


RESEARCH ARTICLE

Open Access



Evolution of ancient satellite DNAs in extant alligators and caimans (Crocodylia, Reptilia)

Vanessa C. Sales-Oliveira^{1†}, Rodrigo Zeni dos Santos^{2†}, Caio Augusto Gomes Goes², Rodrigo Milan Calegari², Manuel A. Garrido-Ramos³, Marie Altmanová^{4,5}, Tariq Ezaz⁶, Thomas Liehr^{7*} , Fabio Porto-Foresti², Ricardo Utsunomia^{2†} and Marcelo B. Cioffi^{1,7†}

Abstract

Background Crocodylians are one of the oldest extant vertebrate lineages, exhibiting a combination of evolutionary success and morphological resilience that has persisted throughout the history of life on Earth. This ability to endure over such a long geological time span is of great evolutionary importance. Here, we have utilized the combination of genomic and chromosomal data to identify and compare the full catalogs of satellite DNA families (satDNAs, i.e., the satellitomes) of 5 out of the 8 extant Alligatoridae species. As crocodylian genomes reveal ancestral patterns of evolution, by employing this multispecies data collection, we can investigate and assess how satDNA families evolve over time.

Results Alligators and caimans displayed a small number of satDNA families, ranging from 3 to 13 satDNAs in *A. sinensis* and *C. latirostris*, respectively. Together with little variation both within and between species it highlighted long-term conservation of satDNA elements throughout evolution. Furthermore, we traced the origin of the ancestral forms of all satDNAs belonging to the common ancestor of Caimaninae and Alligatorinae. Fluorescence in situ experiments showed distinct hybridization patterns for identical orthologous satDNAs, indicating their dynamic genomic placement.

Conclusions Alligators and caimans possess one of the smallest satDNA libraries ever reported, comprising only four sets of satDNAs that are shared by all species. Besides, our findings indicated limited intraspecific variation in satellite DNA, suggesting that the majority of new satellite sequences likely evolved from pre-existing ones.

Keywords Repetitive DNA, Reptiles, Library hypothesis

[†]Ricardo Utsunomia and Marcelo B. Cioffi contributed equally to this paper.

[†]Vanessa C. Sales-Oliveira and Rodrigo Zeni dos Santos contributed equally to this paper.

*Correspondence:

Thomas Liehr

Thomas.Liehr@med.uni-jena.de

¹ Departamento de Genética E Evolução, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil

² Faculdade de Ciências, UNESP, Bauru, São Paulo, Brazil

³ Departamento de Genética, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

⁴ Institute of Animal Physiology and Genetics, Czech Academy of Sciences, 27721 Liběchov, Czech Republic

⁵ Department of Ecology, Faculty of Science, Charles University, 12844 Prague, Czech Republic

⁶ Institute for Applied Ecology, University of Canberra, Canberra, Australia

⁷ Institute of Human Genetics, Jena University Hospital, Friedrich Schiller University, Jena, Germany



Background

Eukaryotic genomes are highly variable in structure and size because of the presence of vast quantities of repetitive DNA [1, 2]. Satellite DNAs (satDNAs) are a common component, accounting for an important part of the genome in most animal and plant genomes (reviewed by [3–5]). In general, a genome has a varied number of satDNA families (the satellitome) [6] with varying nucleotide sequences and genomic abundance [7–13]. Although some examples of small arrays scattered throughout euchromatin have been documented [6, 14–22], these sequences are often found in centromeres and in pericentromeric and subtelomeric heterochromatic areas [4, 10, 23]. More than just “junk DNA” (as for a long time considered), several studies have revealed that satDNAs have a role in a variety of biological processes, including gene regulation [24], centromere function [25], chromatin modulation [26], and spatial chromosomal structure [27–29]. The combination of cytogenetics and genomics studies has proven to be useful in elucidating numerous aspects of genome evolution and organization [30, 31], with particular emphasis on repetitive DNAs [6, 32–34]. Furthermore, due to their tandemly repeated genomic organization, satDNA studies in non-model organisms were boosted in the last few years, especially with the development of several assembly-free pipelines designed for using raw reads [35–38]. In this context, several satDNA catalogs were characterized in a variety of invertebrate and vertebrate species [6, 34, 39–44].

