Origin and spread of the SRY gene on the X and Y chromosomes of the rodent Microtus cabrerae: Role of L1 elements

Juan A. Marchal a,1, Manuel J. Acosta a,1, Mónica Bullejos a, Rafael Díaz de la Guardia b, Antonio Sánchez a,*

a Departamento de Biología Experimental, Facultad de Ciencias Experimentales, Universidad de Jaén, Campus Las Lagunillas s/n, E-23071 Jaén, Spain
b Departamento de Genética, Facultad de Ciencias, Universidad de Granada, E-18071 Granada, Spain

Received 15 June 2007; accepted 19 October 2007
Available online 4 December 2007

Abstract

In the rodent species Microtus cabrerae, males as well as females present several copies of the SRY gene, a single-copy gene located on the Y chromosome in most mammals. Using different PCR approaches, we have characterized the sequence, structure, and organization of the SRY copies and their flanking regions distributed on the X and Y chromosomes of this species. All copies of SRY analyzed, including those from the Y chromosome, proved to be nonfunctional pseudogenes, as they have internal stop codons. In addition, we demonstrated the association of SRY pseudogenes with different fragments of L1 and LTR retroelements in both sex chromosomes of M. cabrerae. Examining the possible origin of SRY pseudogene and retroposons association, we propose that retroposons could have been involved in the mechanism of SRY gene amplification on the Y chromosome and in the transference of the Y-linked SRY copies to the X-chromosome heterochromatin.

Keywords: Microtus cabrerae; SRY gene; Sex chromosomes; Heterochromatin; Retroposon

The SRY gene (sex-determining region Y) is responsible for sex determination in mammals [1,2]. In the presence of this gene, undifferentiated embryonic gonads develop as testes, while in its absence or malfunction, the gonads follow the ovarian pathway [3]. Accordingly, the SRY transcript is detected in the genital ridge during the critical period of gonad differentiation [4,5].

SRY is a single-exon gene with an ORF that encodes a transcription factor with an HMG domain of about 79 amino acids and N- and C-terminal regions of variable length [6,7]. While the HMG-box sequence is highly similar among species, 5’ and 3’ ends are scarcely conserved, and it is very difficult to find tracks of the interspecific sequence alignment [8,9].

SRY is a single-copy gene located in the nonrecombining region of the Y chromosome in most mammals [2,10]. However, rodent species with no Y chromosome, and hence without the SRY gene, have been reported (two species of the genus Ellobius [11] and one from the genus Tokudaia [12]). On the other hand, multiple copies of SRY are present in several rodent species, such as murids [13,14], species of the genus Akodon [15], and several species of Arvicolinae [16,17]. The characteristic features of the SRY gene in the vole Microtus cabrerae are very striking and so far have not been described in another mammal. Our previous work, using Southern blot, PCR, and sequence analyses of the SRY HMG box, demonstrated the presence of multiple (polymorphic) copies of this gene in males but, unexpectedly, also in all normal females of the species [16]. Other related species analyzed, belonging to the genera Microtus, Arvicola, and Pitymys, have multiple (polymorphic or monomorphic) copies of the SRY gene, but they are restricted to the male genome and hence located only on the Y chromosome [17].

In M. cabrerae, as well as in other species of the genus Microtus, the sex chromosomes are of extreme size (called giant) due to the presence of large blocks of constitutive heterochromatin [18]. In M. cabrerae, the whole Xp and the proximal...
region of the Xq are occupied by constitutive heterochromatin, while the Y chromosome is completely heterochromatic except for the tiny Yq, which is euchromatic. Fluorescence in situ hybridization (FISH) experiments revealed that the multiple copies of the SRY in M. cabrerae are located on both X and Y chromosomes [19]. The distribution of the SRY sequences on the sex chromosomes of this species follows an interesting pattern, with most copies spreading on several clusters along the heterochromatic region of the X chromosome and few sequences located in the euchromatic region of the Y chromosome [19].

Previous SRY HMG-box sequence analysis indicated that M. cabrerae SRY copies are probably nonfunctional pseudogenes, as most of them presented internal stop codons [16]. In addition, the development of XX individuals as normal females suggested the nonfunctional condition of X-linked SRY copies. However, due to the male-determining role of the Y chromosome in M. cabrerae, at least one copy of the SRY gene located on this chromosome should be functional and act as the sex-determining gene in this species [16].

