

Coordinate control and variation in X-linked gene expression among female mice

Alex D. Greenwood,^{1,*} E. Michelle Southard-Smith,^{1,**} Andrzej T. Galecki,² David T. Burke^{1,2}

¹Department of Human Genetics, 1150 West Medical Center Drive, University of Michigan, Ann Arbor, Michigan 48109-0618, USA

²Institute of Gerontology, University of Michigan, Ann Arbor, MI 48109, USA

Received: 5 March 1997 / Accepted: 17 June 1997

Abstract. In normal female mammals, one of the two X Chromosome (Chr) homologs per cell is silenced coordinately during early embryogenesis. The genes located on the inactivated X homolog are predicted to be influenced by the same underlying repression mechanism. To test the uniformity of *cis*-acting gene repression, 32 genetically identical F₁ female mice were analyzed for differential expression of homologous alleles at three X-linked genes—*Otc*, *Atp7a* (=Mottled), and *Hprt*. Gene expression was assayed by the single-nucleotide primer extension (SNuPE) method, thereby allowing the three genes to be quantitated from the same RNA sample. Although variable between individual animals, the relative expression of the two alleles (*allelic expression ratio*) of the genes is significantly correlated within each steady-state RNA pool. When examined by animal age (3 months to 12 months), no statistically significant differences were observed in the mean or variance of allelic expression ratio. Together, the results confirm that X inactivation is coordinately controlled and is stable across the early- to mid-adult life span.

Introduction

Mammalian X Chr inactivation is one of the best studied examples of large-scale coordinate gene regulation. The stable inactivation of one X Chr homolog in female mammals maintains the dose of X-linked genes between XX females and XY males (Lyon 1961, 1988). Transcriptional repression of an inactive X occurs over greater than 100 megabases, in *cis*, silencing most X-linked genes. Three observations support the hypothesis of X inactivation acting over long genetic distances: (1) autosomal loci adjacent to X-autosome translocation breakpoints can be inactivated, (2) autosomal genes integrated into the X can be inactivated, and (3) a single genetic locus—the X inactivation center, *Xic*—appears to be the initiation point for inactivation of the entire chromosome (Russell and Montgomery 1970; Krumlauf et al. 1986; Rastan 1983; Migeon 1994; Willard 1995).

X-linked genes that fail to be inactivated have been noted for specific chromosomal regions and in different species (Disteche 1995). For example, the pseudoautosomal region, which is homologous on the X and Y Chrs, contains genes excluded from inactivation. Also, the X inactive specific transcript (*Xist*) is known not to be expressed from an otherwise active X Chr, but is transcribed from the inactive X (Brown et al. 1991; Brockdorff et

al. 1991). The human X Chr contains genes at several non-pseudoautosomal locations that appear to escape *cis*-inactivation; however, only a single mouse gene is similar (Ballabio and Willard 1992; Wu et al. 1994).

Inactivation occurs early in female embryogenesis, leaving a single X Chr active in every cell. Once made, the inactive state decision is mitotically inherited by all of the cell's progeny. Consequently, X inactivation also represents an important model of mitotic cell memory occurring at the level of chromosome organization (Willard 1995). Analysis of X inactivation stability has demonstrated that for some tissue-specific genes such as ornithine transcarbamylase (*Otc*) and Mottled (*Mo* = *Atp7a*), the process of X inactivation may be mitotically reversible, as detected by reactivation of the silent allele (Wareham et al. 1987; Brown and Rastan 1988). In contrast, housekeeping genes such as hypoxanthine phosphoribosyl transferase (*Hprt*) and phosphoglycerate kinase (*Pgk1*) appear to remain stably inactivated with progressive age (Moore et al. 1992; Tsukada et al. 1991). As expected, the X inactivation process appears to be fully reversible in the female germline.