Although related species sharing a common ancestor share the same ancestral library of satDNA families, differential amplification of the different satDNAs, and different variants of each of them, leads to differentiated satDNA profiles in each species [5, 7, 10]. This involves the replacement of some satDNA families by others at specific sites on chromosomes such as centromeres, for example. The rate of change is very rapid, and satDNA sequences represent one of the fastest evolving genomic components. This often leads to high levels of interspecific sequence diversity even within closely related species, exhibiting very different profiles (both quantitative and qualitative) of satDNAs in their genomes [4, 10, 45]. However, satDNA rate of change can be altered (accelerated or slowed) by various factors such as the location and organization of the repeated sequences [46], functional constraints [42, 47–51], biological factors [8, 11], or population and evolutionary factors [41, 52, 53]. In this regard, it is particularly intriguing to investigate why, for some satDNAs, this process is slower than expected and persists over long periods, spanning dozens (or even hundreds) of million years, in the same chromosomal location in the entire group of related species [32, 42, 47, 52–56]. It is also important to examine the effects

of slow rates of molecular and morphological evolution described in a species group [52] on the evolution of the satellitome.

Crocodylians are one of the oldest extant vertebrate lineages, demonstrating evolutionary success and morphological resilience over many millions of years [57]. Extant species have preserved physical and ecological traits for nearly 100 million years, unlike other vertebrates that have undergone significant diversity [58–60]. Crocodylians have a key position in vertebrate phylogeny because, combined with dinosaurs, pterosaurs, and modern birds, they compose the archosaurs, a monophyletic group [61–63]. Crocodylia is classified into three families: Crocodylidae, Gavialidae, and Alligatoridae, with approximately 27 species [64]. The family Alligatoridae is made up of eight species that are divided in four genera: *Melanosuchus*, *Paleosuchus*, and *Caiman*, which belong to the Caimaninae subfamily and *Alligator*, which forms the monogeneric subfamily Alligatorinae. Except for the *Alligator* genus, where *A. mississippiensis* and *A. sinensis* are limited to the Southeastern United States and China, respectively, all the other six species are presently found in South America, being more widespread in Brazil [57, 65].

The karyotypes of all current Alligatoridae species were recently revised using conventional differential staining and up-to-date molecular cytogenetic approaches [66]. Although there is a limited amount of diversity and a certain level of karyotype stasis (with diploid numbers equal to $2n=42$ and $2n=32$ for all Caimaninae (*Caiman*, *Paleosuchus*, and *Melanosuchus*) and Alligatorinae (*Alligator*) species, respectively), their genomic content revealed significant interspecific divergence [66].

Here, we performed the first broadscale comparative analysis of the alligators’ satellitomes. By combining genomic and chromosomal data, we identified and compared the full catalogs of satDNA families (i.e., the satellitomes) of 5 of the 8 extant Alligatoridae species, revealing ancestral patterns of evolution and enabling investigation into how satDNA families evolve over time. The results revealed strong sequence conservatism among Alligatoridae species with very limited diversity of their satDNA library. Furthermore, fluorescence in situ assays in all the 8 extant Alligatoridae species showed that the identical satDNA orthologs can exhibit various hybridization patterns, indicating their high evolutionary dynamics.

Results

Bioinformatic satDNA characterization

After several iterations (*C. yacare*=5, *C. latirostris*=7, *M. niger*=4, *P. trigonatus*=4, and *A. sinensis*=2), we characterized 39 satDNAs in Alligatoridae, where repeat unit lengths ranged from 23 (PtrSat11-23) to 6317nt

(ClaSat02-6317) and the average of their A+T content was 46.9%. Specific features of the satDNAs in each species are summarized in Table 1. The number of iterations performed for each species was a consequence of the results obtained in each round so that when no new tandem repeats were discovered in a given round the analysis was not continued. Thus, for example, in the case of

A. sinensis no tandem sequences were discovered in the third interaction. Iterations using RepeatExplorer2 after TAREAN did not return any characterized satellite DNA for the five species analyzed.

In general, alligators and caimans analyzed here exhibited few satDNAs (minimum of three in *A. sinensis* and maximum of 13 in *C. latirostris*), and also a

Table 1 General features of Alligatoridae satellitomes characterized with TAREAN. SF=superfamily, RUL=repeat unit length, TSI=tandem structure index. Divergence per satDNA was expressed as the percentage of Kimura divergence. SatDNAs that have 50% or more identity belong to the same group

