OCCASIONAL PAPERS

Museum of Texas Tech University

NUMBER 178 3 SEPTEMBER 1998

ORGANIZATION OF REPETITIVE DNA IN THE PRIMITIVE REPTILE *SPHENODON PUNCTATUS*

Ronald A. Van Den Bussche, Jonathan L. Longmire, Laura L. Janecek, and *Robert J. Baker*

Understanding the relationships and evolution of organisms requires an understanding of the patterns of variation in characters that are used to define systematic relationships. Advances in molecular biology have permitted examinations of genome organization providing a wealth of additional taxonomic characters. Repetitive elements such as rDNA, minisatellites, microsatellites, and heterochromatin have been used to resolve both specific and higher taxonomic levels (Arnason et al., 1978; Arnason and Widegren, 1989; Baker et al., 1997; Hamilton et al., 1990, 1992; Longmire et al., 1991; Love and Deininger, 1992; Macgregor and Sessions, 1986; Porter, 1994; Van Den Bussche et al., 1993, 1995, in Press). To provide insights into the abundance and position of repetitive elements in the reptilian genome, we examined the relative copy number and frequency of co-occurrence of repetitive DNA in the tautara (*Sphenodon punctatus*), which is considered by many reptilian systematists to be the most primitive extant reptile (Colbert and Morales, 1991). Using a cosmid genome library we estimate the presence and organization of repetitive DNA in the *Sphenodon* genome. More specifically, we have probed this library with five repetitive elements and genomic DNA from *Crotalus*, *Alligator*, *Iguana*, *Chrysemys*, *Cynoscion*, *Falco*, and *Homo*. This library is archived in the vital tissue collection of the Museum of Texas Tech University.

Repetitive DNA can comprise a large proportion of the genome (Britten and Kohne, 1968; Hake and Walbott, 1980; Flavell et al., 1974; Miklos, 1985; Flavell, 1986; Janecek et al., 1993). The importance of some repetitive elements, such as the role of the ribosomal genes, is well documented (Gerbi, 1985). Some repetitive elements were previously thought to be examples of selfish DNA, such as the proposed precursors of rodent B1s (BC1 RNA) and primate Alus (BC200 RNA), have been interpreted as playing a role in translation in neural tissues (Brosius, 1991; Tiedge et al., 1992; Tiedge et al., 1993). Repetitive DNA also has been postulated to play a role in chromosome evolution (Cooper, 1964; Peacock and Miklos, 1973; Vig, 1982; Bennett, 1984; Hamilton et al., 1990; Hamilton et al., 1992; Wichman et al., 1992), regulation of gene expression (Spofford, 1976), determination of chromosome structure (Flavell, 1983), genomic response to environmental and physiological stimuli (Cullis and Cleary, 1986 *a,b*; Kikuchi et al., 1987; Zheng et al., 1987), and organismal growth (Macgregor and Sessions, 1986). Finally, microsatellite loci have been shown to play a role in several human ge-

netic diseases (Heavne et al., 1992; La Spade et al., 1991; Morgante and Olureri, 1993). Therefore, the presence and organization of repetitive elements as well as the degree of conservation of DNA sequences among diverse taxa may be important for undestanding the evolution of eukaryotic genomes.

MATERIALS AND METHODS

CONSTRUCTION OF COSMID GENOMIC LIBRARY

High molecular weight genomic DNA isolated from a male *Sphenodon punctatus* (TK 27921) following the procedure of Longmire et al. (1991) was used to construct a cosmid library as described by Janecek et al. (1993) and Longmire et al. (1993). In general, this procedure consists of partially digesting genomic DNA with the restriction endonuclease *Sau*3AI and dephosphorylating with calf intestinal alkaline phosphatase. Approximately 0.5 g of dephosphorylated genomic DNA was ligated with 1.0 g of *Bam*HI cloning arms from the cosmid vector sCos-1 (Evans et al., 1989). In vitro packaging was carried out in Giga Pack Gold packaging extracts (Stratagene). Primary infection of *E. coli* host strain DH5 MCR yielded 1.6×10^5 independent recombinants. Average size of inserts in recombinant cosmids was determined by digestion of 20 randomly selected primary clones with *Eco*RI, followed by electrophoresis within a 0.7% agarose gel. Fragment sizes were determined from photographs of ethidium-stained bands and all bands were sized by comparing their mobility with two known size standards (bacteriophage- DNA digested with *Hin*dIII and a 1 kb ladder).

MOLECULAR CHARACTERIZATION OF REPETITIVE ELEMENTS IN THE TUATARA GENOMIC LIBRARY

One thousand seven-hundred and twenty-eight independent clones from the primary library were picked, grown, and archived in 96-well microtiter plates. A replica plater (Sigma Chemical Co.) was used to inoculate nylon membranes (Biodyne B 0.45 micron) with clones from the microtiter plates. Membranes were incubated at 37°C for 7 h on LB agar containing kanamycin (30 g/ml), transferred to LB agar containing kanamycin and chloramphenicol (170 g/ml; Sambrook et al., 1989), and grown at 37°C overnight. DNA was fixed onto membranes by placing the membranes sequentially on blotting pads soaked in 0.4 M NaOH (5 min.), 0.5 M Tris-1.5 M NaCl, pH 7.5 (5 min.), and 2X SSC (5 min.), followed by baking at 80°C for 2 h.

