Regulation of X-Chromosome Inactivation in Development in Mice and Humans

TETSUYA GOTO* AND MARILYN MONK

Molecular Embryology Unit, Institute of Child Health, London WC1N 1EH, United Kingdom

X-CROMOSOME INACTIVATION

General Features

In female mammals, dosage compensation of X-linked genes between males and females occurs by genetic inactivation of one of the two X chromosomes (96). The choice of X chromosome to be inactivated is random in somatic cells, i.e., either the paternally or maternally inherited X chromosome is inactivated in a given cell. Once established, the inactivity is clonally maintained, and spontaneous, unprogrammed reactivation is extremely rare. The inactive X chromosome is distinguished from the active X chromosome by the following characteristics: (i) overall transcriptional inactivation (53) (apart from certain X-linked genes which escape inactivation and the Xist gene); (ii) heterochromatic condensation at interphase of the cell cycle, sometimes visible as the Barr body (6); and (vi) expression of the Xist (X-inactive specific transcript) gene located at the X-inactivation center (15, 17, 19, 20).

These characteristics are associated, causally and/or consequentially, with the process of X-chromosome inactivation (XCI), which occurs early in development. Both the paternally and maternally inherited X chromosomes are active in pre-implantation embryos after fertilization, and XCI occurs sequentially, coupled with cell differentiation, first in the extra-embryonic trophoderm and primitive endoderm of the blastocyst and then in the fetal precursor cells arising from the inner cell mass (ICM) around the time of implantation (113, 158). XCI is a multistep process, comprising (i) counting of the number of X chromosomes with respect to the number of autosomal complements, (ii) choice of a single X chromosome to remain active per autosome complement, (iii) initiation of inactivation of an additional X chromosome(s), (iv) spreading of inactivity along the length of the inactive X chromosome(s), and (v) stabilization and maintenance of inactivity throughout future cell divisions. These steps in the XCI process are discussed further below.

Theoretically, the choice involved in XCI could be either the choice of X chromosome to be active or the choice of X chromosome to be inactive. The observation of a single active X chromosome in cases of X-chromosome aneuploidy, e.g.,
X^X X^X AA and X^X X^X X^A (X^+, active X chromosome; X^-, inactive X chromosome; A, haploid set of autosomes), suggests that one X chromosome is always chosen to be active rather than different numbers of X chromosomes being chosen to be inactive. Cells count the number of autosomes for determining the number of X chromosomes to be active, since tetraploid cells have two active X chromosomes, i.e., X^X X^X X^X X^AAAA.

XCI is initiated from the X-inactivation center (Xic), which is required in cis for the X chromosome to be inactivated (145). The evidence for the existence of the Xic is derived from observations on X-autosome translocations; the X chromosome devoid of a particular segment (putative Xic) does not undergo inactivation (135, 145). Once XCI is initiated, inactivity spreads from the Xic in both directions along the chromosomes.

The spreading process has been probed by using transgenes. Some X-linked transgenes (transgene integrated on the X chromosome) are inactivated on the inactive X chromosome (135, 145), whereas others escape inactivation (49, 164). In addition, the strength of inactivation of the transgene may depend on tissue-specific transgene expression; e.g., a transgene containing the alpha-fetoprotein gene was shown to be inactivated on the inactive X chromosome in somatic tissues but to remain active in the yolk sac (where alpha-fetoprotein is normally expressed) of the developing mouse conceptuses (82). Whether a transgene is inactivated on the inactive X chromosome may depend on its size, site of insertion, mode of insertion, and tissue specificity of expression, as well as on the structure of the transgene itself. On the other hand, it has been shown that X-linked phosphoglycerate kinase-1 (Pgk-1) gene-containing transgenes on autosomes are ubiquitously expressed (100, 132), which suggests that inactivation is not intrinsic to the linked genes themselves. This is a striking contrast to the dosage compensation mechanism in Drosophila, where X-linked genes translocated to autosomes show inactivation (reviewed in reference 95).

Several X-linked genes have been shown to escape XCI in both humans and mice (reviewed in reference 42). Although the mechanism of escape from inactivation is unknown, the phenomenon indicates that spreading of inactivation can be interrupted and then resume and that local control is also involved. More X-linked genes escape XCI in humans than in mice, and this difference is considered to be responsible for the more severe phenotype and reproductive failure in human XO females, known as Turner’s syndrome or haploinsufficiency (reviewed in reference 167). In contrast, XO female mice are fertile, albeit they produce a smaller number of eggs with lower developmental potential compared to those from XX mice (24, 25).

Maintenance (or stabilization) of inactivity may involve multiple mechanisms, such as changes in Xist gene structure or Xist transcription, some function of Xist RNA, X-chromosome DNA methylation, histone H4 hypoacetylation, heterochromatinization, late replication during S phase, and nuclear compartmentalization (reviewed in references 116, 120, and 143). Among these, late replication and heterochromatinization seem to be more universal mechanisms since they are also observed for non-X-linked genes and are found in other species, such as yeast and Drosophila (reviewed in references 59, 62, and 111).

Activity of the X Chromosomes in Female Development

The activity of the X chromosomes throughout female mouse development is illustrated in Fig. 1. It has been shown cytologically (157) and enzymatically (44, 80, 112) that both the maternally and paternally inherited X chromosomes are active in female mouse preimplantation embryos. Then XCI occurs sequentially in development, coupled with cell differentiation (113, 158). XCI first occurs in the trophectoderm and the primitive endoderm which are the first two cell lineages that differentiate, at 3.5 and 4.5 days postcoitum (dpc), respectively (Fig. 2). In these two extraembryonic lineages and their derivative tissues, XCI is nonrandom in that the paternally inherited X chromosome is preferentially inactivated (46, 47, 55, 126, 156, 163) (summarized in Table 1)

In the primitive ectoderm (epiblast), which will give rise to the embryo proper (somatic cells and germ cells) and yolk sac mesoderm, XCI is initiated at around the time of implantation and is random in that either the paternal or the maternal X chromosome is inactivated in a given cell.

In female germ cells, XCI is also random, and it has been shown, by a study of X-inactivation cell mosaicism, that the somatic tissues and the germ cells are derived from a common pool of cells after XCI has occurred and that the germ cells are derived from a sizable pool of precursor cells (104, 105). This original work (115) challenged the popular view that the continuity of the totipotency of the germ line necessarily required that the germ cells be set aside very early in development and from very few cells. It also means that the “germ line” itself includes the early stages of preimplantation development through to the epiblast of the implanted embryo when the primordial germ cells are delineated. A few years later, Monk et al. (117) provided further evidence in support of this view and showed that the ground state, with respect to erasure of epigenetic methylation programs, was not a property of the gametes at the time of fertilization but, rather, a property of the totipotent ICM, epiblast cells and germ cells at the time of implantation.

