

## Retrotransposon *Mys* Was Active During Evolution of the *Peromyscus leucopus-maniculatus* Complex

Rhonda N. Lee,<sup>1</sup> Jacqueline C. Jaskula,<sup>1,\*</sup> Ronald A. Van Den Bussche,<sup>1,\*\*</sup> Robert J. Baker,<sup>2</sup> Holly A. Wichman<sup>1</sup>

<sup>1</sup> Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA

<sup>2</sup> Department of Biological Sciences and The Museum, Texas Tech University, Lubbock, TX 79409, USA

Received: 20 October 1994 / Accepted: 18 November 1994

**Abstract.** *Mys* is a retrovirus-like transposable element found throughout the genus *Peromyscus*. Several *mys* subfamilies identified on the basis of restriction site variation occur in more than one species. The distribution of these subfamilies is consistent with the accepted species phylogeny, suggesting that *mys* was present in the ancestor of *Peromyscus* and has been active through much of the evolution of this genus. Quantitative Southern blot analysis was used to examine the variability of subfamilies in *P. leucopus* and *maniculatus*. We found that subfamilies with phylogenetically narrow distributions were more variable in copy number both within and between species than subfamilies with a broader distribution. Taken together, our data suggest that *mys* has undergone multiple rounds of transposition since the peromyscine radiation, and that five subfamilies have been amplified during the evolution of the *leucopus-maniculatus* species complex.

**Key words:** Transposable element — Retrotransposon — Retrovirus-like element — *mys* — *Peromyscus leucopus* — *Peromyscus maniculatus*

### Introduction

Genomes of all mammals are peppered with reverse-transcribed DNA sequences. In fact, Temin (1985) estimated that the mouse and human genomes contain over 500,000 reverse transcripts. Many of these reverse-transcribed sequences fall into one of several classes of transposable elements common to all mammalian genomes studied to date. These include the SINEs (short interspersed elements) that are generally present in hundreds of thousands of copies per haploid genome (Deininger 1989), the LINEs (long interspersed elements) that are present in tens of thousands of copies (Hutchison et al. 1989), and one or more classes of RLEs (retrovirus-like elements) that may be present in hundreds to thousands of copies. RLEs are sometimes included among the LINEs, but are distinguished from L1 by their strong structural similarity to the retroviruses. However, RLEs vary in their degree of similarity to known retroviruses. Retroviral characteristics found in some or all such elements include LTRs (long terminal repeats), a tRNA binding site inside the upstream LTR, a polypurine stretch inside the downstream LTR, open reading frames with similarity to one or more retroviral genes, and the production of target site duplications upon insertion into the genome. Some RLEs, such as the IAP elements (Intracisternal A-Particle) of *Mus*, exhibit all of these similarities to retroviruses (Kuff and Lueders 1988), and are likely to be recently evolved from defective retroviruses. At the other extreme are the MaLRs (mammalian apparent LTR-retrotransposons) that have LTRs, a polypurine stretch, and produce 5-bp target site duplications but do

\* Present address: Canyon Ranch, Hwy 187, Vanderpool TX 78885, USA

\*\* Present address: Department of Zoology, Oklahoma State University, Stillwater, OK 74078, USA

Correspondence to: H.A. Wichman

not encode retrovirus-like genes and do not have a tRNA binding site. Furthermore, the transcriptional start sites do not appear to be conserved, although the polyadenylation signal is present in the consensus sequence for all families described. MaLRs are present in both primates and rodents, and it has been hypothesized that they are ancient components of the mammalian genome (Smit 1993).

*Mys* is a family of RLEs identified by virtue of its high copy number in the white-footed mouse, *Peromyscus leucopus*, and its absence in the genome of the house mouse, *Mus domesticus* (Wichman et al. 1985). *Mys* elements are intermediate between IAPs and MaLRs in their similarity to retroviruses. The LTRs of *mys* have a putative transcriptional start site and polyadenylation signal. There is a lysine tRNA binding site inside the upstream LTR and a polypurine stretch inside the downstream LTR, and *mys* produces 6-bp target site duplications upon insertion into the genome. Although *mys* has two putative open reading frames, they are on the non-transcribed strand. A relationship of ORF1 to reverse transcriptase has been proposed (Wichman et al. 1985), but the sequence similarity is limited, and the identity of ORF2 remains unclear. Furthermore, *mys* is only 2.8 kb in contrast to an expected 5 kb or greater for a retrovirus. There are additional cross-hybridizing sequences in the *Peromyscus* genome, and in the genomes of closely related cricetid rodents, but the 2.8-kb family appears to be the major one in the species of *Peromyscus* that have been examined in detail. Interestingly, *mys* elements are preferentially accumulated on the sex chromosomes (Baker and Wichman 1990).

*Peromyscus* is an ideal genus in which to study the evolution of a mammalian RLE. This genus of New World rodents has long been a model system for evolutionary and population studies (King 1968; Kirkland and Layne 1989). There are about 57 currently recognized species, ranging from those with geographically restricted distributions to species such as *P. maniculatus* and *leucopus* that have extensive geographic ranges. Despite the fact that they have been extensively studied from an evolutionary perspective, phylogeny for the genus is still controversial. Nevertheless, several species groups have been recognized and are generally accepted, although the relationship of these groups to each other remains unclear (Carleton 1989). Of significance to this study are two species groups: (1) *P. leucopus* and *gossypinus* are generally considered to be sister taxa, and (2) *P. maniculatus* is part of a species complex that also includes *P. melanotis*, *polionotus*, *oreas*, *sitkensis*, *sejurgis*, and perhaps *slevini* (Carleton 1989).