To estimate the relative abundance and composition of repetitive DNA in the *Sphenodon* genome, several probes for families of repetitive DNA known to exist in other vertebrate genomes were hybridized to the *Sphenodon* library. The abundance of dinucleotide microsatellite repeats was evaluated by using the oligonucleotide repeats $(GT)_n$ X $(CA)_n$, $(CT)_n$ X $(GA)_n$, $(AT)_n$ X $(TA)_n$, and $(GC)_n$ X $(CG)_n$. These oligonucleotides were approximately 1.2 kb, purchased from Pharmacia LKB, and represent all possible dinucleotide repeats. Hereafter, these four dinucleotide microsatellites are referred to as $(GT)_n$, $(CT)_n$, $(AT)_n$, and $(GC)_n$. Relative abundance of the rDNA cistron was estimated using a cloned fragment of the 28S subunit (pI19) gene from *Mus musculus* (Arnheim, 1979). The rDNA probe was isolated from the vector, electrophoresed on 0.8% low melting point agarose, and purified from the gel using Prep-A-Gene (BioRad Laboratories). The purified 28S rDNA fragment was gel purified a second time to ensure that all vector DNA was removed from the sample.

To evaluate the overall frequency of repetitive DNA in the cloned fragments, 1 g of *Sphenodon* genomic DNA was labeled by nick translation and hybridized to the library. Because single and low copy DNA will be in low abundance in this sample of DNA, only DNA which is moderately to highly repetitive will produce detectable hybridization signal. Genomic DNAs also were isolated from tissue or blood from *Crotalus atrox* (western diamond back rattlesnake; TK 24446), *Iguana*, *Chrysemys scripta elegans* (red-eared slider; TK 32879), *Cynoscion nebulosus* (spotted sea trout; TK 32873), *Falco peregrinus* (peregrine falcon) and *Homo sapiens* (TK 30732) using the techniques described above. Genomic DNA from *Alligator sinans* was donated by Dr. L. Densmore (Texas Tech University).

Prior to hybridization, membranes were washed for 1 h at 65°C in 0.1X SSC, 0.1% SDS. Prehybridization was carried out at 65°C for 1 h in 6X SSC, 40% formamide (Kodak), 1% SDS, 0.005 M EDTA (pH 8.0), and 0.005 g/ml Carnation evaporated milk. Membranes were hybridized overnight at 42°C in fresh prehybridization solution containing approximately $1X10^6$ cpm/ml probe. All probes were labeled with [α^{32}P]dCTP or [α^{32}P]dATP by nick translation and the nonincorporated label was removed by spin column chromatography (Sambrook et al., 1989). Prior to hybridization, probes were denatured for 10 min at 37°C in 0.1 M NaOH. Following hybridization, membranes were washed once for 15 min in 2X SSC, 0.1% SDS at room temperature and twice for 15 min in 0.1X SSC, 0.1% SDS at 50°C. Washed membranes were autoradiographed at -80°C using Kodak XAR-5 film and two lightning plus intensifying screens. For each probe, all clones were scored on a scale of 0, representing no detectable hybridization, to 3, a completely black spot on an autoradiograph, representing maximum detectable hybridization.

LIMITATIONS AND STRENGTHS OF THE METHODS

It is generally assumed that a library constructed from genomic DNA provides an accurate representation of the genome of the organism from which the DNA originally was isolated. Two factors that can prevent a library from actually being representative of the organism's genome are nonrandom distribution of restriction sites for the enzyme chosen to digest the genomic DNA and methylation of nucleotides within the recognition site of the chosen restriction endonuclease. Whereas it is not possible to eliminate potential sources of error completely, for the following reasons the construction of the *Sphenodon* library should be minimally affected by such factors. First, to eliminate the problem associated with a nonrandom distribution of restriction sites, we performed a partial digestion of the genomic DNA with *Sau*3AI, a restrictioin endonuclease that has a 4 bp recognition sequence. Second, *Sau*3AI was chosen to construct the library because this enzyme is not sensitive to methylation (Stratagene, La Jolla, Ca.).

A second cautionary note must be made regarding our estimates of copy number of repetitive elements. First, copy number estimates for specific elements are always problematic because error can come from several sources. These sources of error for direct examination of copy number can be attributed to DNA loading, stringency of hybridization and washes, as well as the efficiency of transfer of DNA in Southern blot experiments. In this study, we calculated copy number by assuming that each clone contained a single copy of the particular element used as a probe and then extrapolated how many copies would be present in a complete genome based on the amount of the genome the archived clones were estimated to represent. The approach of screening large insert cosmids to provide an estimate of copy number has the potential of providing an underestimate of the total copy number in that any recombinant cosmid could potentially contain more than a single copy of a repetitive sequence. Moreover, when genomic DNA is used as a probe, different clones will hybridize with varying intensity. This difference in intensity could be due either to relative copy number of the repetitive sequence or the relative amount of divergence seen in that family of repetitive DNA. Therefore, copy number estimates from this approach must be considered as minimal. The advantage of using large insert recombinants (such as cosmids), however, is that this method provides a means to detect nonrandom associations of repetitive elements within the genome (Janecek et al., 1993).

