Revista Chilena de Historia Natural
68: 171-183, 1995

Origin and evolution of a major satellite DNA from South American rodents of the genus *Ctenomys* +

Origen y evolución del principal ADN satélite de los roedores sudamericanos del género *Ctenomys*+

MARIA SUSANA ROSSI* [1], C. GUSTAVO PESCE[2], ALBERTO R. KORNBLIHTT[2] and JORGE ZORZOPULOS[3].

GIBE, Departamento de Biología, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (FCEyN, UBA), Ciudad Universitaria, Pabellón 2, 4to. piso, 1428 Buenos Aires, Argentina [1]; INGEBI-FCEyN, UBA [2] and Departamento de Química Biológica FCEyN, UBA [3].
(*) Present address: Instituto de Investigaciones en Ingeniería Genética y Biología Molecular (INGEBI-CONICET). Vuelta de Obligado 2490 2do. piso, 1428 Buenos Aires, Argentina.

+The nucleotide sequence data included in this paper is registrated in the EMBL Gene Bank and DDBJ Nucleotide Sequences under the accession number X69517 (C. porteousi SAT-DNA sequence).

ABSTRACT

This article is about the origin, amplification and evolution of the major satellite DNA of South American rodents known popularly as tuco-tucos (genus *Ctenomys*). This satellite DNA, named RPCS, differentially amplified during the evolution of *Ctenomys* species, which resulted in quantitative and qualitative differences in restriction-periodicity patterns among species. A rolling circle amplification model has been proposed to explain the qualitative and quantitative differences of RPCS sequences. RPCS-related sequences were also found in octodontines, but not in echimyids; that is, they arose in the common ancestor of ctenomyines and octodontines. In situ hybridization experiments showed that RPCS is located in heterochromatic areas of the chromosomes. Finally, RPCS shows unusual features among satellite DNAs: it has a conspicuous retroviral origin. The RPCS monomer has identity with the U3 region of a retroviral LTR, including promoters and enhancers. In addition, DNaseI protection assays showed that this sequences specifically bind transcription factors present in the nucleus of *Ctenomys* cells. The putative function that a retroviral sequence as RPCS may have played in shaping *Ctenomys* genome during evolution is discussed.

Key words: rodents evolution, satellite DNA, retroviral LTRs.

RESUMEN

Este artículo se refiere al origen, amplificación y evolución del principal DNA satélite de los roedores sudamericanos conocidos popularmente como tuco-tucos (género *Ctenomys*). Este ADN satélite, denominado RPCS, se amplificó diferencialmente durante la evolución de las especies de *Ctenomys*, lo que originó diferencias cuantitativas y cualitativas en las periodicidades de restricción entre especies. Se postuló un modelo de amplificación de las secuencias RPCS por círculo rodante para explicar estas diferencias. Secuencias relacionadas a la RPCS se encontraron en octodontinos pero no en echímidos; es decir, se originaron en el ancestro común de octodontinos y ctenominos. Experimentos de hibridación in situ mostraron que las secuencias RPCS están localizadas en regiones heterocromáticas de los cromosomas. Finalmente, la RPCS presenta una característica sorprendente para un ADN satélite: tiene origen retroviral. El monómero de la RPCS tiene identidad con la región U3 de un LTR retroviral, incluidos promotores y enhancers. Además experimentos de protección a la digestión con DNasaI demostraron que esta secuencia une específicamente factores de transcripción presentes en los núcleos de las células de *Ctenomys*. Se discute el hipotético rol que una secuencia de origen retroviral como la RPCS pudo haber jugado durante la evolución del genoma de *Ctenomys*.

Palabras clave: evolución de roedores, ADN satétite, LTRs retrovirales.

INTRODUCTION

South American caviomorph rodents of the genus *Ctenomys* are known popularly as tuco-tuco, or "ocultos" in reference to their hidden habitat. Tuco-tuco species are the most numerous of all fossorial rodents. They are an example of explosive intrageneric radiation that led to about 55 living species during a short period of 1,8 Myr in the Medium Pleistocene (Reig 1989, Reig et al. 1990). *Ctenomys* species show one of the highest rates of karyotypic variation for mammals, with diploid

(Recibido el 5 de Enero de 1993; aceptado el 14 de Junio de 1994)

numbers ranging from 10 (*Ctenomys steinbachi*) to 70 (*Ctenomys dorbignyi*) chomosomes (Reig & Kiblisky 1969, Reig et al. 1990). This highly karyotypic variation led to the hypothesis that explosive species differentiation may be triggered by chomosomal rearrangements (Reig 1989).

Ctenomys chromosomes are also characterized by great variation among species in the amount and location of heterochromatin (for a review see Reig et al. 1992). Therefore, heterochromatin and its molecular constituents, highly repetitive or satellite DNA, seem to behave as a dynamic component of these genomes throughout evolution.

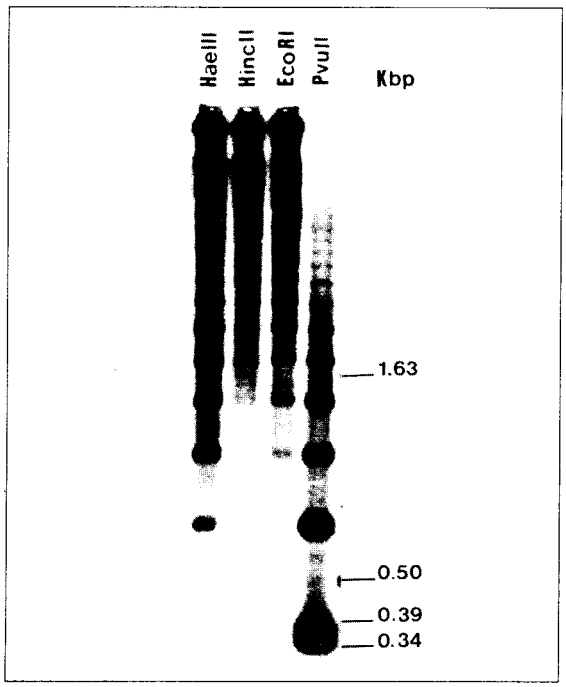

Fig. 1. RPCS sequences are arranged as in a typical satellite DNA. Hybridization patterns of HaeIII, HincII, EcoRI and PuvII-digested DNA of *Ctenomys porteousi* in a Southern-blot experiment. *C. porteousi* DNA was hybridized to labeled RPCS. The hybridization was performed under high stringency conditions. (from Rossi et al. 1990).

Las secuencias RPCS están dispuestas como en un ADN satélite típico. Patrones de hibridación del ADN de *C. porteousi* digerido con las endonucleasas de restricción HaeIII, HincII, EcoRI y PuvII en un experimento de Southern-blot. El ADN de *C. porteousi* se hibridó a la RPCS marçada. La hibridación se desarrolló en condiciones de alta exigencia de identidad. (de Rossi et al.1990).

Taking these considerations into account, as well as the statement by Rose and Doolittle (1983), that "much of the evolutionary history of eukaryotic genome reflects the operation of turnover process involving repetitive DNA sequences", we have studied the evolution of a major satellite DNA in these rodents, named RPCS (repetitive PuvII *Ctenomys s*equence).