Our previous molecular analyses of SRY from M. cabrerae were restricted to the conserved HMG-box sequence (202 bp), so no data regarding the 5' and 3' ends of the gene and its flanking sequences were available. Here, we present the detailed characterization of the sequence, structure, and organization of the SRY copies and their flanking regions distributed on the X and Y chromosomes of this species. These data are of great relevance to gain knowledge of the exceptional condition of the Y chromosomes of this species. These data are of great relevance to gain knowledge of the exceptional condition of the Y chromosomes of this species. These data are of great relevance to gain knowledge of the exceptional condition of the Y chromosomes of this species.

Results

X-linked SRY sequences

Using several overlapping fragments amplified by different PCR strategies, we deduced the complete sequence of the EcoRI 4.5-kb fragment of the M. cabrerae female genome containing the X-linked copies of the SRY gene. Fig. 1a shows the composition of the entire fragment, while Fig. 1b contains a diagram with the organization and structure of all the different female SRY sequences analyzed in the present study (for detailed information about the cloning methods, sequence size, and nomenclature, refer to Material and methods and Table 2). In summary, we characterized three fragments expanding from the HMG box to the 5' end, five fragments expanding from the HMG box to the 3' end, and one fragment containing only SRY sequence.

Sequence alignments of the clones containing the same fragment (SRY sequence and/or flanking regions) demonstrated they are almost identical, with an identity ranging between 95.2 and 100%. The sequence variations observed are due mainly to single base changes and also to small deletions and insertions. Most deletions and insertions are related to length variations in microsatellite repeats.

Analysis of the assembly of these fragments revealed the structure and organization of the EcoRI 4.5-kb genomic fragment containing the SRY sequence. It includes a single copy of the entire SRY gene and complex repeated sequences flanking both the 5' and the 3' ends, composed of different L1 elements and one LTR (only at 5') (Fig. 1a).

For a detailed analysis, we divided the 4.5-kb band into three main regions: upstream 5' SRY region (position 1 to 2517), SRY region (position 2518 to 3688), and downstream 3' SRY region (positions 3688 to 4591) (Fig. 1a).

Upstream 5' SRY region

This region is 2517 bp length and is composed of several L1 fragments, one single LTR, and one unknown sequence (Fig. 1a).

The first L1 fragment (1097 bp) probably belongs to an ancestral rodent L1 family called Lx. In particular, it shares a 66.5% identity with the same fragment of the Lx6 member of this family (sequence available in the Repbase, at URL http://www.girinst.org/repbase/index.html). The Lx family is composed of truncated ancient L1 elements, of approximately 1.5 kb in length on average, containing the 3' end of the ORFII and downstream sequence described in the family Murinae [20]. The Lx6 fragment in M. cabrerae is interrupted by the insertion of a (TG)$_n$ microsatellite repeat (positions 694 to 748).

The second L1 fragment (1107 bp) is 78.1% similar to a fragment of the ORFII of the L1-mur3 element from mouse (sequence available in the Repbase) and 70.9% identical to one L1 element from Microtus kirgisorum (GenBank AF115970) [21]. This L1 sequence is interrupted by a solitary LTR sequence (position 1805 to 2135).

The LTR sequence has 68.8% identity with the mouse LTR element ORR1B2 (ORR, origin region repeat) (sequence available in the Repbase) from the MaLRs family (mammalian apparent LTR-retrotransposon). The solitary LTR (ORR1) elements have excised the original internal sequence as a consequence of homologous recombination between both LTRs and have an average length of 327 bp [22]. The LTR located on the 5' region of M. cabrerae SRY presents a polyadenylation site (AATTAAA) and the motif AAAAGAAA, which is a potential SRY binding site.

Finally, in the 5' region of the SRY sequence there is a fragment of 313 bp with no homology to any DNA sequences available in GenBank database. This sequence also includes a microsatellite of (TG)$_n$ repeats (position 2263 to 2306).

Downstream 3' SRY region

While we have found only one single organization for the 5' region, the 3' region is arranged in four different configurations differing in the composition of the L1 elements and in the site of L1 insertion into the SRY sequence. As a consequence, the lengths of the SRY and L1 sequences vary.

Configuration 1: McaXSRY (3 kb) and McaXSRY (2 kb) sequences (Fig. 1b)

It is 903 bp in length and composed of one unknown sequence plus two small L1 fragments (Fig. 1b).

Downstream of SRY, there is a sequence of 557 bp with no homology within the GenBank database. This fragment includes
seven repeats of the tetranucleotide CAAA, which constitutes five potential binding sites for SRY (AAACAAA).

