



First Description of a Satellite DNA in Manatees' Centromeric Regions

Mirela Pelizaro Valeri¹, Guilherme Borges Dias², Alice Alves do Espírito Santo¹, Camila Nascimento Moreira³, Yatiyo Yonenaga-Yassuda³, Iara Braga Sommer⁴, Gustavo C. S. Kuhn¹ and Marta Svartman^{1*}

¹Laboratório de Citogenômica Evolutiva, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, ²Department of Genetics and Institute of Bioinformatics, University of Georgia, Athens, GA, United States, ³Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil, ⁴Centro Nacional de Pesquisa e Conservação da Biodiversidade Marinha do Nordeste, Instituto Chico Mendes de Conservação da Biodiversidade, Brasília, Brazil

Trichechus manatus and *Trichechus inunguis* are the two Sirenia species that occur in the Americas. Despite their increasing extinction risk, many aspects of their biology remain understudied, including the repetitive DNA fraction of their genomes. Here we used the sequenced genome of *T. manatus* and TAREAN to identify satellite DNAs (satDNAs) in this species. We report the first description of TMA_{sat}, a satDNA comprising ~0.87% of the genome, with ~684 bp monomers and centromeric localization. In *T. inunguis*, TMA_{sat} showed similar monomer length, chromosome localization and conserved CENP-B box-like motifs as in *T. manatus*. We also detected this satDNA in the *Dugong dugon* and in the now extinct *Hydrodamalis gigas* genomes. The neighbor-joining tree shows that TMA_{sat} sequences from *T. manatus*, *T. inunguis*, *D. dugon*, and *H. gigas* lack species-specific clusters, which disagrees with the predictions of concerted evolution. We detected a divergent TMA_{sat}-like homologous sequence in elephants and hyraxes, but not in other mammals, suggesting this sequence was already present in the common ancestor of Paenungulata, and later became a satDNA in the Sirenians. This is the first description of a centromeric satDNA in manatees and will facilitate the inclusion of Sirenia in future studies of centromeres and satDNA biology.

Keywords: tandem repeats, *Trichechus manatus*, *Trichechus inunguis*, chromosome mapping, fluorescent *in situ* hybridization, TAREAN

INTRODUCTION

The order Sirenia encompasses four extant herbivorous aquatic mammals. The Dugongidae family includes the *Dugong dugon* and the Steller's sea cow *Hydrodamalis gigas*, the latter now extinct due to overhunting, and the Trichechidae family includes three manatee species: *Trichechus manatus*, *Trichechus inunguis*, and *Trichechus senegalensis* (Domning, 2018). *Dugong dugon* occurs across coastal waters in the Indo-West Pacific Ocean, and *T. senegalensis* is restricted to the west coast of Africa, making *T. manatus*, the West Indian manatee, and *T. inunguis*, the Amazonian manatee, the only sirenians to occur in the Americas. The West Indian manatee occurs in Caribbean waters and the Atlantic coast ranging from Florida to the northeast of Brazil, and *T. inunguis* is found along the Amazon River basin (Bonvicino et al., 2020).

OPEN ACCESS

Edited by:

Francisco J. Ruiz-Ruano,
University of East Anglia,
United Kingdom

Reviewed by:

Nevenka Mestrovic,
Rudjer Boskovic Institute, Croatia
Isidoro Feliciello,
University of Naples Federico II, Italy

*Correspondence:

Marta Svartman
svartmanm@ufmg.br

Specialty section:

This article was submitted to
Evolutionary and Population
Genetics,
a section of the journal
Frontiers in Genetics

Received: 13 April 2021

Accepted: 30 July 2021

Published: 24 August 2021

Citation:

Valeri MP, Dias GB, Espírito Santo AA, Moreira CN, Yonenaga-Yassuda Y, Sommer IB, Kuhn GCS and Svartman M (2021) First Description of a Satellite DNA in Manatees' Centromeric Regions. *Front. Genet.* 12:694866. doi: 10.3389/fgene.2021.694866

All extant sirenians are considered as vulnerable by the International Union for Conservation of Nature and Natural Resources (IUCN; Deutsch et al., 2008; Keith, 2015; Marmontel et al., 2016; Marsh and Sobotzick, 2019).

The West Indian manatee has two recognized subspecies: *Trichechus manatus latirostris* (Florida Manatee), found in the United States and Gulf of Mexico coasts, and *T. m. manatus* (Antillean manatee), found in the Caribbean, Central and South America. Recent morphological and genetic analyses suggest the need for a revision in the *T. manatus* taxonomy considering the influence of the Amazon River as a barrier to gene flow. These studies showed that the *T. m. manatus* populations from the Caribbean and up to the Amazon River mouth are phylogenetically closer to the populations of *T. m. latirostris* from the United States than to the Brazilian *T. m. manatus* populations south of the Amazon River mouth (Vianna et al., 2006; Barros et al., 2017; Lima et al., 2019, 2021). Hybrids between *T. manatus* and *T. inunguis* have also been reported on the sympatric area at the Amazon River mouth (Vianna et al., 2006; Lima et al., 2019; Luna et al., 2021).

Satellite DNAs (satDNAs) are a type of repetitive DNA found in most eukaryotic genomes. They are arranged as long arrays of tandem repeats with variable unit length, number of copies and chromosome organization. SatDNAs are usually associated with chromosome landmarks such as centromeres, telomeres, and heterochromatic regions. Despite the fact that satDNAs do not encode proteins, they are associated with important biological functions such as formation and maintenance of heterochromatin at telomeres and centromeres, and maintenance of chromosome integrity and genome stability (reviewed in Shapiro and von Sternberg, 2005; Biscotti et al., 2015; Shatskikh et al., 2020). SatDNAs can form higher-order repeat (HOR) units made of multimers with a number of diverged monomers that are tandemly repeated as a set (reviewed in Plohl et al., 2012; Vlahović et al., 2016). HOR organization has been found in several satDNAs, including the alpha centromeric satDNA in humans, and may be relevant to the centromeric function (Sujiwattananat et al., 2015; Sullivan et al., 2017). In addition, satDNAs monomer sequences can present internal repetitions, which may be related with secondary structures relevant to centromeric function (Kasinathan and Henikoff, 2018). Centromeric satDNAs in mammals usually present the CENP-B box, a conserved 17bp region known to be the DNA-binding domain for the centromeric protein B (CENPB), with nine nucleotides (nTTCGnnnnAnnCGGn) composing the most evolutionarily conserved domain (ECD; Muro et al., 1992; Masumoto et al., 2004; Alkan et al., 2011; Kasinathan and Henikoff, 2018). Most satDNAs are under concerted evolution, a process by which new mutations within monomers are quickly homogenized across the repeat family and fixed in reproductively isolated populations, resulting in intraspecific repeat homogeneity but interspecific divergence (Dover, 1982; Plohl et al., 2012; Smalec et al., 2019). Moreover, according to the library model, related species may share a collection of satDNAs sequences with mostly quantitative interspecies differences due to expansion or contraction (even elimination) during the evolution (Fry and Salser, 1977; Meštrović et al., 1998).

