

RETROTRANSPOSON *MYS* IS CONCENTRATED ON THE SEX CHROMOSOMES: IMPLICATIONS FOR COPY NUMBER CONTAINMENT

ROBERT J. BAKER

Department of Biological Sciences and The Museum, Texas Tech University,
Lubbock, TX 79409 USA

AND

HOLLY A. WICHMAN

Department of Biological Science, University of Idaho, Moscow, ID 83843 USA

Abstract.—Chromosomal distribution of the *mys* retrotransposon was examined by in situ hybridization with a biotinylated probe. Thirty-six mice from four species of the *Peromyscus leucopus/maniculatus* complex were examined. *Mys* hybridized to every chromosome in all individuals examined. However, the pattern of hybridization was nonrandom. *Mys* elements were excluded from C-banding regions of the autosomes, and hybridized preferentially to G-bands. The most prominent feature of these hybridizations was the preferential accumulation of *mys* on the X and Y chromosomes of all four species examined. Accumulation of *mys* on the X is incompatible with the hypothesis that selection acting on deleterious mutations is the major mechanism regulating the copy number of this element. Rather, this supports the Langley model for containment of transposable element copy number by unequal exchange during meiosis.

Received March 21, 1990. Accepted April 6, 1990.

*Big fleas have little fleas
upon their back to bite 'em,
Little fleas have lesser fleas
and so, ad infinitum.*

Jonathan Swift

Left unchecked, parasitic sequences such as transposable elements (TE's) have the potential to burden the genome with excess copies. Several alternative mechanisms have been proposed that might limit the copy number of TE's. Each mechanism leads to different predictions as to the distribution of elements on the autosomes versus the sex chromosomes of mammals. Copy number could be controlled by selection acting on the deleterious effects of individual insertions into a vital DNA sequence (Montgomery et al., 1987), by the repression of transposition, as seen for the P element of *Drosophila* (Laski et al., 1986; Rio et al., 1986), or by unequal exchange during meiosis due to recombination between homologous elements at nonhomologous sites (Langley et al., 1989). Results of our experiments to determine if *mys* elements are randomly distributed in the genome of *Peromyscus* permit comparison with published data from *Drosophila* and discriminate among these hypothesized control mechanisms.

A growing number of mammalian TE's have been isolated and characterized, but an understanding of the biology of such elements in mammals lags behind work in eukaryotic model systems such as *Drosophila*, yeast, and maize. Here we examine the genomic distribution of *mys*, a mammalian retrotransposon (Wichman et al., 1985). *Mys* was isolated from the genome of *Peromyscus leucopus*, the white-footed mouse. The element has many features characteristic of a retrovirus, including 343 bp long terminal repeats (LTR's), a six bp target site duplication, a polypurine tract and a lysine tRNA binding site at the LTR junctions, and two open reading frames, the first of which has amino acid similarity to reverse transcriptase. The majority of *mys* elements in *P. leucopus* are 2.8 kb long; there are 500 to 1,000 elements per haploid genome with this conserved structure and size, and an unknown number of other elements that cross-hybridize with these (Wichman et al., 1985). *Mys* transcripts have been detected in brain, liver, heart, kidneys, and testes from *P. leucopus*. Although these transcripts are very heterogeneous in size, polyadenylated RNA corresponding approximately in size to a full-length *mys* was detected in some tissues (Pine et al., 1988). *Mys* ele-

ments are found throughout the genus, but internal restriction site variation defines subfamilies that differ between species of *Peromyscus* with primitive and derived karyotypes (Wichman et al., 1990). This strongly suggests that *mys* elements have been active during the evolution of the genus. Homologous sequences are present in several other closely related cricetid rodents, but are not detected in *Mus* or other mammals examined to date (Wichman et al., 1985; unpubl. data, R.J.B. and H.A.W.).

MATERIALS AND METHODS

In situ hybridization was used to examine the genomic distribution of *mys* in 36 mice. These mice represent four species from the *Peromyscus maniculatus/leucopus* complex, and were from a wide geographic range (See *Specimens Examined*). The probe for these experiments was the 1.3 kb *Eco*R I-*Hind* III internal fragment of *mys*-1 cloned into pUC-19 (Wichman et al., 1990) and labeled by biotinylation, and methods employed for in situ hybridization were essentially those of Moyzis et al. (1987, 1988), except that slides were prepared by flame drying.