The research was mostly focused on the molecular nature of this major satellite DNA, as well as on its biological relevance to *Ctenomys* evolution, particularly its origin and dispersion among its species. Although RPCS shares some characteristics with other satellite DNA sequence (Rossi 1990, Rossi et al. 1990, Rossi et al. 1993b), it shows un unusual feature: it has an important identity with the U3 region of the LTR of the Rous Sarcoma Virus, including promoters and enhancers, which characterized the transcription machinery of retroviruses. Therefore, RPCS most probably arose by integration-deletion of retroviral DNA sequences (Rossi 1990, Rossi et al. 1993b). Such characteristics found in a satellite DNA fires the old debate on the biological significance of these sequences in eukaryotic genomes. We discuss these questions by the end of this review.

Isolation of the monomer of a major satellite DNA of Ctenomys porteousi

Ctenomys porteousi is one of the species in the genus which contains large amounts of heterochromatin according to the C-banding pattern (Massarini et al. 1990). This fact prompted us to search for highly repetitive sequences in this species.

During these studies, we observed that *C. porteousi* genomic DNA digested with the endonuclease PvuII, presented a distinct electrophoretic band about 350 bp and also fainter bands multiples of 350 bp (Rossi 1990, Rossi et al. 1990). This progression of electrophoretic bands suggested that they were arranged in long tandems, as it is usual in satellite DNAs (Southern 1975b).

The smallest PvuII electrophoretic band, named RPCS, was cloned into the SmaI restriction site of the pUC18 plasmid and

used as a probe in subsequent hybridization experiments (Rossi 1990, Rossi et al. 1990). *C. porteousi* genomic DNA was digested with several restriction endonucleases, electrophoresed and transfered onto nylon membranes (Southern-blot technique) (Southern 1975a) and hybridized to labeled RPCS (Fig. 1). Two kind of patterns were observed: one was the PvuII pattern, in which the strongest signal was observed in the 350 bp monomer, decreasing the intensity towards the higher molecular weight multiple bands. This kind of pattern has been previously denominated type A (Hörz & Zachau 1977). Inverted patterns (type B), where the monomer was present in small amounts or not at all, were observed with HaeIII, EcoRI or HincII restriction endonucleases. These results indicated that recognition sequences for PvuII were present in most of the monomers which conform the satellite DNA, while recognition sequences for EcoRI, HaeIII and HincII were present only in a small portion of them.

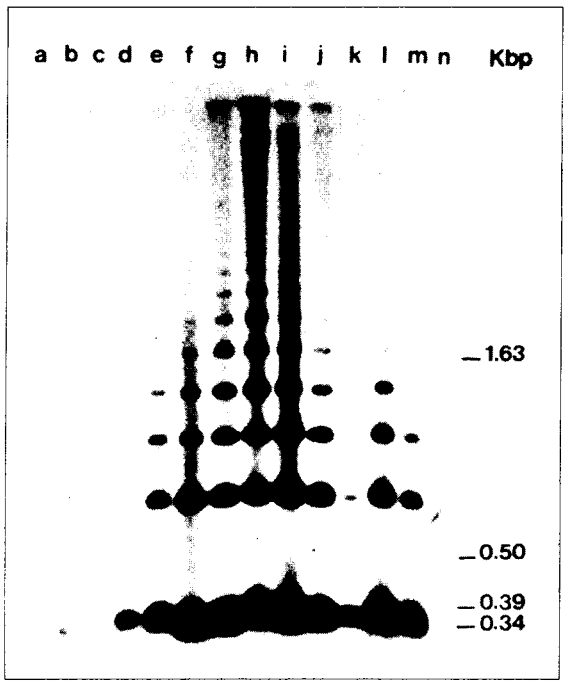

Fig. 2. RPCS is present in several *Ctenomys* species. Southern-blot analysis of RPCS-related sequences in different species of *Ctenomys*, *Mus musculus* and *Calomys musculinus*. All DNAs were cleaved with PuvII. Hybridization to labeled RPCS was performed under high-stringency conditions. Species were as follows: lane a, *M. musculus*; lane b, *C. occultus*; lane c, *C. opimus*; lane d, *C. tuconax*; lane e, *C. talarum*; lane f, *C. latro*; lane g, *C. azarae*; lane h, *C. australis*; lane i, *C. porteousi*; lane j, *C. mendocinus*; lane k, *C. cf. perrensi* (2n=54); lane l, *C. argentinus*; lane m, *C. cf. perrensi* (2n=58); lane n, *Calomys musculinus*. (from Rossi et al.1990).

Las secuencias RPCS están presentes en varias especies de *Ctenomys*. Análisis de Southern-blot de las secuencias relacionadas a la RPCS en diferentes especies de *Ctenomys*, *Mus musculus* y *Calomys musculinus*. Todos los ADNs se digirieron con PvuII. La hibridación con RPSC marcado se desarrolló en condiciones de alta exigencia de identidad. Las especies son las siguientes: carril a, *M. musculus*; carril b, *C. occultus*; carril c, *C. opimus*; carril d, *C. tuconax*; carril e, *C. talarum*; carril f, *C. latro*; carril g, *C. azarae*; carril h, *C. australis*; carril i, *C. porteousi*; carril j, *C. mendocinus*; carril k, *C. cf. perrensi* (2n=54); carril l, *C. argentinus*; carril m, *C. cf. perrensi* (2n=58); carril n, *Calomys musculinus*. (de Rossi et al. 1990).

RPCS and related sequences are present in Ctenomyines and Octodontines

Labeled RPCS was used as probe in hybridization experiments where PuvII-digested genomic DNA from several *Ctenomys* species, *Calomys musculinus* and *Mus musculus* were analyzed (Fig. 2). Hybridization, performed under high-stringency conditions, was detected in every line containing *Ctenomys* DNA, except for the one loaded with *Ctenomys opimus* DNA. RPCS failed to cross-hybridize with DNA from the murid rodent *Mus musculus* and the cricetid rodent *Calomys musculinus* (Fig.2, lines a and n), even if the hybridization conditions were relaxed to allow detection of sequences with 60% identity (not shown).

It has been observed that even in related animal species, satellite DNAs may have important variations in the amount of monomers organized in tandem repeats (e.g. Widegren et al. 1985). This is also the case for *Ctenomys* RPCS satellite DNA. The total amount of RPCS varies among *Ctenomys* species (see also section E). In some species, the longest visible band corresponds to the hexameric or heptameric forms (Fig. 2, lines f and l), while in others, longer polymeric forms are seen (Fig.2, lines h and i). These results suggested that, even though the overall sequence of the RPCS monomer was conserved, amplification reached varying degrees in

different *Ctenomys* species (see also section F). However, for all these species, the observed PuvII-pattern was type A.

As mentioned above, *Ctenomys opimus* was the only species analyzed of the genus whose DNA failed to cross-hybridize with RPCS under high-stringency hybridization conditions. However, under relaxed conditions (60% identity), it was possible to detect a signal in the DNA of this species (Fig.3). Therefore, RPCS-related sequences are present in this species. However the PvuII pattern is an atypical one because monomers are absent, while dimers, trimers, tetramers and pentamers are the only molecular species detected. The observed differences in the sequences and organization of the RPCS satellite betwen *C. opimus* and the remaining species analyzed, suggests that *Ctenomys opimus* may not be very closely related to them.