The rest of this region (346 bp) is composed of two L1 fragments. The first (249 bp) shows 56.6% identity with an L1-erotic sequence (or the ancient Lx2 element) from Cricetus (sequence available in the Repbase). A region of 143 bp from this first L1 fragment has 81.8% of identity with the ORFI of the L1 element (clone Mk23) of M. kirgisorum (GenBank Accession No. AF115974) [21]. The second L1 fragment (97 bp) has 76.3 and 88.7% identity with the ORFII of mouse L1-mur3 and M. kirgisorum L1 element (clone MK1), respectively (GenBank Accession No. AF115970) [21].

**Configuration 2: McaXSRY (0.7 kb) large 3′ SRY sequence (Fig. 1b)**

In this case, the 3′ SRY sequences are flanked, in the same position as observed in Configuration 1, by a small L1 fragment (21 bp) showing 70% identity with L1-mur3. It is also identical to a fragment of 21 bp from the L1-mur3 element described in the upstream 5′ SRY region.

**Configuration 3: McaXSRY (0.7 kb) short 3′ SRY sequence (Fig. 1b)**

Here, the SRY sequence is truncated 433 bp before the 3′ end described for Configurations 1 and 2. The flanking sequence is a fragment of 425 bp from an inverted L1 element, with 77.8 and 69.3% identity with the ORFII of L1-mur3 and L1 from M. kirgisorum, respectively. Also, this L1 fragment is 97.2% identical to 425 bp of the L1-mur3 located upstream of the 5′ SRY region.

**Configuration 4: McaXSRY (0.3 kb) sequence (Fig. 1b)**

In this configuration, the SRY sequences are truncated 420 bp before the 3′ end described for Configurations 1 and 2.
and are flanked by 25 bp of an inverted L1 fragment. It is also identical to a fragment of 25 bp from the L1-mur3 element present upstream of the 5′ SRY region.

**SRY region**

The SRY region extends 1171 bp (from position 2518 to 3688) and contains the coding region and 5′ and 3′ UTR (Figs. 1a and 2).

The 5′ UTR was identified by the comparison of our sequence with the 5′ region described for the SRY gene in the vole species Clethrionomys rufocanus, Eothenomys andersoni, and E. smithii [23,24]. The percentages of identity observed ranged from 79.1 to 81.0%. As described for these species, the SRY sequence from the M. cabrerae female genome also presents the motif TTAAATAACAAAAT, which could represent a degenerate and expanded TATA box (Fig. 2). In addition, the 5′ UTR region contains the repeats (TC)_n(TG)_n, which vary in length among the different clones studied. This microsatellite is conserved in the same position in the SRY genes from mouse [25], C. rufocanus, E. andersoni, and E. smithii [23,24] and in all the species of the Arvicolinae group analyzed by our group (unpublished data).

In the 3′ UTR sequences we found a typical polyadenylation signal (AATAAA) (Fig. 2).

The 603-bp-long ORF included the HMG box and 5′ and 3′ SRY regions. The start codon of the SRY sequence from female M. cabrerae is conserved in all analyzed clones and is also present in the same position in the SRY gene from the above-mentioned vole species. Three putative stop codons are located in SRY sequences at the 3′ end, but only the first codon is conserved in all of the sequences (Fig. 2). Furthermore, all the sequences analyzed presented two internal stop codons (in the 5′ and 3′ regions), indicating that all SRY sequences from the X chromosome characterized here are nonfunctional pseudogenes.

The HMG-box sequences from M. cabrerae female SRY pseudogenes showed high identity with HMG boxes of SRY from mouse (88.6%), human (82.9%), and several Arvicolinae species (95.6 to 98.7%). This similarity in the only conserved region of the gene signifies that all the sequences from the X chromosome analyzed in our study are actually SRY sequences.

**Y chromosome SRY sequences**

The analysis of chromosome-specific SRY sequences could be of importance for a proper understanding of the unique distribution of this gene in this species, so we also attempted the characterization of the SRY copies located on the Y chromosome. To obtain a sample containing DNA only from the Y chromosome and not from the X chromosome we carried out a microdissection of Y chromosomes and used them as the sample in PCR amplification. By this approach, we successfully cloned and characterized three different fragments from the Y chromosome, including just the SRY sequence (McaYSRY (1 kb)), the SRY and 5′ flanking sequences (McaYsRY (1.2 kb)), or the SRY and 3′ flanking sequences (McaYSRY (1.7 kb); Fig. 1c) (for detailed information about the cloning methods, sequence size, and nomenclature, see Material and methods and Table 2).

The sequence and structure of the SRY copies from the Y chromosome are identical to those described above for the X chromosome.