RESULTS

CHARACTERIZATION OF THE SPHENODON *GENOMIC LIBRARY*

Only 57 (3.3%) of the 1,728 cosmid clones screened in this study did not hybridize to any of the 13 probes examined. Of these 57 clones, digestion with the restriction endonuclease *Eco*RI suggested the absence of an insert in three clones. These three colonies apparently did not contain a recombinant cosmid or failed to grow to a density that allowed detection of the recombinant cosmid DNA using a standard miniprep procedure, bringing the actual number of recombinant cosmids screened to 1,725. Minipreped DNAs from the remaining 54 repeat-negative clones were visible on an

ethidium-stained gel when digested with *Eco*RI. Each clone was verified to have the 6.7 kb DNA fragment characteristic of the sCos-1 vector as well as additional bands totaling approximately 35 kb in size.

The size of the *Sphenodon* DNA inserted into 20 randomly-selected primary recombinant cosmids ranged from 25.8 - 50.1 kb, with a mean insert size of 39.5 kb. Based on this mean insert size, the 1,725 clones represented 6.8 X 10^7 bp. Although no estimate for *S. punctatus* genome size exists, Dingerkus (1979) examined the karyotypes of *Sphenodon* and postulated that, due to tandem gene duplications, the genome size of *Sphenodon* is larger than that of crocodilians and turtles. Based on this hypothesis and the observation that the genomes of crocodilians and turtles are approximately 70% that of a typical mammal (Szarski, 1974; Dingerkus, 1979) we can approximate the size of the *Sphenodon* genome as being essentially the same as that of a typical mammal (7.0 pg DNA/cell; Bachmann, 1972). Assuming a haploid genome size of 3 X 10^9 bp, the 6.8 X 10^7 bp represented by the 1,725 recombinant cosmids screened would represent approximately 2.3% of the *S. punctatus* genome.

CHARACTERIZATION OF REPETITIVE DNA IN THE GENOMIC LIBRARY

Table 1 summarizes the number of clones from the *Sphenodon* library that hybridized to each of the 13 probes used in this study. The four dinucleotide microsatellites used in this study varied greatly in their representation in the clones examined. Dinucleotides $(GT)_n$ and $(CT)_n$ hybridized to 1,229 (71.2%) and 543 (31.5%) recombinant cosmids, respectively. For $(GT)_n$, 457 of the clones (37.2%) were scored as maximally hybridizing (score=3) whereas 314 (25.5%) of these clones were scored as 2 and 458 (37.3%) were assigned a score of 1. For $(CT)_n$, 256 of the 543 clones (47.1%) were scored as maximally hybridizing, 108 (19.9%) were assigned a score of 2, and 179 clones (33.0%) were scored as 1. A single recombinant cosmid hybridized to the dinucleotide $(GC)_n$ whereas, no hybridization was detected with either the dinucleotide $(AT)_n$ or the 28S rDNA subunit.

Hybridization of the *Sphenodon* library with *Sphenodon* genomic DNA resulted in some degree of hybridization to 1,586 (91.9%) cosmid clones. Of these positive clones, 377 (23.8%) were scored as maximally hybridizing whereas 744 (46.9%) and 465 (29.3%) were assigned scores of 2 (medium intensity) and 1 (low intensity), respectively. When genomic DNA from representatives of other vertebrate classes was hybridized to the *Sphenodon* library, the percentage of clones producing detectable hybridization showed a reduction in hybridization compared to *Sphenodon* genomic DNA (Table 1). The mean percent hybridization for the class Reptilia is 46.9%, class Reptilia excluding *Sphenodon*, 31.9%; class Pisces, 37.7%; class Aves, 37.7%, and class Mammalia, 20.2%.

Pairwise comparisons of all probes hybridized to the *Sphenodon* genomic library are presented in Table 2. For all pairwise comparisons between *Sphenodon* genomic DNA and the dinucleotides $(GT)_n$ and $(CT)_n$, the occurrence of these probes was as expected based on their individual representation in the library. Additionally, all comparisons between *Sphenodon* genomic DNA and genomic DNA from the other seven vertebrates examined in this study were as expected based on their individual representation in the library. However, all other pairwise comparisons showed significantly higher co-occurrence than expected based on their individual representation in the *Sphenodon* library.

Table 1.— Representation of 13 probes in 1,725 independent clones from a cosmid library constructed from Sphenodon punctatus *genomic DNA. Percentages are shown in parentheses, some values do not total 100 due to rounding error.*

Probe	Negative Clones	Positive clones
rDNA	1725 (100.00)	0 (0.00)
$(AT)_n$	1725 (100.00)	0 (0.00)
$(GC)_n$	1724 (99.88)	1 (0.001)
$(CT)_n$	1178 (68.29)	547 (37.70)
$(GT)_n$	512 (29.68)	1213 (70.31)
Sphenodon DNA	138 (8.00)	1587 (92.00)
Crotalus DNA	739 (42.84)	986 (57.16)
Alligator DNA	1294 (75.01)	431 (24.99)
Iguana DNA	295 (75.05)	430 (24.93)
Chrysemys DNA	1368 (79.30)	357 (20.70)
Cynoscion DNA	1074 (62.26)	651 (37.74)
Falco DNA	1364 (79.07)	361 (20.93)
Homo DNA	1376 (79.77)	349 (20.23)