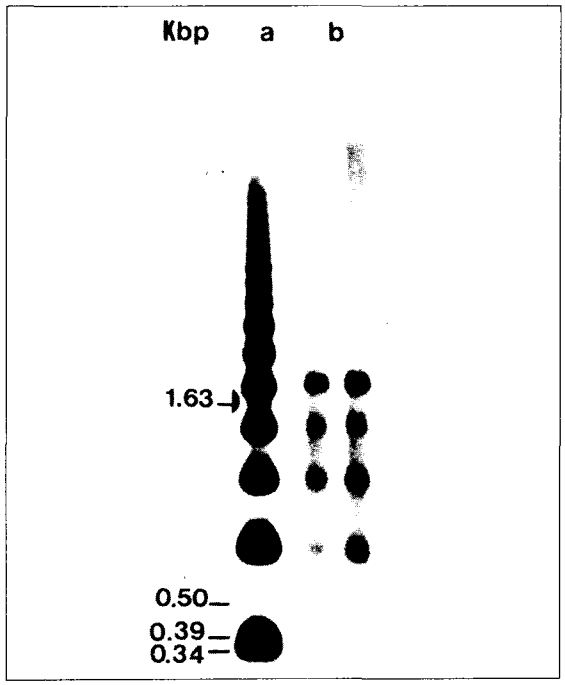

Fig. 3. RPCS-related sequences are present in *Ctenomys opimus*. Comparison between RPCS sequences in *C. porteousi* and *C. opimus* by Southern-blot analysis. DNAs were cleaved with PvuII. Hybridization was performed under low-stringency conditions. Lane a, *C. porteousi*. Lanes b1 and b2, two individuals of *C. opimus*. (from Rossi et al. 1990).

Secuencias relacionadas a la RPCS están presentes en *C. opimus*. Comparación entre las secuencias relacionadas a la RPCS en *C. porteousi* y *C. opimus* en un experimento de Southern-blot. Los ADNs se digirieron con PuvII. La hibiridación fue desarrollada en condiciones de baja exigencia de identidad. Carril a, *C. porteousi*. Carriles b1 y b2, dos individuos de *C. opimus*. (de Rossi et al. 1990).

Ctenomys species are the only living members of the subfamily Ctenomyinae, and are caracterized by advanced adaptation to a herbivorous, subterranean life style. Members of the very closely related subfamily Octodontinae are more generalized, including surface dwellers and burrowers as well. Both subfamilies are included in the Octodontidae family. Reig (1986) suggested that the Octodontidae family is a derived sister taxon of the Echimidae family. Additionally, he proposed that echimyids are the oldest South American caviomorph rodents. With this in mind, we searched for RPCS-related sequences in octodontids as well as in echimyids. Under relaxed hybridization conditions, RPCS cross-hybridized with DNA from two species belonging to the Octodontidae, but failed to cross-hybridize with DNA from a species of the Echimyidae (Fig. 4).

These results seem to indicate that RPCS-related sequences arose after the echimyid divergence, but before the octodontid-ctenomyine divergence, and agree with the close relationship between these two last groups of rodents, postulated by some authors (Ellerman 1940, Pascual et al. 1965, Reig et al. 1990).

RPCS sequences are mainly located in heterochromatic regions of the chromosomes

There appears to be a general correlation between the presence of C-heterochromatic bands and satellite DNA sequences. The tandem arrays of satellite DNA are often located in areas which include heterochromatic centromers and telomers (Pardue and Gall 1970, John and Miklos 1979, Singer 1982, John 1986).

In order to study the chromosomal location of RPCS, in situ hybridization studies using RPCS as probe were carried out in several *Ctenomys* species (Rossi et al. 1995). Three main hybridization patterns were observed: in one of them RPCS sequences were absent in most chromo-

somes or present as weakly signals in the centromeric position. A second pattern showed pericentric RPCS-signals in more than 50% of the chromosomes. The third pattern was characterized by strong hybridization signals in short arms, in about 50% of the chromosomes and centromeric or pericentromeric signals in the others. One of the species with this latter pattern was *Ctenomys porteousi* (Fig. 5). The first pattern was observed in species as *Ctenomys opimus* and *Ctenomys occultus*. In spite of this, conspicuous pericentromeric C-bands were observed in these species (Reig et al. 1992) suggesting that satellite DNA sequences other than RPCS should be present.

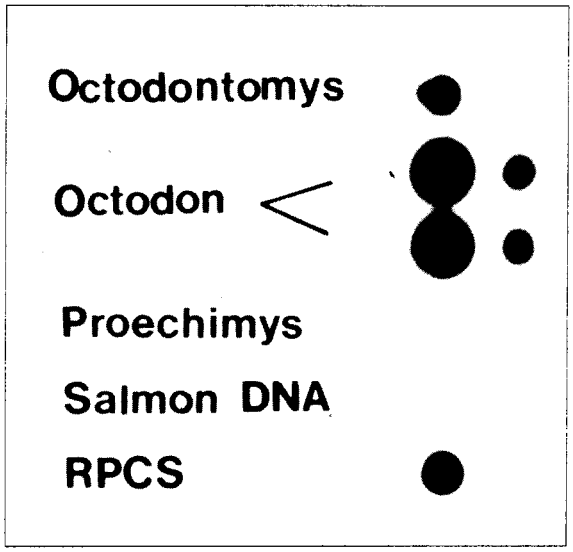

Fig. 4. RPCS-related sequences in octodontids and echimyids. Labeled RPCS was hybridized under relaxed conditions with 1000, 100, and 10 ng DNA of two species belonging to Octodontoidea family (*Octodontomys gliroides* and *Octodon degus*, two individuals), and one species belonging to Echimyidae family (*Proechimys guairae*). Salmon DNA was included as negative control. 1000, 100, and 10 pg RPCS were included as positive control. (from Rossi et al. 1990).

Secuencias relacionadas a la RPCS en octodontinos y equímidos. RPCS marcado se hibirdó en condiciones de baja exigencia de identidad con 1000, 100 y 10 ng de ADN de dos especies pertenecientes a la familia Octodontidae (*Octodontomys gliroides* y *Octodon degus*, dos individuos) y una especie perteneciente a la familia Echimydae (*Proechimys guairae*). DNA de esperma de salmón se incluyó como control negativo. 1000, 100 y 10 pg de RPCS se incluyeron como control positivo. (de Rossi et al. 1990).

With the only exception of *Ctenomys talarum recesus*, the RPCS hybridization patterns of the species of the third group are roughly coincident with the C-banding patterns previously described (Massarini et al. 1990, Reig et al. 1992). In conclusion, most of the RPCS tandem repeats are located in heterochromatic areas of the chromosomes.

Correlation between geographic distribution, quantity and quality of RPCS in Ctenomys *species*

To evaluate the total amount of RPCS sequences in various *Ctenomys* species, genomic DNA was hybridized with labeled RPCS (dot-blot technique). The hybridization was carried out under high-stringency conditions and according to the results, the species were classified in three groups, containing high, low and none RPCS (Rossi 1990, Rossi et al. 1993b). *Ctenomys opimus* was the only species included in the last group. The species with high content of RPCS are: *C. mendocinus*, *C. porteousi*, *C. azarae*, *C. australis*, *C. talarum* and *C. latro*. In the low-RPCS-content group were included *C. cf. perrensi*, *C. tuconax* and *C. occultus*.