In eutherian mammals, the choice of which X Chr homolog is inactivated is random, with the exception of the cells comprising the primitive endoderm and trophectoderm, which preferentially inactivate the paternal X Chr (West et al. 1977). Random choice of X inactivation generates mosaic animals with approximately 50% of cells paternal-derived X active and 50% maternal-derived X active, in all tissues. However, in heterozygous females produced from interstrain or interspecific mouse crosses, consistent non-random X inactivation selection is sometimes observed (that is, not 1:1, maternal:paternal active). Coat color polymorphism, PGK1 allozyme analysis, and Kanda staining of morphologically distinct X Chrs have been used to observe heritable, non-random inactivation patterns (Cattanach and Williams 1972; Johnston and Cattanach 1981; Rastan 1982). Nonrandom X inactivation is controlled, in part, by an X-linked locus termed the X chromosome controlling element (*Xce*). Female animals heterozygous for *Xce* inactivate one X preferentially and give skewed distributions of mosaicism (Cattanach and Rasberry 1989). The molecular basis of the *Xce* effect is not known, but may represent an important component of the X inactivation process.

In this study, 32 (C57BL/6J × *Mus spretus*/Ei) F₁ hybrid female mice were examined for the relative expression of the C57BL/6J- or *Mus spretus*-derived alleles at the *Hprt*, *Otc*, and *Mo* genes in liver RNA. The animals were analyzed in matched cohorts, at 2, 3, 9, and 12 months of age. The complete data set represents a large, quantitative study of coordinate X-linked gene expression across the adult life span, assayed at the level of steady-state mRNA pools. This detailed quantitative analysis provides insight into the stability of coordinate regulation, the limits of age-dependent X-reactivation rates, and the inter-individual variation of X inactivation ratios.

*Present address: University of Munich, Institute of Zoology, Box 202136, Munich D-80021, Germany

**Present address: Laboratory of Genetic Disease Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

Correspondence to: D.T. Burke

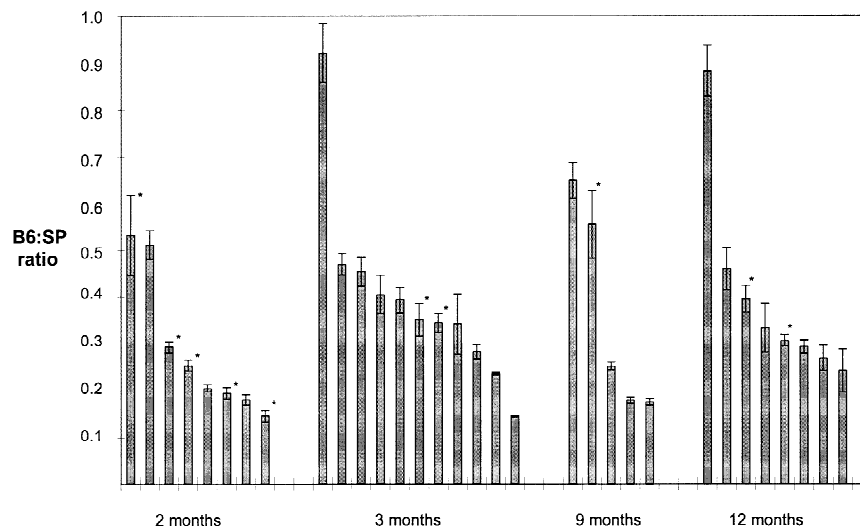


Fig. 1. Summary of *Hprt* SNUPE results for the complete set of F₁ hybrid animal liver RNAs. Allelic expression ratios are defined as relative amount of B6 to SP extension product (B6:SP). Each animal was assayed by SNUPE in triplicate. Values marked with an asterisk indicate the ten randomly chosen animals that were reverse transcribed, PCR amplified, and triplicate SNUPE assayed a second time. Within each age cohort, the animals are presented in decreasing ratio order.

Materials and methods

Experimental animals. Hybrid F₁ animals were generated from C57BL/6J (B6) female by *Mus spretus*/Ei (SP) male matings in a specific pathogen-free environment and fed ad libitum. The parental animals were all obtained directly from The Jackson Laboratory (Bar Harbor, Me., USA) breeding stock over a 12-month period. (B6 × SP)F₁ female mice of known age were euthanized, dissected, and tissues snap frozen in liquid nitrogen. No evidence of cancer or other gross pathology was observed upon dissection for any of the animals.