---

**Fig. 2.** Nucleotide sequence of the SRY pseudogene from the X chromosome (Clone McaXSRY(1.0)-2). The expanded and degenerate TATA box is dot underlined, the start codon is double underlined, the two internal stop codons are in italics and underlined, the three possible stop codons are underlined, the polyadenylation signal is in italics and double underlined, and the HMG box is boxed.
sequences of the X chromosome (Figs. 1a and 2). It should be mentioned that all Y-linked sequences characterized so far are also nonfunctional SRY pseudogenes containing internal stop codons at the same positions as in the X-linked sequences (Fig. 2). The identity between X-and Y-linked SRY sequences ranged from 96.1 to 99.7%, and no chromosome-specific base changes were detected in the alignments. In fact, some X chromosome sequences showed more identity with Y sequences than with sequences from the same chromosome.

Concerning the 5′ and 3′ flanking regions of the SRY copies located on the Y chromosome, they also include the same repeated DNA elements (L1 and LTR fragments) described for the X sequences arranged in the same order (Fig. 1c). In particular, as expected by primer pair positions, at the 5′ end the organization is identical to that found in the McaXSRY (1.5 kb) sequence, while at the 3′ end the organization corresponds to the arrangement described for the McaXSRY (3 kb) and McaXSRY (2 kb) sequences (Fig. 1c).

**Predicted SRY protein**

We have demonstrated that all SRY sequences analyzed in the present study are nonfunctional pseudogenes, as they contain two internal stop codons. Despite the fact that we have not characterized the functional SRY gene, we could infer the main features of the functional copy assumed to exist on the Y chromosome of *M. cabrerae*. If we consider the first TAG conserved in all sequences as the stop codon, then the hypothetical ORF of the SRY gene in this species would be 603 bp long, including the HMG box and 5′ and 3′ coding sequences (Fig. 2). The translation of this ORF, skipping the two internal stop codons, results in a hypothetical protein of 200 amino acids, with a HMG domain of 79 amino acids and N- and C-terminal ends of 39 and 82 amino acids, respectively. Only two sequences (McaXSRY(1.0)-4 and McaXSRY(1.0)-5) would show different protein features; the first presents deletions resulting in a frameshift, while the second has a trinucleotide insertion that changes the protein size (201 amino acids; Fig. 3). The identity between the translated amino acid sequences, excluding the one corresponding to McaXSRY(1.0)-4, varies from 89.2 to 100%.

Comparison of the predicted amino acid sequences of the SRY proteins from *M. cabrerae* with the sequences from other mammals demonstrated that the HMG domain is the most conserved region of the protein, while the N- and C-terminal regions present high degrees of variation affecting both the length and the amino acid composition.

**FISH analysis**

FISH with a probe of 1 kb containing exclusively the whole sequence of the SRY gene (McaXSRY (1 kb)-1; Figs. 1a and b) confirmed the results of Fernández et al. [19], who used only the HMG-box sequence of the gene (202 bp) as a probe. The SRY sequences were located on the euchromatic short arm of the Y chromosome and distributed in several clusters along the heterochromatic block of the X chromosome, which occupied its entire short arm and the proximal region of its long arm (Fig. 4, inset).

FISH with a probe that contained the Lx6 and L1-mur3 repeated sequences flanking the SRY at the 5′ end (Fig. 1a) revealed a wide autosomal location according to its mobile condition, together with a noteworthy distribution on the sex chromosomes. Lx6 and L1-mur3 are located on the euchromatic Y short arm and in the heterochromatin of the X chromosome.

---

Fig. 3. Clustal alignment of the predicted amino acid sequences of the SRY pseudogene from the X (McaXSRY) and Y (McaYSRY) chromosomes. The predicted amino acid sequences were obtained without considering the two internal stop codons. The HMG domain is underlined and stop codons are indicated by asterisks.
following the same distribution pattern as SRY sequences (Fig. 4). This colocalization suggests that SRY sequences are arranged and spread together with L1 elements on both sex chromosomes. In addition, Lx6 and L1-mur3 are also localized on the euchromatic region of the X chromosome (Fig. 4).

Discussion

SRY pseudogenes on the X and Y chromosomes of M. cabrerae

The results we present here clearly confirm the existence of multiple variable copies of SRY in both sex chromosomes of M. cabrerae, as previously described by our group [16,19]. Here, we provide a chromosome-specific characterization of SRY sequences and demonstrate that both the X and the Y chromosome contain several variable SRY copies. The existence of multiple SRY copies on the Y chromosome of M. cabrerae is not surprising as this situation has been described in all the related Arvicolinae rodents studied to date [17] and also in other rodents from the family Muridae [13–15]. However, the presence of SRY sequences on the X chromosome, and hence in all normal females, is so far a unique condition of the species M. cabrerae.