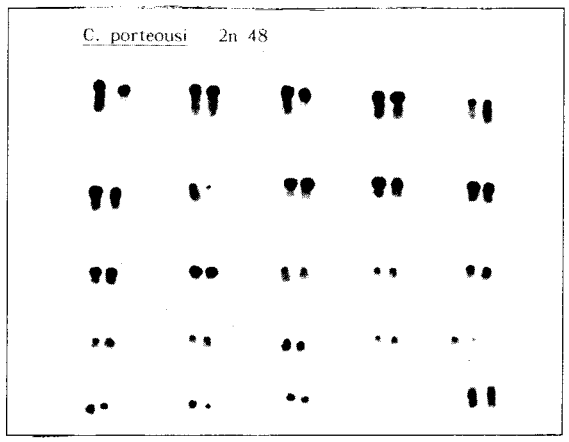

Fig. 5. RPCS are mainly located in heterochromatic regions of the chromosomes. In situ hybridization of labeled RPCS to metaphase chromosomes from bone marrow of *Ctenomys porteousi*. (Rossi et al. 1995)

Las secuencias RPCS están localizadas en regiones heterocromáticas de los cromosomas. Hibridación in situ de RPCS marcado a cromosmas metafásicos de médula ósea de *Ctenomys porteousi*. (Rossi et al. 1995)

Genomic DNA from all these species was also digested with various restriction endonucleases and analyzed by Southern-blot (Rossi 1990, Rossi et al. 1993b) (Fig. 6). Species were grouped according to the resemblance in the periodicities of their restriction patterns. Species that showed identical patterns for all enzymes: *C. mendocinus*, *C. porteousi*, *C. australis*, *C. talarum* and *C. azarae* (Fig.6, lines a to e), were also the ones characterized by their high content of RPCS. Furthermore, these species also share cytological characteristics: for example, although *C. talarum* shows some differences of its own, all of them have the same C and G-banding patterns (Massarini et al. 1990) and the same diploid number (2n=48). All these species, excluding *C. talarum*, have been clustered into them in *mendocinus* group. *C. latro*, however, shows some minor differences in its RPSC restriction periodicities as compared with the remaining species of this group as shown by the Southern-blot analysis. (Fig.6, lane f). It has also a different karyotype (2n=42).

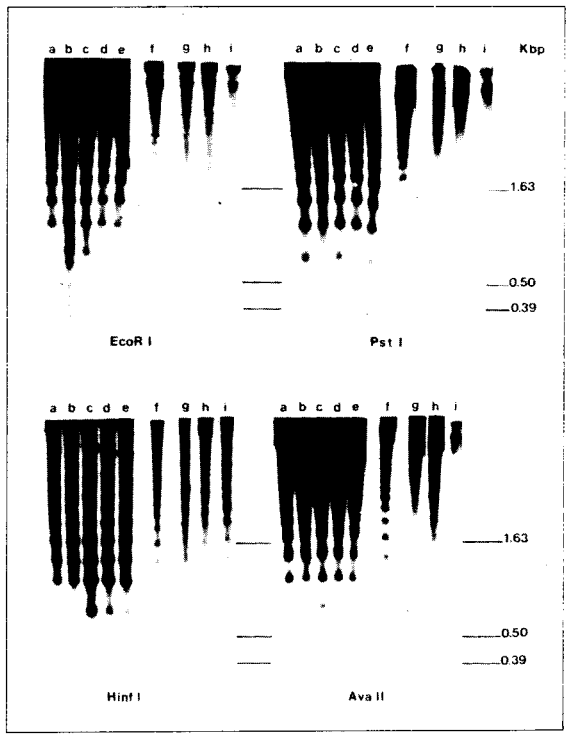

Fig. 6. Patterns of restriction periodicities of various *Ctenomys* species. The DNAs were cleaved with EcoRI, PstI, HinfI and AvaII restriction endonucleases. Hybridization to labeled RPCS was performed under high- stringency conditions in a Southern-blot experiment. The species are as follows: a-*C.mendocinus*, b-*C.porteousi*, c-*C australis*, d-*C.talarum*, e-*C.azarae*, f-*C.latro*, g-*C.cf perrensi* (2n=54), h-*C.tuconax*, i-*C.occultus*. (from Rossi et al. 1993 b).

Patrones de las periodicidades de restricción de varias especies de *Ctenomys*. Los ADNs se digirieron con las endonucleasas de restricción EcoRI, PstI, HinfI y AvaII. RPCS marcado se hibridó en condiciones de alta exigencia de identidad en un experimento de Southern-blot. Las especies son las siguientes: a-*C.mendocinus*, b-*C.porteousi*, c-*C australis*, d-*C.talarum*, e-*C.azarae*, f-*C.latro*, g-*C.cf perrensi* (2n=54), h-*C.tuconax*, i-*C.occultus*. (de Rossi et al. 1993 b).

The group of species characterized by relatively low content of RPCS, *C. cf perrensi*, *C. tuconax* and *C. occultus*, is quite heterogeneous with respect to the Southern-blot analysis (Fig.6, lines g, h and i). These species also display differences in their karyotypes (Ortells et al. 1990, Reig & Kiblisky 1969).

Ctenomys species share an unusual feature among mammals; they have two basic sperm types: symmetric and asymmetric. The symmetric type is very similar to that of other caviomorphs especially the Octodontidae, and may be regarded as the ancestral type (Vitullo et al. 1988). Feito and Gallardo (1982) first described the asymmetric type. They also noticed that species with symmetric sperm type are distributed north to 30° South latitude, while those with asymmetric sperm are distributed south of this latitude (Feito & Gallardo 1982). This association between sperm morphology and geographic distribution was further confirmed (Altuna et al. 1985, Vitullo et al. 1988, Vitullo & Cook 1991).

Figure 7 shows the geographic distribution of the *Ctenomys* species, from which quantity and quality of their RPSC sequences were analyzed. It can be seen that species containing low amounts or none RPCS are distributed north of 30° South latitude. On the other hand, the group of species characterized by high content of RPCS and typical satellite restriction patterns, with the exception of *C. latro*, are distributed South of this latitude. As mentioned before, *C. latro*'s RPCS is very similar in quantity and quality to that of the species with southern distribution, in spite

of its northen location and symmetrical sperm type. This suggests that *C. latro* may be closely related to the stock which colonized the southern area of Argentina.

Amplification of RPCS

The monomer of *Mus musculus* major satellite DNA is 234 base pair (bp) in length and has a hierarchical organization. The 234 bp unit evolved by successive duplication of a basic 9-bp sequence, in addition to accumulation of substitutions, insertions and deletions (Southern 1975 a and b). This

Fig. 7. Geographic distribution of the *Ctenomys* species and their RPCS content. The shady area denotes geographic distribution of the family Ctenomydae according to Vitulo et al. (1988). 0, 0, and 0 denotes high, low and none content of RPCS, respectively. In parenthesis are include the type of sperm and the diploid number. S=symmetrical sperm, A=asymmetrical sperm. (from Rossi et al. 1993 b).

Distribución geográfica de las especies de *Ctenomys* y su contenido en RPCS. El area sombreada indica la distribución geográfica de la familia Ctenomydae de acuerdo a Vitullo et al. (1988). 0, 0 y 0 indica alto, bajo y nulo contenido de RPCS, respectivamente. Entre paréntesis figuran el tipo de espermatozoide y el número diploide. S=espermatozoide simétrico, A=espermatozoide asimétrico. (de Rossi et al. 1993 b).

doesn't seem to be the case for the monomer of *Ctenomys* RPCS, since it has not repetitions of any internal subsequence (Rossi 1990). That is, the original sequence that was amplified was the entire RPCS monomer.