In keeping with a progressive dedifferentiation in early development approaching a ground state in the epiblast cells and the primordial germ cells arising from them, the inactive X chromosome is reactivated in the primordial germ cells at around the time of entry into meiosis, i.e., 12.5 to 13.5 dpc. Following reactivation, both X chromosomes continue to be active throughout oogenesis (66, 81, 104, 114).

Activity of the X Chromosome in Male Development

In males, the single X chromosome is active in all somatic cells. However, it is transcriptionally inactivated during spermatogenesis. At meiosis, the X chromosome pairs with the Y chromosome in the XY body, or the sex vesicle, where the X heterozygotes, the X chromosome carrying Xcde is more likely to be inactivated than the X chromosome carrying Xceb. In Xcde/Xceb heterozygotes, the X chromosome carrying Xceb is more likely to be inactivated than the X chromosome carrying Xcde. In Xcde/Xceb heterozygotes, nonrandom (skewed) XCI is most prominent. Therefore, the relative strengths of the three Xcde alleles, with respect to likelihood of remaining active, are Xcde < Xceb < Xcde. Recently, a fourth allele, Xceb, carried in Mus spretus, has been identified and seems to be stronger than Xcde (34). The strength of the Xce has been inversely correlated with the degree of Xist expression from the inactive X chromosome in that the level of Xist RNA is markedly lower in Mus spretus mice which carry the Xceb allele than in C57BL/10 mice which carry the Xceb allele (17).
and Y chromosomes are sequestered away from the autosomes. In some instances, transcriptional silencing of the X-linked genes during spermatogenesis is compensated by expression of autosomal genes, e.g., the X-linked phosphoglycerate kinase-1 (Pgk-1) gene by the autosomal Pgk-2 gene (102) and the X-linked pyruvate dehydrogenase E1α subunit (Pdha-1) gene by the autosomal Pdha-2 gene (38).

The inactive X chromosome carried in sperm is reactivated soon after fertilization in female preimplantation embryos (44, 80, 112).

**X-Chromosome Inactivation and Genomic Imprinting**

Genomic imprinting is a phenomenon in which expression of a gene locus is determined by its parental origin: either the paternal or the maternal allele is expressed and the other is repressed. Genomic imprinting demonstrates that the parental chromosomal complements are not equivalent with respect to their potential for expression and implies some form of marking (imprint or blueprint) distinguishing certain genes in the egg and in the sperm. These marks must be imposed as modifications to the DNA of the imprinted genes during gametogenesis (reviewed in reference 121). The imprinting mark distinguishing the imprinted gene in either eggs or sperm is perpetuated into early development after fertilization, so that parental alleles are differentially expressed in the progeny embryos and offspring.

Imprinting is a form of transgenerational change in gene expression. A gene may be switched on or off in successive generations as it passes through the male or female parent. Therefore, in the germ line of the progeny, the imprint must be erased to be reestablished once again, depending on the sex of progeny, as the male or female parental imprint in sperm and eggs, respectively. For example, a maternally transmitted allelic imprint in male progeny undergoes erasure and new modification during spermatogenesis to become the paternal allelic imprint.

Genomic imprinting may explain the requirement for both the paternal and maternal sets of chromosomes for normal development. Pronuclear transplantation experiments (97, 103, 154) have shown that neither androgenetic (paternal chromosomes only) nor gynogenetic (maternal chromosomes only) embryos develop to term. Androgenetic embryos show proliferative growth of the extraembryonic tissues and poor growth of the fetus, whereas gynogenetic embryos show poor growth of the extraembryonic tissues and comparatively normal growth of the fetus. In the human, an aberration in embryonic development during pregnancy, known as the hydatidiform mole, is the hyperplastic growth of the extraembryonic trophoblast with reduced or absent fetal growth, which resembles the growth pattern of mouse androgenetic embryos. In fact, the complete form of the hydatidiform mole (no fetal components) is androgenetic in origin, with karyotype 46, XpXp in 96% of cases and karyotype 46, XpY in the remaining 4% of cases (71, 124). The incomplete hydatidiform mole (with fetal components) is, in most cases, due to triploidy with one set of maternal and two sets of paternal genomic complements and karyotype 69, XpXpXp or XpXpY.

The chromosomal regions containing putative imprinted genes have been identified by investigations of abnormal phenotypes and developmental lethality associated with regions of uniparental disomy in progeny derived from breeding mice carrying translocations (33). Once such an imprinted chromosomal region is identified, specific genes can be cloned from the chromosomal region. These specific genes can be tested for imprinted expression by using a sequence polymorphism (often found between species in an interspecific cross) or by observing a parent-of-origin-specific effect of a deletion or mutation created in a candidate imprinted gene. Cattanach and Jones (35) have reported that imprinted chromosomal regions in mice may be limited to only six autosomes (chromosomes 2, 6, 7, 11, 12, and 17). Recently, however, a newly identified chromosomal region, Irlg3, which contains at least one imprinted gene, Irf1, has been mapped to chromosome 9 by the technique of restriction landmark genomic scanning with methyl-sensitive enzymes (RLGS-M) (130). Thus, an increasing number of regions containing imprinted genes have been identified and, to date, around 20 imprinted genes have been identified in humans and mice.

Preferential paternal XCI in mice (46, 47, 55, 126, 156, 163) was one of the first examples of genomic imprinting in mam-
mals and was well characterized before specific imprinted genes were identified. Imprinted XCI depends on the timing of XCI and/or the tissue involved; preferential paternal XCI is observed only in the extraembryonic trophectoderm and the primitive endoderm, which are the first differentiated tissues, delineating at 3.5 and 4.5 dpc, respectively (Fig. 2). This suggests that the imprint for nonrandom XCI is present at these stages of development and is recognized in these tissues and that the imprint is erased by the time XCI occurs in the fetal cells (reviewed in reference 118). Alternatively, the gametic imprint may not be recognized by fetal cells (irrespective of timing).

The imprint for nonrandom XCI should be present on either the paternal or the maternal X chromosome. Shao and Takagi (149) reported that a supernumerary maternally derived X chromosome (Xm) in XmXmXpAmAp (where A refers to a haploid set of autosomes) causes early embryonic death and hypothesized that this is due to two active X chromosomes (XmX Xp AmAp) in the extraembryonic tissues. Their results suggest that the maternally derived X chromosome is resistant to inactivation in those lineages. Recently, Marahrens et al. (98) have shown that embryos inheriting an inactivation-resistant (owing to a deleted Xist locus) X chromosome from the father also die, presumably due to their inability to inactivate the normal maternal X chromosome (hence both Xm and Xp are active) in the extraembryonic tissues. However, Rastan et al. (134) showed that XCI does occur in the extraembryonic tissues of diploid XmXm parthenogenones, which can develop up to the somite stage (72). In addition, XpO embryos are viable (26, 27, 45, 126), suggesting that Xp is not necessarily inactivated in the extraembryonic tissues. Therefore, these relatively imprecise studies do not help at this stage to elucidate the possible mechanisms of the differential marking of parental X chromosomes in gametes and extraembryonic tissues.

