

# Centromeric enrichment of LINE-1 retrotransposons and its significance for the chromosome evolution of Phyllostomid bats

Cibele Gomes de Sotero-Caio · Diogo Cavalcanti Cabral-de-Mello ·  
Merilane da Silva Calixto · Guilherme Targino Valente · Cesar Martins · Vilma Loreto ·  
Maria José de Souza · Neide Santos

Received: 7 May 2017 / Revised: 24 August 2017 / Accepted: 28 August 2017 / Published online: 15 September 2017  
© Springer Science+Business Media B.V. 2017

**Abstract** Despite their ubiquitous incidence, little is known about the chromosomal distribution of long interspersed elements (LINEs) in mammalian genomes. Phyllostomid bats, characterized by lineages with distinct trends of chromosomal evolution coupled with remarkable ecological and taxonomic diversity, represent good models to understand how these repetitive sequences contribute to the evolution of genome architecture and its link to lineage diversification. To test the hypothesis that LINE-1 sequences were important modifiers of bat genome architecture, we characterized the distribution of LINE-1-derived

sequences on genomes of 13 phyllostomid species within a phylogenetic framework. We found massive accumulation of LINE-1 elements in the centromeres of most species: a rare phenomenon on mammalian genomes. We hypothesize that expansion of these elements has occurred early in the radiation of phyllostomids and recurred episodically. LINE-1 expansions on centromeric heterochromatin probably spurred chromosomal change before the radiation of phyllostomids into the extant 11 subfamilies and contributed to the high degree of karyotypic variation observed among different lineages. Understanding centromere

---

Responsible Editor: Rachel J. O'Neill.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10577-017-9565-9>) contains supplementary material, which is available to authorized users.

C. G. de Sotero-Caio (✉) · M. d. S. Calixto · V. Loreto ·  
M. J. de Souza · N. Santos  
Departamento de Genética, Laboratório de Genética e Citogenética  
Animal e Humana, UFPE-Universidade Federal de Pernambuco,  
Av. da Engenharia s/n; Cidade Universitária, Recife, PE  
CEP:50740-600, Brazil  
e-mail: cibeledcaio@gmail.com

C. G. de Sotero-Caio  
Department of Biological Sciences, Texas Tech University,  
Lubbock, TX, USA

D. C. Cabral-de-Mello  
Departamento de Biologia, Grupo de Estudos em Citogenômica e  
Evolução Animal, UNESP-Universidade Estadual Paulista,  
Instituto de Biociências, Rio Claro, SP, Brazil

M. d. S. Calixto  
Centro de Saúde e Tecnologia, Unidade Acadêmica de Ciências  
Biológicas, UFCG-Universidade Federal de Campina Grande,  
Patos, PB, Brazil

G. T. Valente  
Departamento de Bioprocessos e Biotecnologia da Faculdade de  
Ciências Agronômicas, UNESP-Universidade Estadual Paulista,  
Botucatu, SP, Brazil

C. Martins  
Departamento de Morfologia, Laboratório Genômica Integrativa,  
UNESP-Universidade Estadual Paulista, Botucatu, SP, Brazil

architecture in a variety of taxa promises to explain how lineage-specific changes on centromere structure can contribute to karyotypic diversity while not disrupting functional constraints for proper cell division.

**Keywords** non-LTR · transposable element · Phyllostomidae · Chiroptera · karyotype · centromere

### Abbreviations

FISH Fluorescent in situ hybridization  
 LINE Long interspersed element  
 ORF Open reading frame  
 TE Transposable element  
 GIRI Genetic Information Research Institute  
 ML Maximum likelihood

### Introduction

Centromeres are remarkable components of eukaryotic genomes not only for their role in cell division, but also because of their unique structural features (e.g., DNA composition and variability) and their link to chromosome rearrangements and diversity. Studies of mammalian centromeres and adjacent (pericentromeric) regions indicate that their typical centromere contains long stretches of tandemly repeated satellite DNA as core centromeric sequences, with variable monomer length and sequence complexity (Brown and O'Neill 2014; Steiner and Henikoff 2015). Transposable elements (TEs) are non-satellite DNAs that ubiquitously contribute to pericentromeric sequences in mammals. However, these elements are rarely incorporated as major core centromeres, nor are they visibly enriched in these areas relative to euchromatic regions, with the exception of the X chromosome where they are usually enriched along the chromosomal length (Acosta et al. 2008; Brown and O'Neill 2014; Waters et al. 2004). Despite their significant incidence, distribution patterns of TEs on mammalian chromosomes have been poorly investigated. The handful of molecular cytogenetic studies targeting chromosomal mapping of TEs shows that very few species accumulate these elements preferentially in centromeres (Bulazel et al. 2006; Carbone et al. 2012; Ferreri et al. 2011; Kapitonov et al. 1998; Waters et al. 2004). The limited number of studies, however, makes it difficult to determine if centromeric TE accumulation is indeed a rare phenomenon or if this

assumption results from the lack of representation of investigated mammalian species.

Long interspersed elements (LINEs), particularly the family LINE-1, are the most abundant TEs in mammalian genomes (Boissinot and Furano 2001; Furano et al. 2004; Richardson et al. 2015). These class I TEs, or retrotransposons, amplify and reinsert themselves in host genomes via an RNA intermediate. Their “copy and paste” mechanism allows LINEs to accumulate over time to compose large proportions within whole genomes (Deininger and Batzer 2002; Eickbush and Jamburuthugoda 2008; Luan et al. 1993). The non-random insertion of LINE-1 elements is still a topic of controversy. Although the consensus target site for LINE-1 insertion has a prevalence of adenines and thymines (AT-rich), evidence from a few studies suggests that young and active elements do not preferentially insert themselves in specific genomic regions (in terms of AT-rich G-bands) in mammals other than Euarchontoglires (Ovchinnikov et al. 2001; Waters et al. 2004). Other lines of evidence support initial insertion site preference of LINE-1 elements, targeted at AT-rich regions, as well as specific epigenetic states of chromatin (Graham and Boissinot 2006; Zhang and Mager 2012). Despite insertion site controversies, the consensus is that older insertions might be eliminated through recombination and distinct patterns can rise through processes of drift and purifying selection (Ovchinnikov et al. 2001; Parish et al. 2002; Scott et al. 2006; Wichman et al. 1991, 1992). Thus, areas undergoing low recombination rates, as well as regions less subject to detrimental effects of insertions (both features of centromeres), can be enriched with LINE-1, shaping patterns of chromosomal distribution of these elements differently between species. Nevertheless, centromeric regions are unusual in that they promote sequence homogenization through processes such as gene conversion (Shi et al. 2010). It is expected then that when TEs (including LINEs) are exapted as core centromere sequences, amplification and enrichment in this chromosomal region might be a direct result of intrinsic biological features of centromere.

Besides their contribution in shaping genome architecture, LINE-1 elements are regarded as powerful drivers of chromosomal rearrangements and evolution. Because of their repetitive nature and sequence similarity among elements distributed in different chromosomes, LINEs are prone to serve as templates for recombination of homologous sequences between non-homologous chromosomes and to the creation of double-strand breaks

(Deininger and Batzer 2002; Gray 2000; Hedges and Deininger 2007). Thus, unraveling enrichment and distribution of these sequences is key to understand the extant chromosome diversity of lineages. This approach can provide further insight to the underlying processes leading to increased rates of chromosomal evolution in groups with highly rearranged karyotypes relative to closely related taxa that do not present increased karyotypic change. In this context, phyllostomid bats are good models to test the hypothesis that LINE-1 amplification has contributed to chromosomal diversity.