RNA preparation and reverse transcription. Approximately one-half of the liver was used for each animal RNA preparation. Total RNA was prepared by the acid guanidinium thiocyanate/phenol/chloroform method (Chomczynski and Sacchi 1987). Two ethanol precipitations were performed after an initial isopropanol precipitation. All RNA samples were stored at -20°C as suspensions in 70% ethanol.

Single-stranded, poly(dT) primed cDNAs were generated as described by Ausubel and associates (1991), but scaled down to 1 µg of input RNA. Twenty units of avian myeloblastosis virus reverse transcriptase was diluted tenfold in dilution buffer (10% glycerol, 10 mM potassium phosphate, pH 7.4, 0.2% Triton X-100, 2 mM dithiothreitol) and placed on ice 30 min prior to addition to the reactions as recommended by the supplier (Molecular Genetics Resources, Tampa, Fla., USA).

PCR amplification. Gene-specific oligodeoxynucleotide primers for PCR amplification were designed on the basis of published genomic sequence. Each primer pair is specific for a single gene sequence and amplifies across at least one intron/exon junction in the transcript-derived cDNA. All amplification reactions were performed as previously described (Greenwood and Burke 1996).

Single nucleotide primer extension assays (SNUPE). Reverse transcription-PCR products were purified by electroelution into polyethylene glycol (PEG)/Tris-acetate/EDTA buffer wells as described (Zhen and Swank 1993), except that 10% PEG was substituted for 15% PEG. All oligodeoxynucleotides were gel purified before use in the SNUPE assay, and primer extensions were performed as described previously (Greenwood and Burke 1996). Analysis of radioactive events from each allele-specific extension reaction was accomplished with a PhosphorImager system and ImageQuant Software (Molecular Dynamics, Sunnyvale, Calif., USA).

Statistical methods. Allelic expression ratio values were log transformed owing to a non-Gaussian distribution of data in all three gene assays. Among-group means were compared for the three genes by use of multiple analysis of variance methods (MANOVA). The Levene statistic was used to compare among-group variances. Both mean and variance

statistics were calculated at the 95% confidence limit. All calculations and graphical displays were produced with the Statistica computer software package (SAS Institute, Statsoft, Tulsa, Okla., USA).

Results

Allelic expression of *Hprt*. Thirty-two (B6 × SP) F₁ hybrid females of different ages were analyzed for relative expression of *Hprt* from the two X Chrs. For each animal, *Hprt* SNUPE assays of liver RNA were performed at least in triplicate. The allelic expression ratios are given in Figure 1, grouped by animal age cohort. To assess the variation induced by the reverse transcription and PCR steps, ten RNA samples were reverse transcribed a second time, PCR amplified, and SNUPE analyzed for *Hprt* (Figure 1, columns marked with an asterisk). In the ten replicate samples, the ratios obtained fell within the range obtained with the initial reverse transcription and SNUPE assay. It should be noted that although the animals were sacrificed to generate four age cohorts, the liver RNAs were prepared and analyzed in a random order based on animal cage number rather than age. Also, extreme care was taken with all solutions and equipment to reduce the probability of cross-contaminating samples.

The *Hprt* overall mean B6:SP ratio value, taken as the average of the mean experimental ratios obtained for each animal, was 0.37. The result confirms a stronger expression for the SP allelic version of *Hprt*. The expression ratios have a high degree of variability, both within each age group and between age groups, ranging from a B6:SP ratio of 0.14 to 0.92.

Allelic expression of *Otc* and *Mottled*. From the same initial poly(dT)-primed reverse transcriptions used in the *Hprt* analysis, *Otc* and *Mottled* (*Atp7a*) cDNA sequences were amplified and SNUPE was analyzed. In the case of *Otc*, additional experiments comparing the SNUPE ratio values with RNase protection experiments indicated that the SP allele had a consistent and reproducible primer extension bias (not shown). The ratio value for each sample was corrected to account for the assay skewing. The adjusted SNUPE mean allelic ratio of the full set of animals for *Otc* was slightly greater than 1.0, and ranged from 2.5 to 0.3. *Mo* is a low-abundance transcript in liver (Levinson et al. 1994); nevertheless, it was successfully analyzed by the SNUPE method. The overall mean allelic expression ratio (B6:SP) for the *Mo* reaction was 1.1, and ranged from 2.2 to 0.5.

