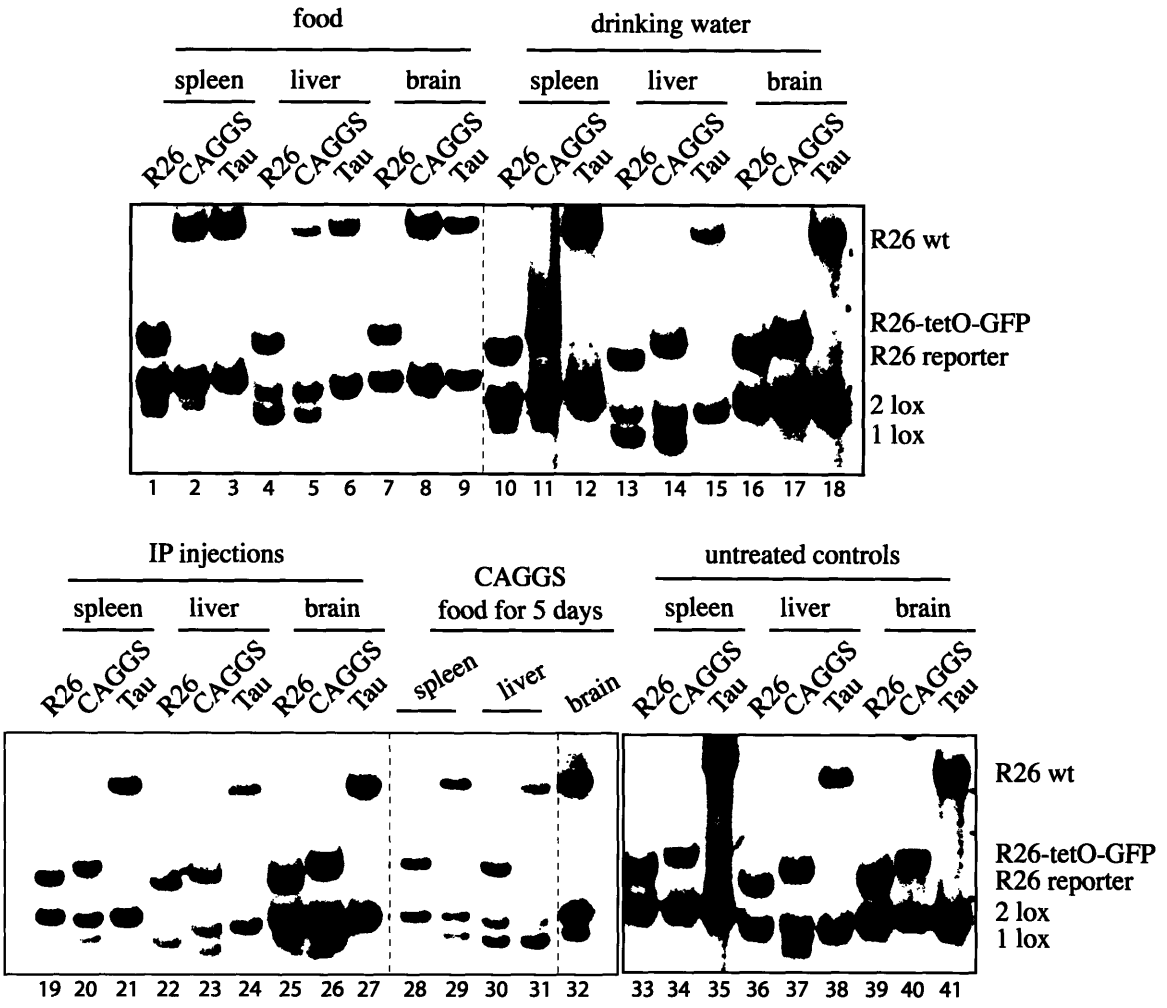


Figure 3. Testing different promoters and different types of drug delivery *in vivo*. *M2rtTA* expression is driven by different promoters as indicated. Doxycycline (dox) was administered to animals carrying an *M2rtTA* transgene and a reporter transgene either in food (0.6 g dox per 100 g soft pellets, prepared daily, lanes 1 - 9); water (0.2 g dox and 5 g sucrose per 100 ml water in light protected bottles, changed every 2 days, lanes 10 -18); or by IP injection (2 mg dox/day as 200 μl of 10 mg dox per 1 ml of 0.05 M Na-phosphate butter pH 7.5, lanes 19 - 27) for 3 days (*CAGGS-M2rtTA* animals) or 5 days (all other animals and *CAGGS-M2rtTA* animals in lanes 28 - 32). DNA was isolated from spleen, liver, and brain, digested with *EcoR*V and analyzed by Southern blotting using the probe indicated in Figure 2. Bands representing the wild type *R26* allele (R26 wt), the reporter transgene (R26 reporter, 1 lox, 2 lox) and an unrelated inducible GPF transgene at the *R26* locus (R26-tetO-GFP) are indicated. To assess the activity of Cre-recombinase, the intensity of the 1 lox band (indicates loop-out) is compared with the 2 lox band.

Discussion

In order to achieve inducible neuron-specific expression of a target gene, several requirements need to be fulfilled. 1) The transactivator must be expressed efficiently in neurons, 2) the drug must be able to cross the blood-brain barrier, and 3) the administration of the drug should have no adverse effect on the animal. We tested a variety of pan-neuronal promoters and found that *Tau-M2rtTA* was expressed in the brain. The other promoters tested, *mNSE* and *rNSE*, did not produce detectable transcripts, possibly due to position effects at the *Col1A1*locus. We have since found that, besides *CAGGS*-driven genes, all other transgenes integrated at *Col1A1* also failed to express (data not shown).