Phyllostomid bats are a remarkable assemblage among higher vertebrates and underwent a series of rapid adaptive radiation events from a common ancestor, beginning ~ 30 mya. Extant lineages (11 subfamilies) present considerable karyotypic diversity, with diploid numbers varying from  $2n = 14$  to 46, as well as varying rates and diverse patterns of chromosome evolution in lineages with similar diversification ages (Baker and Bickham 1980; Martinez et al. 2017). Most of the extreme reshuffling of karyotypes occurred at the base of each subfamily radiation from an ancestral karyotype of  $2n = 46$ ,  $FN = 60$ , mainly through Robertsonian fusions, but also other more complex rearrangements (Baker and Bickham 1980; Pieczarka et al. 2005, 2013; Ribas et al. 2013, 2015; Sotero-Caio et al. 2011, 2013, 2015). Furthermore, bat clades are distinguished by reduced genome size, mainly derived from elimination of repetitive sequences (Baker et al. 1992; Smith et al. 2013). We hypothesize that any unusual accumulation of repeated DNA (including LINEs) will represent potential modifiers of bat genome architecture.

Herein, we describe massive accumulations of a retrotransposon-derived sequence on centromeres of 13 species of phyllostomid bats belonging to lineages characterized by distinct patterns of chromosomal evolution. We discuss our findings in light of overall trends, C-positive heterochromatin distribution, LINE-1 transposition dynamics, and their potential roles for centromere function and spurring of chromosomal change.

## Material and methods

### Sampling and classical cytogenetics

Cytogenetic studies were carried out in 13 phyllostomid bat species captured at different sites across Pernambuco

State, Brazil (Table 1). Voucher specimens are deposited in the mammal collection at the Department of Zoology, Federal University of Pernambuco, Recife, Pernambuco, Brazil. Handling and carrying of animals followed guidelines established by the Animal Care and Use guidelines of ICMBio (Instituto Chico Mendes de Conservação da Biodiversidade), and using the collecting permit number (12264-1). Mitotic chromosomes were obtained from bone marrow cells according to conventional procedures (Baker and Qumsiyeh 1988; Lee and Elder 1980). Slides with metaphase chromosomes of all species were submitted to Giemsa 5% staining. In addition, C-banding was applied to detect constitutive heterochromatin in some species, following previously described methods (Sumner 1972). Idiograms were constructed in Inkscape (<https://inkscape.org>) using previously published descriptions of C-banded chromosomes, supplemented with our data (Araújo 2016; Barros et al. 2009; Lemos Pinto et al. 2012; Santos et al. 2001; Santos and Souza 1998a, b; Varella-Garcia et al. 1989).

### Isolation and sequencing of repetitive sequences

Genomic DNA of three *Artibeus planirostris* individuals was extracted from liver tissue, following standard phenol–chloroform procedures (Sambrook and Russell 2001). Fragments of a LINE-1 element, corresponding to a partial region of the open reading frame I (ORF1), were amplified from the genome of *A. planirostris* by polymerase chain reaction (PCR) using the primers 18SF (5'-CCG CTT TGG TGA CTC TTG AT) and 18SR (5'-CCG AGG ACC TCA CTA AAC CA). Interestingly, the same set of primers were originally designed to amplify the 18S rDNA sequence in fish species (Teixeira et al. 2009) but do not amplify rDNA from bats. To confirm the identity of the amplified fragment, the ~ 900 bp PCR product was cloned, ligated into a plasmid pGEM-T (Promega, Madison, WI, USA), and recombinant constructs were used to transform DH5 $\alpha$  *Escherichia coli*-competent cells. Eighteen positive clones were sequenced using an ABI Prism 3100 automated DNA sequencer (Applied Biosystems, Foster City, CA, USA) with a Dynamic Terminator Cycle Sequencing Kit (Applied Biosystems), as per manufacturers' instructions.

### DNA sequencing and phylogenetic analysis

To confirm sequence identity, the 18 sequenced clones were assembled using Geneious Pro 4.8.5 software (Drummond et al. 2009) relative to four different types

**Table 1** Chromosomal features and sampling locality information of studied species

Taxon	Specimens		Locality	Coordinates	2n	FN	Sex system
	M	F					
Desmodontinae							
<i>Diphylla ecaudata</i>	0	1	Toritama-PE	08° 00' 24"S 36° 03' 24"W	32	60	XY; XX
<i>Diaemus youngii</i>	0	1	Ipojuca-PE	8° 30' 59"S 35° 03' 38.2"W	32	60	XY; XX
Phyllostominae							
<i>Chrotopterus auritus</i>	0	1	Serra dos Cavalos, Caruaru-PE	8° 18' 36"S 36° 00' 00"W	28	52	XY; XX
<i>Phyllostomus elongatus</i>	1	2	Reserva Biológica de Saltinho, Rio Formoso-PE	8° 44' 13" S 35° 10' 11" W and 8° 43' 09" S 35° 11' 02" W	32	58	XY; XX
Glossophaginae							
<i>Glossophaga soricina</i>	1	2	Reserva Natural Brejo, Saloá-PE	8° 57' 05" S 36° 42' 22" W	32	60	XY; XX
Lonchorhinae							
<i>Lonchorhina aurita</i>	1	1	Toritama-PE	8° 00' 24" S 36° 03' 24" W	32	60	XY; XX
Carolliinae							
<i>Carollia perspicillata</i>	2	2	Reserva Biológica de Saltinho, Rio Formoso-PE	8° 44' 13" S 35° 10' 11" W and 8° 43' 09" S 35° 11' 02" W	20/21	36	XY <sub>1</sub> Y <sub>2</sub> ; XX
Stenodermatinae							
<i>Sturnira lilium</i>	2	3	Aldeia, Camaragibe-PE	8° 01' 18" S 34° 58' 52" W	30	56	Neo-XY; XX
<i>Platyrrhinus lineatus</i>	2	3	Reserva Natural Brejo, Saloá-PE	8° 57' 5" S 36° 42' 22" W	30	60	Neo-XY; XX
<i>Artibeus obscurus</i>	2	1	Parque Estadual Dois Irmãos, Recife-PE	8° 7' 30" S 34° 52' 30" W	30/31	56	XY <sub>1</sub> Y <sub>2</sub> ; XX
<i>A. planirostris</i>	3	2	UFPE Campus, Recife-PE	8° 03' 07" S 34° 56' 59" W	30/31	56	XY <sub>1</sub> Y <sub>2</sub> ; XX
<i>A. lituratus</i>	2	1	Reserva Biológica de Saltinho, Rio Formoso-PE	8° 44' 13" S 35° 10' 11" W and 8° 43' 09" S 35° 11' 02" W	30/31	56	XY <sub>1</sub> Y <sub>2</sub> ; XX
<i>A. cinereus</i>	1	2	Estação Ecológica de Tapacurá-PE	8° 0' 13" S 35° 01' 17" W	30	56	Neo-XY; XX

of sequences: type 1 (12 reads; 796 bp), type 2 (two reads; 717 bp), type 3 (two reads; 704 bp), and type 4 (two reads; 659 bp), for simplicity from now on referred to as contigs 1–4, respectively. All sequences were submitted as queries to RepeatMasker against the Repbase database (Smit et al. 1996–2010), located at the Genetic Information Research Institute (GIRI) server (<http://www.girinst.org/repbase/>). A further step of sequence identification was performed by detecting putative open reading frames (ORFs) within these sequences using ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder>) and then using translated

amino acid sequences of identified ORFs as queries on BLASTp searches. The sequences corresponded to a partial ORF-1 region of LINE-1 elements and were deposited in the NCBI database under accession numbers (MF838931–MF838934).