In contrast to mice, it is less clear whether the paternally derived X chromosome is preferentially inactivated in human extraembryonic tissues. Ropers et al. (144) and Harrison (56) claimed that preferential paternal XCI does occur in the human extraembryonic tissues, but Migeon and Do (108) and Migeon et al. (109) reported that there is no evidence for paternal XCI in these tissues. All these studies are based on the glucose-6-phosphate dehydrogenase (G6PD) isozyme assay. Harrison (56) used term placentas heterozygous for G6PD isozymes and showed that the maternal G6PD allele is preferentially expressed in isolated trophoblasts. Recently, we have used a DNA methylation assay to show that preferential pa-

---

**FIG. 2.** Diagrammatic representation of cell differentiation during mouse development. The trophectoderm is the first cell lineage, which differentiates at 3.5 dpc. Then the primitive endoderm differentiates, at 4.5 dpc. In the derivatives of these two extraembryonic cell lineages, XCI is nonrandom, in that the paternally inherited X chromosome is preferentially inactivated (see Table 1 for summary). In contrast, XCI is random in the epiblast, which gives rise to the yolk sac mesoderm, the embryo proper, and the primordial germ cells; i.e., either the paternal or the maternal X chromosome is inactivated in a given cell.

---

**TABLE 1. Summary of studies on paternal XCI in mouse extraembryonic tissues**

<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
<th>Embryonic stage (dpc)</th>
<th>Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takagi and Sasaki (156)</td>
<td>Allocyclic X</td>
<td>8.5</td>
<td>Chorion</td>
</tr>
<tr>
<td>West et al. (163)</td>
<td>PGK-1 isozyme</td>
<td>13.5</td>
<td>Yolk sac (endoderm)</td>
</tr>
<tr>
<td>Freh et al. (46)</td>
<td>PGK-1 isozyme</td>
<td>8.5</td>
<td>Chorionic ectoderm</td>
</tr>
<tr>
<td>Freh and Chapman (47)</td>
<td>PGK-1 isozyme</td>
<td>9.5</td>
<td>Mural trophoblast</td>
</tr>
<tr>
<td>Papaioannou and West (126)</td>
<td>PGK-1 isozyme</td>
<td>12.5 and 17.5</td>
<td>Parietal endoderm</td>
</tr>
<tr>
<td>Harper et al. (55)</td>
<td>PGK-1 isozyme</td>
<td>6.5</td>
<td>Extraembryonic ectoderm</td>
</tr>
<tr>
<td>Harper et al. (55)</td>
<td>PGK-1 isozyme</td>
<td>3.5 + in vitro culture</td>
<td>Outgrowth of blastocyst</td>
</tr>
</tbody>
</table>
ternal XCI also occurs in human trophoblastic cells obtained at 10 to 12 weeks of gestation (50). Our results support the conclusion of Harrison (56) and, in addition, strongly suggest that preferential inactivation of the paternally inherited allele seen in term placentas is not due to cell selection during development against cells with the paternal X chromosome active but is due to primary nonrandom XCI occurring early in development (50).

Many studies have sought the underlying parental allelic differences of imprinted genes. DNA modification by CpG methylation has been most extensively studied, since highly sensitive procedures are available for the detection of this modification throughout development. It is also clear that DNA methylation plays an important role in the regulation of gene expression and, moreover, that it is subject to significant changes during development (117; reviewed in references 118, 121, 139, 141).

X-CHROMOSOME INACTIVATION AND DNA METHYLATION

DNA Methylation of the Mammalian Genome

DNA methylation is an epigenetic modification in that it is superimposed on the DNA base sequence and affects the genetic potential of that sequence. In mammals, DNA methylation is found as 5-methylcytosine (5mC) occurring at CpG dinucleotides (Fig. 3A). Theoretically, CpG and CpC dinucleotides should be found at equal frequencies in bulk genomic DNA; i.e., the CpG/CpC ratio should be 1. However, overall, CpG dinucleotides (unmethylated and methylated CpG combined) are found at only 20 to 25% of the expected frequency (43, 48). This lower than expected frequency of CpG dinucleotides is considered to be due to a tendency to mutation transition of the 5mC residue to thymine by deamination at position 4 of the carbon ring (i.e., from CpG to TpG) throughout evolution (Fig. 3A).

Approximately 70% of all CpGs are methylated (reviewed in references 12, 89, and 138). The significance of DNA methylation as a mechanism of gene regulation has been controversial since well-researched systems, such as Drosophila and Caenorhabditis elegans, are known to regulate gene transcription without detectable DNA methylation being present (152, 162). Nevertheless, recent experiments have shown that DNA methylation is essential for normal mouse development (91, 93), but is due to primary nonrandom XCI occurring early in development (69, 117, 150; see below). This loss in methylation is associated with a marked decrease in the DNA methylation level early in development (69, 117, 150; see below). This loss in methylation is associated with a marked decrease in the DNA methylation level early in development (69, 117, 150; see below). This loss in methylation is associated with a marked decrease in the DNA methylation level early in development (69, 117, 150; see below).

FIG. 3. Methylation of cytosine residues in the mammalian genome. (A) Structure of the cytosine residue and its methylated derivative, 5-methylcytosine, which is methylated at the 5 position of the carbon ring of the cytosine residue. The structure of the thymine residue is also shown, to illustrate that mutational deamination at the 4 position of 5-methylcytosine results in thymine. (B) After replication, daughter strands of fully methylated DNA are hemimethylated (reaction 3) and the original pattern of DNA methylation is maintained by the DNA methyltransferase enzyme activity in preimplantation embryos (reaction 4). De novo methylation (reaction 1) is also considered to be mediated by the DNA methyltransferase, although the efficiency of de novo methylation is low.
it methylates any CpG dinucleotide accessible due to an open chromatin structure, e.g., during DNA replication or displacement of nucleosomes.

During development, dynamic demethylation and de novo methylation occur (69, 117, 150). Figure 4 illustrates these changes in DNA methylation early in development, as shown by Monk et al. (117). Gametic DNA is globally demethylated, descending toward the ground state of absence of methylation and erasure of the bulk of the original gametic methylation patterns of paternal and maternal genetic complements by the blastocyst stage. Stage- and tissue-specific methylation patterns are created by de novo methylation during the process of implantation and gastrulation, whereas demethylation in the germ line continues at this stage.

The genome-wide demethylation occurring in preimplantation embryos is thought to be a process of erasure of the methylation patterns associated with differential programming of oogenesis and spermatogenesis (117; reviewed in reference 121) and is correlated with the decrease in activity of the DNA MTase (119). However, the process may be more complicated in that demethylation at preimplantation stages seems to be site specific; certain CpG sites within imprinted genes are more resistant to demethylation so that they persist until the blastocyst stage (16, 51, 70, 150). We could view a genetic imprint as the differential parental gametic modification which survives erasure during preimplantation development and thus are perpetuated into the soma. Reprinted from reference 121 with permission of the publisher.