Each of the two-gene expression combinations was analyzed for coordinate X Chr regulation. Figure 2 presents the non-

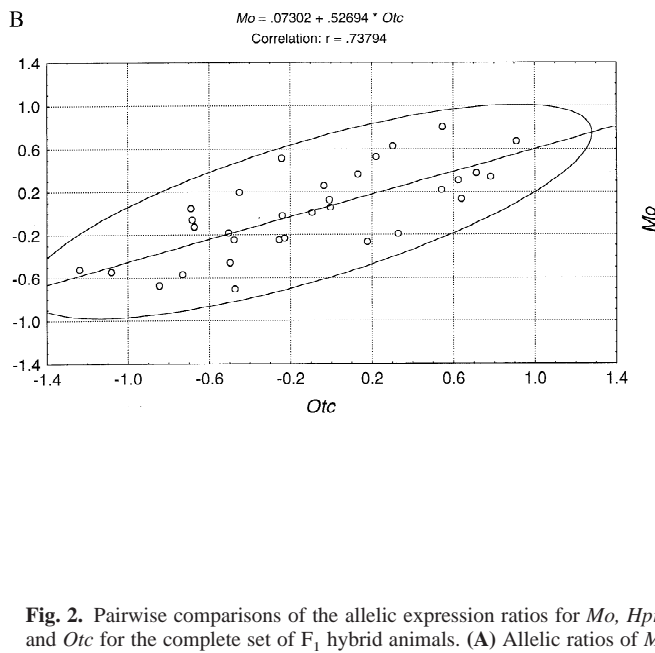
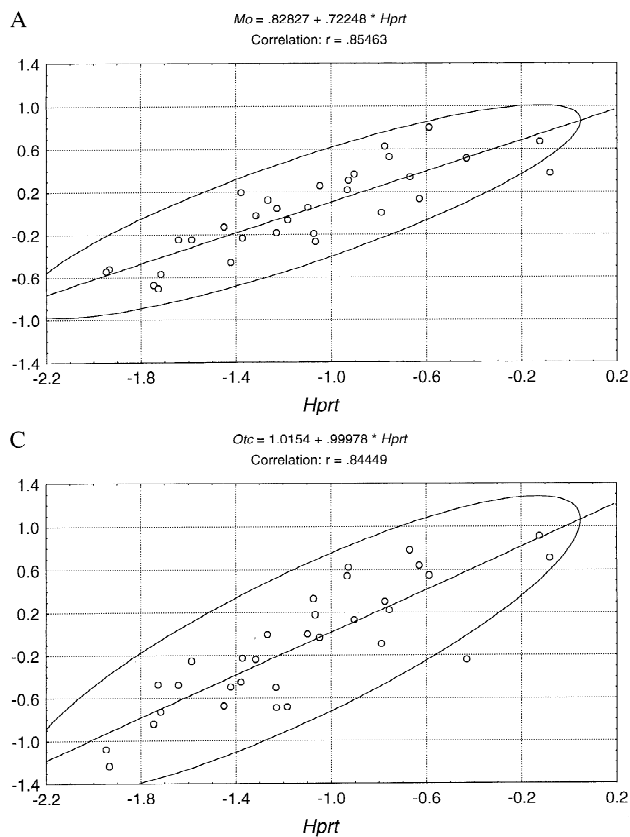


Fig. 3. B6:SP allelic expression ratios for *Hpvt*, *Otc*, and *Mo*, analyzed by age cohort. The mean and standard deviation for each age and gene are indicated. Triangles (\blacktriangle) represent *Hpvt* data, squares (\blacksquare) represent *Otc* data, and circles (\bullet) represent *Mo* data. Numerical values of each mean are given. No statistically significant changes ($p < 0.05$) were observed between age cohorts for means or variance.

parametric statistical analysis of the three comparisons. The scatterplots of the complete set of animals, with a best fit line, are shown for *Hpvt* versus *Otc*, *Hpvt* versus *Mo*, and *Otc* versus *Mo*. The stability of mean allelic ratios is reinforced by the strong correlation of *Hpvt* and *Otc* data sets, $R = 0.84$, $p < 0.0001$. *Mo* also demonstrated pronounced correlation with *Hpvt* and *Otc*: $R = 0.85$, $p < 0.0001$, for *Hpvt*, and $R = 0.74$, $p < 0.0001$, for *Otc*.