Specimens Examined.—A total of 36 mice were examined by in situ hybridization. *Peromyscus leucopus*: Arkansas, Craighead Co., Jonesboro (1 male, 1 female); Texas, Garza Co., 16 mi S, 5 mi E Post (3 males); Oklahoma, Kiowa Co., 0.8 mi W, 1.5 mi S Mountain View (4 males, 1 female); Pottawatomie Co., 5.9 mi E, 2.5 mi N Tecumseh (7 males, Fig. 1C, D); McIntosh Co., 2.2 mi E Raiford (3 males, 1 female); Seminole Co., 3.5 mi E Seminole (2 females); Mexico, Quintana Roo Isla Cozumel, 20.3 km SE San Miguel (1 male). *Peromyscus maniculatus*: Canada, British Columbia, Vancouver Island, Lady Smith (2 males); Arkansas, Craighead Co., Jonesboro (1 male); California, Napa Co., Bothe-Napa State Park (1 male, Fig. 1A); Texas, Castro Co., 5.5 mi S, 2.5 mi W Dimmit (3 females, Fig. 1B); Mexico, Nuevo Leon, Ejido San Francisco (1 female). *Peromyscus gossypinus*: Arkansas, Bradley Co., 8 mi NW Warren (1 male, 1 female). *Peromyscus polionotus*: *Peromyscus* Genetic Stock Center, University of South Carolina, Columbia (1 male, 1 female). Care of animals and protocols used

were approved by the Animal Use and Care Committee of Texas Tech University.

RESULTS

As expected for a retrovirus-like element, the *mys* element hybridized to every chromosome in all species of *Peromyscus* examined. However, hybridization was not evenly distributed over the karyotype. *My*s elements were excluded from autosomal C-banding regions (Fig. 1C, D). A banding pattern was apparent on many of the chromosomes, and we were able to confirm for several chromosomes that there was a greater intensity of hybridization in the G-bands. The most prominent feature of these hybridizations was the preferential accumulation of *mys* elements on the X and Y chromosomes relative to that observed on the autosomes (Fig. 1A–C). This accumulation occurred not only in the G-bands, but also the C-banding regions and both tips of the X, and was found in all 36 *P. leucopus*, *gossypinus*, *maniculatus*, and *polionotus* examined. Evidence that the two chromosomes with the greatest accumulation of *mys* were the X and Y is as follows: In all females ($N = 11$) the two elements with the greatest accumulation were X sized and had the major banding patterns of the X chromosome (Committee for Standardization of Chromosomes of *Peromyscus*, 1977) when counter-stained with DAPI. In males ($N = 25$) only one of the two chromosomes with the greatest accumulation was X sized and appropriately banded. The other chromosome with excessive accumulation appeared to have no homologue and was the size of the Y described for that species.

DISCUSSION

The distribution of *mys* elements on the chromosomes leads to the conclusion that *mys* insertion and/or removal is a non-random process. Sequence analysis indicates that the target site for *mys* insertion is A+T-rich (Wichman et al., 1985; Pine et al., 1988). Because G-bands are known to be A+T-rich relative to R bands (Bernardi et al., 1985), it was anticipated that *mys* elements might preferentially insert into G-bands. Furthermore, satellite DNA from a number of *Peromyscus* species bands below the main band DNA on a cesium-chloride gradient