For phylogenetic analyses, all sequences were first used as queries in a Blastn search against the NCBI nucleotide collection and the hits with E-values  $\leq 2.89 \times 10^{-174}$  and length coverage  $> 50\%$  (a total of 102 hits) were selected to further analysis. Redundant sequences ( $\geq 90\%$  identity) were clustered using CD-Hit (Li and Godzik 2006). A total of

32 sequences were aligned using Muscle (Edgar 2004a, b), and the alignment was manually edited (Online Resource 1; [ESM\\_1.fa](#)). Evolutionary model selection was made using jModelTest (Posada 2008) with Akaike information criterion (AIC) (Akaike 1974). Phylogenetic reconstruction was performed using maximum likelihood (ML) and Bayesian inference by Phyml (Guindon et al. 2009) and Beast (Drummond and Rambaut 2007), respectively, both located in CIPRES Science Gateway (Miller et al. 2010). The approximated likelihood ratio (aLRT (SH2-like)) reliability test was adopted, and values were supported by posterior probabilities by Bayesian analysis. The final tree was analyzed and edited in Figtree v1.3.1 software (Drummond and Rambaut 2007).

### Fluorescence in situ hybridization

Fluorescence in situ hybridization (FISH) was performed by combining the pool of plasmids containing distinct variants of the isolated retrotransposons to be used as probes. Plasmid DNA was labeled by nick translation with biotin 14-dATP (BioNick<sup>TM</sup> Labeling System) (Invitrogen), and FISH procedures followed the protocol proposed by Pinkel et al. (1986) with slight modifications as follows: mitotic chromosomes were denatured in 70% formamide/0.6× SSC (ph 7.0) for 33–36 s at 62 °C; hybridization mixtures (containing 100 ng of denatured probe, 10% dextran sulfate, 2× SSC, and 50% formamide in a final volume of 30 µl) were dropped on slides and hybridization was performed overnight (18 h) at 37 °C in moist chamber; post-hybridization washes were carried out at 72 °C in 2× SSC for 5 min. Detection of probe was carried out with avidin-fluorescein isothiocyanate (FITC) conjugate (Sigma, St Louis, MO, USA). Chromosome preparations were counterstained with propidium iodide (0.2%) and mounted with Vectashield (Vector, Burlingame, CA, USA). Images were captured with an Olympus DP71 digital camera coupled to a BX61 Olympus microscope and were optimized for brightness and contrast using Adobe Photoshop CS2. Between eight and 38 FISH images were recorded for each specimen, except *A. planirostris*, *D. youngii*, and *D. ecaudata*, for which three, three, and six images were, respectively, recorded.

## Results

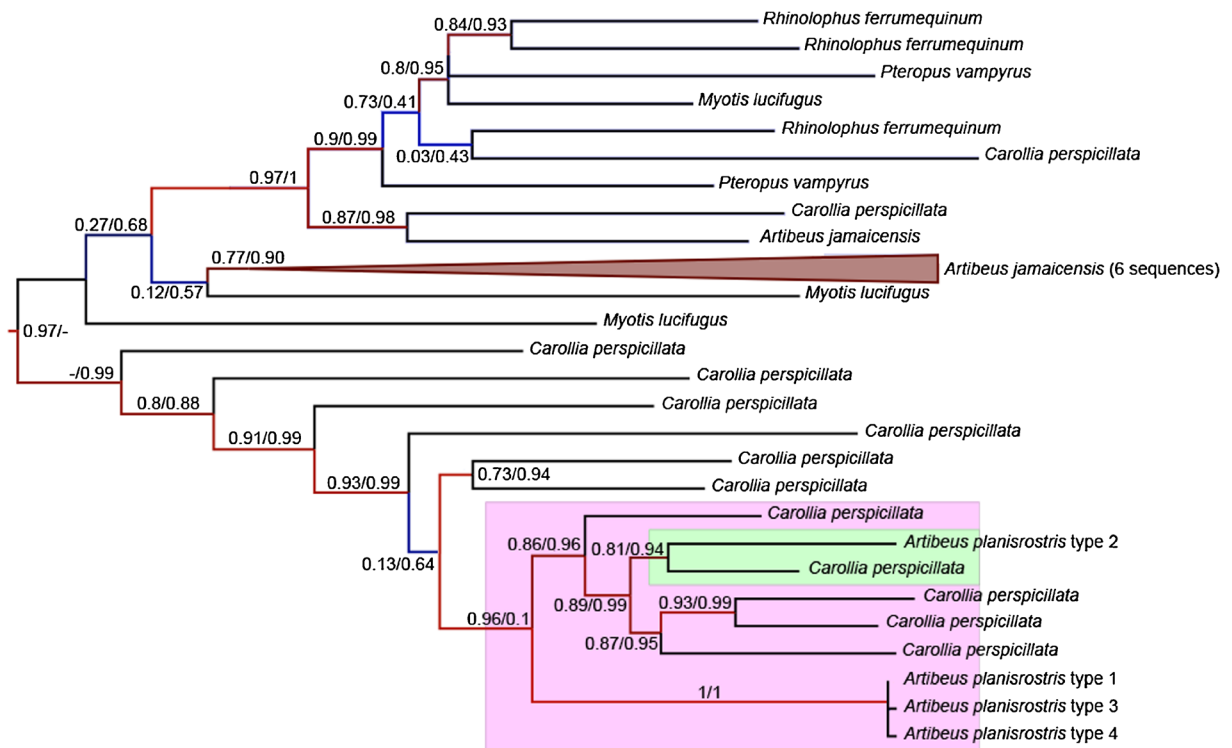
### Characterization of LINE-1 probes

Sequence similarity among the four *A. planirostris* contigs was high: three of the contigs (1, 3, and 4) are more similar, with at least 98.58% sequence similarity to each other. Contig 2 was the most divergent, sharing 79.94% identity of sequence with contig 4. RepeatMasker analysis revealed that these elements present ~ 70% sequence similarity to the consensus sequence of the lineage 4 of a LINE-1 non-LTR retrotransposon of the megabat *Pteropus vampyrus* (L1-4\_PVa). Further investigation showed that the clones span a 686 bp region of *P. vampyrus* LINE-1, corresponding to most of its ORF1 (Yang et al. 2014). Probabilistic-based phylogeny placed the four sequenced contigs among other chiropteran LINE-1 sequences (Fig. 1). Contigs 1, 3, and 4 cluster together; however, contig 2 presented higher similarity to LINE-1 sequences from another phyllostomid species, *Carollia perspicillata*. All sequences from *A. planirostris* are more closely related to those from *C. perspicillata* than to those of the congener *A. jamaicensis*.

### Heterochromatin and LINE-1 distribution on phyllostomid chromosomes

Diploid (2n) and fundamental (FN) numbers, as well as sex system of all species are presented in Table 1 and are consistent with previously reported karyotypic descriptions (O'Brien et al. 2006). Heterochromatin occurs at the pericentromeric region in all species, with some additional heterochromatic blocks (interstitial, terminal, and/or entire arms) as particularities of distinct species. Schematic representations of C-positive regions are shown on Figs. 2 and 3, and representative C-banding images are given in Online Resource 2; [ESM\\_2.jpg](#).