Age-related changes. Multivariate analysis of variance (MANOVA) did not reveal statistically significant differences in the SNUPE means between age cohorts for *Hpvt*, *Otc*, or *Mo* (Wilks lambda = 0.82, Rao's R = 0.59, $p = 0.799$). Although a minor trend toward increasing mean allelic ratios is observed for *Mo*, it is clear that for *Hpvt*, *Otc*, and *Mo* the mean and variance is stable (Fig. 3). The Levene test for homogeneity of variance revealed no statistically significant differences in the variance be-

tween age groups for *Hpvt* or *Otc*. *Mo* did reach a nominal statistical significance ($p = 0.02$) for age-dependent change in variance.

Discussion

Quantitative allelic ratio assays. Prior to this study, the simultaneous analysis of steady-state mRNA levels from multiple X-linked genes has not been performed in a large population of genetically identical animals. The analysis of relative proportion of X Chr homolog expression requires assay methods that are quantitative, consistent between different genes, and amenable to processing large numbers of RNA samples. In the single nucleotide primer extension (SNUPE) procedure, mRNA transcripts originating from the two alleles are distinguished by single-nucleotide polymorphisms identified within the transcribed sequence. SNUPE assays for detection of B6 and SP alleles were developed for *Hpvt*, *Mo*, and *Otc* (Greenwood and Burke 1996; Singer-Sam et al. 1992a), and for each of the three gene-specific assays the response is linear across a range of B6:SP allelic ratios from 1:50 to 50:1. Consequently, the assay is useful within the expected range of allelic ratios from female F_1 hybrid animals.

Coordinate control of three X-linked loci across adult life. Quantitative expression assays of the three X-linked genes confirmed the prediction of regulated coordinate control on X Chr homologs. In each pairwise gene comparison, statistically robust correlation coefficients were observed for the full population (Fig. 2). This effect is independent of the absolute values observed for the genes.

The strong correlation across the four age groups (Fig. 3) is suggestive of highly stable coordinate X inactivation across adult life. Previous experimental evidence has suggested that reactivation of the silent homolog of X-linked genes occurs progressively

with age for some genes (Wareham et al. 1987; Brown and Rastan 1988; Tsukada et al. 1991; Moore et al. 1992). In these experiments, reactivation was measured for individual genes, in animals with a constitutively active, translocated X Chr (T16H) as one homolog. The method of analysis differed for each gene and study, with most assays occurring at the level of protein product or gross phenotype. In contrast, the current work uses a molecular assay that is essentially identical for all genes and directly gauges steady-state mRNA pools.

In this study involving interspecific F_1 hybrid female mice, age-dependent X reactivation might be observed as a progressive change in the mean allelic expression ratio toward 1. In the context of a "strong" *Xce* allele paired with a weak allele, the initial ratios of allelic transcripts in the RNA pool will be skewed in favor of the X Chr bearing the stronger allele (for example, such an effect is seen in the *Hprt* data). For example, if one were to examine 100 cells (200 X Chrs) with an initial ratio of 35 cells with the "weak" X Chr active exclusively and 65 cells with the "strong" X Chr active, the initial neonatal expression ratio would be 0.53. With X reactivation occurring, the ratio of "weak:strong" at 12 months would be approximately 0.8 and 0.6, calculating 5% and 1% probability of an X Chr reactivating per month, respectively. This effect is the result of an increasing number of cells having two Xs active. By chance, the most common event is reactivation of a "weak" chromosome. In the limit, the allelic expression ratio would plateau at 1.0 when all 200 X Chrs are active. Similarly, the variance in a population of animals would decrease as the ratio approached 1.0, by the same mechanism.