FIG. 4. Changes in DNA methylation in early development. Sperm DNA is overall more methylated than egg DNA but less methylated than somatic DNA. During preimplantation development, gamete-specific patterns of methylation are erased by the genome-wide demethylation, tending toward a ground state of absence of methylation in the ICM of the blastocyst and in the primordial germ cells. Around the time of implantation and gastrulation, stage- and tissue-specific patterns of DNA methylation are created de novo. In the extraembryonic tissues, the methylation level is globally lower than that in somatic tissues. It is hypothesized that demethylation continues in the germ line until the onset of establishment of new egg- or sperm-specific patterns of methylation, including the differences in methylation between the gametes that will survive as imprints in the next generation. Imprinting can be viewed as differences in gametic DNA modification (differences between sperm and egg) that survive erasure during preimplantation development and thus are perpetuated into the soma. Reprinted from reference 121 with permission of the publisher.

Three related forms of DNA MTase (Mr., 190,000, 175,000, and 150,000) have been identified to date, and a detailed characterization has suggested that they are proteins encoded by a single gene with different extents of deletion of the carboxyl terminus (10). Li et al. (91) introduced targeted mutations into the DNA MTase gene to investigate the role of the enzyme and DNA methylation during development. Mice carrying the mutated allele(s) of the DNA MTase gene die at around mid-gestation, depending on the severity of deficiency of the enzyme activity (91). Investigations of expression patterns of imprinted genes, insulin-like growth factor 2 (Igf2), Igf2 receptor (Igf2r), and H19, in these DNA MTase-deficient mice have shown that all three genes are aberrantly expressed and that the extents of demethylation of normally methylated CpG sites are different among these genes (93). H19, which is normally expressed only from the maternal allele (7), shows biallelic expression, and the differentially methylated region on the paternal allele is demethylated. Igf2, which is normally expressed only from the paternal allele (41), is not expressed from either allele. Igf2r, which is normally expressed only from the maternal allele (5), shows an unchanged level of expression and methylation (at the CpG island of the region 2 on the expressed maternal allele) in embryos with a less severe type of mutation of the DNA MTase but shows no detectable expression and substantial demethylation in a more severe type of mutation. These results clearly show the biological significance of DNA methylation in development and in the regulation of...
gene expression and imprinting. They also indicate that the sensitivity to demethylation is site and/or gene specific.

**DNA Methylation, CpG Islands, and Gene Expression**

Approximately 30% of all CpG dinucleotides of the mammalian genome are unmethylated, and these unmethylated CpG dinucleotides make up only 1 to 2% of the whole genome. Unmethylated CpG dinucleotides are often found as a cluster (12, 48); if the cluster is 0.5 to 2 kb long, GC-rich (65% or more), and of the expected frequency of CpGs (CpG/GpC ratio of ca. 1; no CpG suppression), it is called a CpG island (reviewed in reference 12). The CpG island is often associated with the 5' region and/or the promoter region of a gene, and its methylation pattern is correlated with gene activity for X-linked genes and imprinted genes (reviewed in references 12, 13, and 139).

In principle, CpG islands are not methylated (they are therefore also called methylation-free islands). For example, CpG islands of ubiquitously expressed housekeeping genes are unmethylated. Some tissue-specific genes also have CpG islands, and these are normally unmethylated in both expressing and nonexpressing tissues (reviewed in reference 12). Exceptions are observed for X-linked genes on the inactive X chromosome (reviewed in references 12 and 116) and certain tissue-specific genes, such as the α-globin gene and the Thy-1 gene, in vitro-cultured cells (3). Since CpG islands are usually associated with the promoter region and the 5' end of a gene, binding of transcription factors and the chromatin structure around the transcription start site may inhibit the access of the DNA MTase and protect the CpG islands from de novo methylation. This could be the case for housekeeping genes and tissue-specific genes in the expressing tissue. However, for tissue-specific genes in nonexpressing tissues, other mechanisms, such as binding of a tissue-specific repressor protein or packaging into a closed chromatin structure, must be operative to maintain the methylation-free status.

The repressive effect of DNA methylation on the regulation of gene expression involves the formation of an inactive chromatin structure (28, 75, 76), which is thought to be mediated by proteins which specifically bind to methylated DNA. To date, in addition to the DNA MTase, several such proteins have been identified (Table 2); methylated DNA-binding protein (MDBP) (63), methyl-CpG binding proteins, McCP1 (106) and McCP2 (90, 107), MDBP-2-H1 (a member of the histone H1 family) (68, 127), the protein which binds to the mouse sex-limited protein (Slp) promoter (166), and the methylation-dependent protein which binds to an Xist promoter sequence in mice (64). MDBP requires high sequence specificity (77) and McCP1 binds only to long DNA fragments with more than 15 methylated CpG dinucleotides (106). By contrast, McCP2 is able to bind at a single methyl-CpG site (90) and is essential for mouse embryonic development (161). However, the level of expression of McCP2 is very low in undifferentiated embryonic stem (ES) cells (107), and these cells can survive with no McCP2 proteins (161), indicating that McCP2 is required during or after differentiation. The methylation-dependent protein binding to the Xist promoter, identified in our laboratory, is sequence specific and preferentially binds to the methylated promoter sequence (64; see below).

**DNA Methylation and X-Chromosome Inactivation**

The X chromosome contains a large number of housekeeping genes. As described above, CpG islands of housekeeping genes are, in general, unmethylated. However, CpG islands of most X-linked genes, including housekeeping genes, are methylated on the inactive X chromosome, and this DNA methylation is involved in the maintenance of inactivity of genes on the inactive X chromosome (reference 116 and references therein). It has been well established that genes on the inactive X chromosome are at least partially reactivated in cultured cells by treatment with an inhibitor of the DNA MTase, 5-azacytidine (reviewed in reference 116 and references therein).

Although DNA methylation is considered to play a role in the maintenance of inactivity, it is controversial whether it is causally involved in initiation and spreading of inactivity along the X chromosome in development. It has been reported that the timing of DNA methylation for the Pgk-1 gene and the G6PD gene is coincident with XCI (52, 153). An earlier report suggested that methylation was a later consequence of genetic inactivation, since methylation of the Hprt gene occurs several days after XCI in development (94). The discrepancy of these results may be explained by (i) the difference in the nature and sensitivity of assays used (PCR-based or Southern blotting)—the Southern blot procedure used for the detection of methylation of the Hprt gene would have missed a minority of unmethylated alleles whereas the PCR procedure used for the Pgk-1 gene and the G6PD gene detected a minority of methylated alleles; (ii) the purity of the tissues examined—the Southern blot approach used whole conceptuses with hypo-methylated extraembryonic tissues present; and/or (iii) the fact that methylation may occur at different times for specific genes located at different loci on the X chromosome; interestingly, methylation of the G6PD gene, which is located between Pgk-1 and Hprt, occurs slightly after methylation of the Pgk-1 gene but much earlier than that of the Hprt gene (although these data were determined by different procedures), suggesting that methylation spreads from the Xic in linear fashion throughout the X chromosome (52).
THE Xist GENE

Characterization

An X-linked regulatory gene, Xist (X-inactive specific transcript; Xist in the mouse, XIST in the human), was identified by the unique characteristic that it is expressed exclusively from the inactive X chromosome (15, 17, 19) and is thus implicated in the XCI process itself. Further evidence for a regulatory role is that the Xist gene maps to the Xic (15, 17, 20). In this review, unless specifically stated as the mouse Xist gene or the human XIST gene, the term “Xist” is used to indicate both the mouse and the human genes.