Table 2.— Pairwise comparison of the 10 probes producing detectable hybridization in the 1,825 cosmid clones from the genomic library of Sphenodon. *Two probes [rDNA, $(AT)_n$] that failed to hybridize to any of the clones and the $(GC)_n$ in probe, which hybridized to a single cosmid clone, are not included. Above the diagonal: expected number for each pair of probes based on total occurrence in the 1,725 clones examined. Chi-square values at the 0.05 level are denoted by *. Below diagonal: observed numbers for each pair of probes; + indicates pairs of probes observed more often than expected.*

	$(CT)_n$	$(GT)_n$	*Sphenodon*	*Crotalus*	*Alligator*	*Iguana*	*Chrysemys*	*Cynoscion*	*Falco*	*Homo*
$(CT)_n$	----	384.1	502.50	212.2	136.5	136.2	113.0	206.20	114.2	110.6
		(12.7)*	(0.6)	(93.4)*	(28.6)*	(74.7)*	(35.1)*	(88.2)8	(58.6)*	(13.3)*
$(GT)_n$	454	----	1114.1	692.2	302.6	301.9	250.6	456.9	253.3	245.0
	+		(0.5)	(44.6)*	(48.7)*	(40.2)*	(38.7)*	(76.7)*	(33.2)*	(42.2)*
Sphenodon	520	1138	----	905.5	395.7	394.9	327.8	597.9	331.6	320.5
				(1.2)	(1.5)	(1.1)	(0.4)	(3.4)	(1.7)	(1.3)
Crotalus	483	868	938	----	245.9	245.4	203.7	371.4	206.0	199.2
	+	+			(117.7)*	(114.5)*	(88.5)*	(162.5)*	(73.5)*	(106.6)*
Alligator	199	424	420	416	----	107.3	89.0	162.3	90.0	87.1
	+	+		+		(324.8)*	(339.9)*	(348.3)*	(239.9)*	(545.2)*
Iquana	237	412	416	413	294	----	88.8	161.9	89.9	86.9
	+	+		+	+		(349.5)*	327)*	(218.6)*	(389.8)*
Chrysemys	176	349	339	338	263	265	----	134.4	74.7	72.1
	+	+		+	+	+		(287.4)*	(227.6)*	(439.4)*
Cynoscion	341	644	643	617	400	392	331	----	136	131.5
	+	+		+	+	+	+		(212.5)*	(340.2)*
Falco	196	345	355	329	237	230	205	306	----	72.9
	+	+	+		+	+	+	+		(288.6)*
Homo	149	347	341	345	305	271	250	343	218	----
	+	+		+	+	+	+	+	+	

DISCUSSION

Most estimations of the distribution and relative copy number of repetitive elements within eukaryotes have been conducted either by direct experimental methods or by surveying DNA sequences deposited in major data banks (Hamada et al., 1982; Moyzis et al., 1989; Beckmann and Weber, 1992; Moran, 1993). Recently however, the relative abundance of total repetitive DNA and the four dinucleotide microsatellites have been estimated for a number of eukaryotes by screening cosmid libraries with large (≥35 kb) inserts (Stallings et al., 1991; Janecek et al., 1993; Longmire, 1993; Porter, 1994; Baker et al., 1995; Van Den Bussche et al., 1995). Because these studies used nearly identical techniques for the estimation of the relative frequency of total repetitive DNA and the four dinucleotide microsatellites, it allows for a direct comparison of the relative contribution of these components to these diverse eukaryotic genomes. Additionally, because a large number of independent cosmids were screened for all studies, it is possible to calculate 95% confidence intervals for our estimation of the relative abundance of these various classes of repetitive DNA in each genome to determine statistical significance using standard statistical parameters (Baker et al., 1995; Van Den Bussche et al., 1995).

TOTAL REPETITIVE DNA

When the *Sphenodon* library was probed with *Sphenodon* genomic DNA, approximately 8% of the 1,725 clones produced no detectable levels of hybridization. These clones represent the single to low copy DNA in the *Sphenodon* genome. The remaining 92% of the clones produced varying levels of hybridization and represent the highly repetitive, middle repetitive, and low repetitive DNA families. Although independent assessments on the relative contribution of these various classes of DNA in the *Sphenodon* genome based on reassociaton kinetics do not exist, Baker et al. (1995) have shown that estimates of these classes of DNA based on library screenings with genomic DNA produce nearly identical estimates to those obtained by reassociation kinetics for the upland cotton genome. Therefore, we conclude that the *Sphenodon* genome is comprised of approximately 8% unique and low copy sequences, 29.3% low copy repetitive DNA sequences, 46.9% middle repetitive DNA sequences, and finally, 23.8% highly repetitive sequences.