Divergence between units of a satellite DNA have been accounted for the occurrence of restriction-periodicities patterns type A and B. Type A for a gradual mutation of a restriction site present in the original amplified unit, and type B, for restriction sites which arose later during the satellite evolution. In the case of RPCS, we found that, with the exception of the species of the *mendocinus* group, restriction periodicities were very different among species (Fig.6). For instance, in some species, the restriction pattern for a given endonuclease was of type B (see, for instance Fig. 6, line f); while in other species the same endonuclease did not digested RPCS (Fig.6, lines i for EcoRI, PstI and Ava II). These differences cannot be accounted mainly by overall nucleotide divergence among the monomer of RPCS in these species, since the hybridization conditions used in these experiments allow only recognition of sequences having identities of more than 90%. The most plausible explanation is that the proportion of monomers with a given restriction site(s) varies among species. That is, in some species the monomer that was preferentially amplified contained recognition site(s) for some enzime(s), while sites others than that particular one (or even none) were amplified in other species.

These differences in the RPCS restriction periodicities observed in *Ctenomys* species, cannot be explained either by major divergence in the monomer sequence or by random processes, such as unequal crossover (see also Walsh 1987). A model involving local amplification restricted to a few monomers containing different restriction sites, agrees better with the restriction periodicities profiles of RPCS. Local amplification of satellite DNA sequences has already been postulated by Hörz and Zachau (1977). We suggested the following model for RPCS satellite DNA amplificacion (Rossi et al. 1990): in the

precursor of RPCS there were sequences that by single point mutation could become recognition sites for PuvII or other restriction endonucleases. This precursor was present in some oligomeric form in the common ancestor of ctenomyines and octodontines. In this protosequence there was also a subsequence which could potentially act as a replication origin. In relation to this last issue, Rossi (1990) found that in the monomer of RPCS, there were present sequences with 70% identity with the ARS (autonomous replication sequence), regarded as putative chromosome replication origin (Huberman et al. 1988, Diffllley & Stillman 1990).

In the ctenomyines branch appeared a mutation which created a PuvII site in one monomer of the RPCS protosequence (Fig. 8A). The monomer containing this PuvII was amplified by unequal crossing-over or other amplification process in the ancestor of the *Ctenomys* lineage. This amplification continued after the initiation of the explosive cladogenesis in *Ctenomys*, but its rate was different for different lineages (see also E). We postulated that most amplification was the consequence of sporadic formation of circular DNA structures by intrastrand exchange, followed by rolling-circle replication (for a discussion on why the unequal crossing-over process may not be sufficient for subsequent amplifications, see Walsh 1987). Reinsertions after amplifications ocurred mostly at homologous sites and

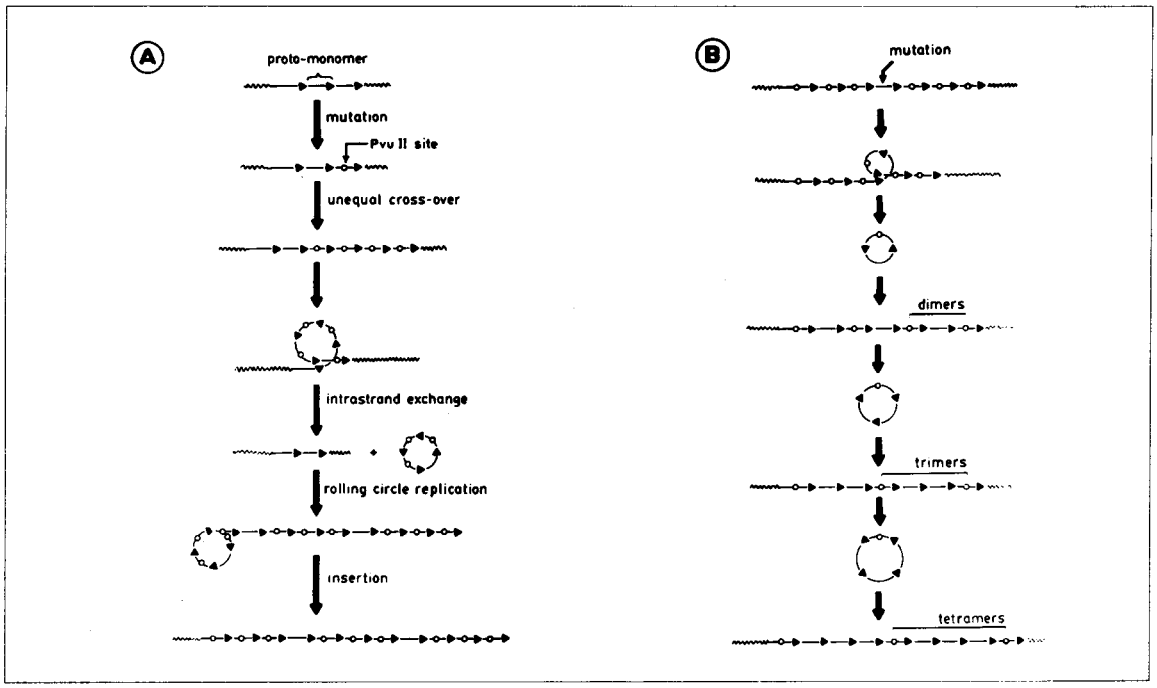

Fig. 8. Amplification model of RPCS by rolling-circle. Scheme explaining the RPCS amplification process, based on the model proposed by Walsh (1987). A, Ancestral sequence with a PvuII site amplified by unequal crossing-over or other amplification mechanisms. This tandem array undergoes an intrastrand exchange resulting in the formation of a circular plasmid carrying monomers with the PvuII site. This plasmid is replicated through a rolling-circle mechanism which therefore amplifies a particular section of the original tandem. Finally, the amplified sequence is reinserted into the chromosomal DNA into the original tandem. B, Periodicities generated by local amplification of monomers containing mutations which affect restriction patterns. (from Rossi et al. 1990).

Modelo de amplificación de la RPCS por círculo rodante. Esquema explicativo del proceso de amplificación basado en un modelo propuesto por Walsh (1987). A, Una secuencia ancestral conteniendo un sitio PuvII se amplifica por intercambio desigual u otros mecanismos de amplificación. Esta forma oligomérica sufre una recombinación intracatenaria, resultando en la formación de un plásmido circular que lleva monómeros con sitio PuvII. Este plásmido se replica mediante un mecanismo de círculo rodante que de esta forma amplifica una porción particular del tandem original. Finalmente la secuencia amplificada se reinserta en el ADN cromosomal en el tandem original. B, Periodicidades generadas por la amplificación local que afectan los patrones de restricción. (de Rossi et al. 1990).

occasionally at nonhomologous sites which became hot spots for future reinsertions. The ocurrence of dimers, trimers and multimers of the monomer was the result of mutations that supressed the PvuII sites of some monomers, followed by amplification (Fig.8B). Amplification of monomers containing mutations that created other restriction sites (i.e. EcoRI or PstI sites) was also the result of this same local amplification process. The rolling-circle amplification model for satellite DNA has been previously proposed by Walsh (1987), and we think that is the one that best fits our results.

RPCs most probably arose by integration-deletion of retroviral DNA sequences

The nucleotide sequence of the 348-bp RPCS monomer revealed extensive homology with the U3-region of LTR (long terminal repeat) from Rous sarcoma virus (Rossi 1990). Retroviruses have transcriptional and replicational regulatory elements within the U3 region of their LTRs (Fields & Knipe 1990). These consensus sequences are also present in the RPCS monomer (Fig. 9) (Rossi 1990, Rossi et al. 1993a). The polypurine (pp) region is involved in the initiation of the DNA plus-strand synthesis (Varmus 1982). The sequences C/EBP (CCAAT/enhancer binding protein), enhancer core, CCAAT and NF1 (nuclear factor 1), boxed in Figure 9, are well characterized DNA consensus sequences found in many eukaryotic and/or viral-transcriptional enhancers. All of them bind interacting factors (Gluzman & Sherik 1983, Ryden & Beemon 1989).