The considerable sequence identity among our four LINE-1 fragments, as well as the high similarity to other bat elements, allowed us to combine them as a single probe for in situ hybridization experiments. Under low stringency, it is expected that these probes detect any enrichment of the target LINE-1 sequence on phyllostomid chromosomes. We found different patterns of distribution of the LINE-1 ORF1 fragment on phyllostomid bat autosomes (Figs. 2, 3, and 4): (1) intense signal spanning the centromeric region and extending into the



**Fig. 1** Cladogram generated using maximum likelihood and Bayesian inference for 32 partial LINE-1 OFR1 sequences from bats, including the four types of contigs isolated from *A. planirostris* (types 1–4, sequenced in this study), and 28 remainder sequences from GenBank. Note that not all sequences from *Artibeus* cluster together. The support values beside nodes

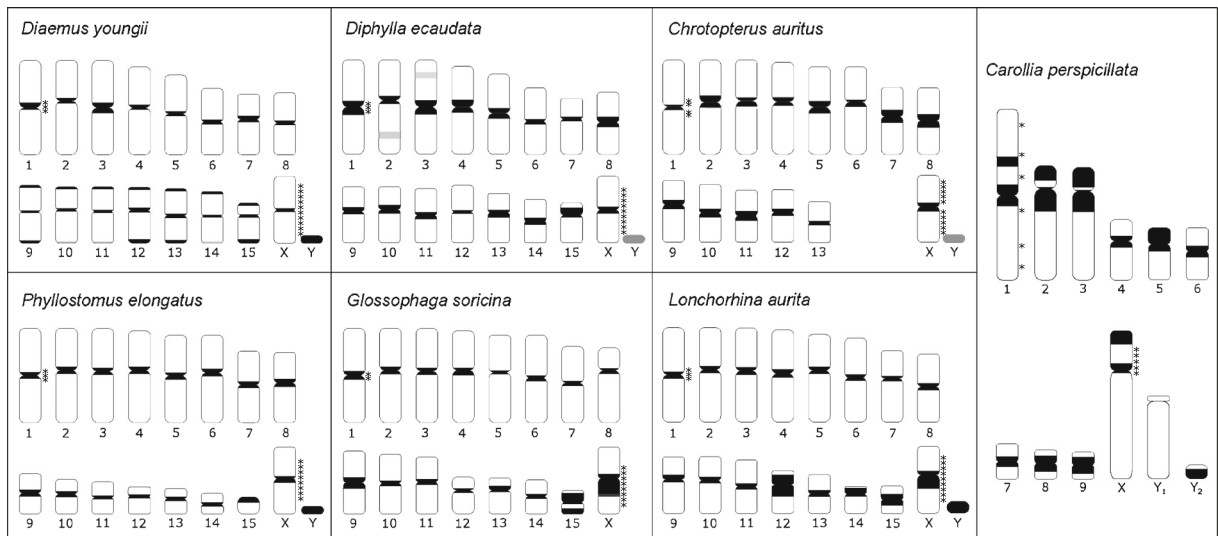
correspond to the approximate likelihood ratio test (aLRT, left) and posterior probability (right); ‘-’ indicates no support or absence of a branch in either of the phylogenetic inference methods. The branch colors correspond to level of branch support (blue, low scores; red, high scores). Scale bar, average of substitutions per site = 0.04

pericentromere (most species); (2) signal restricted to the small, centric region (physical primary constriction) with fainter signal extending into the pericentromere (*A. obscurus*, and *A. cinereus*); (3) signal restricted to pericentromeric regions and absent from centromeres (*Chrotopterus auritus*); (4) no evident enrichment in centromeres and faint enrichment in other chromosomal regions (*C. perspicillata*). It is noteworthy that, although most autosomes within the complement of one species fit one of these specific patterns, outlier chromosomes belonging to distinct categories were also present. Examples are the pericentromeric enrichment on one large metacentric autosome of *A. obscurus* (with absence of signal in the primary constriction), and absence of enrichment on particular autosomes of the species *G. soricina*, *A. obscurus*, and *A. cinereus*. The variation between chromosomes of a single species might, however, be derived from differential chromosome condensation together with reduced copy number of LINE-1 sequences

in particular chromosome regions, not detected by FISH resolution.

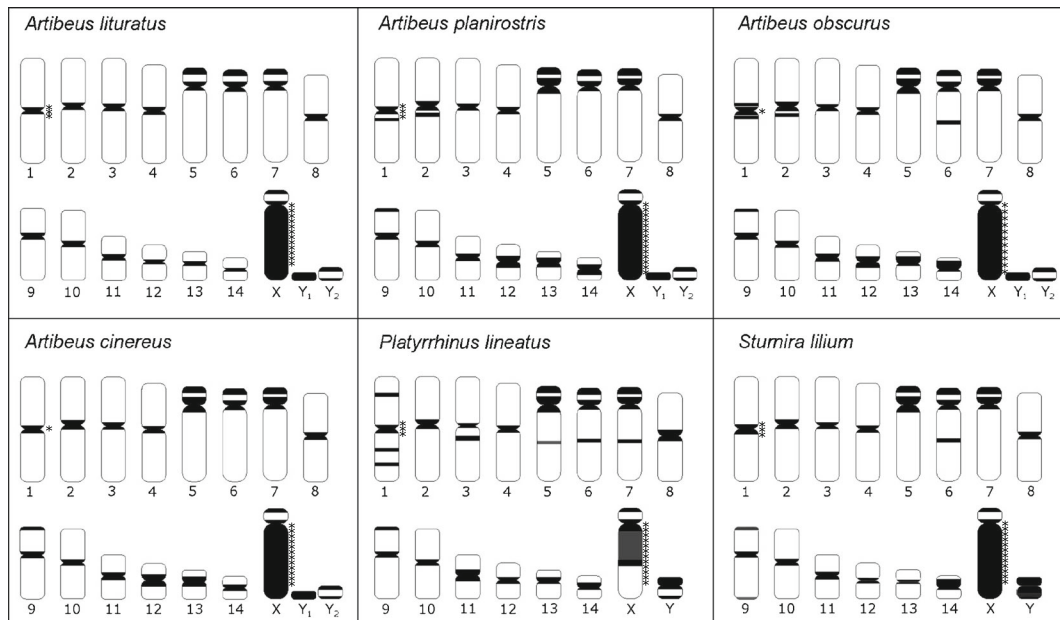
Furthermore, as expected, X chromosomes of all species were enriched with LINE-1 signal throughout their length, to the exception of distal-most regions. Species with X-autosome translocations, such as all Stenodermatinae and *Carollia perspicillata*, show variation of LINE-1 signal in the composite X. In these cases, only the original X chromosome shows euchromatic enrichment of the LINE-1 probes (Figs. 2, 3, and 4).