The transcript is a large, polyadenylated RNA molecule, 17 kb for the human XIST transcript and 15 kb for the mouse Xist transcript, but there is no extended open reading frame for protein translation, suggesting that Xist functions as an RNA molecule (18, 21). In keeping with this idea, Xist RNA is localized in the nucleus (18, 21) and RNA FISH (fluorescent in situ hybridization) studies have shown that Xist RNA is associated with the inactive X chromosome at interphase (21, 36, 86, 125).

The human and mouse Xist genes consist of eight and six exons, respectively (18, 21) (Fig. 5a), with conserved repetitive sequences in exons 1 and 6 (18, 21) (Fig. 5b). The most con-
served repetitive sequence resides at the 5′ end of the gene in exon 1 (nucleotides [nt] 350 to 770 for Xist and 292 to 713 for Xist). This region (Fig. 5b, region A) consists of nine and eight repeats of consensus sequence, 5′-GCCCATC GGGGCCTCG GATA CC TGCG-3′, in the human and mouse Xist genes, respectively (18, 21). This repeat shows homology to a sequence of Xisir1, which is essential for the transfer of Xisir1 to the cytoplasm (78), and to a sequence of the latency-associated transcript of herpes simplex virus, which is required for efficient reactivation of the latent herpes simplex virus (14). It has also been reported that a protein binds to this repetitive sequence in human XIST RNA (23). Hendrich et al. (58) determined the DNA sequence of this repetitive region in mice carrying different Xce alleles, since the level of Xist expression correlates with strength of Xce (17). The major difference in DNA sequence identified is the number of adenine (A) residues between the fourth and fifth repeat sequence; the XceX allele (BALB/c and C3H mice) has 24 uninterrupted A residues, the XceY control of 6). Further evidence for the role of DNA methylation in the male is induced to be expressed in the DNA MTase-deficient mice; the normally silent, with DNA MTase-deficient ES cells, and random XCI occurs when they are induced to be expressed in undifferentiated ES cells. However, the level of Xist expression correlates with preferential inactivation of the paternal X chromosome in these tissues (73) (see below). The active Xce allele on the inactive paternal X chromosome is unmethylated, whereas the inactive allele on the active maternal X chromosome is methylated (123) (Fig. 6).

**Xist expression in extraembryonic tissues.** In the extraembryonic tissues, Xist expression occurs only from the paternal allele (imprinted), and this correlates with preferential inactivation of the paternal X chromosome in these tissues (73) (see below). The active Xce allele on the inactive paternal X chromosome is unmethylated, whereas the inactive allele on the active maternal X chromosome is methylated (123) (Fig. 6).

**Xist expression in germ cells.** Xist expression in germ cells is also correlated with the presence of the inactive X chromosome. In female primordial germ cells, Xist expression ceases at around 12.5 to 13.5 dpc (101), when the oocyte enters meiosis and the inactive X chromosome is reactivated (114). It has been hypothesized that the promoter region and the 5′ end of exon 1 of the Xist gene are methylated during oogenesis (4, 168; see below).

In male germ cells, Xist is expressed during spermatogenesis (101, 142, 147), when the single X chromosome is inactive, although the presence of the Xist gene and hence the expression of Xist in preimplantation embryos do not appear to be required for silencing of the X chromosome in spermatogenesis (98) (see below). Since the Xist gene itself is not required for production of viable, fertile sperm, the role of DNA methylation for the regulation of Xist expression in spermatogenesis may not be significant (although the methylation pattern created in spermatogenesis is, in fact, important for imprinted Xist expression in preimplantation embryos [see below]). It has been shown that Xist expression during spermatogenesis can be detected, by reverse transcriptase PCR, at as early as 3 days postpartum (dpp) (73) whereas demethylation of the promoter and the 5′ end of exon 1 is only detected between 9 and 21 dpp by Southern blot analysis (123). However, a later study involving a PCR-based methylation assay has shown that demethylation at the 5′ end of the first exon occurs even before birth, at between 18.5 and 21.5 dpc (4). The discrepancy with respect to the timing of demethylation detected by the two different methods (Southern blot and PCR-based methods) is, as described above, probably due to the difference in the purity of the samples and the sensitivity of the methods, and demethylation in spermatogenesis may be occurring throughout the perinatal period.

**Xist expression in preimplantation embryos.** In mouse preimplantation embryos, Xist expression is imprinted in that the paternal allele is expressed whereas the maternal allele is repressed (73). This preferential paternal allele expression precedes preferential (imprinted) paternal X-inactivation in the extraembryonic tissues, suggesting a causal role for the early Xist expression in the imprinting of XCI. The mark for imprinted Xist expression is either erased or not acted upon when both the maternal and paternal alleles are expressed around the time of implantation, correlating with the timing of when random XCI occurs in the fetal precursor cells.

DNA methylation is considered to play a crucial role in the regulation of imprinted Xist expression in mouse preimplantation embryos. This is discussed below.
Mechanism of Imprinting of Mouse Xist

DNA methylation. The imprinted Xist expression in preimplantation embryos must be controlled by a gametic imprint (differential marking in sperm and eggs), since analyses of androgenetic and gynogenetic embryos show Xist expression only in the androgenetic embryos (74, 84, 85). It has been proposed that DNA methylation is involved in the regulation of imprinted expression in preimplantation embryos (4, 168). The promoter region (168) and the 5' end of the first exon (4) are differentially methylated at certain CpG sites, which are hypomethylated in sperm and methylated in oocytes (Fig. 6). More importantly, the methylation of the maternal Xist allele survives the genome-wide demethylation occurring in preimplantation development (4, 117, 168). This strongly suggests that the differential methylation in sperm and oocytes is a candidate gametic imprint for the repression of the maternal allele of Xist, leading to preferential paternal allele expression in preimplantation embryos. Later experiments have suggested that methylation of the promoter region in oocytes is mosaic, as seen for undifferentiated ES cells (169), and recent investigations involving bisulfite genomic sequencing supported these findings (124a) (Fig. 6). Moreover, this sequence is important for transcription in its unmethylated form (64). Therefore, it is hypothesized that this repressor protein is also present in preimplantation embryos and is responsible for imprinted Xist expression, i.e., repression of the maternal allele and expression of the paternal allele, by binding to the mosaically methylated promoter sequence of the maternal Xist allele. This hypothesis is further supported by comparison of the mouse and human Xist promoter sequences and of their expression patterns in preimplantation embryos, but this is discussed below.