MICROSATELLITE SEQUENCES

Investigations into the frequency and distribution of the four dinucleotide microsatellites in various eukaryotic genomes have revealed that such repeats are distributed widely in the genome with little tendency to cluster together and that considerable variation exists among eukaryotes for microsatellite copy number (Hamada et al., 1982; Stallings et al., 1991; Janecek et al., 1993; Baker et al., 1995; Van Den Bussche et al., 1995). For example, $(GT)_n$ sequences vary from approximately 30,000 in the cow to 100,000 in *Mus* (Hamada et al., 1982). Hybridization of the *Sphenodon* cosmid library with oligonucleotide probes for the dinucleotide microsatellites $(GT)_n$ and $(CT)_n$ resulted in detectable levels of hybridization to 70% and 38% of the clones, respectively (Table 1). Assuming that 2.3% of the genome was examined, the *Sphenodon* genome contains approximately 54,000 $(GT)_n$ and 24,000 $(CT)_n$ repetitive sequences (Table 3). As with other studies, the dinucleotide microsatellites $(GC)_n$ and $(AT)_n$ are rare in the *Sphenodon* genome in that hybridization to the 1,725 recombinant cosmids resulted in hybridization of the $(GC)_n$ probe to a single cosmid and no detectable hybridization with $(AT)_n$ (Table 1). Table 3 lists the relative frequency of the dinucleotide microsatellites $(GT)_n$ and $(CT)_n$ along with 95% confidence limits for several reptilian and mammalian genomes. As can be seen from this table, considerable variation exists within both Reptilia and Mammalia for the relative copy number of the dinucleotide microsatellites $(GT)_n$ and $(CT)_n$.

An alternative way of comparing the abundance of dinucleotide microsatellites in various genomes which takes into account the variation in genome size is by comparing apparent interspersion frequencies. As with the estimates of relative abundance, considerable variation exists for interspersion frequencies in both Mammalia and Reptilia. Interspersion frequencies of $(GT)_n$ have been estimated as one $(GT)_n$ repeat every 106 kb in *Macrotus* (Van Den Bussche et al., 1995), 54 kb in humans (Moyzis et al., 1989), 40 kb in *Peromyscus* (Janecek et al., 1993),

Table 3.— Estimated copy number of 95% confidence interval for the relative abundance of the dinucleotide microsatellite repeats $(GT)_n$ and $(CT)_n$ for six vertebrate genomes based on screening cosmid libraries.

Taxon	$(GT)_n$	$(CT)_n$	Genome Size
Sphenodon	54,106 ± 1,639	23,937 ± 1,681	3.00 X 10^9 bp
Holbrookia[1]	16,733 ± 2,447	43,656 ± 3,573	4.63 X 10^9 bp[4]
Crotalus[1]	68,852 ± 2,476	36,743 ± 2,438	2.59 X 10^9 bp[4]
Cnemidophorus[1]	20,160 ± 2,090	3,456 ± 939	2.49 X 10^9 bp[4]
Macrotus[2]	22,239 ± 1,488	20,646 ± 1,460	2.40 X 10^9 bp[2]
Peromyscus[3]	75,049 ± 1,210	50,832 ± 1,851	3.00 X 10^9 bp[3]

[1] Porter (1994)
[2] Van Den Bussche et al. (1998)
[3] Janecek et al. (1993)
[4] estimated genome size based on closely related taxa (Porter, 1994)

21 kb in *Rattus* (Stallings et al., 1991), and 18 kb in *Mus* (Stallings et al., 1992). Although comparable data currently are not available for diverse reptiles, we can calculate apparent interspersion frequencies from the work of Porter (1994). For the three squamate reptilies, *Cnemidophorus neomexicanus*, *Holbrookia maculata*, and *Crotalus atrox*, the apparent interspersion frequency of $(GT)_n$ is one repeat sequence every 204 kb, 278 kb, and 38.4 kb, respectively, whereas within the *Sphenodon* genome we detected one $(GT)_n$ every 56 kb. This suggests that the distribution of $(GT)_n$ sequences in *Crotalus* and *Sphenodon* are more similar to that seen in mammals than the distribution of this dinucleotide microsatellite in *Cnemidophorus* or *Holbrookia*.

Although fewer genomes have been examined for the dinucleotide repeat $(CT)_n$, patterns similar to $(GT)_n$, albeit slightly reduced, have been documented. For example, the apparent interspersion frequency of $(CT)_n$ is one repeat sequence every 115 kb and 59 kb in the haploid genomes of the mammalian species *M. waterhousii* (Van Den Bussche et al., 1995) and *P. leucopus* (Janecek et al., 1993), respectively. Within reptilian genomes, the apparent interspersion frequency of $(CT)_n$ sequences is one repeat sequence every 1,200 kb in *Cnemidophorus neomexicanus*, 108 kb in *Holbrookia maculata* and 72 kb in *Crotalus atrox*. Based on screening the *Sphenodon* library with the $(CT)_n$ oligonucleotide, we predict an interspersion frequency of one $(CT)_n$ sequence every 105 kb. Therefore, the distribution of $(CT)_n$ sequences in *Holbrookia*, *Crotalus*, and *Sphenodon* are more similar to the distribution seen in mammals than to the distribution seen in *Cnemidophorus*. Comparative results from this small sample of diverse reptiles and mammals are interpreted as indicating that the percentage of repetitive DNA in the genome, as well as copy number and interspersion frequencies of the microsatellites $(GT)_n$ and $(CT)_n$, are highly variable.