Placing the slightly different patterns of retroelement accumulation into a phylogenetic framework (Online Resource 3; [ESM\\_3.tif](#)) has not yielded any particular trend among different subfamilies. For instance, the degree of signal intensity (lower or higher abundance of centromeric LINE-1) does not seem to correlate with age of the lineage or degree of chromosomal reorganization. Therefore, a link between rapid karyotypic change and centromeric turnover cannot be readily established from our mapping results. Finally,



**Fig. 2** Schematic representation of C-banding and LINE-1 FISH patterns on phyllostomid species from the subfamilies Desmodontinae, Phyllostominae, Glossophaginae, Lonchorhinae, and Carollinae. Dark regions correspond to constitutive heterochromatin blocks. Gray shading corresponds to fainter heterochromatin distribution. Schematic representation of chromosomes and bands were compiled from this study and those of Araújo (2016), Barros et al. (2009), Bass (1978), Morielle-Versute et al. (1992), Santos

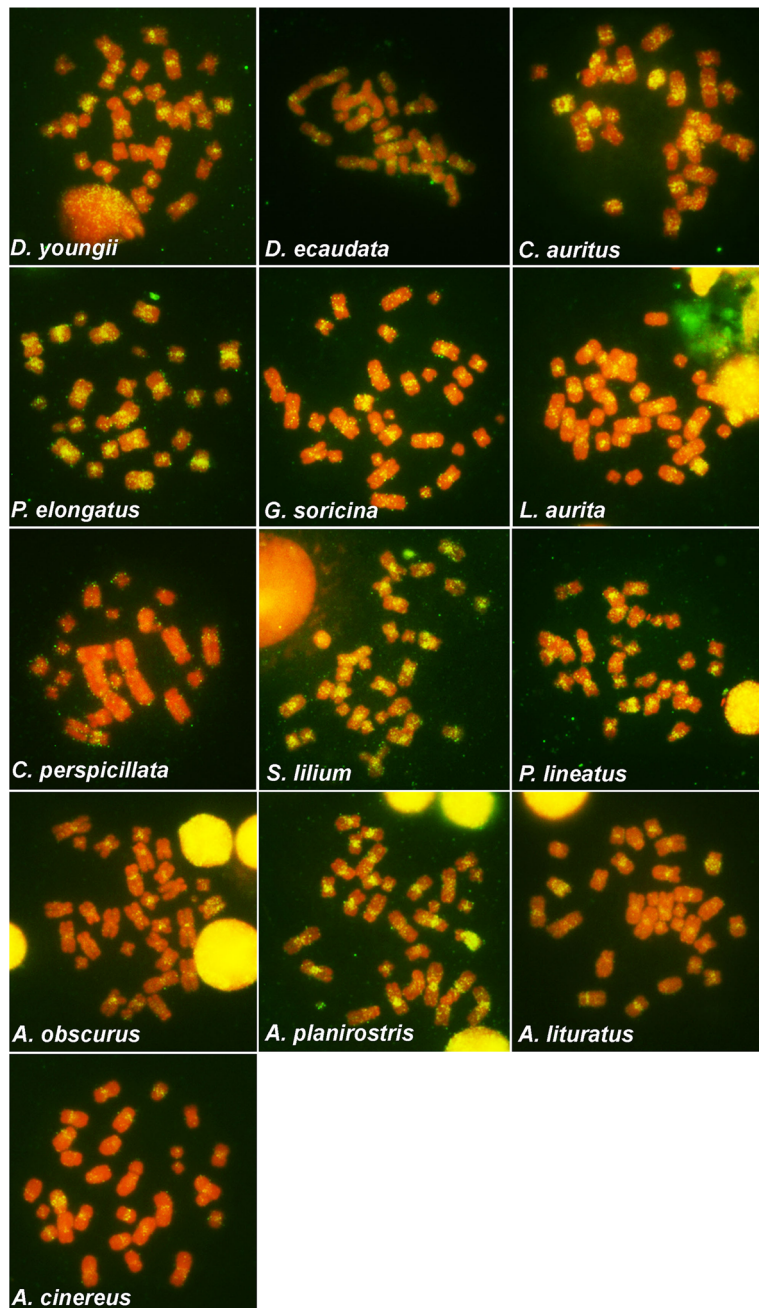
et al. (2001), Santos and Souza (1998b), Sotero-Caio et al. (2011), Varella-Garcia et al. (1989), Volleth et al. (1999). Asterisks to the side of the first autosome correspond to the prevalent location pattern of LINE-1 FISH signals. Note that *C. perspicillata* presented no particular centromeric or euchromatic enrichment. Asterisks to the right side of the X chromosome correspond to a schematic representation of the visualized LINE-1 FISH signals



**Fig. 3** Schematic representation of C-banding and LINE-1 FISH patterns on phyllostomid species from the subfamily Stenodermatinae. Dark regions correspond to constitutive heterochromatin blocks. Schematic representation of chromosomes and bands were compiled from this study and those of Lemos Pinto

et al. (2012), Santos and Souza (1998a), Varella-Garcia et al. (1989). Asterisks to the side of the first autosome correspond to the prevalent location pattern of LINE-1 FISH signals. Asterisks to the right side of the X chromosome correspond to a schematic representation of the visualized LINE-1 FISH signals

**Fig. 4** In situ hybridizations using a fragment of LINE-1 ORF1 as probes on mitotic metaphases of phyllostomid bats



integration of C-banding and FISH did not provide evidence for collocation of distal or interstitial heterochromatin blocks and LINE-1 signal (Figs. 2 and 3). Nevertheless, faint LINE-1 signals on non-centromeric or pericentromeric areas can be observed on autosomes of all species, but without particular enrichment patterns (Fig. 4).

## Discussion

Exclusive centromeric enrichment of TEs is an unusual phenomenon in mammals

Our results have shown that the centromeric regions of phyllostomid species are highly populated by LINE-1



elements. Most of available physical mapping data from mammals indicates that exclusive centromere enrichment of TEs is rare and is one more unusual feature of phyllostomid genomes. Absence of LINE-1 signal at centromeres is often reported, particularly in a number of rodent species, by far the best investigated group by FISH (Acosta et al. 2008; Baker and Wichman 1990; Dobigny et al. 2004; Gauthier et al. 2010; Marchal et al. 2006; Rebuzzini et al. 2009; Vieira-da-Silva et al. 2016), but also in other taxa, including xenarthrans, afrotherians, as well as ungulate interspecific hybrids (Dobigny et al. 2006; Waters et al. 2004). In a few groups, centromeric signals have been observed, but not as the sole location of retroelement enrichment. More specifically, chromosomal arms present non-random accumulation of retroelements, and additional signals are present on the centromeres of few to multiple autosomes. Thus, usually retroelement enrichment follows chromosome- and species-specific patterns in most species (Bonifácio et al. 2012; Liu et al. 2010; Marchal et al. 2006; Vieira-da-Silva et al. 2016; Waters et al. 2004). Contrasting with available examples, the pattern observed for most phyllostomid bats, with exclusive centromeric and/or pericentromeric signals, better resembles the massive retrotransposon expansion found in gibbons and marsupials (Alkan et al. 2011; Bulazel et al. 2006, 2007; Carbone et al. 2012; Ferreri et al. 2011). Carbone and colleagues have shown that a gibbon-specific composite non-autonomous retrotransposon (3'-L1-AluS-VNTR-Alu-like-5' or LAVA) has not only undergone a massive accumulation in centromeres, but is also more associated than expected by chance to cell division-related genes (Carbone et al. 2012, 2014). On the other hand, an endogenous retrovirus (Kangaroo Endogenous Retrovirus; KERV) has been amplified in a lineage-specific manner at macropodid marsupials centromeres (Ferreri et al. 2011).