Functional Studies of Xist and X-Chromosome Inactivation Using Xist Gene Deletions and Transgenesis

The role played by Xist in XCI has been extensively investigated in the past few years, although it is still unknown exactly how this gene functions.

Most reports on the role of XIST in humans are descriptive rather than the results of experimental manipulation. Migeon et al. (110) have shown that an X chromosome lacking the XIST locus due to a spontaneously occurring deletion is not inactivated, suggesting that XIST is required in cis for the X chromosome to be inactivated. However, it has also been reported that the continued presence of XIST is not required to maintain the inactivity of the X chromosome in leukemic cells of a female patient (133) and in cultured mouse/human so-
heterozygous for the (128) deleted the minimal promoter region and the first exon been reported (98, 128) (summarized in Table 3). Penny et al. activation in the differentiating XX ES cells. Their results show 1o ft h e seem that the deletion (the promoter region and 7 kb of exon the X chromosome to be inactivated. In addition, it would might be expected to select the one intact counting function, if it were present in the deleted region, since other-wise we would expect XCI to occur. This contrasts with the findings with Xist-containing yeast artificial chromosome (YAC) transgenes in mice, where multiple copies of transgene Xics seem to be counted as a separate Xic (61, 86, 87) (see below). It is possible that the duplicated region spanning the human XIST gene lacks an essential sequence required for XCI and therefore is not counted as two intact Xics.

In mice, more direct experimental strategies have been used to investigate the functional role of Xist; these include (i) experiments involving knockout (deletion) of the Xist gene and (ii) transgenic experiments with the Xist gene. The findings reported in published work are summarized in Tables 3 and 4, respectively.

Xist knockout mutations. Two knockout experiments have been reported (98, 128) (summarized in Table 3). Penny et al. (128) deleted the minimal promoter region and the first exon of the Xist gene to create a null mutation in XX ES cells which are heterozygous for the Xce allele. They showed that the X chromosome bearing the null mutation does not undergo inactivation in the differentiating XX ES cells. Their results show that the intact Xist gene and/or Xist RNA is required in cis for the X chromosome to be inactivated. It would seem that the deletion (the promoter region and 7 kb of exon 1 of the Xist gene) does not contain the sequence used for counting, i.e., selection of one active X chromosome, since the counting function, if it were present in the deleted region, might be expected to select the one intact Xist locus on the normal X chromosome and choose this X chromosome to be active.

Marahrens et al. (98) replaced exons 1 to 5 of the Xist gene with the neomycin resistance gene, creating a null mutation but leaving the promoter region and the 5′ end of the first exon intact in male XY ES cells. In these cells, Xist RNA is not produced. However, it was confirmed, by RNA FISH with a probe hybridizing with the intact region of the first exon, that the remaining promoter was functional in the transcription of truncated Xist-neomycin resistance gene hybrid RNA and to a similar extent to that in wild-type XY ES cells. These authors have shown that (i) the Xist-deficient X chromosome is not inactivated; (ii) the Xist structural gene is not required to produce viable, fertile sperm, and (iii) paternal transmission of the mutated Xist is lethal, suggesting that Xist is necessary in cis for the paternally inherited X chromosome to be inactivated in the extraembryonic tissues.

The results of Marahrens et al. (98) also suggest that the action of Xist is, in fact, mediated in cis by the transcribed Xist RNA and that the mere act of transcription at the Xist locus is not sufficient. This contrasts with a report of deletion of the H19 gene, in which transcription of either the endogenous H19 structural gene, or the replaced luciferase reporter gene, driven by the intact H19 promoter results in the normal pattern of imprinting of the neighboring Igf2 gene (148).

Xist-containing transgenes. Transgene experiments (summa- rized in Table 4) have also supported a significant role for Xist in XCI. In male (XY karyotype) ES cells, an autosomally integrated 450-kb Xist-containing YAC shows some properties of the Xic in that the YAC transgene can inactivate limited chromosomal regions surrounding its integration site (86, 87). However, the inactivation spreads only up to 50 centimorgans from the integration site (87). Again with male ES cells, Her- zing et al. (61) have shown that an autosomally integrated YAC, which contains only the Xist gene and 9 kb of 5′- and 6 kb of 3′-flanking sequences, is sufficient to show the properties of the Xic in that the transgene can inactivate a downstream reporter gene in the YAC construct. In both studies, multiple copies of the YAC transgenes are integrated in an autosomal in male XY ES cells (Table 4). The level of Xist expression correlates with the copy number in the study of Lee et al. (86) but not in that of Herzing et al. (61). The difference may be explained by the actual copy numbers (3 to 24 versus 2 to 8 copies) and the sizes and structures of the YAC transgenes (450 versus 35 kb) (Table 4). It is possible that a larger trans- gene unit will be better insulated from the regulation on tran- scription imposed by the surrounding sequences.

Lee et al. (86) and Herzing et al. (61) have both reported that the Xist-containing YAC transgenes used in their studies also contain the information required for counting because the endogenous Xist allele on the single, active X chromosome in

<table>
<thead>
<tr>
<th>Study</th>
<th>Deletion*</th>
<th>Studied in</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penny et al. (128)</td>
<td>7 kb of exon 1 and 36-bp promoterb</td>
<td>ES cells and chimeric embryos</td>
<td>Differentiated ES cells: X bearing the mutant Xist remains active and X bearing the normal Xist is inactivated in 60% of cells; both Xs are active in the remaining cells. Chimeric embryos: X bearing the mutant Xist remains active, and X bearing the normal Xist is inactivated in the ES cell-derived cells.</td>
</tr>
<tr>
<td>Marahrens et al. (98)</td>
<td>Exons 1–5c</td>
<td>Mice</td>
<td>Germ cells: male with the mutant Xist is able to produce viable, fertile sperm. Somatic tissues: X bearing the mutant Xist is not inactivated, but dosage compensated by inactivation of X bearing the normal Xist; cells with both Xs active may be eliminated during development. Extraembryonic tissues: suggested that lethal when Xp bears the mutant Xist, possibly due to inability to inactivate Xp; viable when Xm bears the mutant Xist</td>
</tr>
</tbody>
</table>

* In both studies, the Xist gene is replaced with the neomycin resistance gene by homologous recombination.
* An Xist deletion was created on 129-derived X chromosome in a female XX ES cell line (PGK12.1).
* The deletion was created on 129-derived X chromosome in a male XY ES cell line (J1).
TABLE 4. Summary of Xist transgenic experiments