Although numerous studies have elucidated some aspects of the evolutionary and population biology of SINEs and LINEs, similar studies of mammalian RLEs are lacking. For example, it has been suggested for both SINEs and LINEs that one or a few elements ("masters") act as the templates to produce most additional

elements (Deininger et al. 1992). It is not known whether there are also masters for RLEs, or whether all or many RLEs have an equal probability of producing new copies. It has been proposed that LINEs have been components of the mammalian genome since the mammalian radiation (Hutchison et al. 1989), and that they have remained active to the present (Dombroski et al. 1991). Additionally, the tempo of transposition is thought to be uneven for SINEs and LINEs (Hutchison et al. 1989), but is unknown for RLEs.

In this paper we begin to examine some of these questions for the retrovirus-like element, *mys*. Specifically, we will ask whether there is evidence that *mys* has remained active during speciation of *Peromyscus*. We will address this question by examining the distribution of restriction-site-defined *mys* subfamilies based on the phylogeny of *Peromyscus*. Examining transposable elements in a phylogenetic context may allow us to establish the origin of new subfamilies in evolutionary time. Additionally, we will examine several of these subfamilies in more detail in two wide-ranging species, *P. leucopus* and *maniculatus*, and we will also look for evidence of recent activity of specific subfamilies in these two species.

## Materials and Methods

Frozen tissues or DNA for this study were donated from the frozen tissue collection at The Museum, Texas Tech University, and by Kimberly Nelson, Pennsylvania State University. *P. oreas* was donated by Ira Greenbaum, Texas A&M. Tk, Gk, and H prefixes on specimen numbers cross-reference laboratory records and tissues to voucher specimens deposited in The Museum, Texas Tech University, Texas A&M, and the Museum of Comparative Zoology, Harvard University, respectively. DNA was prepared from tissue by the standard phenol extraction or by the method of Longmire et al. (1988). Concentrations of genomic DNA stocks were determined using the Hoechst dye method (Labarca and Paigen 1980) on an LS30 luminescence spectrometer (Perkin Elmer, Norwalk, CT).

Specimens examined: *P. alstoni*: Tk13194, Tlaxcala, Mexico. *P. atwateri*: Tk27738, Tk29509. *P. banderanus*: Tk19658. *P. boylii*: Tk32541. *P. crinitus*: Tk26308. *P. gossypinus*: Tk24227. *P. leucopus*: Tk11712, Ontario, Canada; Tk20975, Tk27680, Tk28977, Tk20973, Maine; H139, New Hampshire; Tk33791, New York; H185, H256, H359, H369, H395, H397, H401, H408, Massachusetts; Tk33886, West Virginia; Tk24940, Georgia; Tk31881, North Dakota; Tk33807, Kansas; Tk30070, Tk33834, Oklahoma; Tk33761, Tk33765, Tk33757, Tk33739, 20-314-323, Texas; Tk19777, Tk29581, Tk33878, New Mexico; Tk27653, Tk27150, Tk27148, Tk27127, Tk26335, Mexico. *P. maniculatus*: Tk24386, Texas; Tk22437, Tk24156, Idaho; Tk13649, North Carolina; Tk25739, Tk25740, Tk29798, Maine; Tk25398, Iowa; Tk27135, Tk27532, Tk27553, Colorado; Tk25401, Tk25403, Nevada; Tk13748, Tk13751, Distrito Federal, Mexico. Tk28643, Zacatecas, Mexico. Tk13404, Tk25418, Tk26231, California. *P. melanophrys*: Tk26237. *P. melanotis*: Tk13756, Tk13757, Tk19828. *P. mexicanus*: Tk34520, Tk22523. *P. oreas*: Gk1938. *P. pectoralis*: Tk30615. *P. perfulvus*: Tk19583. *P. pirrensis*: Tk26910. *P. polionotus*: Tk24230, Tk24231. *P. sitkensis*: Tk21897. *P. thomasi*: Tk20608. *P. truei*, Tk21858.

For Southern hybridization analysis, a 1- $\mu$ g sample of genomic DNA was digested with the appropriate restriction enzyme according to

manufacturer's instruction (USB, Cleveland, OH) and electrophoresed on a 22-cm-long 0.8% or 1% Synergel-agarose gel (Diversified Biotech, Newton Centre, MA), transferred to Magnagraph nylon transfer membrane (MSI, Westboro, MA), and immobilized by UV cross-linking. In the quantification study, an *EcoRI* digest of each genomic DNA was done in triplicate, and one digest for each genomic DNA was electrophoresed on three separate gels. All digests were performed using the same stock of DNA, and in each replicate all digests were electrophoresed, Southern blotted, and hybridized at the same time.