Two hypotheses best explain centromeric accumulation of retroelements: first, in the library hypothesis (Salser et al. 1976), centromeric satellite evolution is tightly linked to the availability of different classes of repeat in pericentromeric heterochromatin. Core centromere and satellite evolution accompanies the expansion of repeat sequences already available in the pool of heterochromatin, resulting in extremely divergent prevalent centromere sequences even among closely related species. Similarly, LINE-1 expansion on phyllostomid bat centromeres might have derived from expansion of elements previously

inserted in heterochromatin. Independent events of expansion of LINE-1s can explain the different patterns (presence or absence of LINE-1 in the primary constriction) observed in different lineages. Furthermore, *C. perspicillata* was the only species here investigated without co-location of centromeric heterochromatin and LINE-1 probes, suggesting that these sequences were either lost in the process of centromere homogenization, or present sparsely and underrepresented in heterochromatin to be recovered by FISH resolution.

In the second hypothesis, centromere enrichment of retrotransposons might result from an imbalance in the control of epigenetic mechanisms responsible for keeping repetitive DNA expansion in check (O'Neill et al. 1998). This latter hypothesis explains how chromosomal enrichment can result from the increased activity of these elements, but does not necessarily explain why an exclusive centromeric enrichment is observed. Thus, factors other than deregulation of TE amplification must work in concert to shape exclusive centromeric distribution of LINE-1s on phyllostomid bat centromeres. Finally, despite some variability in enrichment of LINE-1 sequences in centromeres, these elements seem to be usually depleted from non-centromeric heterochromatin on mammalian autosomes. This observation is counterintuitive, since deregulation of retroelement activity can potentially lead to enrichment in virtually any chromosome region, which in turn, can promote heterochromatin formation in otherwise euchromatic locations (Acosta et al. 2008; Richards and Elgin 2002).

Further studies are required to explain the centromeric enrichment of LINE-1 ORF1 on phyllostomid bat chromosomes. Hypotheses to be tested include the following: (1) a partial LINE-1 ORF1 sequence was incorporated into the monomeric unit of the centromeric satellite through transposition or recombination; (2) there is preferential insertion of LINE-1s at centromeres; (3) there is rapid removal of LINE-1 from euchromatin as opposed to the centromeric heterochromatin. All these scenarios would require a further step of LINE-1 amplification through mechanisms of centromere homogenization such as gene conversion and concerted evolution, which should also be investigated.

Potential role of centromeric LINE-1 for the chromosomal evolution of phyllostomid bats

Despite the indication that centromeres are the key regions involved in chromosomal reshuffling, little

information is available on how centromere composition and architecture might influence chromosomal variation in phyllostomid bats. Herein, we provided evidence that LINE-1 retrotransposon sequences might constitute important target sites for non-homologous chromosome exchange at centromeres and that (peri)centromeric sequence turnover might be responsible for the high diversity of karyotypes of phyllostomid bats. Regardless of the underlying mechanisms for centromeric enrichment, our data suggest an influence of these retroelements in the evolution of chromosomal architecture of bats and the potential of these sequences as important karyotype modifiers.

Studies on other mammals suggested that the turnover of centromere sequence composition parallels chromosome evolution and might be linked to rapid karyotypic change. Examples include gibbons and macropodid marsupials, which, interestingly, have experienced rapid chromosomal evolution together with centromeric enrichment of retroelements (Bulazel et al. 2007; Wichman et al. 1991). In the case of phyllostomid bats, all analyzed species present highly derived karyotypes from the proposed  $2n = 46$  ancestral condition for the family, and all but *C. perspicillata* show centromere signal using LINE-1 probes (see also Sotero-Caio et al. 2015).

Although Phyllostomidae presents increased rates of chromosomal evolution when compared to other bats, all species investigated here present fixed karyotypes with no chromosomal polymorphisms among populations. Furthermore, their current karyotypes have reached fixation at least five mya (Baker and Bickham 1980; Baker et al. 2012), which means that presently there is no intense karyotypic reshuffling within these species. On the other hand, no extant lineages but the basal subfamily Macrotinae has retained the ancestral,  $2n = 46$  karyotype for the family, suggesting that intense chromosome reshuffling might have been driven by similar mechanisms in a wave-like fashion throughout the family evolution. Our finding that species belonging to different lineages share the centromeric enrichment of LINES indicates that TE expansions and centromeric accumulation might have built a genomic landscape of the common ancestor of all 10 subfamilies that was important in generating the current karyotypic variation patterns. LINE-1 mapping on *Macrotus* species will help testing this hypothesis. Our results also suggest that LINE-1 expansions might no longer be happening and the observed patterns are molecular evidence of past expansions. Therefore, we hypothesize that the

evolution of phyllostomid bats has paralleled defects in epigenetic repression of TEs, which also resulted in centromeric expansion of LINE-1 elements. These events were important in reshaping the karyotypic composition of bats from different subfamilies, but are latent in the present, reflecting the low number of species with chromosomal polymorphisms.

## Conclusions

Phyllostomid bats present uncommon centromere architecture when compared to other mammals, evidenced by a high degree of LINE-1 retroelement accumulation in centromeric and pericentromeric regions. We hypothesize that a combination of reduced TE repression with intense euchromatin removal created a differential TE landscape on the pericentromeric heterochromatin and allowed further centromeric enrichment of LINE-1s through centromere homogenization processes. In addition, our data suggest that LINE-1 accumulation occurred early in the radiation of phyllostomids and might have recurred episodically. Thus, LINE-1 expansions on centromeric heterochromatin probably contributed to the karyotypic variation observed among different phyllostomid lineages. The species investigated herein present stable karyotypes for at least 5 million years, indicating that centromeric retroelements might be in check and maintained by epigenetic centromeric features responsible for chromosomal rearrangement stability.

**Acknowledgements** The authors are grateful to Cirlene Maria da Silva, Izaquiel Santos de Andrade, and Mario Ferreira da Silva (in memoriam) for assistance in the field work. Authors are also grateful for the useful comments by David A. Ray (Texas Tech University). The study was supported by Fundação de Amparo à Pesquisa do Estado de Pernambuco (FACEPE), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## Compliance with ethical standards

Handling and carrying of animals followed guidelines established by the Animal Care and Use guidelines of ICMBio (Instituto Chico Mendes de Conservação da Biodiversidade), and using the collecting permit number (12264-1)