When RPCS monomer was incubated with nuclear factors of *Ctenomys porteousi*, and the protected sites analyzed by digestion with DNAsaI (footprinting), it was observed three protected regions (Fig. 10). RPCS I and III overlap almost completely with the E1 and E2 from Figure 9, respectively. RPCS II contains in the complementary strand the consensus of the binding transcription factor C/EBP (Pesce et al.

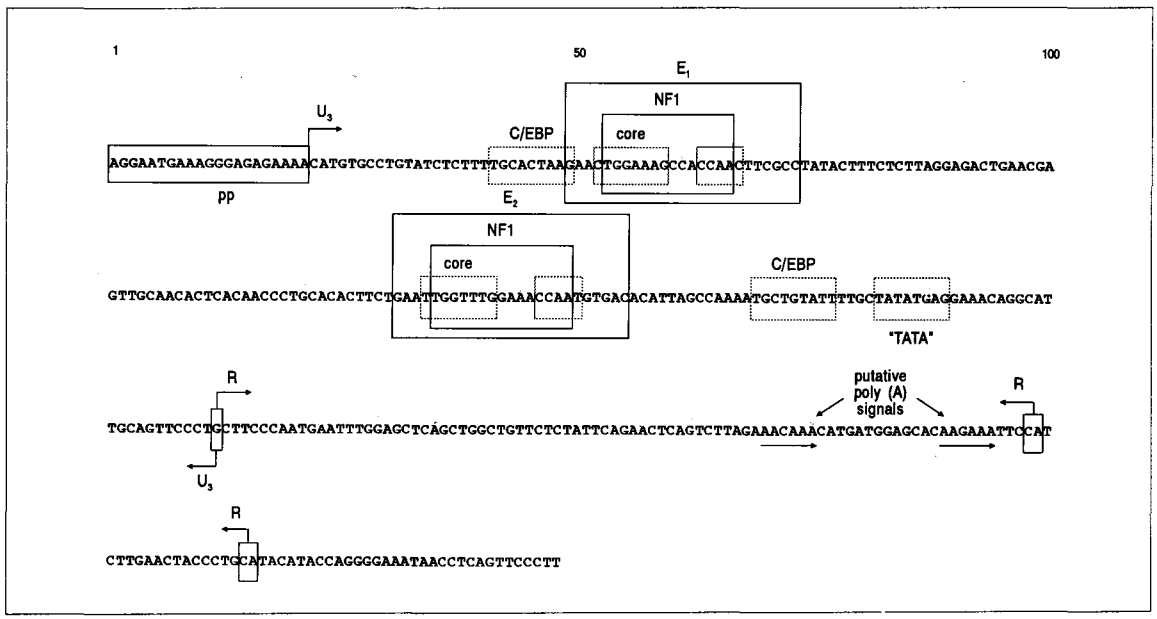

Fig. 9. RPCS has a retroviral origin. Binding sites for various transcription factors as well as a polypurine region (pp) and a TATA-box, which characterized the LTRs (long terminal repeats) of the retroviruses, are indicated. E1 and E2 denotes enhancers regions. In order to highlight the ressemblance between the monomer of the RPCS with an LTR, the nucleotide sequence is shown here in a different order from that of Figure 10: nucleotide 1 corresponds to nucleotide 113 of Figure 10. (from Rossi et al. 1993).

La SRPC tiene un origen retroviral. Se indican los sitios de unión de varios factores de transcripción así como la región de polipurinas (pp) y el TATA-box que caracterizan los LTRs (long terminal repeats) de los retrovirus. E1 y E2 indican regiones enhancer. Para resaltar la similitud entre la RPCS y un LTR, la secuencia nucleotídica se muestra aquí en orden distinto al que se muestra en la Figura 10: el nucleótido 1 corresponde al 113 de la Figura 10. (de Rossi et al. 1993).

1994). That is, the enhancers sequences present in the RPCS monomer specifically bind transcription factors.

A TATA-box, the target site for RNA polymerase, is present in the U3 region about 25 bp upstream from the start of transcription (Varmus 1982). In RPCS, a TATA-box is also present, and according to its position, a putative transcription initiation region has been boxed (Fig.9). This transcription iniciation point is by definition the 5' limit of the R region in retroviral LTRs.

The signal of polyadenilation (AATAAA) is located downstream from the transcription start of retroviruses (Varmus 1982). For several eukaryotic and retroviral genes, the dinucleotide CA is located at the polyA addition site, about 20 bp of the polyadenilation signal (Weiss et al. 1985). According to these locations, the AACAAA and the AAGAAA hexanucleotides and the CA dinucleotides present downstrean of them in RPCS (Fig. 9), resemble retroviral polyadenilation motifs.

All these facts strongly suggest that the RPCS monomer was acquired from the U3-R region of a retroviral genome. This is the first reported case of a satellite DNA with retroviral origin and also this is the first case that has been demonstrated that a satellite DNA specifically binds transcription factors.

DISCUSSION

The original events of the RPCS *Ctenomys* satellite DNA evolution involved the insertion of a retroviral genome in the common ancestor of ctenomyines and octodontines, that diverged more than 10 Myr ago (Reig et al. 1990). These ancestral sequences suffered a moderate amplification in the octodontine's branch. These sequences were probably inherited by the *Ctenomys* branch. In this branch, RPCS entered in two different evolutionary pathways: 1) In *Ctenomys opimus* the amplification did not continued, and perhaps part of the tandem repeats was lost by deletion. 2) In most (if not all) of the other *Ctenomys* species a great amplification occurred. Henceforth, this great amplification started most probably in the common ancestor of these *Ctenomys* species, after the divergence of the *Ctenomys opimus* branch. The great amplification process of RPCS occurred after the divergence of *C. opimus* from the remaining species.

The great amplification in the *Ctenomys* lineage involved mainly monomers carring PuvII restriction sites, resulting in the type

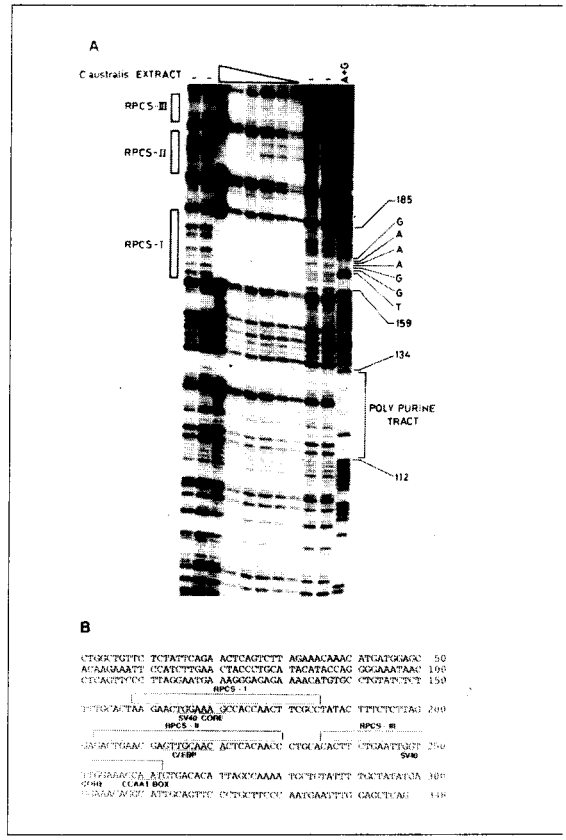

Fig. 10. Transcription factors binds the monomer of RPCS. Nuclear protein extracts from *Ctenomys porteousi* were incubated with labeled RPCS with ^{32}P. Then the DNA were digested with the endonuclese DNaseI and electrophoresed. The protected regions (RPCS-I, II and III) are visualized as "windows" in the gel. Nucleotide 113 corresponds to nucleotide 1 from Figure 9. (Pesce et al. 1994).