EXTENT OF SEQUENCES SHARED BETWEEN GENOMES

The degree to which repetitive sequences are shared among eukaryotic genomes can provide valuable genetic markers for systematic studies (Arnason et al., 1978; Arnason and Widegren, 1989; Hamilton et al., 1990, 1992; Love and Deininger, 1992; Van Den Bussche et al., 1993). Janecek et al. (1993) hybridized a *P. leucopus* library with radioactively labeled genomic DNA from the harvest mouse (*Reithrodontomys*), *Mus*, and human. They found the percentage of the library clones hybridizing with these genomic DNAs showed a trend of decreasing percent hybridization with increasing phylogenetic distance. Moreover, they determined that as the phylogenetic distance increased between genomic DNA in the library and probe DNA, an increasing number of cross hybridizing clones were accounted for by dinucleotide microsatellites.

In addition to multiple representatives of the class Reptilia, we probed the *S. punctatus* cosmid library with

radioactively labeled genomic DNA from representatives of three other vertebrate classes (Table 1). As expected, the greatest hybridization to the *S. punctatus* library was observed with the *S. punctatus* genomic DNA as a probe. Only 8% of the 1,725 clones did not hybridize when radioactively labeled *Sphenodon* genomic DNA was used as a probe.

In general, hybridization to the *S. punctatus* genomic library with genomic DNA from the other seven taxa listed in Table 1 showed a phylogenetic pattern, in which the percentage of recombinant cosmids that hybridized decreased with increasing phylogenetic distance. For example, when genomic DNA from *Crotalus* was used as a probe, 57% of the *S. punctatus* cosmid clones produced some level of hybridization. However, when genomic DNA from turtle, falcon, or human was used as a probe, only 20% of the clones produced detectable levels of hybridization (Table 1). The only probe that does not fit this general trend was genomic DNA from the spotted sea trout (*Cynoscion*), which hybridized to 37% of the *S. punctatus* cosmid clones. This value is higher than that detected with either alligator or iguana genomic DNA (Table 1).

Based on hybridizations of the *P. leucopus* cosmid library with genomic DNA from phylogenetically divergent sources, Janecek et al. (1993) found that as the phylogenetic distance between source of probe DNA and the DNA in the cosmid library increased, the presence of microsatellites accounted for a greater than expected percentage of the sequences that hybridized to a specific probe. In addition, as the level of hybridization of clones decreased for the entire library, the percentage of microsatellite positive clones increased. With the exception of using *Sphenodon* genomic DNA as the source for a probe, there was a statistically significant relationship between the individual clones that hybridized to genomic DNA and the presence of dinucleotide microsatellites in the recombinant cosmids (Table 2). These results are compatible with the observation of Janecek et al. (1993) that with increasing phylogenetic distance between probe and library DNA, dinucleotide microsatellites in the genome account for a disproportionately high number of positive hybridizations due to the ubiquitous presence of microsatellite sequences within vertebrate genomes.

Considerable effort currently is being devoted to understanding the organization of the human genome. Although these efforts will provide valuable information for understanding the organization and distribution of repetitive DNA and genetic diseases in humans, it is only through comparative studies on the organization of diverse genomes that we will be able to understand the organization of the eukaryotic genome and how it evolves. This study provides insight into the frequency, distribution, and organization of repetitive DNA in the reptilian genome. However, to broaden our knowledge of organization of the eukaryotic genome and the role that repetitive DNA may play in genome organization and evolution, we ultimately will need additional comparative data from taxonomically diverse eukaryotes. Studies of this type will ultimately broaden our knowledge of the structural and functional constraints affecting the organization and evolution of the eukaryotic genome.

ACKNOWLEDGMENTS

Cheryl A. Schmidt and Lara Wiggins critically reviewed earlier versions of this manuscript. This work was supported by a National Science Foundation grant to R. J. Baker.

LITERATURE CITED

Arnason, U., I. F. Purdom, and K. W. Jones, 1978. Conservation and chromosomal localization of DNA satellites in balenopterid whales. Chromosoma 66:141-159.

Arnason, U. and B. Widegren, 1989. Composition and chromosomal localization of cetacean highly repetitive DNA with special reference to the blue whale, *Balaenoptera musculus*. Chromosoma 98:323-329.

Arnheim, N. 1979. Characterization of mouse ribosomal gene fragments purified by molecular cloning. Gene 7:83-96.

Bachmann, K. 1972. Genome size in mammals. Chromosoma 37:85-93.

Baker, R. J., J. L. Longmire, M. Maltbie, M. J. Hamilton, and R. A. Van Den Bussche. (1997) DNA synapomorphies for a variety of taxonomic levels from a cosmid library from the New World bat *Macrotus waterhousii.* Syst. Biol. 46: 579-589.

Baker R. J., J. L. Longmire, and R. A. Van Den Bussche, 1995. Organization of repetitive elements in the upland cotton genome (*Gossypium hirsutum*). J. Heredity.86:178-185.

Beckmann, J. S. and J. L. Weber, 1992. Survey of human and rat microsatellites. Genomics 12:627-631.

Bennet, M. D. 1984. The genome, the natural karyotype, and biosystematics. *In* Plant Biosystematics (Vickery, R.K., ed.). Academic Press, Ontario, Canada.

Britten, R. J. and D. E. Kohne, 1968. Repeated sequences in DNA. Science 61:529-540.

Brosius, J. 1991. Retroposons—Seeds of evolution. Science 251:753.

Colbert, E. H. and M. Morales. 1991. Evolution of the Vertebrates. John Wiley & Sons, Inc., New York.