We have demonstrated that SRY sequences on both sex chromosomes include the complete ORF of the gene and 5' and 3' UTR regions. However, all analyzed copies of SRY, including those from the Y chromosome, proved to be nonfunctional pseudogenes, as they contain two internal stop codons in their coding sequences. Despite the pseudogene condition of all SRY copies characterized from the Y chromosome, it might be assumed that at least one functional sequence of this gene exists on it, as this chromosome is male determining in this species [18].

The SRY sequences located on the X chromosome should always have been inactive; otherwise females carrying an X chromosome containing a functional SRY gene would have developed as sterile males, as occurs with female mice transgenic for the SRY gene [3], and hence this chromosome would have been lost from the population. The most likely scenario to explain the inactivation of the SRY copies transferred from the Y to the X chromosome is that they were already pseudogenes, as most Y-linked SRY sequences are also nonfunctional genes sharing the internal stop codons with the X-linked sequences. On the other hand, the insertion into the X chromosome heterochromatin could have been another mechanism to achieve gene inactivation in case transferred sequences included a functional SRY copy.

From the SRY pseudogenes analyzed, we inferred the molecular sequence and structure of the functional SRY gene of this species. The features of this gene are very similar to those observed in other Microtus species (our unpublished data) and would make the M. cabrerae SRY gene more similar to human and Bovidae SRY [2,26] than to mouse Sry [1].

A previous molecular study of the 5' region of the SRY gene demonstrated that sequence homology upstream of the start codon is maintained for only 400–600 bp comparing related mammals [27], which hampers the identification of regulatory elements. However, we could identify in this region an expanded degenerate TATA box, which has also been characterized in other rodent species and has been considered a transcriptional regulatory element [4,23,24]. In addition, in the 5' region the M. cabrerae SRY pseudogenes present the microsatellite (TC)n(TG)n, which has been related to the formation of secondary structures that could reduce the translational efficiency of the mouse Sry gene [4].

SRY–retroposons association

One important finding from our analyses is the association between SRY pseudogenes, L1, and LTR retroelements in both sex chromosomes of M. cabrerae. FISH experiments also demonstrated that SRY and L1 sequences (Lx6-L1-mur3) colocalize on the Y chromosome euchromatic short arm and along the X chromosome heterochromatic block.

The LTR sequence flanking M. cabrerae SRY pseudogenes could be a member of the MalRs family (mammalian apparent LTR–retrotransposons), a type of retroelement usually located as solitary LTRs in the UTRs and flanking regions of genes [22]. Solitary LTR sequences associated with the 5' and 3' ends of some genes could either present bidirectional promoter activity, being capable of promoting transcription of upstream and downstream genes [28], or act as enhancer elements or transcription terminators [29]. Hence, LTR elements associated with M. cabrerae SRY pseudogenes could conceivably have some of these activities. Nevertheless, we do not know yet whether SRY pseudogenes are actually transcribed in this species.

The association of SRY pseudogenes with truncated L1 sequences could have happened by chance, due to the abundance of L1 sequences in the mammalian genome, since most L1 elements are functionally inactive because of truncation, rearrangement, or mutation [30]. Complex L1 fragments associated with genes and pseudogenes have been previously described [31–33]. In addition, after exhaustive sequence analyses, we have observed

Fig. 4. FISH of M. cabrerae male metaphase with flanking 5' L1 sequences (probe Lx6-L1-mur3) and with a probe that contained exclusively SRY sequences (probe McaXSRY (1 kb)) (inset).
diverse L1 and LTR fragments flanking the SRY gene in chimpanzee, human, and mouse. However, a noteworthy question that remains elusive is whether such association with retro-elements was the situation of the (functional) ancestral SRY gene in *M. cabrerae* or was randomly achieved during the process of pseudogene formation and amplification.

**Amplification of SRY sequences in the Y chromosome**

Gene amplification on the Y chromosome has been a frequent event during mammalian evolution with several genes, such as *Ube1Y*, *ZFY*, *RBMY*, and *TSFY*, presenting multiple copies [34–37]. Concerning *Microtus* species, Bullejos et al. [16] proposed that SRY amplification on the Y chromosome occurred before the evolutionary split of the genera *Microtus*, *Pitmys*, *Arvicola*, and *Chionomys* (all of them members of the Arvicolinae group), as all the species analyzed in this group have multiple SRY copies on the Y chromosome. However, it could not be excluded that gene amplification occurred independently on the Y chromosome of each species, as demonstrated for the SRY and ZFY genes in several species of Old World rodents [13].