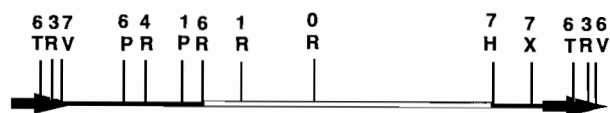
Hybridizations were performed at 58°C with  $5 \times 10^7$  cpm  $^{32}\text{P}$ -labeled probe. For the phylogenetic study the prehybridization and hybridization buffers were  $4 \times$  SSCP and  $1 \times$  Denhardt's solution; for the quantification study the buffers were  $5 \times$  SSCP and  $1.25 \times$  Denhardt's solution. ( $1 \times$  SSCP is 120 mM sodium chloride, 15 mM sodium citrate, 20 mM sodium phosphate, pH 7.0.) Initial studies were carried out using *mys*-internal as a hybridization probe (Wichman et al. 1990). ORF2, a subfragment of *mys*-internal consisting of a 450-bp *Mbol*-*HindIII* fragment of *mys*-1 (Wichman et al. 1985), was used for all quantification studies. For subfamilies in which the structural relationship to *mys* is known, ORF2 is completely included in each analyzed fragment. This allows a more accurate estimate of copy number and yields results that are comparable between subfamilies. The probes were prepared by random prime labeling (USB, Cleveland, OH) with [ $^{32}\text{P}$ ]-dCTP (DuPont, Boston, MA).

To estimate the number of copies of *mys* represented by a fragment on a Southern blot, a standard was loaded on each gel subjected to radioanalytic imaging. The standard used in all experiments was the *PstI* fragment of *mys*-1 (Fig. 1) representing 1,000 copies per haploid genome. To calculate the concentration of the *mys*-*PstI* standard, the haploid genome size of *Peromyscus* was estimated at  $3 \times 10^9$  bp and *mys*-*PstI* at 2,500 bp. In these experiments, the radioactive membrane was scanned on a radioanalytic imaging detector (Ambis 100, San Diego, CA) for 2 h (phylogenetic study) or 6 h (quantification study). The AMBIS 100 detects  $\beta$ -emissions directly from the membrane, translating counts into a composite computer image. Each band on the computer image was analyzed as a rectangular object drawn to approximately the size of the fragment, and the resulting net counts for each band were used to calculate copy number of elements in that fragment using the standard as a reference. Average copy number for each fragment and corresponding standard error were determined from the replicates. In determining standard error  $n = 3$ , except calculations of *P. leucopus* (OK) and *maniculatus* (IA) copy number, for which  $n = 2$ .

## Results

Restriction enzymes for these studies were chosen based on the known maps of *mys* elements from *P. leucopus* (Fig. 1). When Southern blot analysis is carried out on total genomic DNA using *mys*-internal (Wichman et al. 1990) or *mys* ORF2 (Wichman et al. 1985) as a probe, discrete fragments are detectable. The presence of such fragments indicates that a subset of *mys*-hybridizing sequences in the genome have a common restriction pattern. Under conditions used in this study, such subsets are only seen as discrete fragments when they contain at least a few hundred copies of *mys*-hybridizing sequences. A subset of elements yielding a discrete fragment in genomic Southern blot analysis is referred to here as a restriction-site-defined subfamily of *mys*. The size of most such fragments can be predicted from the composite restriction map (Figs. 1 and 2), although some fragments are novel.

In a pilot study, genomic Southern blot analysis was



**Fig. 1.** Composite restriction-site map of seven cloned *mys* elements. Restriction sites are: *T* = *PstI*, *R* = *EcoRI*, *V* = *EcoRV*, *P* = *PvuII*, *H* = *HindIII*, *X* = *XhoI*. Numbers above restriction sites indicate the frequency of that site in seven elements (Wichman et al. 1985). The open bar on the map represents the 1.2-kb *mys*-internal probe. The black arrows represent the LTRs. The *EcoRI* site in the middle of the map (frequency = 0) is suggested by the Southern blot analysis.

carried out on nineteen species of *Peromyscus* (*alstoni*, *attwateri*, *boylli*, *crinitus*, *gossypinus*, *leucopus*, *maniculatus*, *melanophrys*, *melanotis*, *mexicanus*, *oreas*, *pectoralis*, *perfulvus*, *polionotus*, *sitkensis*, *truei*, *pirrensis*, *thomasi*, *banderanus*). Although there was strong hybridization to the *mys* probe in all species, in five species (*boylli*, *crinitus*, *pirrensis*, *thomasi*, *banderanus*) we were unable to resolve the subfamilies of interest as distinct fragments with most enzymes selected based on the known *mys* restriction maps from *P. leucopus*. Because these species could not be reliably scored, they were excluded from further analysis. In addition, insufficient DNA was available for *P. perfulvus* to include it in subsequent analyses, but in all preliminary studies it exhibited results indistinguishable from those for *P. melanophrys*.