<table>
<thead>
<tr>
<th>Study</th>
<th>Transgene</th>
<th>Integration</th>
<th>Copy no.</th>
<th>Studied in:</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al. (86)</td>
<td>450-kb YAC</td>
<td>Autosomal: Tg Xist promoter</td>
<td>3–24</td>
<td>ES cells and fibroblasts</td>
<td>Tg Xist is expressed exclusively (53–76% of cells) or mutually (24–44% of cells) with endogenous Xist upon differentiation. Expression of Tg Xist represses the downstream reporter gene. Tg Xist RNA coats an autosomal region flanking the Tg Xist in cis.</td>
</tr>
<tr>
<td>Lee and Jaenisch (87)</td>
<td>450-kb YAC</td>
<td>Chromosome 12 in male ES cells</td>
<td>24</td>
<td>ES cells and fibroblasts</td>
<td>Tg Xist RNA coats the autosomal region in cis. Tg Xist inactivates a region of 50 cM around the integration site and causes the Ch.12 Tg to be late replicated and hypomethylated.</td>
</tr>
<tr>
<td>Herzing et al. (61)</td>
<td>35-kb YAC</td>
<td>Autosomal: Tg Xist promoter</td>
<td>2–8</td>
<td>Transgenic mice</td>
<td>Endogenous Xist expression is induced in 30% of male differentiating ES cells. Expression of Tg Xist represses the downstream reporter gene.</td>
</tr>
<tr>
<td>Heard et al. (57)</td>
<td>460-kb YAC</td>
<td>Autosomal by pronuclear injection</td>
<td>1–2</td>
<td>Transgenic mice</td>
<td>Tg Xist is not expressed in males and females. Tg Xist in methylated. Reporter gene downstream of Tg Xist is expressed.</td>
</tr>
<tr>
<td>Matsuura et al. (99)</td>
<td>350-kb YAC</td>
<td>Autosomal Y by pronuclear injection</td>
<td>3–4</td>
<td>Transgenic mice</td>
<td>Autosomal: Tg Xist is not expressed in males (female not tested) and is methylated. Y-linked: Tg Xist is expressed, and some copies are hypomethylated. Mice are phenotypically normal.</td>
</tr>
<tr>
<td>Hendrich et al. (60)</td>
<td>0.2–1.2-kb promoter</td>
<td>Autosomal by pronuclear injection</td>
<td>NR</td>
<td>Transgenic mice</td>
<td>Tg Xist promoter is active in 9.5–13.5-dpc embryos in 9 of 11 lines to variable extents. DNA methylation was not investigated.</td>
</tr>
<tr>
<td>Goto et al. (51)</td>
<td>233-bp promoter</td>
<td>Autosomal by pronuclear injection</td>
<td>1–71</td>
<td>Transgenic mice</td>
<td>Tg Xist promoter is inactive and methylated in somatic tissues but active and hypomethylated in the testis. Activity of Tg Xist promoter in preimplantation embryos is correlated with the extent of methylation in sperm.</td>
</tr>
</tbody>
</table>

* In all studies, transgenes were integrated at a single site.
* cM, centimorgan; Tg Xist, transgene Xist; Ch.12 Tg, chromosome 12 with the Xist transgene integrated.
* NR, not reported.

The XY ES cells is chosen to be up-regulated in a proportion of the cells following induction of differentiation. Lee et al. (86) have reported that the YAC transgene induces cell death in the male ES cells, probably due to activation of the endogenous Xist gene and inactivation of the single X chromosome in these cells.

Transgenic mice harboring Xist-containing YAC transgenes have been created by Heard et al. (57) and Matsuura et al. (99) (Table 4). Heard et al. (57) reported that the Xist gene in an autosomal integrated 460-kb YAC transgene construct is not expressed but that a lacZ reporter gene, which is ligated downstream of the Xist in the YAC, is expressed. They concluded that essential sequences for Xist expression might be absent in their construct and therefore that the construct could not cause inactivation of the region surrounding the integration site. Similarly, Matsuura et al. (99) reported that the Xist gene in an autosomal integrated 350-kb YAC transgene construct is not expressed. However, in one transgenic line, in which a 240-kb YAC transgene (due to spontaneous deletion of the 110-kb 3′-region of the 350-kb YAC) is integrated in a heterochromatic region of the long arm of the Y chromosome, the Xist gene is expressed. This indicates that the 240-kb Xist-containing YAC transgene has the information required for expression of the Xist gene in vivo in certain circumstances. This contradicts the conclusion by Heard et al. (57). These large Xist-containing YAC transgenes include the 9-kb 5′-flanking and 6-kb 3′-flanking sequences shown to be sufficient for initiating inactivation of the region surrounding the autosomal integration site in differentiating ES cells (61), but neither expression of the transgene Xist nor inactivation of the surrounding autosomal region was observed with the autosomally integrated transgenes in adult mice.

The differences between the results obtained with ES cells by Lee et al. (86) and Herzing et al. (61) and those obtained with adult transgenic mice by Heard et al. (57) and Matsuura et al. (99) may be explained by the differences in the copy number and/or by the differences in the methods of introduction of the YACs into ES cells and oocytes. For example, pronuclear microinjection of a large YAC transgene may alter the nature of the Xist gene and its flanking sequences, resulting in permanent silencing of Xist. It is also possible that in the transgenic mice, the cells in which the transgene expresses Xist and initiates inactivation of the autosomal genes around the integration site were eliminated or that fetuses where this occurred died in utero. In the case of the Y-linked transgene (99), which is the only YAC transgene expressed in transgenic mice, the relatively smaller size (240 kb) and/or the structure of the integration site may have allowed expression. Alternatively, the region around the integration site may not contain essential genes.

Although these knockout and YAC transgene studies indicate that the Xist gene and the 15 kb of surrounding sequences...
have some properties of the Xic, i.e., counting, spreading, and maintenance, it remains to be investigated how Xist expression itself is regulated. To this end, further transfection and transgene studies have been carried out with Xist promoter-reporter gene constructs (see below).

**Xist promoter methylation and function.** Deletion and mutation analyses of Xist promoter-reporter gene constructs in transfection assays have defined the Xist minimal promoter sequence: nucleotide (nt) sequences −81 to +1 for the mouse Xist (129) and −93 to +31 for the human XIST (60). Both the mouse Xist and the human XIST minimal promoters contain the binding sequences for ubiquitous transcription factors, TATA box-binding protein, and Sp1 (60, 79, 129). An additional binding site for the initiation protein YY1 has also been identified within the human XIST minimal promoter (60). Transient transfection assays have shown that the Xist minimal promoter is active to a similar extent in male and female cultured somatic cells (60, 129). These results indicate that repression of Xist (on the single active X chromosome) in male somatic cells is not due to the absence of transcription factors but is due to the presence of some mechanism to silence the Xist gene.