SRY amplification on the Y chromosome of *M. cabrerae* might have been favored due to the association with L1 elements. Unequal sister chromatid exchange or 3′ transduction, both mediated by L1 sequences, could have been the amplification mechanisms responsible. In fact, gene duplication and pseudogene formation by unequal exchange mediated by L1 elements have been previously described in human olfactory receptor genes [33], in the primate γ-globin gene [31], and in zinc-finger protein genes from mouse [32]. Also, gene duplication by 3′ transduction of L1-associated genes has been demonstrated [38,39].

A third possible mechanism of amplification, even if ancestral *M. cabrerae* SRY sequences on the Y were not associated with L1 elements, could have been retrotransposition. In a scenario in which SRY expression and elevated L1 activity coexist, as occurs in germ-line cells, an SRY transcript could be retroposed by the L1 machinery and integrated between L1 sequences. In fact, it has been demonstrated that L1 can act in trans, mobilizing transcribed sequences that generate processed pseudogenes [40]. Unfortunately, SRY is an intronless gene and therefore we cannot prove this hypothesis by the absence of introns in the retroposed pseudogene sequence. Once the SRY gene was retroposed and integrated among L1 sequences, it could also be subsequently amplified by L1-mediated unequal exchange or 3′ transduction, as mentioned above.

**Transference and amplification of SRY sequences on the X chromosome**

A difficult question to explain is the mobilization of SRY sequences from the Y to the X chromosome might be the above-mentioned L1-mediated retrotrotransposition. An alternative explanation that could be proposed is an ectopic homologous or nonhomologous recombination caused by the L1 elements flanking the SRY sequence on the Y chromosome. Genetic rearrangement due to ectopic recombination mediated by L1 elements has been described in human [41] and *Drosophila* [42]. Transference between nonrecombining regions of sex chromosomes by ectopic gene conversion has been demonstrated in Zfy genes in Felidae [43]. In this respect, our molecular and FISH analyses showed that L1 sequences associated with SRY pseudogenes are distributed in *M. cabrerae* along the entire X chromosome.

Once a cluster of SRY pseudogenes was transferred to the X chromosome heterochromatin, it could spread along the entire heterochromatic block as a consequence of unequal exchanges, probably caused by L1 association. Alternatively, the original transferred SRY sequences could have been amplified and spread along the X heterochromatin after a consecutive process of deletion and amplification, normally involving repetitive DNA sequences. The heterochromatic regions of the sex chromosomes of *M. cabrerae* are frequently implicated in chromosomal rearrangements [18].

According to the proposed mechanisms to explain SRY mobilization, SRY sequences from the Y chromosome might have been transferred to other regions in autosomes or in euchromatin of the X chromosome. However, the damaging effect of insertions into these regions could have favored their elimination by negative selection. By contrast, SRY insertion into the X heterochromatin could have been easily preserved as no gene damage occurred.

Finally, the high sequence homology existing between X-linked and Y-linked SRY pseudogenes suggests that SRY transference from the Y to the X chromosome was a recent evolutionary event. Alternatively, concerted evolution through ectopic gene conversion between X- and Y-linked SRY copies could homogenize the SRY sequences. This mechanism of sequence homogenization has been proposed to occur in Felidae between Zfy and Zfx, both genes located on the nonrecombining regions of sex chromosomes [43].

**Materials and methods**

**Genomic DNA extraction**

*M. cabrerae* male and female genomic DNAs were extracted from several tissue samples fixed in 70% ethanol or from immortalized cultured cells following the standard phenol–chloroform procedure [45].

**Cloning and sequencing of the complete SRY gene from *M. cabrerae***

Our cloning strategies are based on previous analyses of the SRY gene of *M. cabrerae* from [16]. This study demonstrated that when male and female genomic DNA is digested with EcoRI and analyzed by Southern blot, the SRY copies are present on a single band of approximately 4.5 kb. As most SRY copies are located on the X chromosome and few of them are located on the Y chromosome, to clone X-linked copies, we used female genomic DNA as sample; whereas for cloning Y-linked copies, we used DNA from microdissected Y chromosomes.
Cloning SRY sequences from the X chromosome