Factores de transcripción se unen al monómero de la RPCS. Un extracto de proteínas nucleares de *Ctenomys porteousi* se incubó con RPCS marcado con 32P. Luego el ADN se digirió con la endonucleasa DNasaI y se sometió a electroforesis. Las regiones protegidas (RPCS-I, II y III) se ven como "ventanas" en el gel. El nucleótido 113 corresponde al nucleótido 1 de la Figura 9. (Pesce et al. 1994).

A pattern, that characterizes *Ctenomys* species with the exception of *Ctenomys opimus* (Fig.2). Although the amplification reached varying degrees in different groups of species, the overall nucleotide sequence of the RPCS monomer was conserved during this process. However, when restriction endonucleases others than PuvII were employed, dramatical differences in the RPCS periodicities were found among all species, except for the species of the *mendocinus* group. The process that resulted in the differences in the RPCS periodicities among *Ctenomys* species could be interpreted through a local amplification model involving rolling-circle replication (Fig.8) (Rossi et al. 1990).

The possible functions of satellite DNAs are still obscure. Different components of these sequences evolved at different rates (i.e. Arnason & Widegren 1984). For example, a case of conspicuous conservation is the major satellite DNA of cetacean families, whose 1,740-bp repeat has been maintained in lineages that separated more than 40 Myr ago (Widegren et al. 1985). Although sequences-homogenization mechanisms as "molecular drive" (Dover 1982, Dod et al. 1989) or selfish DNA hypothesis (Orgel & Crick 1980) cannot be ruled out, it is likely that major components of satellite DNAs affect the spatial organization of chromosomes in meiosis and mitosis, and therefore evolve under certain selective pressures.

Ctenomys RPCS satellite has a relatively recent origin, and it is difficult to assure the putative mechanisms involved in its conservation. RPCS tandem repeats behave like a typical satellite DNA in several characteristics. However, RPCS is the only case of a satellite DNA sequence with a conspicuous retroviral origin (Rossi et al.1993a), and though binding of non-histone proteins to a *Drosophila* satellite DNA has been reported (Doshi et al. 1991), RPCS is the only case of a highly repetitive DNA which has been demostrated to bind transcription factors. That is, this satellite DNA has the potential capacity for regulation of the transcriptional activity.

As shown by in situ hybridization experiments, most of the copies of RPCS are clustered in heterochromatic regions of the chromosomes (Fig.5) (Rossi et al.1995), and therefore they might not be normally accessible for transcription factors. This fact does not rule out the possiblility that isolated copies of RPCS dispersed in different locations of the genome could have regulatory transcriptional activity. Furthermore, it cannot rule out the possibility that in particular developmental stages or tissues, transcription factor could bind to the RPCS, modifying the pattern of gene activity or gene commitment (Villarreal 1991).

A retroviral sequence as RPCS, with intrinsic capacity for amplification and relocation in the genome, might influence the expression of neighboring genes, as is the case for endogenous proviruses and related intracisternal A-type particles (Kongsuwan et al. 1989, Leslie et al. 1991). The endogenous retroviral elements have been detected in the genomic DNA of many vertebrate species (see La Mantia et al. 1992 and references therein). Endogenous elements most likely entered the germ line of their hosts through retroviral infection of germ lines or germ cells progenitors. Since then, they have persisted as stable integrated, vertically transmitted proviruses, but they have largely diverged from each other as well as from their infectious progenitors. Some endogenous retroviral sequences do not contain a full complement of retroviral sequences and in the majority of the cases, these elements are defective in their capacity to encode functional proteins (La Mantia et al. 1992). However, many of them have been shown to be transcriptionally active in tissues and cell lines (Ting et al. 1992).

LTR sequences are considered to be the primary determinants of the regulation of endogenous retroviral expression. Morever, the nature of the LTR transcriptional regulatory elements influences to a great extent the ability of retroviral elements to affect expression of nearby cellular genes. In mouse, insertions of retroviral elements near or within normally regulated genes resulted in their constitutive expression (La Mantia et al. 1992, and references therein). Conversely, the androgen responsiveness

of the mouse Sex-limited protein gene seems to be conferred by a provirus-like LTR inserted 2 kbp upstream of the gene (Stavenhagen et al. 1988). It is worth noting that in this latter case the phenotype is imposed by a highly defective element that has maintained functional LTR sequences. It have recently been described a low repeated family of human endogenous retroviral elements (ERV9), whose expression is mainly detectable in undifferentiated embryonal carcinoma cells (La Mantia et al. 1992). This endogenous retroviral sequence was demonstrated to be capable to drive the expression of a reporter gene. Insertions of a retroviral element has also been correlated with a switch from pancreatic to parotid expression of the human amylase gene (Ting et al. 1992).

It seems therefore that endogenous retroviral sequences can represent an important reservoir of potentially functional promoter sequences that could affect the pattern of expression of a number of cellular genes and could provide a significant mechanism for rapid acquisition of diversity in evolution. In a group as the ctenomyines, in which cladogenesis was explosive, changes in gene regulation may be even more important to the process of speciation than gradual accumulation of structural variation.

Further research would determine whether the regulation of particular genes or block of genes has been taking over during the speciation process of *Ctenomys* by the RPCS and if this has resulted in phenotypic differences among species.

ACKNOWLEDGMENTS

This paper is dedicated to the dear memory of Professor Osvaldo A. Reig, who first percived the relevance of a unique evolutionary model as *Ctenomys*. This work has been supported by a grant from the University of Buenos Aires (EX 144) to M.S.R. and A.R.K.

LITERATURE CITED

ALTUNA CA, A NOVELLO & EP LESSA (1985) Notas sobre la morfología espermática de *Ctenomys* rionegrensis (Rodentia, Octodontidae) del Uruguay. Brenesia 24: 397-401.

ARNASON U & B WIDEGREN (1984) Different rates of divergence in highly repetitive DNA of cetaceans. Hereditas 101: 171-177.

DIFFLEY JFX & B STILLMAN (1990) The iniciation of chromosme DNA replication in eukaryotes. Trends in Genetics 12: 427-432.

DOSHI P, S KAUSHAL, C BENYAJATI & CHI WU (1991) Molecular analysis of the responder satellite DNA in *Drosophila melanogaster*: DNA bending, nucleosome structure and Rsp-binding proteins. Molecular Biology and Evolution 8: 721-741.

DOD B, E MOTTEZ, E DESMARAIS; F BONHOME & G ROIZES (1989) Concerted evolution of light satellite DNA in genus *Mus* implies amplification and homogenization of large blocks repeats. Molecular Biology and Evolution 6: 478 491.

DOVER GA (1982) Molecular drive: a cohesive model of species evolution. Nature 299: 111-117.

ELLERMAN KR (1940) Rodents other than Muridae. In: The families and genera of living rodents:Vol.1. British Museum (Natural History) London.

FIELDS F & D KNIPE (1990) (eds) Virology: 463-467.

FEITO R & M GALLARDO (1982) Sperm morphology of the Chilean species of *Ctenomys* (Octodontidae). Journal of Mammalogy 63: 658-661.