Four sets of genomic Southern blots were carried out on the remaining 13 species of *Peromyscus* for the phylogenetic analysis. The nine distinct fragments scored in Southern blot analysis are shown in Fig. 2. *PstI* yielded the predicted 2.5-kb fragment (referred to as T2.5) and a novel 0.7-kb fragment (T0.7). *EcoRV* yielded the predicted 2.5-kb fragment (V2.5). *EcoRI* yielded five fragments of 2.5, 2.0, 1.8, 1.5, and 1.3 kb (E2.5 to E1.3), of which 2.5, 2.0, and 1.8 kb are predicted from the composite map. *PvuII* plus *HindIII* yielded the predicted 1.56-kb fragment (PH1.56). Species were scored as positive for a fragment when it was detectable above the background smear by radioanalytic imaging. The distribution of restriction-site-defined subfamilies is superimposed on a phylogenetic tree of these species in Fig. 3. This phylogenetic tree is based on the karyotypic evolution of the genus (Stangl and Baker 1984), except that members of the *leucopus*-*maniculatus* species complex are shown as a monophyletic group. Two *mys* subfamilies, E2.5 and E1.3, were detected in all 13 species examined, and a third *mys* subfamily, V2.5, was detected in all species except *P. pectoralis*, *mexicanus*, and *attwateri*. Five *mys* subfamilies, T2.5, T0.7, E1.8, E2.0, and PH1.56, were identified only in species within the *leucopus*-*maniculatus* species complex (*P. leucopus*, *gossypinus*, *maniculatus*, *melanotis*, *polionotus*, *oreas*, *sitkensis*).

To examine the variation in restriction-site-defined subfamilies of *mys* within species, we examined individ-

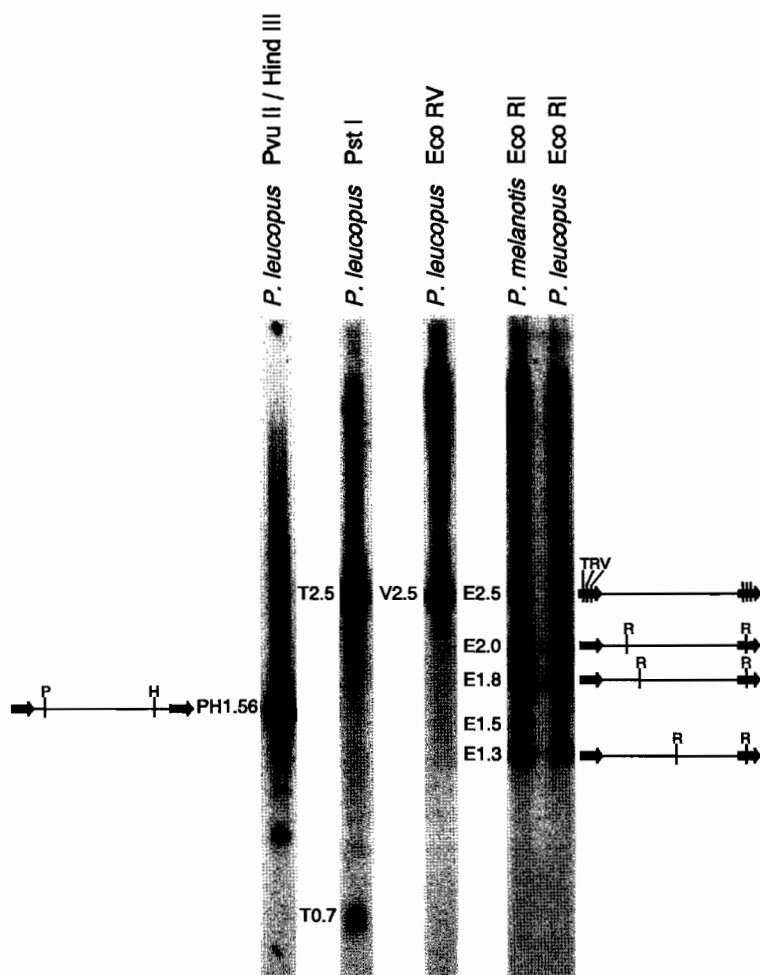


Fig. 2. Restriction-site-defined subfamilies of *mys*. Southern blot analysis of total genomic DNA digested with restriction endonucleases as indicated and probed with ORF2 reveals the nine subfamilies that were the subject of analysis. Each fragment scored is identified with a letter(s) referring to the restriction enzyme(s) used and a number referring to the size of the fragment. The restriction maps indicate the restriction sites within *mys* giving rise to each subfamily where this is known. The arrows in each map indicate the LTRs flanking the element.

uals from across the range of the two most widely distributed species of *Peromyscus*. *P. maniculatus* is distributed across most of North America, and *P. leucopus* occurs from central Maine and Canada to Cozumel, Mexico, and from the east coast to the Central Plains (Hall 1981). Examples are shown in Fig. 4. Although this technique cannot detect low levels of *mys* transposition, it could potentially uncover major transpositional bursts in these species that might be evident as new subfamilies of *mys*-related elements. With exception of the two cases discussed below, we detected only subtle quantitative differences in the hybridization patterns between individuals from across the species range. In some cases, this variation reflected minor, but reproducible, copy number differences in some restriction-site-defined subfamilies, and in some cases reflected variation in DNA loadings.