Experimentally, the allelic expression ratio data show no statistically significant change across age, in either mean or variance. The variation observed in the early age cohorts, however, tends to obscure any age-dependent effects that may be occurring. Post-hoc power calculations indicate that for this study design, approximately 70 animals per age cohort would be required to detect changes in the mean *Hprt* allelic ratio of >28% with power of 80%, $\alpha = 0.05$. Within the current data set, a high degree of X Chr reactivation could have occurred without being detected. However, the strong age-independent correlation of allelic expression between genes analyzed suggests that X inactivation is stable at least to 12 months of age.

Inter-individual variation. The observed inter-individual variation in X homolog inactivation ratio is somewhat surprising, given that the animals are genetically identical. The use of liver tissue as the sample for the study may have some influence on variation, since the number of early embryonic precursors may be relatively small. However, previous studies of PGK1 allozyme expression for *Xce* effects indicated no significant difference between tissue types, including liver (West et al. 1977). Consequently, the allelic expression ratio values may be presumed to reflect general X inactivation homolog choice occurring throughout the animal.

The variation between animals could result from stochastic factors operating on the X Chr inactivation center at the time of homolog inactivation choice. The X inactive specific transcript (*Xist*) has recently been demonstrated to be required for initiation of X inactivation (Penny et al. 1996), and its control is a likely candidate for the target of the inactivation-choice developmental program. A 2.5-kb cluster of tandem repeats 15 kb downstream of the *Xist* gene varies in methylation status and has been proposed to be concordant with the strength of the *Xce* allele (Courtier et al. 1995). If this, or a similar, DNA sequence element is the genetic target of *Xce* and other components of the inactivation system, variation in the methylation status of the element may provide a source for the stochastic response. The degree of methylation in an embryonic progenitor cell at the decision point for X inactivation may vary over some range, yet the X inactivation choice outcome must be all-or-none. Consequently, minor differences in the degree

of methylation could lead to significant variation in X inactivation ratios. A similar minor variation in binding of chromosomal organizational proteins at the critical decision time-point may have the same effect.

The data in this study provide an accurate assessment of X inactivation patterns in hybrid animals. The animals are phenotypically normal, genetically identical, and contain no X-autosome translocations. Future studies on X inactivation are encouraged to employ biological samples of comparable uniformity, and to assay multiple X-linked genes. The quantitative measurement of the allelic transcript pools of specific genes is a powerful method for determining subtle changes in gene regulation. The SNUPE technique for assessing relative allelic expression values should be applicable to the study of other X-linked or autosomal imprinted genes that vary in allelic expression during development, pathogenesis, or over the life span.

Acknowledgments. The authors thank Anne U. Jackson and Matthew R. Nelson for assistance with statistical analysis of data and Dr. Richard A. Miller for helpful discussions. The study was supported by a graduate training fellowship from the University of Michigan Institute of Gerontology and by the National Institutes of Health (grant R01-AG11249). D.T. Burke is the recipient of support from the Searle Scholars Program of the Chicago Community Trust.