GLUZMAN Y & T SHENK (eds) (1983) Enhancers and eukaryotic gene expression. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

HÖRZ W & HG ZACHAU (1977) Characterization of distinct segments in mouse satellite DNA by restriction endonucleases. European Journal of Biochemistry 73: 383-392.

HUBERMAN JA, J ZHU, LR DAVIS & C NEWLON (1988) Close association of a DNA replication origin and ARS element on chromosome III of the yeast *Saccharomyses cerevisiae*. Nucleid Acid Research 16: 6373-6387.

JOHN B & GLG MIKLOS (1979) Functional aspects of satellite DNA and heterochromatin. International Reviews of Cytology 58: 1-114.

JOHN B (1986) The biology of heterochromatin. In: Verma RS (ed):101. Heterochromatin. Cambridge, University Press.

KONGSUWAN K, J ALLEN & JM ADAMS (1989) Expression of Hox-2.4 homeobox gene directed by proviral insertion in a myeloid leukemia. Nucleic Acids Research 17: 1881-1892.

LA MANTIA G, B MAJELLO, A DI CRISTOFANO, M STRAZZULLO, G MINICHIOTTI & L LANIA (1992) Identification of regulatory elements within the minimal promoter region of the human endogenous ERV9 proviruses: accurate transcription initiation is controlled by an Inr-like element. Nucleic Acids Research. 16: 4129-4136.

LESLIE KB, F LEE & JW SCHRADER (1991) Intracisternal A-type particle-mediated activations of cytokine genes in a murine myelomonocytyc leukemia: generation of functional cytokine mRNAs by retroviral splicing events. Molecular Cell Biology 11: 5562-5570.

MASSARINI AI, MA BARROS, MO ORTELLS & OA REIG (1990) Chromosomal polymorphism and small karyotypic differenciation in a group of *Ctenomys* species from Central Argentina (Rodentia, Octodontidae). Genetica 83: 131-144.

ORGEL LE & FHC CRICK (1984) Selfish DNA: the ultimate parasite. Nature 284: 604-607.

ORTELLS MO, J CONTRERAS & OA REIG (1990). New *Ctenomys* karyotypes (Rodentia, Octodontidae) from north-eastern Argenina and from Paraguay confirm the chromosomal multiformity of the genus. Genetica 82: 189-201.

PARDUE ML & JG GALL (1970) Chromosomal location of mouse satellite DNA. Science 168: 1356-1358.

PASCUAL R, J PISANO & EJ ORTEGA (1965) Un nuevo octodontinae (Rodentia, Caviomorpha) de la formación Epecuén (Plioceno Medio) de Hidalgo (Provincia de La Pampa). Ameghiniana 4: 19-30.

PESCE CG, MS ROSSI, AF MURO, OA REIG, J ZORZOPULOS & AR KORNBLIHTT (1994) Binding of nuclear factors to a satellite DNA of retroviral origin with marked differences in copy number among species of the rodent *Ctenomys*. Nucleic Acid Reasearch 22: 656-661.

REIG OA & P KIBLISKY (1969) Chromosome multiformity in the genus *Ctenomys* (Rodentia, Octodontidae). Chromosoma. 28: 211-244.

REIG OA (1989) Chromosomal repatterning as a causative factor in cases of explosive speciation. In: Fontdevila A (ed): 246-289. Evolutionary biology of transient unstable populations. Springer, Berlin.

REIG OA, C BUSH, MO ORTELLS & J CONTRERAS (1990) An overview of evolution, systematics, population biology and speciation in *Ctenomys*. In: Nevo E & OA Reig (eds) Evolutionary Biology of Subterranean Mammals: 71-96 Alan R. Liss, New York.

REIG OA, AI MASSARINI, MO ORTELLS, MA BARROS, S TIRANTI & F DYZENCHAUZ (1992) New karyotypes and C-banding patterns of the subterranean rodents of the genus *Ctenomys* (Caviomporpha, Octodontidae) from Argentina.Mammalia. 56(4): 603-623.

ROSE MR & WF DOOLITTLE (1983) Molecular biological mechanism of speciation. Science 220:187-279.

ROSSI MS (1990) Organización, amplificación y evolución de un ADN satélite del género *Ctenomys* (Octodontidae, Rodentia) Tesis doctoral. Universidad de Buenos Aires. 102 pp.

ROSSI MS, OA REIG & J ZORZOPULOS (1990) Evidence of rolling-circle replication in a major satellite DNA from the South American rodents of the genus *Ctenomys*. Molecular Biology and Evolution 7: 340-350.

ROSSI MS, CG PESCE, OA REIG, AR KORNBLIHTT & J ZORZOPULOS (1993a) Retroviral like-features of the major satellite DNA of the South American rodents of the genus *Ctenomys*. DNA Sequences. 3: 379-381.

ROSSI MS, OA REIG & J ZORZOPULOS (1993b) A major satellite DNA from the South American rodents of the genus *Ctenomys*: Quantitative and qualitative differences in species with different geographic distribution. Zeitschrift für Säugetierkunde, International Journal of Mammalian Biology 58: 244-251.

ROSSI MS, CA REDI, G VIALE, AI MASSARINI & E CAPANNA (1995) Chromosomal distribution of the major satellite DNA of South American rodents of the genus *Ctenomys*. Cytogenetics and Cell Genetics. 69: 179-184.

RYDEN TA & K BEEMON (1989) Avian retroviral long terminal repeats bind CCAAT enhancer binding protein. Molecular Cell Biology. 9: 1155-1164.

SINGER MF (1982) Highly repeated sequences in mammalian genomes. International Review of Cytology 76: 67-112.

SOUTHERN, EM (1975a) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology 98: 503-517.

SOUTHERN, EM (1975b) Long range periodicity in mouse satellite DNA. Journal of Molecular Biology 94: 61-69.

STAVENHAGEN JB & DM ROBINS (1988) An ancient provirus has imposed androgen regulation on the adjacent mouse sex-limited protein gene. Cell 55: 247-254.

TING CH, MP ROSEMBERG, CM SNOW LC SAMUELSON & MH MEISLER (1992) Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. Genes & Development 6: 1457-1465.

VARMUS HE (1982) Form and function of retroviral proviruses. Science 216: 812-820.

VILLARREAL L (1991) Relationships of eukaryotic DNA replication to committed gene expression: general theory for gene control. Microbiological Reviews 3: 512-542.

VITULLO AD, ERS ROLDAN & MS MERANI (1988) On the morphology of spermatozoa of tuco-tucos, *Ctenomys* (Rodentia: Ctenomyidae): New data and its implications for the evolution of the genus. Journal of Zoology London 215: 675-683.

VITULLO AD & JA COOK (1991) The role of sperm morphology of spermatozoa of tuco-tucos, *Ctenomys* (Rodentia: Ctenomyidae): New data and its implications for the evolution of the genus. Zeitschrift für Saügetierkunde, International Journal of Mammalian Biology. 56: 359-364.

WEISS R, N TEICH, H VARMUS & J COFFIN (eds) (1985). Molecular biology of Tumor Viruses: RNA tumor viruses:part 2. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory.

WIDEGREN B, U ARNASON & G AKUSJARVI (1985) Characteristics of a conserved 1579-bp highly repetitive component in killer whale *Orcinus orca*. Molecular Biology and Evolution 2: 411-419.

WALSH, JB (1987) Persistence of tandem arrays: implication for satellite and simple-sequence DNA. Genetics 115: 553-567