Another aspect of satDNAs evolution is their relationship with mobile elements, since there are several examples of satDNAs derived from transposons and retrotransposons in plants and animals (reviewed in Meštrović et al., 2015).

The repetitive DNA fraction of manatees' genomes has been poorly studied, especially in the case of satDNAs. We used the sequenced genome of *T. manatus* and the TAREAN (Novák et al., 2017, 2020) pipeline to explore the satDNAs present in this genome. Herein, we describe for the first time the centromeric satDNA of the West Indian manatee, which we found to be also present in the Amazonian manatee, the dugong, and in the extinct Steller's sea cow. We characterized this sequence *in silico* and mapped it in *T. manatus* and *T. inunguis* chromosomes. In addition, we investigated the presence of the TMA sat sequence in mammals outside the order Sirenia, which allowed us to establish a rough timeline for its origin.

MATERIALS AND METHODS

De novo Identification of Satellite DNAs

In order to identify satDNAs in manatees, we used whole-genome sequencing data from *T. m. latirostris* (accession number SRR328416) available in the *National Center for Biotechnology Information* – NCBI and the TAREAN pipeline (Novák et al., 2017). The first step of this pipeline is a graph-based clustering, which performs all to all similarity comparisons of DNA sequencing reads, resulting in clusters of those reads derived from repetitive elements. Then, it examines the presence of circular or globula-like graph structures to identify potential tandem repeats, classified as putative high or low confidence satellites. The raw Illumina reads (~100bp long) used in this analysis were randomly sampled by TAREAN, comprising ~2.4% (870,965 reads) of the 3.67pg estimated genome size (Kasai et al., 2013). The reads that make up each cluster are partially assembled into contigs that were used for repeat annotation with the CENSOR web server (Kohany et al., 2006) that contains a collection of Mammalia repeats from RepBase, updated in 08-24-2020 (Bao et al., 2015). The single potential tandem repeat cluster (13) with globula/ring-like structure was analyzed in detail through similarity searches against the *T. manatus* reference genome (accession GCA_000243295.1) using the BLASTn tool with default parameters (Altschul et al., 1990) to verify if the sequence is a tandem repeat. In addition to the annotation using the CENSOR web server, this cluster was annotated through BLASTn similarity searches against the whole nucleotide collection (nr/nt).

The identified satDNA sequence was characterized regarding its genome proportion, monomer length, AT content, and presence of internal direct or inverted duplications. The satDNA genome proportion was estimated by TAREAN. TAREAN tries to improve the assembly process by applying a k-mer-based approach to obtain a less fragmented monomer consensus, but it restricts itself to the 50% most prevalent k-mers in a cluster. For this reason, we chose the whole-genome assembly resource as a more representative sample of TMA sat diversity. The most common sequence (MCS) of TMA sat was generated using Geneious (prime version 2020.2.4) with a 25% threshold and 66 monomeric sequences retrieved from the reference genome, previously aligned with the muscle aligner

implemented in MEGA X. The MCS was used to estimate monomer length, AT content, and presence of internal repetitions. The last feature was also conducted in the Geneious software using the diagonal plot method in high sensitive mode, with window size of 50bp and identity threshold of 60%.

We searched for the presence of TMA_{sat} in the two other Sirenia species with a sequenced genome available in NCBI, *D. dugon* (under accession numbers of assembled genome GCA_015147995.1 and raw Illumina reads DRR251525) and the extinct *H. gigas* (under accession numbers of assembled genome GCA_013391785.1 and raw Illumina reads SRR12067498). First, we used TMA_{sat} sequence as query in BLASTn similarity searches against these assembled genomes. In addition, we also used the raw Illumina reads (~150bp long) and TAREAN to identify TMA_{sat} in these genomes. The analyzed reads were randomly sampled by TAREAN totalizing 1,038,927 in *D. dugon* and 570,097 in *H. gigas*. The MCS of TMA_{sat} in *D. dugon* and *H. gigas* were generated using monomeric sequences retrieved from the reference genome after BLAST searches, totalizing 50 sequences from *D. dugon* and 40 from *H. gigas*. The TMA_{sat} MCS in *T. inunguis* was obtained using the five cloned sequences obtained by PCR. All MCS were generated as described previously for *T. manatus*.

Biological Samples

We used biological samples of *T. manatus* and *T. inunguis* to determine TMA_{sat} chromosomal distribution and investigate its presence in *T. inunguis*, whose genome has not been sequenced. Skin sample from a male *T. manatus* captured at Porto de Pedras/AL, Brazil (−9.164167 and −35.294444) in 2019 was provided by CEPENE/ICMBio (SISBIO 60829-2) and used for fibroblast culture. A fibroblast cell line from a male *T. inunguis* established in 1998 was provided by Dr. Yatiyo Yonenaga-Yassuda, from the University of São Paulo, and was previously analyzed by Assis et al. (1998). Chromosome spreads from fibroblast cultures were obtained according to Stanyon and Galleni (1991) and genomic DNAs were extracted with the Wizard Genomic DNA Purification Kit (Promega).

PCR Amplification, Cloning, and Sequencing of Satellite DNAs

TMA_{sat} was amplified by PCR from the *T. inunguis* genomic DNA using primers designed from the satDNA consensus sequence (estimated by TAREAN) as follow: TMA_{sat}-F CTCCTTCAAGCTGCTTAACT and TMA_{sat}-R GGGAACCTTACTTGTCTGCT. The PCR cycling conditions were as follows: 95°C – 3 min, 35 cycles: 95°C – 30 s; 55°C for 30 s; 72°C – 1 min; and 72°C – 3 min for final elongation. The PCR product corresponding to monomer size was excised from the agarose gel and purified using the Illustra GFX PCR DNA and Gel Band Purification Kit. The purified products were ligated into the pGEM-T Easy vector (Promega) and used in the transformation of *Escherichia coli* XL1-BLUE strain electrocompetent cells (Phonetrutria). Five recombinant colonies of TMA_{sat} were sequenced (access numbers MW272776–MW272780) with the ABI3130 platform (Applied Biosystems).

Fluorescence *in situ* Hybridization

Fluorescent *in situ* hybridization (FISH) was performed using the TMA_{sat} cloned (MW272776) sequence as probe on metaphase spreads of *T. manatus* and *T. inunguis*. FISH was performed with 200 ng of biotin-labelled probes, following (Valeri et al., 2020). The analyses and image acquisition were performed under a Zeiss Axioimager 2 epifluorescence microscope equipped with a CCD camera and with the AxioVision software (Carl Zeiss MicroImaging, Jena, Germany), respectively.

In silico Characterization of satDNAs

DNA polymorphisms and nucleotide diversity along the satDNA sequences were analyzed using the software DnaSP 6.12.03 (Rozas et al., 2017) with the same monomer sequences used to generate the MCS from *T. manatus*, *D. dugon*, and *H. gigas*. In this analysis, the monomer sequences were previously aligned with the muscle method (Edgar, 2004) implemented in MEGA X and the window length and step size were set for 10 and 1bp, respectively. Windows were classified as conserved or variable if they exhibited more than two SDs below or above the nucleotide average variability, respectively.