We were able to show that the drug crosses the blood-brain barrier but that high a level of transactivator expression is necessary to achieve induction of target gene expression. Only *CAGGS-M2rtTA* expression was high enough to induce Cre-expression in the brain. However, the drug caused serious side effects in animals carrying *M2rtTA* driven by the high-expressing ubiquitous *CAGGS* promoter. We attribute the side effects to squelching in peripheral tissues, an effect that has been described previously for other transactivators expressed at high levels (Gill and Ptashne 1988; Wright et al. 1991). Reducing the doxycycline concentration to 1/100 of the original dose did not alleviate the problem, which limits the use of *CAGGS-M2rtTA* for *in vivo* experimentation.

References

Feil, R., J. Brocard, B. Mascrez, M. LeMeur, D. Metzger, and P. Chambon. 1996. Ligand-activated site-specific recombination in mice. *Proc Natl Acad Sci U S A* **93:** 10887-10890.

Forss-Petter, S., P.E. Danielson, S. Catsicas, E. Battenberg, J. Price, M. Nerenberg, and J.G. Sutcliffe. 1990. Transgenic mice expressing beta-galactosidase in mature neurons under neuron-specific enolase promoter control. *Neuron* **5:** 187-197.

Friedrich, G. and P. Soriano. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5:** 1513-1523.

Gill, G. and M. Ptashne. 1988. Negative effect of the transcriptional activator GAL4. *Nature* **334:** 721-724.

Gossen, M., S. Freundlieb, G. Bender, G. Muller, W. Hillen, and H. Bujard. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268:** 1766-1769.

Mucke, L., E. Masliah, W.B. Johnson, M.D. Ruppe, M. Alford, E.M. Rockenstein, S. Forss-Petter, M. Pietropaolo, M. Mallory, and C.R. Abraham. 1994. Synaptotrophic effects of human amyloid beta protein precursors in the cortex of transgenic mice. *Brain Res* **666:** 151-167.

Niwa, H., K. Yamamura, and J. Miyazaki. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108:** 193-199.

Rossi, F.M. and H.M. Blau. 1998. Recent advances in inducible gene expression systems. *Curr Opin Biotechnol* **9:** 451-456.

Soriano, P. 1999. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21:** 70-71.

Tucker, K.L., M. Meyer, and Y.A. Barde. 2001. Neurotrophins are required for nerve growth during development. *Nat Neurosci* **4:** 29-37.

Wright, A.P., I.J. McEwan, K. Dahlman-Wright, and J.A. Gustafsson. 1991. High level expression of the major transactivation domain of the human glucocorticoid receptor in yeast cells inhibits endogenous gene expression and cell growth. *Mol Endocrinol* **5:** 1366-1372.