One difference was observed when genomic DNA from a mouse from Distrito Federal, Mexico, identified in the field as *P. maniculatus* (Tk19828), was digested with *EcoRI* and probed with *mys*-internal. In addition to the four fragments that are observed in all members of the *P. leucopus* and *maniculatus* species groups, the E1.5 fragment was observed in this mouse. Upon examination of the voucher specimen, this mouse was found to be *P. melanotis*. Thus these fragments may prove to be useful

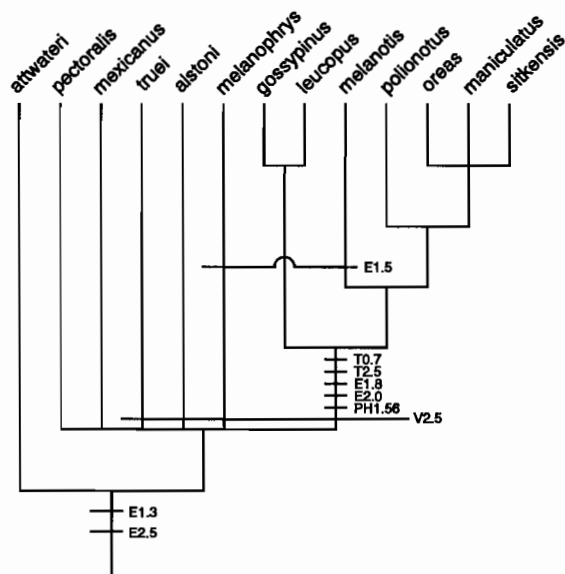


Fig. 3. Distribution of restriction-site-defined subfamilies of *mys* among 13 species of *Peromyscus*. The topology of the phylogenetic tree is based on karyotypic evolution in the genus. The emergence of each subfamily is deduced from species in which it could be detected by radioanalytic image analysis under standardized conditions.

*P. maniculatus*  
Eco RI

*P. leucopus*  
Eco RI

*P. leucopus*  
Pst I

*P. leucopus*  
Pvu II - Hind III

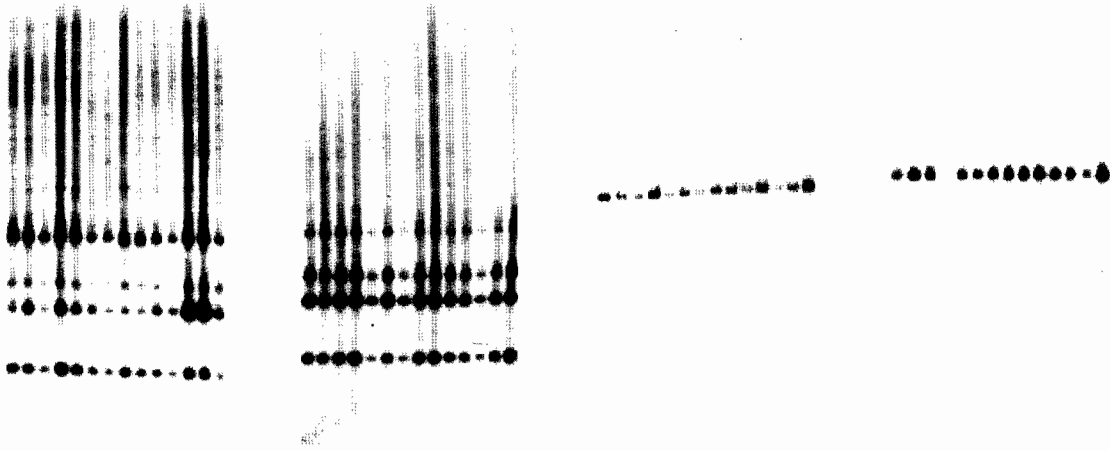


Fig. 4. Southern blot analysis of restriction-site-defined subfamily of *mys* in two wide-ranging species. This figure shows representative portions of autoradiograms resulting from Southern blot analysis of total genomic DNA from *P. leucopus* and *maniculatus* from across their species ranges. The probe is as indicated in Fig. 1.

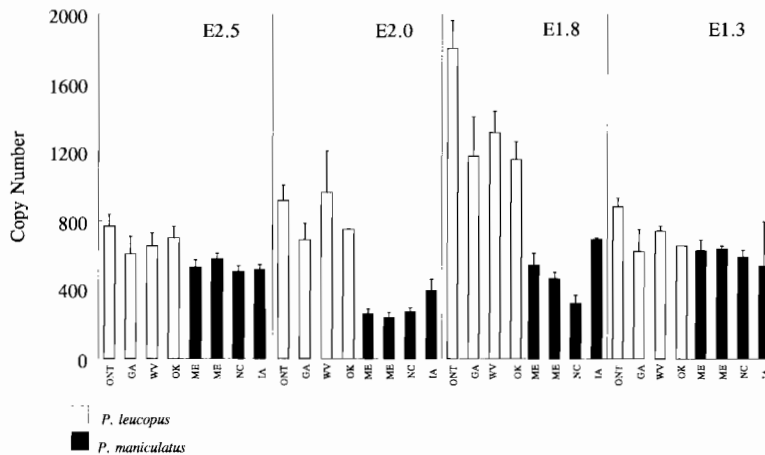
diagnostic markers for species identification at the molecular level. A second difference was observed when *P. leucopus* genomic DNA digested with *Hind*III + *Pvu*II was probed with *mys*-internal. The predicted 1.56-kb fragment, present in all other samples of *P. leucopus* and *P. maniculatus*, was absent in the Maine representative. We examined three additional specimens from the same locality and detected variation among these individuals for the presence of this fragment. Two individuals appeared to be either missing this fragment or showing it at greatly reduced intensity while the other two appeared to have this fragment in similar copy number to all other populations examined. Based on Southern blot analysis, we determined that it is the *Hind*III site, and not the *Pvu*II site, that appeared to be absent in these *P. leucopus* from Maine. These mice were originally trapped in Penobscott County, Maine, in 1986. Examination of 98 additional *P. leucopus* taken from the same location in 1993 failed to yield any mice with the *Hind*III-minus genotype (unpublished data). This finding remains enigmatic, and we lack material to carry out a more detailed analysis.