Monomer sequences of TMA_{sat} from *T. manatus*, *T. inunguis*, *D. dugon*, and *H. gigas* were aligned with the muscle method implemented in MEGAX and used for the construction of a neighbor-joining tree. These sequences were the same used to obtain the MCS, totalizing 161 sequences, including 66 from *T. manatus*, five from *T. inunguis*, 50 from *D. dugon* and 40 from *H. gigas*. The neighbor-joining tree was obtained using MEGA X with 500 bootstrap replicates and the final tree was visualized in iTOL v4.3.3¹ (Letunic and Bork, 2019). We also used the same set of sequences to estimate the inter- and intra-specific nucleotide divergence (number of base substitutions per site), as well as the average nucleotide divergence over all pairwise sequence comparisons using MEGA X.

We searched for any putative CENP-B box in TMA_{sat} MCS from *T. manatus*, *T. inunguis*, *D. dugon*, and *H. gigas* using the 17bp sequence containing the ECD (nTTCGnnnnAnnCGGGn; Masumoto et al., 2004) and CENP-B box sequences of *Loxodonta africana*/*Dasyypus novemcinctus* (CTTTGCCGAGAACGGAG; Alkan et al., 2011). This search was conducted in the Geneious software in global pairwise alignment mode and 51% similarity cost matrix.

To investigate the presence of TMA_{sat} in other mammals, we utilized the MCS from *T. manatus* as query in BLASTn similarity searches against Mammalia (NCBI:txid40674) wgs database excluding Sirenia (NCBI:txid9774; search date 06-07-2021). The flanking regions of TMA_{sat} similarity hits were analyzed with the CENSOR web server (Kohany et al., 2006) containing the Mammalia RepBase library (updated in 06-14-2021; Bao et al., 2015). To better analyze these hits with the TMA_{sat} consensus sequence, we compared their sequences using dotplots and pairwise alignments in the Geneious software.

¹<https://itol.embl.de/>

RESULTS

In silico Identification and satDNA Analyses

The only potential satDNA identified (with low confidence) by TAREAN in the *T. manatus* genome was represented by the cluster 13. This sequence was analyzed in detail through

similarity searches against the *T. manatus* reference genome (accession GCA_000243295.1) using the BLASTn tool with default parameters (Altschul et al., 1990). Despite being classified by TAREAN with low confidence, we verified this sequence tandemly repeated at least 25 times on assembled contigs of *T. manatus*. These repeats comprise 0.87% of the genome of *T. manatus* with monomer length of ~684bp estimated by

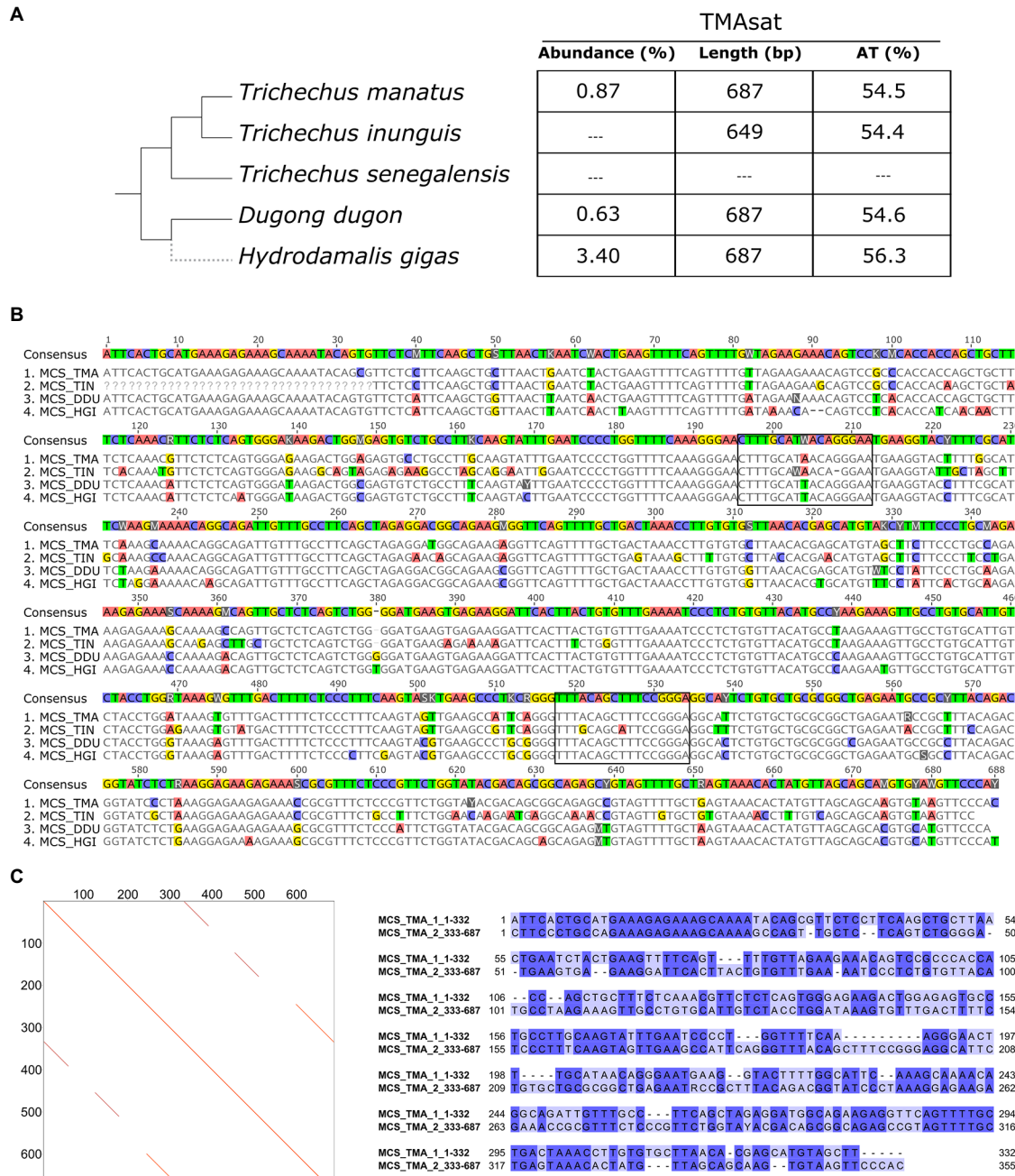


FIGURE 1 | (A) TMA sat abundance, monomer length and AT content of the monomers for the analyzed species. **(B)** Alignment of TMA sat MCS from *Trichechus manatus*, *Trichechus inunguis*, *Dugong dugon*, and *Hydrodamalis gigas* showing the two putative CENP-B box like motifs. **(C)** Dot plot comparison of the TMA sat MCS sequence from *T. manatus* against itself and pairwise alignment of TMA sat position 1–332 against 333–687, with 55.8% of DNA sequence identity.