Cooper, K. W. 1964. Meiotic conjunctive elements not involving chiasmata. Proc. Natl. Acad. Sci. USA 52:1248-1255.

Cullis, C. A. and W. Cleary, 1986*a*. DNA variation in flax tissue culture. Can. J. Genet. Cytol. 28:247-251.

_____. 1986*b*. Rapid varying DNA sequences n flax. Can. J. Genet. Cytol. 28:252-259.

Dingerkus. G, 1979. Chordate cytogenetic studies: an analysis of their phylogenetic implications with particular reference to fishes and the living coelacanth. Occas. Papers Calif. Acad. Sci. 134:111-127.

Evans, G. A., K. Lewis, and B. E. Rothenberg, 1989. High efficiency vectors for cosmid microcloning and genomic analysis. Gene. 79:9-20.

Flavell, R. B. 1983. Repeated sequences and genome architecture. *In* Structure and Function in Plant Genomes, (Ciferri, O. and L. Dure, eds.). Plenum Press, New York. 1-14

Flavell, R. B. 1986. Repetitive DNA and chromosome evolution in plants. Phil. Trans. R. Soc. Lond. B, 312:227-242.

Flavell, R. B., M. D. Bennett, J. B. Smith, and D. B. Smith, 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem. Genet. 12:257-269.

Gerbi, S. A. 1985. Evolution of ribosomal DNA. *In* Molecular Evolutionary Genetics, (MacIntyre, R.J., ed.). Plenum Publishing Corp. 419-517.

Hake, S., and V. Walbot, 1980. The genome of *Zea mays*, its organization, and homology to related grasses. Chromosoma 79:251-270.

Hamada, H., M. G. Petrino, and T. Kakunaga, 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA 79:6465-6469.

Hamilton, M. J., R. L. Honeycutt, and R. J. Baker, 1990. Intragenomic movement, sequence amplification and concerted evolution in satellite DNA in *Reithrodontomys*: Evidence from in situ hybridization. Chromosoma 99:321-329.

Hamilton M. J., G. Hong, and H. A. Wichman, 1992. Intragenomic movement and concerted evolution of satellite DNA in *Peromyscus*: Evidence from in situ hybridization. Cytogenet. Cell Genet. 60:40-44.

Heavne, C. M., S. Ghosh, and J. A. Todd, 1992. Microsatelllites for linkage analysis of genetic traits. Trends In Genet. 8:288-294

Janecek, L. L., J. L. Longmire, H. A. Wichman, and R. J. Baker, 1993. Genome organization of repetitive elements in the rodent, *Peromyscus leucopus*. Mammal. Genome. 4:374-381.

Kikuchi, S., F. Takaiwa, and K. Oono, 1987. Variable copy number DNA sequences in rice. Mol. Gen. Genet. 210:373-380.

La Spade, A. R., E. M. Wilson, D. B. Lubahn, A. E. Hardings, and K. H. Fishbeck, 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature, 352:77-79.

Longmire, J. L. 1993. Distribution and organization of repetitive DNA sequences on human chromosome-16. Ph.D. dissertation. Texas Tech University, Lubbock.

Longmire, J. L., R. E. Ambrose, N. C. Brown, T. J. Cade, T. L. Maechtle, W-S. F. Seeger, P. Ward, and C. M. White, 1991. Use of sex-linked minisatellite fragments to investigate genetic differentiation and migration of North American populations of the Peregrine Falcon (*Falco peregrinus*). *In* DNA figerprinting: Approaches and applications, (Burke,T., G. Dolf , G.A. Jefferys , and R. Wolff, eds.). Birkhauser Press, Brasil. 217-229.

Longmire, J. L., N. C. Brown, L. J. Meinke, M. L. Campbell, K. L. Albright, J. J. Fawcett, E. W. Campbell, R. K. Moyzis, C. E. Hildebrand, G. A. Evans, and L. L. Deaven, 1993. Construction and characterization of partial digest libraries made from flow sorted human chromosome 16. Genetic Analysis, Techniques and Applications. 10:69-76.

Love, J. and P. Deininger, 1992. Characterization and phylogenetic significance of a repetitive DNA sequence from whooping cranes (*Grus americana*). The Auk, 109:73-79.

Macgregor, H. C. and S. K. Sessions, 1986. The biological significance of variation in satellite DNA and heterochromatin in newts of the genus *Triturus*: an evolutionary perspective. Phil. Trans. R. Soc. Lond. B. 312:243-259.

Miklos, G. L. G. 1985. Localized highly repetitive DNA sequences in vertebrate and invertebrate genomes. *In*: Molecular Evolutionary Genetics, (MacIntyre, R.J., ed.). Plenum Press. New York. 241-321.

Moran, C. 1993. Microsatellite repeats in pig (*Sus domestica*) and chicken (*Gallus domesticus*) genomes. J. Heredity. 84:274-280.

Morgante, M. and A. M. Oluveri, 1993. PCR-amplified microsatellites as markers in plant genetics. The Plant J. 3:175-182.

Moyzis, R. K., D. C. Tomey, J. Meyne, J. M. Buckingham, J. R. Wu, C. Burkes, K. M. Sirotkin, and W. B. Goad, 1989. The distribution of interspersed repetitive DNA sequences in the human genome. Genomics 4:237-289.