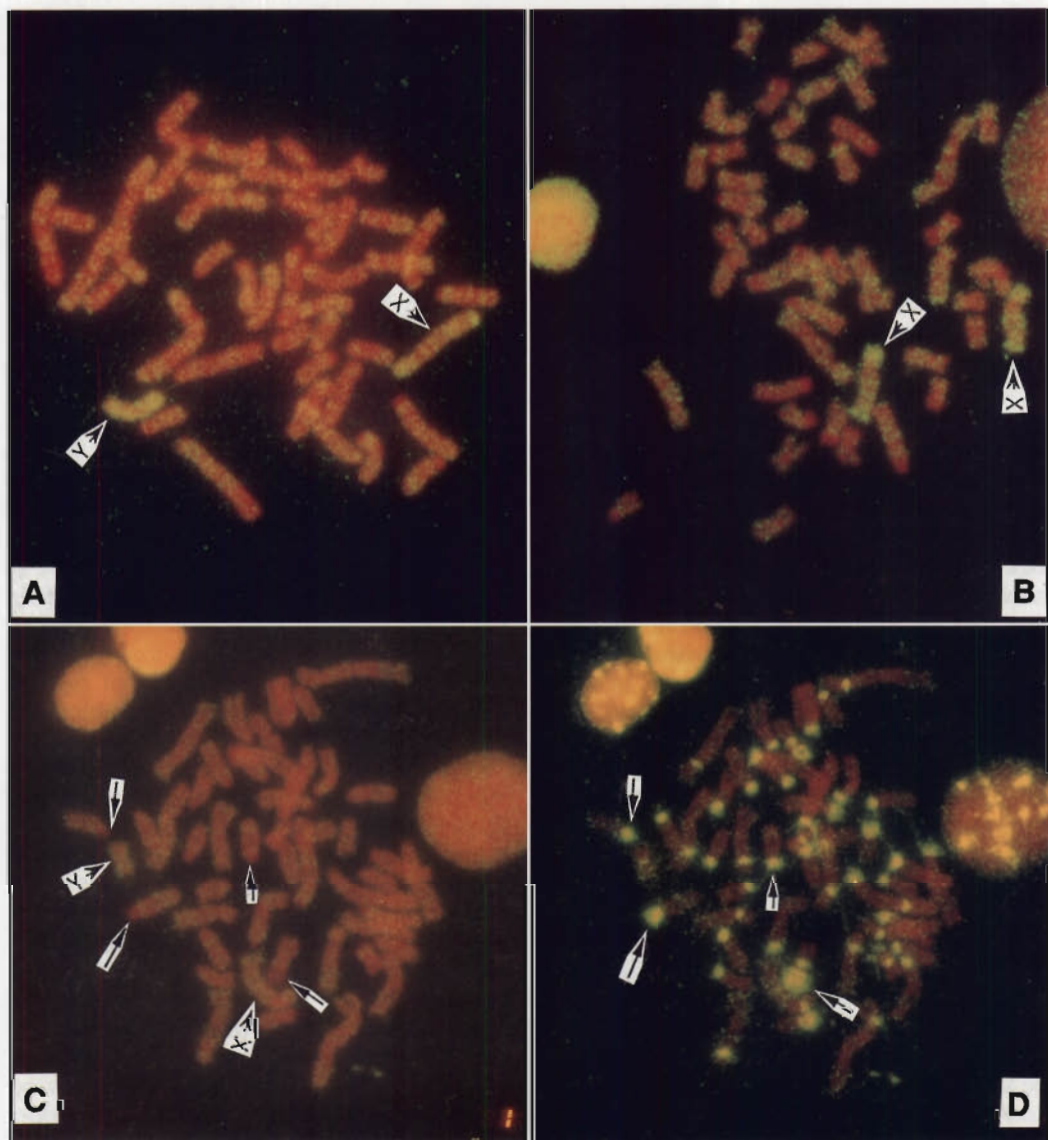


FIG. 1. Karyotypes of a male (A) and female (B) *Peromyscus maniculatus* in situ hybridized with *mys* showing the preferential accumulation of the TE on the sex chromosomes. (C) Karyotype of a male *P. leucopus* in situ hybridized with *mys* showing the regions that appear to be free of the TE's. Four representative regions are marked by white arrows. (D) Same karyotype shown in Figure 1C sequentially in situ hybridized with a cloned probe homologous to satellite DNA from *P. leucopus* to show that most areas that appear to be free of *mys* strongly hybridize to this heterochromatic sequence.

(unpubl. data, H.A.W.), providing evidence that the autosomal heterochromatin is G+C-rich, so that a scarcity of target sites may explain the exclusion of *mys* from autosomal C-banding regions. However, there is no evidence that the greater accumulation of *mys* on the X and Y chromosomes re-

flects preferential insertion into the sex chromosomes; rather, it may be the result of non-random removal of TE's from some regions of the genome. Therefore, accumulation of *mys* on the X and Y may provide insight into the mechanism of containment of transposon copy number.

It is not surprising that the Y chromosome, which carries few functional genes, can tolerate a large accumulation of TE's. However, the X chromosome is expressed in the hemizygous condition and is the most conserved chromosome in the mammalian karyotype in terms of its structure and gene order (Haldane, 1927; Ohno, 1967; O'Brien et al., 1985). It seems paradoxical that the X chromosome accumulates more TE's than do the autosomes.