Data from the previous experiment suggest that some restriction-site-defined subfamilies varied in copy number both within and between species. To explore this variation in more detail, the *Eco*RI-defined *mys* subfamilies of *P. leucopus* and *maniculatus* from different geographic locations were analyzed by quantitative Southern blot analysis using an Ambis radioanalytic imaging system. The total number of *mys* copies accounted for by the *Eco*RI-defined subfamilies was approximately 3,600 in *P. leucopus* and 1,900 in *maniculatus*. We found that

E2.5 and E1.3 were less variable in copy number both within and between species than E1.8 and E2.0 (Fig. 5).

## Discussion

Retrotransposons move by reverse transcription of an RNA intermediate and insertion of this DNA copy into the genome at a new site. This mechanism of movement results in an increase in copy number, so that such elements may be present in hundreds to hundreds-of-thousands of copies per haploid genome. The relationship of these copies to each other will depend largely on the nature of the template for transposition. On one hand, if every copy has an equal probability of serving as a template for retrotransposition, it has been reasoned that elements will quickly diverge (Deininger and Batzer 1993). This divergence will be more rapid than for stable sequences in the genome because the high error rate of reverse transcriptase will be elaborated with each round of retrotransposition. Given this scenario, randomly selected elements from one species may be as distant from each other as they are from randomly selected elements from a distinct but related species. At the other extreme, if a single template serves as the master element that gives rise to all new copies, elements from within a species may be more closely related to each other than they are to elements from a distinct but related species. Such a relationship between sequences has been termed "concerted evolution," a concept that refers to a pattern of greater similarity of sequences within than between species rather than to the mechanisms by which such simi-



**Fig. 5.** Estimated copy number of *EcoRI*-defined *mys* subfamilies in *Peromyscus leucopus* and *maniculatus*. Bars indicate the standard error of the mean. *P. leucopus* used in this analysis were Tk11712 (Ontario), Tk24904 (Georgia), Tk33834 (Oklahoma), and Tk33886 (West Virginia); *P. maniculatus* used were Tk13649 (North Carolina), Tk25398 (Iowa), Tk25740 (Maine), and Tk29798 (Maine).

larity arises (Dover 1986). Intermediate scenarios, with several active templates or a series of active templates, may give other patterns of similarity between elements in the genome. For some elements such as LINES and SINES, it is generally accepted that there are few active templates for retrotransposition at any given period of evolutionary history (Hardies et al. 1986; Willard et al. 1987; Britten et al. 1988; Shen et al. 1991; Deininger et al. 1992). Although the details are still the subject of debate, both types of elements are clearly evolving in concert in the mammalian genomes examined to date. Similar questions have not been widely addressed for RLEs, and few cross-species comparisons have been made. Our data do not resolve this issue, but the punctuated appearance of restriction-site-defined *mys* subfamilies during peromyscine evolution suggests that only a subset of *mys* elements may be transpositionally active during some periods of evolution. Further investigation at the DNA sequence level is needed to determine how many *mys* elements are active at any time and whether elements of one subfamily form a more homogeneous group in terms of sequence similarity and age than members of different subfamilies.

Before going on to specific aspects of our data, several points need to be clarified. First, not all restriction-site-defined subfamilies examined in this study are independent of each other. For enzymes yielding multiple restriction-site-defined *mys* subfamilies, each subfamily is independent because any given element will only be present in one fragment in the Southern blot. For example, an element that belongs to the E1.8 subfamily cannot at the same time belong to the E2.5 subfamily. However, subfamilies defined by *different* restriction enzymes are not independent; i.e., the T2.5 subfamily may contain members of the E2.5, E2.0, E1.8, E1.5, and E1.3 subfamilies, as well as members that belong to none of these. In some cases multiple subfamilies may actually reflect a single class of elements. Second, the distribution of subfamilies on the phylogenetic tree (Fig. 3) suggests that the number of subfamilies is considerably greater in the *P. leucopus* and *maniculatus* species complex than in

other species surveyed. It is somewhat premature to draw such a conclusion because the large number of subfamilies observed in this species complex may be subject to ascertainment bias based upon the choice of enzymes from *P. leucopus* restriction maps. Third, in the analysis of variation within *P. leucopus* and *maniculatus*, mice from each species were selected from different locations to increase the probability of finding within-species differences if such differences exist. Although several geographic locations were sampled, it is premature to interpret differences between mice as geographic variation because we do not yet know the amount of variation within populations. Finally, estimation of copy numbers of repeated elements is always problematic because error can come from several sources (e.g., DNA loading, hybridization stringency, efficiency of DNA transfer in Southern blotting, etc.). Every effort was made to carefully control these variables, and consistency within the experiments reported here is good. However, we find that our estimates of total copy number and copy numbers for restriction-site-defined subfamilies may vary more than twofold when determined by different methods (Wichman et al. 1985; Janecek et al. 1993) and all these numbers should thus be considered very rough estimates of copy number.