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Acosta MJ, Marchal JA, Fernández-Espartero CH, Bullejos M, Sánchez A (2008) Retroelements (LINEs and SINES) in vole genomes: differential distribution in the constitutive heterochromatin. *Chromosom Res* 16:949–959
- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723
- Alkan C, Cardone MF, Catacchio CR, Antonacci F, O'Brien SJ, Ryder OA, Purgato S, Zoli M, Della Valle G, Eichler EE (2011) Genome-wide characterization of centromeric satellites from multiple mammalian genomes. *Genome Res* 21:137–145
- Araújo SEM (2016) Citogenômica comparativa de morcegos da família Phyllostomidae na Amazônia. Thesis, Universidade Federal do Pará
- Baker RJ, Bickham JW (1980) Karyotypic evolution in bats: evidence of extensive and conservative chromosomal evolution in closely related taxa. *Syst Zool* 29:239–253
- Baker RJ, Qumsiyeh MB (1988) Methods in chiropteran mitotic chromosomal studies. In: Kunz TH (ed) *Ecological and behavioral methods for the study of bats*. Smithsonian Institution Press, Washington, DC, pp 425–435
- Baker RJ, Wichman HA (1990) Retrotransposon *mys* is concentrated on the sex chromosomes: implications for copy number containment. *Evolution* 44:2083–2088
- Baker RJ, Maltbie M, Owen JG, Hamilton MJ, Bradley RD (1992) Reduced number of ribosomal sites in bats: evidence for a mechanism to contain genome size. *J Mammal* 73:847–858
- Baker RJ, Bininda-Emonds ORP, Mantilla-Meluk H, Porter CA, Van Den Bussche RA (2012) Molecular timescale of diversification of feeding strategy and morphology in new world leaf-nosed bats (Phyllostomidae): a phylogenetic perspective. In: Gunnell GF, Simmons NB (eds) *Evolutionary history of bats: fossils, molecules and morphology*. Cambridge Studies in Molecules and Morphology—New Evolutionary Paradigms. Cambridge University Press, Cambridge, pp 385–409
- Barros HMDR, Sotero-Caio CG, Santos N, Souza MJ (2009) Comparative cytogenetic analysis between *Lonchorhina aurita* and *Trachops cirrhosus* (Chiroptera, Phyllostomida). *Genet Mol Biol* 32:748–752
- Bass RA (1978) Systematics of the Desmodonitinae and Phyllonycterinae (Chiroptera: Phyllostomatidae) based on G-band chromosomal homologies. Dissertation, Texas Tech University
- Boissinot S, Furano AV (2001) Adaptive evolution in LINE-1 retrotransposons. *Mol Biol Evol* 18:2186–2194
- Bonifácio HL, da Silva VM, Martin AR, Feldberg E (2012) Molecular cytogenetic characterization of the Amazon River dolphin *Inia geoffrensis*. *Genetica* 140:307–315
- Brown JD, O'Neill RJ (2014) The Evolution of Centromeric DNA Sequences. eLS. doi:<https://doi.org/10.1002/9780470015902.a0020827>.pu
- Bulazel K, Metcalfe C, Ferreri GC, Yu J, Eldridge MD, O'Neill RJ (2006) Cytogenetic and molecular evaluation of centromere-associated DNA sequences from a marsupial (Macropodidae: *Macropus rufogriseus*) X chromosome. *Genetics* 172:1129–1137
- Bulazel KV, Ferreri GC, Eldridge MD, O'Neill RJ (2007) Species-specific shifts in centromere sequence composition are coincident with breakpoint reuse in karyotypically divergent lineages. *Genome Biol* 8:1
- Carbone L, Harris RA, Mootnick AR, Milosavljevic A, Martin DI, Rocchi M, Capozzi O, Archidiacono N, Konkel MK, Walker JA (2012) Centromere remodeling in *Hoolock leuconedys* (Hylobatidae) by a new transposable element unique to the gibbons. *Genome Biol Evol* 4:760–770
- Carbone L, Harris RA, Gnerre S, Veeramah KR, Lorente-Galdos B, Huddleston J, Meyer TJ, Herrero J, Roos C, Aken B (2014) Gibbon genome and the fast karyotype evolution of small apes. *Nature* 513:195–201
- Deininger PL, Batzer MA (2002) Mammalian retroelements. *Genome Res* 12:1455–1465
- Dobigny G, Ozouf-Costaz C, Waters PD, Bonillo C, Coutanceau J-P, Volobouev V (2004) LINE-1 amplification accompanies explosive genome repatterning in rodents. *Chromosom Res* 12:787–793
- Dobigny G, Waters PD, Robinson TJ (2006) Absence of hypomethylation and LINE-1 amplification in a white × black rhinoceros hybrid. *Genetica* 127:81–86
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
- Drummond A, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2009) Geneious v4. 8. 2009. doi: <http://www.geneious.com>
- Edgar RC (2004a) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinf* 5:113
- Edgar RC (2004b) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Eickbush TH, Jamburuthugoda VK (2008) The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res* 134:221–234
- Ferreri GC, Brown JD, Obergefell C, Jue N, Finn CE, O'Neill MJ, O'Neill RJ (2011) Recent amplification of the kangaroo endogenous retrovirus, KERV, limited to the centromere. *J Virol* 85:4761–4771
- Furano AV, Duvernell DD, Boissinot S (2004) L1 (LINE-1) retrotransposon diversity differs dramatically between mammals and fish. *Trends Genet* 20:9–14
- Gauthier P, Hima K, Dobigny G (2010) Robertsonian fusions, pericentromeric repeat organization and evolution: a case study within a highly polymorphic rodent species, *Gerbillus nigeriae*. *Chromosom Res* 18:473–486
- Graham T, Boissinot S (2006) The genomic distribution of L1 elements: the role of insertion bias and natural selection. *Biomed Res Int*. <https://doi.org/10.1155/JBB/2006/75327>
- Gray YH (2000) It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. *Trends Genet* 16:461–468
- Guindon S, Dufayard J, Hordijk W, Lefort V, Gascuel O (2009) PhyML: fast and accurate phylogeny reconstruction by maximum likelihood. *Infect Genet Evol*. Elsevier Science, Amsterdam, pp 384–385
- Hedges D, Deininger P (2007) Inviting instability: transposable elements, double-strand breaks, and the maintenance of genome integrity. *Mutat Res Fund Mol Mech Mut* 616:46–59