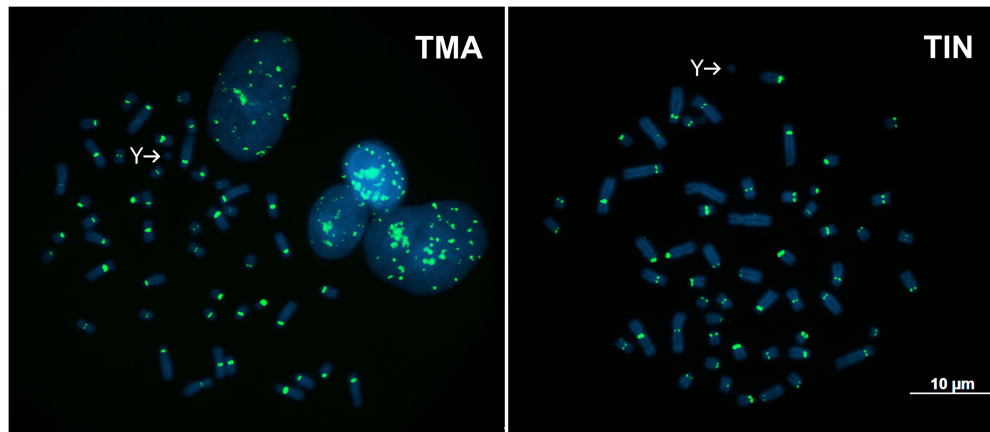


FIGURE 2 | Metaphases of *T. manatus* (TMA) and *T. inunguis* (TIN) after FISH with the TMA sat probe. Y chromosomes without signals of TMA sat are indicated.

TAREAN. The consensus sequence generated by TAREAN (Supplementary Figure 1) did not show similarity with any known repetitive DNA from the mammalian RepBase collection (Bao et al., 2015). We named this new satDNA as TMA sat (for *T. manatus* satellite).

The TMA sat MCS from *T. manatus* was generated from an alignment of 66 monomers manually isolated from the assembled reference genome (Supplementary Figure 2; Supplementary File 1). It showed monomer length of 687 bp and 54.5% of AT content (Figures 1A,B). The dotplot of TMA sat against itself revealed a segment repeated twice inside TMA sat, from position 1 to 332 and 333 to 687 (Figure 1C). A pairwise alignment of the two segments of TMA sat, 1–332 and 333–687 bp, showed that they are related but quite divergent, with only 55.8% identity (Figure 1C). A detailed investigation in the assembled contigs showed that the TMA sat unit of ~687 bp is organized in higher-order structure, mostly alternating the segments TMA sat1 (1–332) and TMA sat2 (333–687). However, we found one case of TMA sat1 dimer (accession NW_004443969.1 position 56,989–75,023 bp), few cases of TMA sat2 dimer (accessions NW_004444053.1; NW_004444936.1; and NW_004444425.1) and in one contig (accession NW_004443969.1) three, six and 10 tandemly repeated units of TMA sat2.

Genomic Distribution of TMA sat in the Genus *Trichechus*

TMA sat was amplified by PCR from *T. inunguis* genomic DNA, and the resulting PCR products showed a similar monomer length of ~647 bp (Supplementary Figure 3). The PCR product was cloned and sequenced in order to confirm that it was indeed homologous to TMA sat. The MCS based on the cloned sequences showed similar AT content and 89.6% of identity (Figures 1A,B; Supplementary Figure 4). A selected TMA sat cloned sequence was labeled and used as probe in FISH on chromosomes of both *T. manatus* and *T. inunguis*. TMA sat showed centromeric localization in *T. manatus* ($2n=48$) and *T. inunguis* ($2n=56$), mapping to the centromeres of all chromosomes, except the Y (Figure 2; Supplementary Figure 5).

TMA sat localization is compatible with the CBG-banding pattern in both species, which reveals centromeric heterochromatin in all chromosomes (Assis et al., 1998; Gray et al., 2002), with the exception of the Y.

TMA sat in Other Sirenia

Besides *T. manatus*, there are two additional Sirenia species with sequenced genomes available: *D. dugon* and the extinct *H. gigas*, both belonging to the Dugongidae family. A search for TMA sat sequences on the assembled contigs of these species revealed the presence of tandemly repeated TMA sat sequences. TAREAN returned with high confidence one cluster of a putative satDNA with 685 bp length in both species, cluster 8 in *D. dugon* and cluster 3 in *H. gigas*, which contained homologous sequences to TMA sat (Figures 1A,B). In *D. dugon*, cluster 8 represents 0.63% of the genome and the MCS generated from the 50 monomers retrieved from the assembled genome is 687 bp long with 54.6% of AT content (Supplementary Figure 6; Supplementary File 2). In addition, we found evidence of other HOR configurations rather than alternating TMA sat1 and TMA sat2 in *D. dugon*: a dimer of TMA sat1 (BMBL01107524.1 and BMBL01079760.1), a dimer of TMA sat2 (BMBL01112453.1 and BMBL01093845.1), four (BMBL01013125.1), five (BMBL01107524.1) and six (BMBL01055248.1) tandemly repeated copies of TMA sat2.

In *H. gigas*, cluster 3 comprises 3.4% of the genome and the MCS based on 40 monomers from the reference genome is 687 bp long and has 56.3% of AT content (Supplementary Figure 7; Supplementary File 3). In this species, the most frequent TMA sat organization is the alternating segments of TMA sat1 and TMA sat2, and we only found one dimer of TMA sat2 (JACANZ010402190.1).

The sliding window analysis of nucleotide variability of this satDNA in *T. manatus*, *D. dugon*, and *H. gigas* revealed the presence of conserved and variable regions within the monomers (Figures 3A–C). However, we did not have access to biological samples of *D. dugon* or *H. gigas* to map TMA sat on their chromosomes. The monomeric TMA sat sequences from *T. manatus*, *T. inunguis*, *D. dugon*, and *H. gigas* were aligned and