SINES and LINES have proven to be good markers for evolution at various taxonomic levels (Perna et al. 1992; Kass et al. 1992; Murata et al. 1993). Our data suggest that RLEs may also be useful evolutionary markers. The phylogenetic distribution of restriction-site-defined *mys* subfamilies appears to be compatible with the accepted phylogenetic relationships among species of *Peromyscus*. Eight of the nine subfamilies examined in this study are distributed in a manner that is not inconsistent with the accepted relationships among species. E1.3 and E2.5 occur in all 13 species, whereas V2.5 occurs in ten species. There is strong support for a clade including both the *P. leucopus* and *maniculatus* species groups. Five of the nine *mys* subfamilies, T2.5, T0.7, E1.8, E2.0, and PH1.56, were seen only in species of this complex (*P. leucopus*, *gossypinus*, *maniculatus*, *melanotis*, *poliono-*

*tus, oreas, sitkensis*). The close relationship between members of the *leucopus* and *maniculatus* groups is also supported by biochemical data (Avisé et al. 1974, 1979). The ninth subfamily, E1.5, appears to have arisen or been amplified twice within the species examined: once in *P. melanotis* and independently in *P. melanophrys*. We have also observed this subfamily in *P. crinitus* and *P. perfulvus*. *P. perfulvus* is closely related to *P. melanophrys*; however, *P. crinitus* is quite distantly related to these species. It is possible that this is an ancient subfamily that was amplified independently in these species. Alternatively, these apparently similar fragments may represent distinct *mys* amplifications. It will be interesting to determine how closely related these E1.5 subfamilies may be at the sequence level.

Our original phylogenetic analysis of *mys* included five more distantly related species. Although there was strong hybridization to the *mys* probe in all species, we were unable to resolve distinct fragments with most enzymes selected based on the known *mys* restriction maps from *P. leucopus*. Detailed analysis of *mys* elements from one of these more distantly related species may allow us to select enzymes that will define subfamilies that span the genus. However, the phylogenetically distinct distribution of most subfamilies suggests that *mys* has undergone multiple rounds of transposition during the peromyscine radiation. Furthermore, it appears that different subsets of the elements in the genome contribute to these bouts of transposition. Five subfamilies (T2.5, T0.7, E1.8, E2.0, and PH1.56) have been amplified during the evolution of the *leucopus-maniculatus* species complex. The most convincing evidence of recent *mys* activity comes from the difference in copy numbers of *EcoRI*-defined subfamilies. E2.5 and E1.3 have a wide phylogenetic distribution (Fig. 3). These subfamilies show little variation in copy numbers in mice from different geographic locations in both *P. leucopus* and *maniculatus*, and copy numbers for these subfamilies do not differ dramatically between these species (Fig. 5). On the other hand, E2.0 and E1.8 are more restricted in their phylogenetic distribution, being found in detectable levels only in the *leucopus-maniculatus* complex (Fig. 3). These subfamilies differ significantly in copy number between mice of the same species and even more dramatically between species (Fig. 5). Furthermore, these two subfamilies appear to have been active in the common ancestor of *P. leucopus* and *maniculatus*, with higher transpositional levels persisting in *P. leucopus*.

A similar study of restriction-site-defined LINE subfamilies was carried out in *Peromyscus* by Kass et al. (1992). Their study included some species used in our analysis, *P. leucopus*, *gossypinus*, *maniculatus*, *polionotus*, and *truei*, as well as two species not included in our analysis, *P. difficilis*, a close relative of *P. truei*, and *P. californicus*, their outgroup. They found subfamilies that united each of the three species groups in their analysis, as well as subfamilies that united all species examined.

Direct comparison of these studies is not possible because of differences in species examined, hybridization stringency, and methods by which enzymes were chosen, but these data suggest that both *mys* and LINE have been active during the evolution of this genus.

Mechanisms for generating new genetic variability are critical to evolutionary success. One such potential mechanism is the transposition of mobile elements (Furano et al. 1994). RLEs in particular have been proposed to provide raw material for the evolution of regulatory novelty (McDonald 1990). On the other hand, high levels of transposition might lead to extinctions of local populations or even species. It remains to be determined what roles transposable elements have played in the evolution of peromyscine rodents.

**Acknowledgments.** We thank Kimberlyn Nelson and Ira Greenbaum for donating some of the tissues and DNA for this research. This work was supported by grants from the National Institutes of Health (GM 39727 to H.A.W.) and National Science Foundation (BSR-90-06797 to R.J.B.). R.A.V.D.B. was supported by a postdoctoral fellowship from the Alfred P. Sloan Foundation. R.N.L. was supported in part by a Sundquist Undergraduate Research Award. Collection of mice from Maine in 1993 was supported by a grant from the National Science Foundation (DEB-9220520 to H.A.W.) that also supported J.C.J.'s efforts on this project.