References

- Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K (1991) *Current Protocols in Molecular Biology*. (New York, NY: Greene Publishing Assoc. and Wiley-Interscience)
- Ballabio A, Willard HF (1992) Mammalian X-chromosome inactivation and the *XIST* gene. *Curr. Opin. Genet. Dev.* 2, 439–447
- Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, McCabe VM, Norris DP, Penny GD, Patel D, Rastan S (1991) Conservation of position and exclusive expression of mouse *Xist* from the inactive X chromosome. *Nature* 351, 329–331
- Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349, 38–44
- Brown S, Rastan S (1988) Age-related reactivation of an X-linked gene close to the inactivation centre in the mouse. *Genet Res* 52, 151–154
- Cattanach BM, Rasberry C (1991) Identification of the *Mus spretus Xce* allele. *Mouse Genome* 89, 565–566
- Cattanach BM, Williams CE (1972) Evidence of non-random X chromosome activity in the mouse. *Genet Res* 19, 229–240
- Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162, 156–159
- Courtier B, Heard E, Avner P (1995) *Xce* haplotypes show modified methylation in a region of the active X chromosome lying 3' to *Xist*. *Proc Natl Acad Sci USA* 92, 3531–3535
- Disteche C (1995) Escape from X inactivation in mouse and human. *Trends Genet* 11, 17–22
- Greenwood AD, Burke DT (1996) Single nucleotide primer extension: quantitative range, variability, and multiplex analysis. *Genome Res* 6, 336–348
- Johnston PG, Cattanach BM (1981) Controlling elements in the mouse IV. Evidence of non-random X-inactivation. *Genet Res* 37, 151–160
- Konecki DS, Brennan J, Fuscoe J, Caskey CT, Chinault AC (1982) Hypoxanthine-guanine phosphoribosyltransferase genes of mouse and Chinese hamster: construction and sequence analysis of cDNA recombinants. *Nucleic Acids Res* 10, 6763–6775
- Krumlauf R, Chapman VM, Hammer RE, Brinster R, Tilghman SM (1986) Differential expression of alpha-fetoprotein genes on the inactive X chromosome in extraembryonic and somatic tissues of a transgenic mouse line. *Nature* 319, 224–226
- Levinson B, Vulpe C, Elder B, Martin C, Verley F, Packman S, Gitschier J (1994) The *mottled* gene is the mouse homologue of the Menkes disease gene. *Nature Genet* 6, 369–373
- Lyon MF (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* 190, 372–373
- Lyon MF (1988) X-chromosome inactivation and the location and expression of X-linked genes. *Am J Hum Genet* 42, 8–16

- Migeon BR (1994) X-chromosome inactivation: molecular mechanisms and genetic consequences. *Trends Genet* 10, 230–235
- Moore TF, Sheardown SA, Whittingham DG (1992) Failure to detect age-related reactivation at the *Hprt* locus. *Mouse Genome* 90, 687–689
- Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N (1996) Requirement for *Xist* in X chromosome inactivation. *Nature* 379, 131–137
- Rastan S (1982) Primary non-random X-inactivation caused by controlling elements in the mouse demonstrated at the cellular level. *Genet Res* 40, 139–147
- Rastan S (1983) Non-random X chromosome inactivation in mouse X-autosome translocation embryos—location of the X inactivation center. *J Embryol Exp Morphol* 78, 1–22
- Russell LB, Montgomery CS (1970) Comparative studies on X-autosome translocations in the mouse II: inactivation of autosomal loci, segregation, and mapping of autosomal breakpoints in five T(X:1)s. *Genetics* 64, 281–312
- Singer-Sam J, LeBon JM, Dai A, Riggs AD (1992a) A sensitive, quantitative assay for measurement of allele-specific transcripts differing by a single nucleotide. *PCR Methods Appl* 1, 160–163
- Tsukada M, Wada Y, Hamade N, Masuda H, Koisumi A (1991) Stable lyonization of X-linked *Pgk-1* gene during aging in normal tissues and tumors of mice carrying Searle's translocation. *J Gerontol* 46, B213–B216
- Wareham KA, Lyon MF, Glenister PH, Williams ED (1987) Age related reactivation of an X-linked gene. *Nature* 327, 725–727
- West JD, Frels WI, Chapman VM, Papaioannou VE (1977) Preferential expression of the maternally derived X chromosome in the mouse yolk sac. *Cell* 12, 873–882
- Willard HF (1995) The sex chromosomes and X-chromosome inactivation. In: *The Metabolic and Molecular Basis of Inherited Disease*, CR Scriver, AL Beaudet, WS Sly, D Valle, eds. (New York, N.Y.: McGraw-Hill), pp 717–737
- Wu J, Salido EC, Yen PH, Mohandas TK, Heng HHQ, Tsui L-C, Park J, Chapman VM, Shapiro LJ (1994) The murine *Xe169* gene escapes X-inactivation like its human homologue. *Nature Genet* 7, 491–496
- Zhen L, Swank RT (1993) A simple and high yield method for recovering DNA from agarose gels. *BioTechniques* 14, 894–898