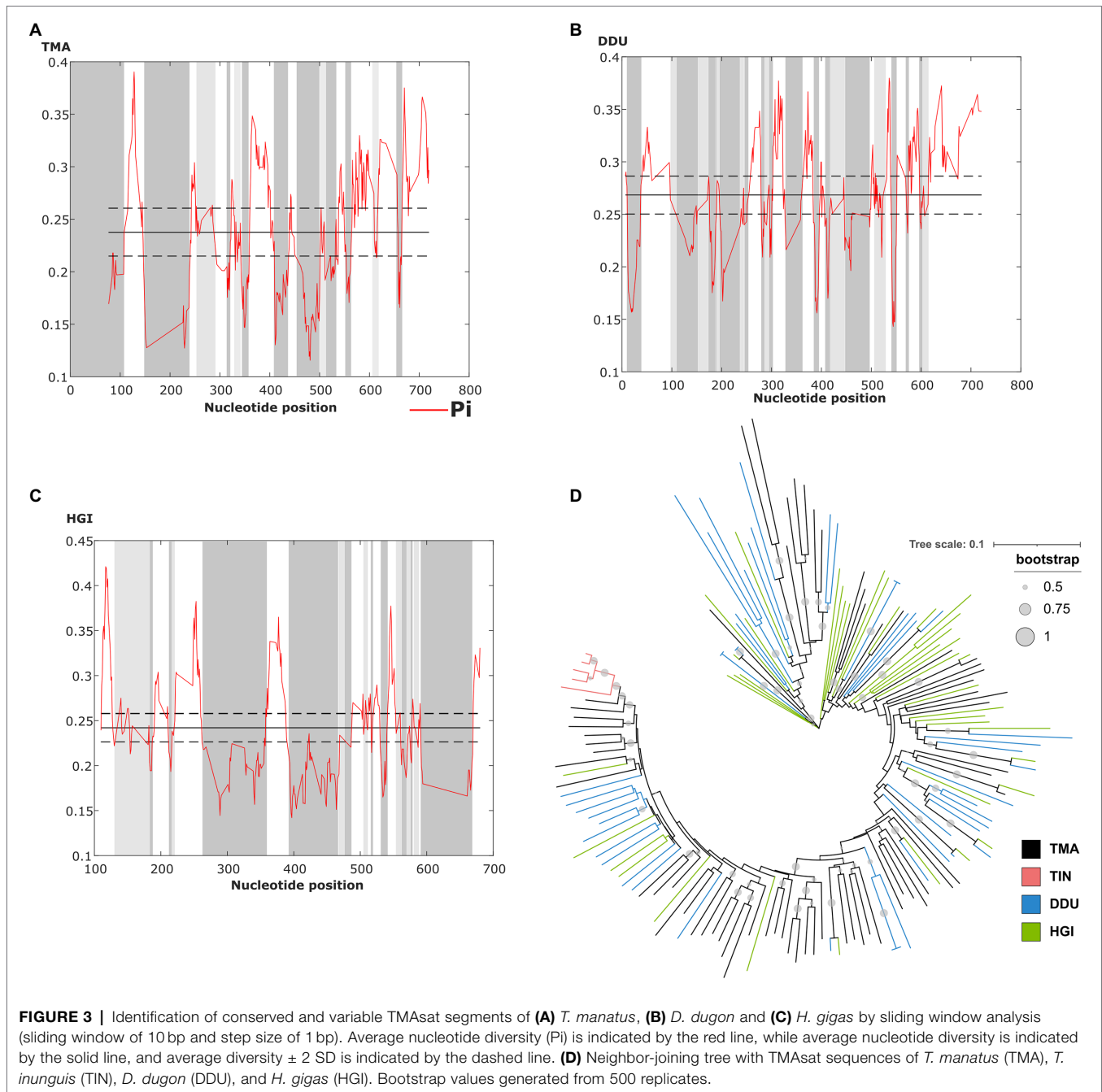


TABLE 1 | The two putative CENPB box-like motifs identified in the MCS of TMA from *T. manatus* (TMA), *T. inunguis* (TIN), *D. dugon* (DDU), and *H. gigas* (HGI).

	Position 196–212 bp	Position 518–534 bp
ECD	NTTCGNNNNANNCGGGN	NTTCGNNNNANNCGGGN
<i>L. africana</i> and <i>D. novemcinctus</i>	CTTTGCCGAGAACGGAG	CTTTGCCGAGAACGGAG
<i>T. manatus</i>	CTTTGCATAACAGGGAA	TTTACAGCTTTCCGGGA
<i>T. inunguis</i>	CTTTGCAWAACAGGAAT	TTTGCAGCATTCCGGGA
<i>D. dugon</i>	CTTTGCATTACAGGGAA	TTTACAGCTTTCCGGGA
<i>H. gigas</i>	CTTTGCATTACAGGGAA	TTTACAGCTTTCCGGGA

Conserved nucleotides in the evolutionarily conserved domain (ECD) are shown in red/highlighted, and conserved nucleotides compared with *L. africana* and *D. novemcinctus* motif (Alkan et al., 2007) other than the ECD domain are shown in blue.

used to construct a neighbor-joining tree, which did not reveal any species-specific clustering (Figure 3D). We also estimated the inter and intraspecific nucleotide divergence (Supplementary Table 1), as well as the average divergence over all sequence pairs ($d=0.34$). As expected from the Neighbor Joining results, TMA sat intraspecific diversity was not lower than interspecific diversity, except in *T. inunguis* ($d=0.06$). The low diversity of TMA sat sequences in *T. inunguis* may be due to the low number of sequences used in the analysis and the use of PCR.

CENP-B Box Is Present in TMA sat

The CENP-B box is a 17 bp region conserved among mammalian centromeric satDNAs and known to be the DNA-binding domain for the centromeric protein CENPB. We searched for putative CENP-B box-like motifs within TMA sat MCS from *T. manatus*, *T. inunguis*, *D. dugon*, and *H. gigas*, and found two putative motifs MCS (Figure 1B). The first is located in position 196 to 212 bp, matching best with the CENP-B box sequence found in *L. africana* and *D. novemcinctus*. The second putative motif was found in position 518–534 bp. Both motifs display 5–6 identical nucleotides to ECD out of nine in all Sirenia species (Table 1). The two putative CENP-B box motifs were present in a conserved segment as indicated by the sliding window analysis of nucleotide variability among satDNA monomers from *T. manatus* and *D. dugon* (Figures 3A,B). In *H. gigas*, these motifs spanned both conserved and variable regions of the TMA sat monomer (Figure 3C).

The CENP-B box-like motifs found in positions 196–212 bp of TMA sat from *T. manatus* (CTTTGCATAACAGGGAA) and *T. inunguis* (CTTTGCAWAACA-GGAAT) shared 14 out of the 17 nucleotides with each other. In *D. dugon* and in *H. gigas* the CENPB-box-like motif was the same (CTTTGCATTACAGGGAA) and shared 15 out of 17 nucleotides with *T. manatus*. Six bases in *T. manatus*, *D. dugon* and *H. gigas* and five in *T. inunguis*, out of the nine from the ECD were conserved. The second putative motif (position 518–534 bp) showed six out of nine identical bases to the ECD in the four analyzed species. *T. manatus*, *D. dugon*, and *H. gigas* shared an identical second motif (TTTACAGCTTTCCGGGA), whereas *T. inunguis* differed in two nucleotides (TTTGCAGCATTCCGGGA).

TMA sat in Other Mammals

We investigated the presence of TMA sat in other mammals using the MCS from *T. manatus* as query in similarity searches against Mammalia sequences in the wgs database from NCBI excluding Sirenia. The total number of returned hits was 13 distributed in four species (Supplementary Table 2). With a cut off for query cover equal or greater than 30%, we found four hits in the African elephant (*L. africana*) and four hits in the Asian elephant (*Elephas maximus*). In addition to African and Asian elephants, *Procavia capensis* and *Heterohyrax brucei* appeared in the hits with query covers smaller than 30%. Looking closer into these contigs from *L. africana*, *E. maximus*, *Procavia capensis*, and *Heterohyrax brucei*, we verified few sequences in tandem (maximum of 18) with the repetition unit comprising roughly one TMA sat HOR monomer. The small number of hits found

suggests that this sequence is not a typical satDNA in these taxa, but is instead a repetition related to a transposable element. Indeed, nine out of 10 TMA sat arrays were flanked by LINE-1 in *L. africana*, and 15 out of 16 in *E. maximus*.