## References

- Avisé JC, Smith MH, Selander RK (1979) Biochemical polymorphism and systematics in the genus *Peromyscus*. VII: Geographic differentiation in members of the *truei* and *maniculatus* species groups. *J Mamm* 60:177-192
- Avisé JC, Smith MH, Selander RK, Lawlor TE, Ramsey PR (1974) Biochemical polymorphism and systematics in the genus *Peromyscus*. V: Insular and mainland species of the subgenus *Haplomylo-mys*. *Syst Zool* 23:226-238
- Baker RJ, Wichman HA (1990) Retrotransposon *mys* is concentrated on the sex chromosomes: Implications for copy number containment. *Evolution* 44:2083-2088
- Britten RJ, Baron WF, Stout DB, Davidson EH (1988) Sources and evolution of human *Alu* repeated sequences. *Proc Natl Acad Sci USA* 85:4770-4774
- Carleton MD (1989) Systematics and evolution. In: Kirkland Jr GL, Layne JN (eds) *Advances in the study of Peromyscus (Rodentia)*. Texas Tech University Press, Lubbock, pp 7-141
- Deininger PL, Batzer MA (1993) Evolution of retrotransposons. *Evol Biol* 27:157-196
- Deininger PL, Batzer MA, Hutchison III CA, Edgell MH (1992) Master genes in mammalian repetitive DNA amplification. *Trends Genet* 8:307-311
- Deininger PL (1989) SINEs: Short interspersed repeated DNA elements in higher eucaryotes. In: Berg DE, Howe MH (eds) *Mobile DNA*. Am Soc Micro, Washington, DC, pp 619-636
- Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH (1991) Isolation of an active human transposable element. *Science* 254:1805-1808
- Dover G (1986) Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends Genet* 2:159-165
- Furano AV, Hayward BE, Chevret P, Catzeflis F, Usdin K (1994) Amplification of the ancient murine Lx family of long interspersed

- repeated DNA occurred during the murine radiation. *J Mol Evol* 38, 18–27
- Hall ER (1981) *The mammals of North America*. John Wiley & Sons, New York
- Hardies SC, Martin SL, Voliva CF, Hutchison III CA, Edgell MH (1986) An analysis of replacement and synonymous changes in the rodent L1 repeat family. *Mol Biol Evol* 3:109–125
- Hutchison III CA, Hardies SC, Loeb DD, Shehee WR, Edgell MH (1989) LINEs and related retroposons: long interspersed repeated sequences in the eucaryotic genome. In: Berg DE, Howe MH (eds) *Mobile DNA*. Am Soc Micro, Washington, DC, pp 593–617
- Janecek LL, Longmire JL, Wichman HA, Baker RJ (1993) Genome organization of repetitive elements in the rodent, *Peromyscus leucopus*. *Mamm Genome* 4:374–381
- Kass DH, Berger FG, Dawson WD (1992) The evolution of coexisting highly divergent LINE-1 subfamilies within the rodent genus *Peromyscus*. *J Mol Evol* 35:472–485
- King JA (1968) *Biology of Peromyscus (Rodentia)*. American Society of Mammalogists, Stillwater, Oklahoma
- Kirkland Jr GL, Layne JN (1989) *Advances in the study of Peromyscus (Rodentia)*. Texas Tech University Press, Lubbock
- Kuff EL, Lueders KK (1988) Structure and evolution of retroviruses with homology to the transposable intracisternal A-particle of mice. In: Kingsman AJ, Kingsman SM, Chater KF (eds) *Transposition*. Society for General Microbiology Symposium, 43rd ed. Cambridge University Press, Cambridge, pp 248–270
- Labarca C, Paigen K (1980) A simple, rapid, and sensitive DNA assay procedure. *Anal Biochem* 102:344–352
- Longmire JL, Lewis AK, Brown NC, Buckingham JM, Clark LM, Jones MD, Meincke LJ, Meyne J, Ratliff RL, Ray FA, Wagner RP, Moyzis RK (1988) Isolation and molecular characterization of a highly polymorphic centromeric tandem repeat in the family Falconidae. *Genomics* 2:14–24
- McDonald J (1990) Macroevolution and retroviral elements. *Bio-science* 40:183–191
- Murata S, Takasaki N, Saitoh M, Odada N (1993) Determination of the phylogenetic relationships among Pacific salmonids using short interspersed elements (SINEs) as temporal landmarks of evolution. *Proc Natl Acad Sci USA* 90:6995–6999
- Perna NT, Batzer MA, Deininger PL, Stoneking M (1992) *Alu* insertion polymorphism: a new type of marker for human population studies. *Hum Biol* 64:641–648
- Shen MR, Batzer MA, Deininger PL (1991) Evolution of the master *Alu* gene(s). *J Mol Evol* 33:311–320
- Smit AFA (1993) Identification of a new, abundant superfamily of mammalian LTR-transposons. *Nucleic Acids Res* 21:1863–1872
- Stangl Jr FB, Baker RJ (1984) Evolutionary relationships in *Peromyscus*: congruence in chromosomal, genic, and classical data sets. *J Mamm* 65:643–654
- Temin HM (1985) Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retroposons, and retrotranscripts. *Mol Biol Evol* 2:455–468
- Wichman HA, Potter SS, Pine DS (1985) *Mys*, a family of mammalian transposable elements isolated by phylogenetic screening. *Nature* 317:77–81
- Wichman HA, Payne CT, Reeder TW (1990) Intrageneric variation in repetitive sequences isolated by phylogenetic screening of mammalian genomes. In: Clegg MT, O'Brien SJ (eds) *Molecular evolution*. Wiley-Liss, New York, pp 153–160
- Willard C, Nguyen T, Schmid CW (1987) Existence of at least three distinct *Alu* subfamilies. *J Mol Evol* 26:180–186