We have investigated a 233-bp fragment of the mouse Xist promoter (nt position −220 to +13) by using a reporter construct, incorporating the minimal promoter sequence linked to the firefly luciferase gene in both microinjected mouse embryos and transgenic mice (51). To investigate the role of DNA methylation in Xist expression, we have created six lines of transgenic mice (Table 4). In all six lines, the 233-bp promoter fragment is inactive and methylated in somatic tissues in both males and females. However, in the male germ cells in all six lines, the promoter transgene is active and hypomethylated, although the degree of hypomethylation is variable among the lines.

Hendrich et al. (60) have recently reported a transgenic-mouse study of three different human XIST promoter fragments (nt −1187 to +31, −221 to +31, and −211 to +767) linked to the luciferase reporter gene. They found that the promoter fragments are active to variable extents in both male and female 9.5- to 13.5-dpc embryos in 9 of 11 transgenic lines. Their results contrast with those obtained with the mouse Xist transgenes by Goto et al. (51) and others (57, 99) in that the human XIST promoter transgenes are active in most transgenic mouse embryos whereas the mouse Xist transgenes are inactive in all cases of autosomally integrated transgenes (51, 57, 99). It is possible that the presence of the exogenous (transgene) mouse Xist promoter interferes with some function of the endogenous mouse Xist promoter in development, and therefore that the transgene Xist promoter is preferentially inactivated and methylated (51, 57, 99) or cells with an active transgene Xist promoter are eliminated during development. In the case of the human XIST promoter transgene in transgenic mice, such an interference may not occur between the human XIST promoter and the mouse Xist promoter.

In our study, we specifically addressed whether imprinted expression of the endogenous Xist gene in preimplantation embryos is controlled by DNA methylation (51). In microinjection experiments, we have shown that the 233-bp Xist promoter fragment (−220 to +13) is sufficient to drive transcription of the luciferase reporter gene at the two-cell stage when microinjected into one-cell embryos. This promoter activity is repressed by in vitro methylation of the construct with site-specific DNA methylases prior to microinjection, consistent with the hypothesis that DNA methylation regulates the expression of the endogenous Xist gene in preimplantation embryos. To further investigate the role of DNA methylation as a gametic imprint for imprinted Xist expression, the activity and degree of methylation of the transgene were studied in the male germ cells of the testis and in morula-stage embryos in the established transgenic mouse lines. We have shown that the degree of methylation of the transgene Xist promoter in the germ cells (and sperm) does not correlate with the promoter activity in these cells in the testis (the promoter is active even when it is methylated) but does correlate with the activity in morula-stage embryos following transmission of the transgene from the father. Only when the transgene is extensively hypomethylated in sperm is it active, after fertilization, in transgenic morulae (51). In addition, we have found that certain CpG sites within the transgene Xist promoter are more resistant to the genome-wide demethylation occurring between the eight-cell and blastocyst stages than are neighboring CpG sites, suggesting that methylation of CpG sites within the transgene Xist promoter can be locally and independently regulated. Significantly, one of the transgene promoter CpG sites surviving the genome-wide demethylation present in the sequence 5′-CGGCCGCGG-3′ (−44 to −36), which is differentially methylated in sperm and eggs (168) and whose methylation is recognized by the methylation-dependent repressor protein identified by Huntriss et al. (64). These results strongly support our hypothesis that the gametic methylation pattern governs the imprinted expression of the endogenous Xist gene in pre-implantation embryos.

**Human XIST Regulation in Comparison with Mouse Xist Regulation**

Although the mouse and human Xist genes are positionally and structurally conserved, it is not yet known whether the two genes share all aspects of their functions in XCI. For example, the continued presence of the human XIST gene is not required for maintenance of the inactivity of the established X chromosome (22, 133), but it is not known whether this is also the case in mice. A transgene containing the mouse Xist gene and a 15-kb flanking sequence, integrated on an autosome, is counted as a separate Xic in differentiating ES cells (61), but the duplicated XIST-containing chromosomal regions, on the same X chromosome, are not counted separately in humans (122).

A most striking difference in the characteristics of the mouse and human Xist genes is their expression patterns in preimplantation embryos. In mouse preimplantation embryos, Xist expression is imprinted, with the paternal allele being preferentially expressed (73). By contrast, in human preimplantation embryos, XIST expression is not demonstrably imprinted, in that both the paternally and maternally inherited alleles are expressed (40, 137). As to the molecular basis underlying imprinted mouse XIST expression in preimplantation embryos, we have strong evidence that the promoter sequence 5′-CGGCCGCGGG-3′ (−44 to −36) is one of the key CpG sites, whose methylation in gametes controls the promoter activity in preimplantation embryos (51, 64, 168) (see above). In humans, where XIST expression in preimplantation embryos is not imprinted (40, 137), this mouse promoter sequence is not conserved but is replaced by 5′-GGCCCGCCCTT-3′ (−45 to −37), thus containing no methylatable CpG sites (60) (for detailed sequence comparison, refer to reference 60).

Plenge et al. (131) have recently reported that a single-base-pair mutation within this sequence, from 5′-GGCCCGCCCTT-3′ to 5′-GGCACCCTCTT-3′, in a short XIST promoter-reporter construct causes a two- to fivefold decrease in promoter activity in a transfection assay and that the X chromosome carrying the mutation is preferentially inactivated (skewed XCI) in so-
matic cells of females from two unrelated families. Their results indicate that this promoter sequence in humans not only is required for normal transcription but also affects some step(s) (choice, initiation, spreading or maintenance) of XCI in development.

In mice, Huntriss et al. (64) have shown that mutations of the corresponding mouse sequence, in its unmethylated form, similarly result in a three- to fivefold reduction in promoter activity in cultured cells. Given the result in humans as obtained by Plenge et al. (131), it is not difficult to see that a change in promoter activity, induced by either base pair mutation or DNA methylation, affects the XCI process in mouse development.

The role played by human XIST in preferential paternal XCI in the extraembryonic tissues is not clear. Since XIST expression in human preimplantation embryos is not detectably imprinted, it is not yet known whether preferential paternal XCI in the human extraembryonic tissues is preceded by preferential expression of the paternal XIST allele, as it is in mice (40, 50, 56, 137).

CONCLUSION AND FUTURE DIRECTIONS

We describe fundamental features of XCI and recent advances in molecular analyses and characterization of the Xic, with particular emphasis on the Xist gene. The process of XCI provides an excellent model system for investigation of the regulation of gene activity. XCI shows a variety of features of gene regulation, such as DNA methylation, histone modification, replication timing, formation of inactive chromatin structure, compartmentalization, and genomic imprinting. Studies on the Xist gene will eventually shed light on the mechanisms of XCI. Future studies will involve the identification of functionally important regions of the Xist gene and Xist RNA, dissection of genomic loci surrounding the Xist gene within and around the Xic, characterization of higher-order gene structure around the Xist gene, and elucidation of the subnuclear organization of the inactive X chromosome, in association with Xist RNA, the nuclear membrane, and as yet unidentified proteins, within the nucleus.

ACKNOWLEDGMENTS


X-CHROMOSOME INACTIVATION