Langley et al. (1989) proposed unequal recombination between homologous TE's at non-homologous sites as a mechanism for regulating TE copy number. It is assumed that unequal recombination can occur only between two elements that are sufficiently close to each other on homologous chromosomes. They predicted that if unequal recombination is correlated with homologous recombination, then areas with reduced recombination will have an accumulation of TE's relative to those areas with greater recombination frequency. In *Drosophila*, meiotic crossing over occurs only in females, so that in any population two-thirds of the X chromosomes and one-half of the autosomes are undergoing recombination. The unequal exchange model would predict that in *Drosophila* the number of TE's on the X should be slightly reduced relative to the number on the autosomes (Langley et al., 1989). The frequency of meiotic recombination is also greatly reduced at the base and tip of *D. melanogaster* chromosomes (Lindsley and Sandler, 1977), leading to the prediction that there should be an accumulation of elements at the base and tips of chromosomes relative to the middle (Langley et al., 1989; Charlesworth and Lapid, 1989). Of three elements examined by Montgomery et al. (1987), only *412* showed a deficiency of insertions on the entire X relative to the autosomes, although in a later analysis the distribution of *roo* in the middle of the X also fits the predictions of the unequal exchange model (Langley et al., 1989). Both Langley et al. (1989), in a study of the distribution of the *roo* element on the X chromosome and the autosomes, and Charlesworth and Lapid (1989), in a study of the distribution of 10 families of TE's on the X chromosome, observed a sig-

nificant accumulation of elements at the base of chromosomes but not at the tip. Although the data from *Drosophila* do not offer unequivocal evidence for the unequal exchange model in every case, this model fits the available data better than either the null hypothesis of a random distribution of elements, or the hypothesis that copy number is regulated by natural selection against the deleterious effects of insertional mutations.

In mammals, where there is recombination in both males and females, two-thirds of the X chromosomes and all autosomes are undergoing recombination, with a limited region of recombination between the X and Y in some species. All other factors being equal, the unequal exchange model would predict the greatest accumulation of elements on the Y chromosome, and a pronounced accumulation on the X relative to the autosomes. This is the pattern seen in the distribution of the *mys* element in four species of *Peromyscus*.

We interpret the distribution of *mys* on the X chromosome as incompatible with the hypothesis that selection acting on individual deleterious insertions is the major mechanism controlling the copy number of this element. If this mechanism is the primary force in regulating the abundance of the *mys* element in *Peromyscus*, then the X chromosome should accumulate TE insertions at a reduced level relative to the autosomes, because the genes of the X are expressed in the hemizygous state, and deleterious mutations would be more easily selected out of the population (Haldane, 1927). This accumulation of elements on the X and Y offers no support for a model of control of copy number by repression of transposition, as there is no reason to believe that repression would be more effective on insertions into the autosomes than into the sex chromosomes. Finally, if copy number is unregulated, there is no a priori reason to anticipate greater accumulation of elements on the sex chromosomes. Thus data from both *Drosophila* and *Peromyscus* are most consistent with the unequal exchange model for the containment of TE copy number.

Similar to retrovirus-like elements, LINE elements have an open reading frame with

amino acid similarity to reverse transcriptase and are thus considered to be transposable (Loeb et al., 1986; Hattori et al., 1986). If the model of TE copy number containment by unequal recombination of elements on the homologous chromosomes is accurate, then LINE elements might also be expected to accumulate preferentially on the X chromosome. This is supported by the observation of Korenberg and Rykowski (1988) that in situ hybridization of LINE elements to the human X chromosome was greater than expected.

The accumulation of these TE's on the X chromosome of mice and humans is in contrast to the observation that in *Drosophila melanogaster* the number of TE's on the X chromosome either does not differ significantly from the number of elements on the autosomes, or is reduced relative to the autosomes (Montgomery et al., 1987; Langley et al., 1989). The unequal exchange model leads to different predictions for the distribution of TE's on the sex chromosomes of *Drosophila* and mammals, and yet empirical data are more consistent with these disparate predictions than with other models that have been proposed for containment of TE copy number.

An overview of these data suggests that the unequal exchange model may be broadly applicable to the containment of TE copy number in eukaryotes. The variation in crossing over between meiotic systems, such as parthenogenic and haploid/diploid systems, and the variation in the frequency of crossing over between different regions of the same karyotype provide naturally occurring experimental designs for further testing of the unequal exchange model.

ACKNOWLEDGMENTS

We thank M. Hamilton, S. Edson, C. Porter, and M. Bayouth for laboratory assistance, R. Honeycutt, J. Meyne, M. L. Pardue, and W. Modi for assistance in developing the in situ protocols used here, and T. Reeder for isolating the satellite DNA probe. This research was supported by grants from the National Science Foundation, BSR-86-00646, the National Institutes of Health, GM38727, and the Idaho State Board of Education.