- Kapitonov VV, Holmquist GP, Jurka J (1998) L1 repeat is a basic unit of heterochromatin satellites in cetaceans. *Mol Biol Evol* 15:611–612
- Lee M, Elder F (1980) Yeast stimulation of bone marrow mitosis for cytogenetic investigations. *Cytogenet Genome Res* 26:36–40
- Lemos Pinto M, Silva Calixto M, Souza MJ, Araújo APT, Langguth A, Santos N (2012) Cytotaxonomy of the subgenus *Artibeus* (Phyllostomidae, Chiroptera) by characterization of species-specific markers. *Comp Cytogenet* 6:17
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658–1659
- Liu Y, Ye J, Fu B, Ng B, Wang J, Su W, Nie W (2010) Molecular cytogenetic characterization of the genome organization of the 6-banded armadillo (*Euphractus sexinctus*). *Cytogenet Genome Res* 132:31–40
- Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* 72:595–605
- Marchal J, Acosta M, Bullejos M, Puerma E, De La Guardia RD, Sánchez A (2006) Distribution of L1-retrotransposons on the giant sex chromosomes of *Microtus cabreræ* (Arvicolidae, Rodentia): functional and evolutionary implications. *Chromosom Res* 14:177–186
- Martinez P, Jacobina U, Fernandes R, Brito C, Penone C, Amado T, Fonseca C, Bidau C (2017) A comparative study on karyotypic diversification rate in mammals. *Heredity* 118:366–373
- Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, p 1–8
- Morielle-Versute E, Taddei V, Varella-Garcia M (1992) Chromosome banding studies of *Chiropterus Auritus* (Chiroptera: Phyllostomidae). *Rev Bras Gen* 1:569–573
- O'Brien SJ, Menninger JC, Nash WG (2006) Atlas of mammalian chromosomes. John Wiley & Sons, Hoboken
- O'Neill RJW, O'Neill MJ, Graves JAM (1998) Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* 393:68–72
- Ovchinnikov I, Troxel AB, Swergold GD (2001) Genomic characterization of recent human LINE-1 insertions: evidence supporting random insertion. *Genome Res* 11:2050–2058
- Parish D, Vise P, Wichman H, Bull J, Baker R (2002) Distribution of LINEs and other repetitive elements in the karyotype of the bat *Carollia*: implications for X-chromosome inactivation. *Cytogenet Genome Res* 96:191–197
- Pieczarka JC, Nagamachi CY, O'Brien PC, Yang F, Rens W, Barros RMS, Noronha RC, Rissino J, de Oliveira EH, Ferguson-Smith MA (2005) Reciprocal chromosome painting between two south American bats: *Carollia brevicauda* and *Phyllostomus hastatus* (Phyllostomidae, Chiroptera). *Chromosom Res* 13:339–347
- Pieczarka JC, Gomes AJ, Nagamachi CY, Rocha DC, Rissino JD, O'Brien PC, Yang F, Ferguson-Smith MA (2013) A phylogenetic analysis using multidirectional chromosome painting of three species (*Uroderma magnirostrum*, *U. bilobatum* and *Artibeus obscurus*) of subfamily Stenodermatinae (Chiroptera-Phyllostomidae). *Chromosom Res* 21:383–392
- Pinkel D, Straume T, Gray J (1986) Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *PNAS* 83:2934–2938
- Posada D (2008) jModelTest: phylogenetic model averaging. *Mol Biol Evol* 25:1253–1256
- Rebuzzini P, Castiglia R, Nergadze SG, Mitsainas G, Munclinger P, Zuccotti M, Capanna E, Redi CA, Garagna S (2009) Quantitative variation of LINE-1 sequences in five species and three subspecies of the subgenus *Mus* and in five Robertsonian races of *Mus musculus domesticus*. *Chromosom Res* 17:65–76
- Ribas TF, Rodrigues LR, Nagamachi CY, Gomes AJ, Benathar TC, Yang F, Ferguson-Smith MA, Pieczarka JC (2013) Two new cytotypes reinforce that *Micronycteris hirsuta* Peters, 1869 does not represent a monotypic taxon. *BMC Genet* 14:119
- Ribas T, Rodrigues L, Nagamachi C, Gomes A, Rissino J, O'Brien P, Yang F, Ferguson-Smith M, Pieczarka J (2015) Phylogenetic reconstruction by cross-species chromosome painting and G-banding in four species of Phyllostomini tribe (Chiroptera, Phyllostomidae) in the Brazilian Amazon: an independent evidence for monophyly. *PLoS ONE* 10:e0122845
- Richards EJ, Elgin SC (2002) Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108:489–500
- Richardson SR, Doucet AJ, Kopera HC, Moldovan JB, Garcia-Perez JL, Moran JV (2015) The influence of LINE-1 and SINE retrotransposons on mammalian genomes. *Microbiol spectr* 3:1165–1208
- Salser W, Bowen S, Browne D, El-Adli F, Fedoroff N, Fry K, Heindell H, Paddock G, Poon R, Wallace B (1976) Investigation of the organization of mammalian chromosomes at the DNA sequence level. *Fed Proc* 35:23–35
- Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. ColdSpring-Harbour Laboratory Press, New York
- Santos N, Souza MJ (1998a) Use of fluorochromes chromomycin A3 and DAPI to study constitutive heterochromatin and NORs in four species of bats (Phyllostomidae). *Caryologia* 51:265–278
- Santos N, Souza MJ (1998b) Characterization of the constitutive heterochromatin of *Carollia perspicillata* (Phyllostomidae, Chiroptera) using the base-specific fluorochromes, CMA3 (GC) and DAPI (AT). *Caryologia* 51:51–60
- Santos N, Fagundes V, Yonenaga-Yassuda Y, Souza MJ (2001) Comparative karyology of Brazilian vampire bats *Desmodus rotundus* and *Diphylla ecaudata* (Phyllostomidae, Chiroptera): banding patterns, base-specific fluorochromes and FISH of ribosomal genes. *Hereditas* 134:189–194
- Scott L, Kuroiwa A, Matsuda Y, Wichman H (2006) X accumulation of LINE-1 retrotransposons in *Tokudaia osimensis*, a spiny rat with the karyotype XO. *Cytogenet Genome Res* 112:261–269
- Shi J, Wolf SE, Burke JM, Presting GG, Ross-Ibarra J, Dawe RK (2010) Widespread gene conversion in centromere cores. *PLoS Biol* 8:e1000327
- Smit A, Hubley R, Green P (1996–2010) RepeatMasker Open-3.0

- Smith JD, Bickham JW, Gregory TR (2013) Patterns of genome size diversity in bats (order Chiroptera). *Genome* 56:457–472
- Sotero-Caio CG, Pieczarka JC, Nagamachi CY, Gomes AJB, Lira TC, O'Brien PCM, Ferguson-Smith MA, Souza MJ, Santos N (2011) Chromosomal homologies among vampire bats revealed by chromosome painting (Phyllostomidae, Chiroptera). *Cytogenet Genome Res* 132:156–164
- Sotero-Caio CG, Volleth M, Gollahon LS, Fu B, Cheng W, Ng BL, Yang F, Baker RJ (2013) Chromosomal evolution among leaf-nosed nectarivorous bats-evidence from cross-species chromosome painting (Phyllostomidae, Chiroptera). *BMC Evol Biol* 13:276
- Sotero-Caio CG, Volleth M, Hoffmann FG, Scott L, Wichman HA, Yang F, Baker RJ (2015) Integration of molecular cytogenetics, dated molecular phylogeny, and model-based predictions to understand the extreme chromosome reorganization in the Neotropical genus *Tonatia* (Chiroptera: Phyllostomidae). *BMC Evol Biol* 15:1
- Steiner FA, Henikoff S (2015) Diversity in the organization of centromeric chromatin. *Curr Opin Genet Dev* 31:28–35
- Sumner A (1972) A simple technique for demonstrating centromeric heterochromatin. *Exp Cell Res* 75:304–306
- Teixeira W, Ferreira I, Cabral-de-Mello DC, Mazzuchelli J, Valente G, Pinhal D, Poletto A, Venere P, Martins C (2009) Organization of repeated DNA elements in the genome of the cichlid fish *Cichla kelberi* and its contributions to the knowledge of fish genomes. *Cytogenet Genome Res* 125:224–234
- Varella-Garcia M, Versute EM, Taddei VA (1989) A survey of cytogenetic data on Brazilian bats. *Rev Bras Genét* 12:761–793
- Vieira-da-Silva A, Adegá F, Guedes-Pinto H, Chaves R (2016) LINE-1 distribution in six rodent genomes follow a species-specific pattern. *J Genet* 95:21–33
- Volleth M, Klett C, Kollak A, Dixkens C, Winter Y, Just W, Vogel W, Hameister H (1999) ZOO-FISH analysis in a species of the order Chiroptera: *Glossophaga soricina* (Phyllostomidae). *Chromosom Res* 7:57–64
- Waters PD, Dobigny G, Pardini AT, Robinson TJ (2004) LINE-1 distribution in Afrotheria and Xenarthra: implications for understanding the evolution of LINE-1 in eutherian genomes. *Chromosoma* 113:137–144
- Wichman HA, Payne C, Ryder O, Hamilton M, Maltbie M, Baker R (1991) Genomic distribution of heterochromatic sequences in equids: implications to rapid chromosomal evolution. *J Hered* 82:369–377
- Wichman HA, Van Den Bussche RA, Hamilton MJ, Baker RJ (1992) Transposable elements and the evolution of genome organization in mammals. *Genetica* 86:287–293
- Yang L, Brunsfeld J, Scott L, Wichman H (2014) Reviving the dead: history and reactivation of an extinct L1. *PLoS Genet* 10:e1004395
- Zhang Y, Mager DL (2012) Gene properties and chromatin state influence the accumulation of transposable elements in genes. *PLoS One* 7:e30158