DISCUSSION

The TMA sat, reported herein for the first time, was the only putative satDNA found in our analysis, comprising less than 1% of the *T. manatus* genome and mapping to the centromeric regions of all chromosomes, except the Y. The TMA sat could be absent or undetectable by FISH due to low copy number or sequence divergence on the Y chromosome. In *T. inunguis*, we confirmed the presence of TMA sat by PCR and FISH and despite the two species having different karyotypes ($2n=48$ and $2n=56$, respectively), TMA sat displayed the same chromosome localization (Figure 2). This could be related to the recent ~1.34 million years ago (Mya) divergence time between the species (de Souza et al., 2021).

We also detected the TMA sat in *D. dugon* and *H. gigas* with similar monomer length, comprising 0.63 and 3.4% of the genomes, respectively. The different genome proportion found in *T. manatus* (Illumina HiSeq; 150x genome coverage), *D. dugon* (Illumina NovaSeq6000; 64x genome coverage) and *H. gigas* (Illumina NovaSeq; 11x genome coverage) could be due to different genome coverage and/or sequencing platforms used for each species, and may not reflect real interspecific variation. This is especially true in the case of the extinct *H. gigas*, whose DNA source for genome sequencing is a petrous bone from a specimen who probably died during the 1760s (Sharko et al., 2021), and thus the abundance estimates need to be taken with caution.

Although there are slight differences within the MCS from each species, the Neighbor Joining analysis does not indicate intraspecific homogeneous monomers. Only the monomers from *T. inunguis* were grouped together, probably due to the low number of sequences used in the analysis or biased PCR amplification with the selected primers. Nevertheless, we cannot discard a species-specific TMA sat sequence in *T. inunguis* since some mutations are present in all or almost all five sequences and are absent or present in just few monomers outside the species. West Indian and Amazonian manatees present a recent divergence time (de Souza et al., 2021) and an incomplete reproductive isolation (Vianna et al., 2006; Lima et al., 2019), which could contribute to the TMA sat high interspecific homogeneity observed. Overall, the species-specific mutations of the group are probably not yet fixed, despite the ~46.83 Mya estimated split of Trichechidae and Dugongidae, thus lacking species-specific sequences as reflected in the neighbor joining tree (Figure 3D), which disagrees with the predictions of concerted evolution. This process, which has been described for many satDNAs, promotes fast sequence homogenization within a species or population, resulting in much higher interspecific than intraspecific differences (Plohl et al., 2012). Although interspecific satDNA sequence conservation is unexpected according to the concerted evolution model, interspecific homogeneity of centromeric satDNAs was observed in other mammalian groups, like in rodents from the *Peromyscus* genus

(Smalec et al., 2019), in four squirrel monkeys (*Saimiri* genus; Valeri et al., 2020) and in two species of two-toed sloths from the genus *Choloepus* (Sena et al., 2020). In all these cases, a possible centromeric function was hypothesized. Moreover, the library model of satDNA evolution relies on the preexistence of a satDNA collection in related species, with the differences observed among the species mostly due to amplification-contraction events of these sequences pool, and does not imply in rapid sequence changes (Plohl et al., 2009). These could be the case of TMA sat evolution if considering the monomer variants as independent amplification-contraction units.

In addition to the centromeric localization in *T. manatus* and *T. inunguis*, we detected the CENPB-box like motif, another centromeric feature, twice in the TMA sat sequences of all four Sirenia species. In *T. manatus* and in *D. dugon*, both putative CENPB-boxes were located in conserved segments of TMA sat. Even though the CENPB-box-like motif found in TMA sat does not present all the nine nucleotides of the ECD, we cannot exclude its functional activity. Among *Peromyscus* species, the CENPB-box-like motifs found in the centromeric satDNA had between four and six conserved bases out of nine ECD nucleotides. It has been suggested that a divergent motif sequence may be required for functional activity in this group (Smalec et al., 2019), which could also be the case for manatees and the dugong. Divergent motif sequences have also been observed in the centromeric satDNAs of the African elephant (*L. africana*), nine-banded armadillo (*D. novemcinctus*; Alkan et al., 2011) and in the two-toed sloths of the genus *Choloepus* (Sena et al., 2020).

The only genomes outside Sirenia in which the TMA sat sequence was found were those of the Order Proboscidea (elephants) and Hyracoidea (hyraxes), that together with Sirenia are reunited in Paenungulata, a subgroup of the Superorder Afrotheria (Foley et al., 2016). With only a few hits (with the short arrays mostly flanked by the transposable element L1), the TMA sat sequence is probably not a typical satDNA in these species. TMA sat in Sirenian probably evolved from these ancestral sequences still found in elephants and hyraxes, which could be the basis for both TMA sat1 and TMA sat2.

In the tree sirenians with sequenced genome, the most frequent organization of TMA sat arrays was the alternating TMA sat1 and TMA sat2 form. In the few exceptions, we found more consecutive TMA sat2 units than TMA sat1. Other satDNAs were found organized as a composite of two related units, mostly in the alternating form as TMA sat. This is the case of S1a-S1b in European brown frogs (Felicciello et al., 2006) and Tcast1a-Tcast1b in the red flour beetle *Tribolium castaneum* (Felicciello et al., 2011, 2014), in which the rolling circle amplification followed by substitutions by homologous recombination were proposed to explain the origin of the composite a-b arrays.

The sequenced genomes we used were generated from short reads (average 100–150bp) that do not cover the total length of the monomeric unit of TMA sat, resulting in an assembly that may not represent well the long satDNA arrays. Further analyses with Southern blot and dot blot experiments as well as long-reads sequencing may help clarify the overall organization of repeats in the genome and within the long satDNA arrays. As an example, Vondrak et al. (2020) using ultra-long nanopore reads found nine

out of 11 putative satDNA sequences derived from short tandem arrays located within LTR-retrotransposons that occasionally expanded in length, and just two organized in long arrays typical of satDNA. In addition, the long-reads sequencing approach proved a valuable contribution in determining the origin of the satDNAs. Several satDNAs from plants and animals derived from tandem amplification of internal segments of TEs (Dias et al., 2015; Meštrović et al., 2015; Vondrak et al., 2020), as was the case of TMA sat described herein, that could be L1 related.

In conclusion, we reported for the first time the centromeric satDNA in the West Indian manatee, which seems to be present across Sirenia, a group with all extant species under threat of extinction. TMA sat monomers from *T. manatus*, *T. inunguis*, *D. dugon*, and *H. gigas* lack species-specific sequences, contradicting the predictions of concerted evolution. The TMA sat-like ancestral sequence is present in other Paenungulata, such as elephants and hyraxes, suggesting that TMA sat suffered an expansion within Sirenia less than ~69 Mya (Foley et al., 2016; de Souza et al., 2021), after the divergence of Sirenia from Proboscidea and Hyracoidea.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ncbi.nlm.nih.gov/genbank/>, MW272776–MW272780.