Olmo, E., T. Capriglione, and G. Odierna, 1989. Genome size evolution in vertebrates: Trends and constraints. Comp. Biochem. Physiol. 92B:447-453.

Peacock, W. J. and G. L. G. Miklos, 1973. Meiotic drive in *Drosophila*: New interpretation of segregation distorter and sex chromosome systems. Adv. Genet. 17:361-409.

Porter, C. A. 1994. Organization and chromosomal location of repetitive DNA sequences in three species of squamate reptiles. Chrom. Res. 2:263-273.

Sambrook, J., E. F. Fritsch, and T. Maniatis, 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, NY.

Spofford, J. 1976. Position effect variegation in *Drosophila*. *In*: The genetics and biology of *Drosophila*, (Novitski, A.M., ed.). Academic Press, New York. 955-1018.

Stallings, R. L., A. F. Ford, D. Nelson, D. C. Torney, C. E. Hildebrand, and R. K. Moyzis, 1991. Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. Genomics 10:807-815.

Stallings, R. S., N. A. Dogget, D. Callen, S. Apostolou, L. Z. Chen, J. K. Nancarrow, S. A. Whitmore, P. Harris, H. Michison, M. Breuning, J. J. Saris, Fickett, M. Cinkoski, D. C. Torney, C. E. Hildebrand, and R. K. Moyzis, 1992. Evaluation of a cosmid contig physical map of human chromosome-16. Genomics 13:1031-1039.

Szarski, H. 1974. Cell size and nuclear DNA content in vertebrates. Internat. Rev. Cytol. 23:459-467.

Tiedge, H., W. Chen, and J. Brosius, 1993. Primary structure, neural-specific expression, and dendritic location of human BC200 RNA. J. Neurosci. 13:2382.

Tiedge, H., U. C. Drager, and J. Brosius, 1992. Murine BC1 RNA in dendritic fields of the retinal inner plexiform layer. Neuroscience Letters. 141:136

Van Den Bussche, R. A., J. L. Longmire, and R. J. Baker, 1995. How bats achieve a small C-value: Frquency of repetitive DNA in *Macrotus*. Mamm. Genome. 6:521-525.

Van Den Bussche, R. A., R. J. Baker, H. A. Wichman, and M. J. Hamilton, 1993. Molecular phylogenetic of Stenodermatini bat genera: Congruence of data from nuclear and mitochondrial DNA. Mol. Biol. Evol. 10:944-959.

Van Den Bussche, R. A., J. L. Hudgeons, and R. J. Baker, In Press. Phylogenetic accuracy, stability, and congruence: relationships within and among the New World bat genera *Artibeus, Dermanura,* and *Koopmania. In* Bats: Phylogeny, Morphology, Echolocation, and Conservation Biology (Kunz, T. H. and P. A. Racey, eds.). Smithsonian Institution Press.

Vig, B. K. 1982. Sequence and centromere separation: role of centromeric heterochromatin. Genetics. 107:795-806.

Wichman, H. A., R. A. Van Den Bussche, M. J. Hamilton, and R. J. Baker, 1992. Transposable elements and the evoluton of genome organization in mammals. Genetica 86:287-293.

Zheng, K. L., S. Castiglione, M.G. Biasini, A. Biroli, C. Morandi, and F. Salsa, 1987. Nuclear DNA amplification in cultured cells of *Oryza sativa* L. Theor. Appl. Genet. 74:65-70.

Addresses of Authors:

RONALD A. VAN DEN BUSSCHE

Department of Zoology
430 LSW, Oklahoma State University
Stillwater, OK 74078
e-mail: ravdb@okway.okstate.edu

LAURA L. JANECEK

Savannah River Ecology Lab
SRS Building 737 A
Aiken, SC 29802
e-mail: janecek@srel.edu

JONATHAN L. LONGMIRE

HRL-1, LS-2, Mail Stop M880
Los Alamos National Laboratory
Los Alamos, NM 87545
e-mail: longmire@telomere.lanl.gov

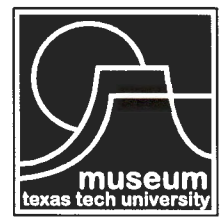
ROBERT J. BAKER

Department of Biological Sciences and
Museum of Texas Tech University
Mail Stop 43131
Lubbock, TX 79409-3131
e-mail: RJBaker@ttu.edu

PUBLICATIONS OF THE MUSEUM OF TEXAS TECH UNIVERSITY

It was through the efforts of Horn Professor J Knox Jones, as director of Academic Publications, that Texas Tech University initiated several publications series including the Occasional Papers of the Museum. This and future editions in the series are a memorial to his dedication to excellence in academic publications. Professor Jones enjoyed editing scientific publications and served the scientific community as an editor for the Journal of Mammalogy, Evolution, The Texas Journal of Science, Occasional Papers of the Museum, and Special Publications of the Museum. It is with special fondness that we remember Dr. J Knox Jones.

Institutional subscriptions are available through the Museum of Texas Tech University, attn: NSRL Publications Secretary, Box 43191, Lubbock, TX 79409-3191. Individuals may also purchase separate numbers of the Occasional Papers directly from the Museum of Texas Tech University.



ISSN 0149-175X

Museum of Texas Tech University, Lubbock, TX 79409-3191