The strategies for cloning SRY sequences located on the X chromosome of *M. cabrerae* are summarized in Fig. 5. First, anchored PCR procedures were designed to clone the 5′ and 3′ sequences flanking the SRY HMG box (Fig. 5a). Briefly, female *M. cabrerae* genomic DNA was restricted with EcoRI and separated by gel electrophoresis. Fragments of 4–5 kb were eluted from the gel, purified, and ligated into EcoRI-restricted pUC19 plasmid to generate a size-selected plasmid library. Next, four anchored PCRs combining a specific primer for the SRY HMG box (primer SRY-A or SRY-F) and a primer derived from the pUC19 sequence (M13-R or F) from the pUC19 sequence. Only one of the two possible cloning orientations is represented. (b) Nested-inverse PCR on circularized 4.5-kb bands containing the SRY sequences. Boxes represent the HMG-box and the 5′ sequences characterized with the anchored PCR strategy. Primers used were SRY-A and-E from the HMG box and SRY-Inv1 and-Inv2 from the 5 sequences.

**Cloning SRY sequences from the Y chromosome**

The previous characterization of the X-linked SRY sequences facilitated the analysis by PCR of the SRY copies located on the Y chromosome. However, due to the high number of copies present on the X chromosome, the accurate amplification of Y-linked SRY sequences required the use of microdissected Y chromosomes isolated from male metaphase plates as sample. For each PCR, 20 Y chromosomes were microdissected with an inverted microscope (Zeiss Axiovert 200) with a sterile glass needle attached to a mechanical micromanipulator (Eppendorf TransferMan NK 2), as was described previously [46]. Next, PCR with the primer pair SRY-X1/SRY-T (Table 1) and using the same conditions described for female genomic DNA led to the amplification of a 1-kb band that corresponded to complete Y-linked SRY sequences. Four clones derived from this band were sequenced and analyzed (Table 2).

To investigate whether the SRY copies from the Y chromosome are flanked by the same repeated-DNA sequences as the X-linked copies, we amplified the 5′ and 3′ flanking sequences from microdissected Y chromosomes. We designed new primer pairs for the amplification of the 5′ flanking regions (SRY-inv3 and SRY-F) and the 3′ flanking regions (SRY-E and SRY-L13) (Fig. 1a and Table 1). The PCR cycling conditions were as follows: 1 min at 95°C, 1 min at 57°C, and 30 s at 72°C (30 cycles). In both PCRs, only one single band of expected size was amplified. Both bands were cloned and several clones sequenced (Table 2).

PCRs amplifications described above were made using the Expand Long Template PCR System (Roche) according to the manufacturer’s instructions. Sequences were determined using the CEQ DTSQ Quick Star Kit from Beckman Coulter and sequence reactions were analyzed in a Beckman Coulter automated sequencer (Model CEQ 2000XL).

**Sequence analysis**

Pair-wise sequence alignments and multiple alignments were performed with the program ClustalW 1.6 [47]. Sequence-homology searches were made in GenBank using BLASTN 2.2.2 with default parameters [48]. Repeated DNA sequences were screened using the program RepeatMasker version 3.1.6 (A.F.A. Smit, R. Hubley, and P. Green, http://www.repeatmasker.org).

**Table 1**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRY-A</td>
<td>5′-GTCAGGAGGCCTCCATAGCAT-3′-</td>
</tr>
<tr>
<td>SRY-F</td>
<td>5′-ATGGTGAGTTGTTGTTC-3′-</td>
</tr>
<tr>
<td>M13-R</td>
<td>5′-GTAAGAGACGAGCCAGT-3′-</td>
</tr>
<tr>
<td>M13-F</td>
<td>5′-CAGAAGACACAGCTATGAC-3′-</td>
</tr>
<tr>
<td>SRY-E</td>
<td>5′-TTCCCAAAGGAGCAGAGACTG-3′-</td>
</tr>
<tr>
<td>SRY-inv1</td>
<td>5′-TACCATGTAACGTAGTCTTT-3′-</td>
</tr>
<tr>
<td>SRY-inv2</td>
<td>5′-ACCCACCAATTGGCCAAAATTCA-3′-</td>
</tr>
<tr>
<td>SRY-X1</td>
<td>5′-TCCCAATCGGAATGTTACT-3′-</td>
</tr>
<tr>
<td>SRY-T</td>
<td>5′-TGAGGACTCAATAGAATGAG-3′-</td>
</tr>
<tr>
<td>SRY-inv3</td>
<td>5′-CTAGGCTGTAATGAGCTTG-3′-</td>
</tr>
<tr>
<td>SRY-P</td>
<td>5′-TATGAGCATAATCATATAT-3′-</td>
</tr>
<tr>
<td>SRY-L13</td>
<td>5′-TCTTATACCCAGGTATGTTCA-3′-</td>
</tr>
</tbody>
</table>
Table 2
Summary of the DNA samples, cloning strategy, and primer pairs used for amplification of the fragments containing SRY sequences and 5’ and 3’ flanking region of M. cabrerae