ETHICS STATEMENT

The animal study was reviewed and approved by SISBio/ICMBio permit 60829-2.

AUTHOR CONTRIBUTIONS

MV and GD conceived and designed the experiments, analyzed the data, and contributed to writing – original draft preparation. MV and AE performed the experiments. CM, YY-Y, and IS obtained the materials for molecular and cytological analyses. MV, GD, AE, GK, and MS contributed to writing – review and editing. GK and MS contributed to supervision and project administration. MS contributed to funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

MV and AE received fellowships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). MS and GK are supported by fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; Processes 310433/2018-5 and 308386/2018-3, respectively). This work was funded with grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

ACKNOWLEDGMENTS

We thank the Programa de Pós-Graduação em Genética, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, for their support in the execution and publishing of this work.

REFERENCES

- Alkan, C., Cardone, M. F., Catacchio, C. R., Antonacci, F., O'Brien, S. J., Ryder, O. A., et al. (2011). Genome-wide characterization of centromeric satellites from multiple mammalian genomes. *Genome Res.* 21, 137–145. doi: 10.1101/gr.111278.110
- Alkan, C., Ventura, M., Archidiacono, N., Rocchi, M., Sahinalp, S. C., and Eichler, E. E. (2007). Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data. *PLoS Comput. Biol.* 3, 1807–1818. doi: 10.1371/journal.pcbi.0030181
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Assis, M. F. L., Best, R. C., Barros, R. M. S., and Yonenaga-Yassuda, Y. (1998). Cytogenetic study of *Trichechus inunguis* (Amazonian manatee). *Braz. J. Genet.* 11, 41–50.
- Bao, W., Kojima, K. K., and Kohany, O. (2015). Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* 6:11. doi: 10.1186/s13100-015-0041-9
- Barros, H. M. D. R., Meirelles, A. C. O., Luna, F. O., Marmontel, M., Cordeiro-Estrela, P., Santos, N., et al. (2017). Cranial and chromosomal geographic variation in manatees (Mammalia: Sirenia: Trichechidae) with the description of the Antillean manatee karyotype in Brazil. *J. Zool. Syst. Evol. Res.* 55, 73–87. doi: 10.1111/jzs.12153
- Biscotti, M. A., Olmo, E., and Heslop-Harrison, J. S. (2015). Repetitive DNA in eukaryotic genomes. *Chromosom. Res.* 23, 415–420. doi: 10.1007/s10577-015-9499-z
- Bonvicino, C. R., Viana, M. C., de Oliveira, E. H. C., Emin, R., Silva Junior, J. D. S., de Sousa, M. E. M., et al. (2020). Distribution of south American manatees, *Trichechus manatus* Linnaeus, 1758 and *T. inunguis* (Natterer, 1883) (Sirenia: Trichechidae). *Bol. Mus. Para. Emílio Goeldi, Ciênc. Nat.* 15, 573–599. doi: 10.46357/bcnaturais.v15i3.246
- de Souza, É. M. S., Freitas, L., da Silva Ramos, E. K., Selleghin-Veiga, G., Rachid-Ribeiro, M. C., Silva, F. A., et al. (2021). The evolutionary history of manatees told by their mitogenomes. *Sci. Rep.* 11:3564. doi: 10.1038/s41598-021-82390-2
- Deusch, C. J., Self-Sullivan, C., and Mignucci-Giannoni, A. (2008). Data from: *Trichechus manatus*, West Indian Manatee. *IUCN Red List Threat. Species*, e.T22103A9356917. doi: 10.2305/IUCN.UK.2008.RLTS.T22103A9356917.en
- Dias, G. B., Heringer, P., Svartman, M., and Kuhn, G. C. S. (2015). Helitrons shaping the genomic architecture of drosophila: enrichment of DINE-TR1 in α - and β -heterochromatin, satellite DNA emergence, and piRNA expression. *Chromosom. Res.* 23, 597–613. doi: 10.1007/s10577-015-9480-x
- Domning, D. P. (2018). "Sirenian evolution," in *Encyclopedia of Marine Mammals*. eds. B. Würsig, J. G. M. Thewissen and K. M. Kovacs (London, United Kingdom: Academic Press), 856–859.
- Dover, G. (1982). Molecular drive: a cohesive mode of species evolution. *Nature* 299, 111–117. doi: 10.1038/299111a0
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Feliciello, I., Akrap, I., Brajković, J., Zlatar, I., and Ugarković, D. (2014). Satellite DNA as a driver of population divergence in the red flour beetle *Tribolium castaneum*. *Genome Biol. Evol.* 7, 228–239. doi: 10.1093/gbe/evu280
- Feliciello, I., Chinali, G., and Ugarković, D. (2011). Structure and population dynamics of the major satellite DNA in the red flour beetle *Tribolium castaneum*. *Genetica* 139, 999–1008. doi: 10.1007/s10709-011-9601-1
- Feliciello, I., Picariello, O., and Chinali, G. (2006). Intra-specific variability and unusual organization of the repetitive units in a satellite DNA from *Rana dalmatina*: molecular evidence of a new mechanism of DNA repair acting on satellite DNA. *Gene* 383, 81–92. doi: 10.1016/j.gene.2006.07.016
- Foley, N. M., Springer, M. S., and Teeling, E. C. (2016). Mammal madness: is the mammal tree of life not yet resolved? *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150140. doi: 10.1098/rstb.2015.0140
- Fry, K., and Salser, W. (1977). Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12, 1069–1084. doi: 10.1016/0092-8674(77)90170-2
- Gray, B. A., Zori, R. T., McGuire, P. M., and Bonde, R. K. (2002). A first generation cytogenetic ideogram for the Florida manatee (*Trichechus manatus latirostris*) based on multiple chromosome banding techniques. *Hereditas* 137, 215–223. doi: 10.1034/j.1601-5223.2002.01657.x
- Kasai, F., O'Brien, P. C. M., and Ferguson-Smith, M. A. (2013). Afrotheria genome; overestimation of genome size and distinct chromosome GC content revealed by flow karyotyping. *Genomics* 102, 468–471. doi: 10.1016/j.ygeno.2013.09.002
- Kasinathan, S., and Henikoff, S. (2018). Non-B-form DNA is enriched at centromeres. *Mol. Biol. Evol.* 35, 949–962. doi: 10.1093/molbev/msy010
- Keith, L. D. (2015). Data from: *Trichechus senegalensis* (African Manatee). *IUCN Red List Threat. Species*, e.T22104A97168578. doi: 10.2305/IUCN.UK.2015-4.RLTS.T22104A81904980.en
- Kohany, O., Gentles, A. J., Hankus, L., and Jurka, J. (2006). Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinform.* 7:474. doi: 10.1186/1471-2105-7-474
- Letunic, I., and Bork, P. (2019). Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- Lima, C. S., Magalhães, R. F., Marmontel, M., Meirelles, A. C., Carvalho, V. L., Laverge, A., et al. (2019). A hybrid swarm of manatees along the Guianas coastline, a peculiar environment under the influence of the Amazon River plume. *An. Acad. Bras. Ciênc.* 91:e20190325. doi: 10.1590/0001-3765201920190325
- Lima, C. S., Magalhães, R. F., and Santos, F. R. (2021). Conservation issues using discordant taxonomic and evolutionary units: a case study of the American manatee (*Trichechus manatus*, Sirenia). *Wildl. Res.* 48:385. doi: 10.1071/WR20197
- Luna, F. O., Beaver, C. E., Nourisson, C., Bonde, R. K., Attademo, F. L. N., Miranda, A. V., et al. (2021). Genetic connectivity of the west Indian manatee in the southern range and limited evidence of hybridization with Amazonian manatees. *Front. Mar. Sci.* 7, 1–15. doi: 10.3389/fmars.2020.574455
- Marmontel, M., de Souza, D., and Kendall, S. (2016). Data from: *Trichechus inunguis*. *IUCN Red List Threat. Species*, e.T22102A43793736. doi: 10.2305/IUCN.UK.2016-2.RLTS.T22102A43793736.en
- Marsh, H., and Soltzick, S. (2019). Data from: Dugong Dugon (amended version of 2015 assessment). *IUCN Red List Threat. Species* 2019, e.T22104A97168578. doi: 10.2305/IUCN.UK.2015-4.RLTS.T22104A81904980.en
- Masumoto, H., Nakano, M., and Ohzeki, J. I. (2004). The role of CENP-B and α -satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. *Chromosom. Res.* 12, 543–556. doi: 10.1023/B:CHRO.0000036593.72788.99
- Meštrović, N., Mravinac, B., Pavlek, M., Vojvoda-Zeljko, T., Šatović, E., and Plohl, M. (2015). Structural and functional liaisons between transposable elements and satellite DNAs. *Chromosom. Res.* 23, 583–596. doi: 10.1007/s10577-015-9483-7
- Meštrović, N., Plohl, M., Mravinac, B., and Ugarković, D. (1998). Evolution of satellite DNAs from the genus *palorus*-experimental evidence for the "library" hypothesis. *Mol. Biol. Evol.* 15, 1062–1068. doi: 10.1093/oxfordjournals.molbev.a026005
- Muro, Y., Masumoto, H., Yoda, K., Nozaki, N., Ohashi, M., and Okazaki, T. (1992). Centromere protein B assembles human centromeric α -satellite DNA at the 17-bp sequence, CENP-B box. *J. Cell Biol.* 116, 585–596. doi: 10.1083/jcb.116.3.585
- Novák, P., Neumann, P., and Macas, J. (2020). Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nat. Protoc.* 15, 3745–3776. doi: 10.1038/s41596-020-0400-y