LITERATURE CITED

- BERNARDI, G., B. OLOFSSON, J. FILIPSKI, M. ZERIAL, J. SALINAS, G. CUNY, M. MEUNIER-ROTIVAL, AND F. RODIER. 1985. The mosaic genome of warm-blooded vertebrates. *Science* 228:953-957.
- CHARLESWORTH, B., AND A. LAPID. 1989. A study of ten families of transposable elements on X chromosomes from a population of *Drosophila melanogaster*. *Genet. Res. Camb.* 54:113-125.
- COMMITTEE FOR STANDARDIZATION OF CHROMOSOMES OF *Peromyscus*. 1977. Standardized karyotype of deer mice, *Peromyscus* (Rodentia). *Cytogenet. Cell Genet.* 19:38-43.
- HALDANE, J. B. S. 1927. A mathematical theory of natural and artificial selection. Part V. Selection and mutation. *Proc. Camb. Philos. Soc.* 23:838-844.
- HATTORI, M., S. KUHARA, O. TAKENAKA, AND Y. SAKAKI. 1986. L1 family of repetitive DNA sequences in primates may be derived from a sequence encoding a reverse transcriptase-related protein. *Nature* 321:625-628.
- KORENBERG, J. R., AND M. C. RYKOWSKI. 1988. Human genome organization: Alu, lines and the molecular structure of metaphase chromosome bands. *Cell* 53:391-400.
- LANGLEY, C. H., E. MONTGOMERY, R. HUDSON, N. KAPLAN, AND B. CHARLESWORTH. 1989. On the role of unequal exchange in the containment of transposable element copy number. *Genet. Res.* 52: 223-235.
- LASKI, F. A., D. C. RIO, AND G. M. RUBIN. 1986. Tissue specificity of *Drosophila* P element transposition is regulated at the level of mRNA splicing. *Cell* 44:7-19.
- LINDSLEY, D. H., AND L. SANDLER. 1977. The genetic analysis of meiosis in female *Drosophila melanogaster*. *Philos. Trans. R. Soc. London, B* 277:295-312.
- LOEB, D. D., R. W. PADGETT, S. C. HARDIES, W. R. SHEHEE, M. B. COMER, M. H. EDGELL, AND C. A. HUTCHISON III. 1986. The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons. *Mol. Cell. Biol.* 6:168-182.
- MONTGOMERY, E., B. CHARLESWORTH, AND C. H. LANGLEY. 1987. A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. *Genet. Res.* 49:31-41.
- MOYZIS, R. K., L. ALBRIGHT, M. F. BARTHOLDI, L. S. CRAM, L. L. DEAVEN, C. E. HILDEBRAND, N. E. JOSTA, J. L. LONGMIRE, J. MEYNE, AND T. S. ROBINSON. 1987. Human chromosome-specific repetitive DNA sequences: Novel markers for genetic analysis. *Chromosoma* 95:375-386.
- MOYZIS, R. K., J. M. BUCKINGHAM, L. S. CRAM, L. L. DEAVEN, M. D. JONES, J. MEYNE, R. L. RATLIFF, AND J. R. WU. 1988. A highly conserved repetitive DNA sequence (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc. Nat. Acad. Sci. USA* 85:6622-6626.
- O'BRIEN, S. J., H. N. SEUANEZ, AND J. E. WOMACK. 1985. On the evolution of genome organization in mammals, pp. 519-589. *In* R. J. MacIntyre (ed.),

- Molecular Evolutionary Genetics. Plenum Publ Corp., N.Y.
- OHNO, S. 1967. Sex Chromosomes and Sex Linked Genes. Springer-Verlag, N.Y.
- PINE, D. S., E. C. BOUREKAS, AND S. S. POTTER. 1988. *Mys* retrotransposons in *Peromyscus leucopus* and transgenic *Mus musculus*. Nucl. Acids Res. 16:3359-3373.
- RIO, D. C., F. A. LASKI, AND G. M. RUBIN. 1986. Identification and immunochemical analysis of biologically active *Drosophila* P element transposase. Cell 44:21-32.
- WICHMAN, H. A., C. P. PAYNE, AND T. W. REEDER. 1990. Intrageneric variation in repetitive sequences isolated by phylogenetic screening of mammalian genomes, pp. 153-160. In M. Clegg and S. J. O'Brien, (eds.), Molecular Evolution. Alan R. Liss, Inc., N.Y.
- WICHMAN, H. A., S. S. POTTER, AND D. S. PINE. 1985. *Mys*, a family of mammalian transposable elements isolated by phylogenetic screening. Nature 317:77-81.

Corresponding Editor: J. Bull