<table>
<thead>
<tr>
<th>Sample</th>
<th>Cloning strategy</th>
<th>Primers used</th>
<th>SRY region</th>
<th>Band size (bp)</th>
<th>Nomenclature</th>
<th>Clones analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female genomic DNA</td>
<td>Anchored PCR</td>
<td>SRY-F, M13-F or SRY-F, M13-R</td>
<td>Partial SRY, 5’ flanking sequences</td>
<td>1482</td>
<td>McaXSRY (1.5 kb)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1st PCR: SRY-A, SRY-inv2</td>
<td>Partial SRY, 5’ and 3’ flanking sequences</td>
<td>3215</td>
<td>McaXSRY (3 kb)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2nd PCR: SRY-E, SRY-inv1</td>
<td></td>
<td>2158</td>
<td>McaXSRY (2 kb)</td>
<td>2</td>
</tr>
<tr>
<td>Female genomic DNA</td>
<td>Inverse-nested PCR</td>
<td>SRY-A, SRY-inv2</td>
<td>Partial SRY, 3’ flanking sequences</td>
<td>715</td>
<td>McaXSRY (0.7 kb) large</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1st PCR: SRY-A, SRY-inv2</td>
<td>Partial SRY, 3’ flanking sequences</td>
<td>686</td>
<td>McaXSRY (0.7 kb) short</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2nd PCR: SRY-E, SRY-inv1</td>
<td></td>
<td>277</td>
<td>McaXSRY (0.3 kb)</td>
<td>4</td>
</tr>
<tr>
<td>Female genomic DNA</td>
<td>Inverse-nested PCR</td>
<td>SRY-A, SRY-inv2</td>
<td>Partial SRY, 3’ flanking sequences</td>
<td>1064</td>
<td>McaXSRY (1 kb)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2nd PCR: SRY-E, SRY-inv1</td>
<td>Partial SRY, 5’ flanking sequences</td>
<td>1220</td>
<td>McaYSRY (1.2 kb)</td>
<td>3</td>
</tr>
<tr>
<td>Female genomic DNA</td>
<td>Partial</td>
<td>SRY-X1, SRY-T</td>
<td>SRY sequence</td>
<td>1677</td>
<td>McaYSRY (1.7 kb)</td>
<td>5</td>
</tr>
<tr>
<td>Y-chromosome</td>
<td></td>
<td>SRY-P, SRY-inv3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>micrdissected</td>
<td></td>
<td>1st PCR: SRY-A, SRY-inv2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y-chromosome</td>
<td></td>
<td>2nd PCR: SRY-E, SRY-inv1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>micrdissected</td>
<td></td>
<td>SRY-E, SRY-L13'</td>
<td>Partial SRY, 3’ flanking sequences</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y-chromosome</td>
<td></td>
<td>SRY-X1, SRY-T</td>
<td>SRY sequence</td>
<td>1064</td>
<td>McaXSRY (1 kb)</td>
<td>4</td>
</tr>
<tr>
<td>micrdissected</td>
<td></td>
<td>SRY-P, SRY-inv3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The amplified regions, the size of the amplified PCR product, the name assigned, and the number of analyzed clones are indicated.

Chromosome preparation and fluorescence in situ hybridization

Permanent fibroblast cell lines from M. cabrerae were used to obtain chromosomes, as previously reported [44]. The FISH procedure for the M. cabrerae SRY HMG box [19] was as previously described, using probes PCR-labeled with biotin-16-dUTP (Roche). The sequences used as probes were McaXSRY (1 kb)-1 (containing exclusively SRY sequences) and a fragment of 1638 bp containing L1 sequences (575–121).

Acknowledgments

The authors express their gratitude to Junta de Castilla–La Mancha for capture permits for M. cabrerae. This work was supported by the Spanish Ministerio de Educación y Ciencia through Projects BOS2002-04150-C02-01, BOS2002-04150-C02-02, and CGL2006-05308/BOS (cofunded by the European Regional Development Fund) and by the Junta de Andalucía throughout the program “Ayudas a Grupos de Investigación,” Group CVI 220.

References

cabraeae: a vole species with multiple SRY copies in males and females, Genome 45 (2002) 600–603.