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.694866/full#supplementary-material>

- Novák, P., Robledillo, L. Á., Koblížková, A., Vrbová, I., Neumann, P., and Macas, J. (2017). TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res.* 45:e111. doi: 10.1093/nar/gkx257
- Plohl, M., Meštrović, N., and Mravinac, B. (2012). "Satellite DNA evolution," in *Repetitive DNA. Genome Dyn. Vol. 7*. ed. M. A. Garrido-Ramos (Basel, Karger), 126–152.
- Plohl, M., Petrović, V., Luchetti, A., Ricci, A., Šatović, E., Passamonti, M., et al. (2009). Long-term conservation vs high sequence divergence: the case of an extraordinarily old satellite DNA in bivalve mollusks. *Hereditas* 104, 543–551. doi: 10.1038/hdy.2009.141
- Roza, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248
- Sena, R. S., Heringer, P., Valeri, M. P., Pereira, V. S., Kuhn, G. C. S., and Svartman, M. (2020). Identification and characterization of satellite DNAs in two-toed sloths of the genus *Choloepus* (Megalonychidae, Xenarthra). *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-76199-8
- Shapiro, J. A., and von Sternberg, R. (2005). Why repetitive DNA is essential to genome function. *Biol. Rev.* 80, 227–250. doi: 10.1017/S1464793104006657
- Sharko, F. S., Boulygina, E. S., Tsygankova, S. V., Slobodova, N. V., Alekseev, D. A., Krasivskaya, A. A., et al. (2021). Steller's sea cow genome suggests this species began going extinct before the arrival of Paleolithic humans. *Nat. Commun.* 12, 1–8. doi: 10.1038/s41467-021-22567-5
- Shatskikh, A. S., Kotov, A. A., Adashev, V. E., Bazylev, S. S., and Olenina, L. V. (2020). Functional significance of satellite DNAs: insights from drosophila. *Front. Cell Dev. Biol.* 8:312. doi: 10.3389/fcell.2020.00312
- Smalec, B. M., Heider, T. N., Flynn, B. L., and O'Neill, R. J. (2019). A centromere satellite concomitant with extensive karyotypic diversity across the *Peromyscus* genus defies predictions of molecular drive. *Chromosom. Res.* 27, 237–252. doi: 10.1007/s10577-019-09605-1
- Stanyon, R., and Galleni, L. (1991). A rapid fibroblast culture technique for high resolution karyotypes. *Bolletino di Zool.* 58, 81–83. doi: 10.1080/11250009109355732
- Sujiwattanarat, P., Thapana, W., Srikulnath, K., Hirai, Y., Hirai, H., and Koga, A. (2015). Higher-order repeat structure in alpha satellite DNA occurs in New World monkeys and is not confined to hominoids. *Sci. Rep.* 5, 1–10. doi: 10.1038/srep10315
- Sullivan, L. L., Chew, K., and Sullivan, B. A. (2017). α satellite DNA variation and function of the human centromere. *Nucleus* 8, 331–339. doi: 10.1080/19491034.2017.1308989
- Valeri, M. P., Dias, G. B., Moreira, C. N., Yonenaga-Yassuda, Y., Stanyon, R., Silva Kuhn, G. C., et al. (2020). Characterization of satellite DNAs in squirrel monkeys genus *Saimiri* (Cebidae, Platyrrhini). *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-64620-1
- Vianna, J. A., Bonde, R. K., Caballero, S., Giraldo, J. P., Lima, R. P., Clark, A., et al. (2006). Phylogeography, phylogeny and hybridization in trichechid sirenians: implications for manatee conservation. *Mol. Ecol.* 15, 433–447. doi: 10.1111/j.1365-294X.2005.02771.x
- Vlahović, I., Glunčić, M., Rosandić, M., Ugarković, Đ., and Paar, V. (2016). Regular higher order repeat structures in beetle *Tribolium castaneum* genomex. *Hereditas* 104, 543–551. doi: 10.1093/gbe/evw174
- Vondrak, T., Ávila Robledillo, L., Novák, P., Koblížková, A., Neumann, P., and Macas, J. (2020). Characterization of repeat arrays in ultra-long nanopore reads reveals frequent origin of satellite DNA from retrotransposon-derived tandem repeats. *Plant J.* 101, 484–500. doi: 10.1111/tpj.14546

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Valeri, Dias, Espírito Santo, Moreira, Yonenaga-Yassuda, Sommer, Kuhn and Svartman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.