Species	SF	satDNA	RUL	A+T (%)	Abundance (%)	Divergence (%)	TSI	Group
<i>Caiman yacare</i>	1	CyaSat01-41	41	63.4	2.211	8.79	0.91	1
		CyaSat02-40	40	50.0	0.212	8.35	0.98	3
	2	CyaSat03-60	60	41.7	0.099	18.50	0.98	2
		CyaSat04-41	41	58.5	0.084	21.64	0.89	1
	2	CyaSat05-200	200	43.5	0.062	8.67	0.96	2
		CyaSat06-286	286	57.3	0.061	4.29	0.61	1
<i>Caiman latirostris</i>	1	ClaSat01-41	41	46.3	1.077	11.70	0.93	1
		ClaSat02-6317	6317	51.0	0.733	7.94	0.87	
	2	ClaSat03-183	183	40.4	0.268	9.30	0.87	2
		ClaSat04-536	536	26.1	0.262	4.83	0.92	
		ClaSat05-40	40	47.5	0.207	9.51	0.87	3
	1	ClaSat06-1063	1063	53.3	0.181	9.09	0.74	1
	1	ClaSat07-320	320	53.1	0.129	8.40	0.89	1
	1	ClaSat08-800	800	55.9	0.129	14.84	0.99	1
	1	ClaSat09-285	285	46.5	0.074	3.90	0.95	1
	2	ClaSat10-60	60	40	0.069	17.79	0.4	2
	1	ClaSat11-547	547	51.7	0.069	9.81	0.69	1
	2	ClaSat12-24	24	33.3	0.053	16.21	0.82	2
		ClaSat13-398	398	32.7	0.042	8.61	0.41	
<i>Melanosuchus niger</i>	1	MniSat01-41	41	58.5	1.582	8.99	0.87	1
		MniSat02-246	246	55.7	0.442	5.54	0.74	1
	1	MniSat03-60	60	40.0	0.240	17.03	0.96	2
		MniSat04-40	40	52.5	0.110	11.23	0.97	3
	1	MniSat05-248	248	55.6	0.045	10.56	0.57	1
	1	MniSat06-41	41	56.1	0.032	18.32	0.90	1
<i>Paleosuchus trigonatus</i>	1	PtrSat01-41	41	41.5	1.092	05.02	0.91	1
	2	PtrSat02-24	24	37.5	0.705	10.34	0.98	2
	1	PtrSat03-147	147	55.1	0.594	10.02	0.95	1
	1	PtrSat04-99	99	51.5	0.332	04.07	0.72	1
	1	PtrSat05-41	41	53.7	0.328	12.86	0.92	1
	2	PtrSat06-100	100	45	0.252	9.69	0.97	2
	1	PtrSat07-41	41	56.1	0.155	16.09	0.96	1
	3	PtrSat08-214	214	28	0.14	01.05	0.57	3
		PtrSat09-490	490	31.6	0.112	4.29	0.96	
	3	PtrSat10-40	40	45	0.112	12.41	0.38	3
	1	PtrSat11-23	23	52.2	0.028	13.70	0.05	1
<i>Alligator sinensis</i>		AsiSat01-1717	1717	50.6	0.234	13.39	0.95	
		AsiSat02-60	60	38.3	0.154	3.51	0.99	
		AsiSat03-96	96	34.4	0.02	9.25	0.93	3

small diversity in the within- and between-species level. Intraspecific cases with similarity greater than 50% and less than 80% were classified as the same superfamily, while interspecific cases with similarity greater than 50% were placed in the same group, which was further subdivided into four distinct ones.

Based on sequence alignments, four main groups of satDNAs were identified showing at least 50% of similarity that encompassed satDNAs from at least four species, named here as group 1 ($N=19$ satDNAs shared among Caimaninae), group 2 ($N=8$ satDNAs shared among Caimaninae), group 3 ($N=6$ satDNAs shared among Alligatoridae), and group 4, with satDNAs from two species ($N=2$ satDNAs shared among Caimaninae) (Additional file 1: Fig. S1; Additional file 2: Table S1; Table 1), the remaining 4 satDNAs did not show any similarity with other sequences (Table 1). This classification helped us to delimit the origin of some satDNAs and follow their diversification patterns in each species.

We also performed one comparative RepeatExplorer2 run, inputting reads from all the analyzed species into a single dataset. Results obtained corroborated our previous analyses and satDNAs from groups 1–4 were found, as well as other tandem repeats not classified within these groups (Additional file 2: Table S2). A general clustermap considering the genomic abundances of satDNAs exhibiting a maximum of 20% of divergence in each species was also generated and is in accordance with the phylogenetic relationships (Fig. 1). As expected by theory, the more distant two species are, the fewer satDNAs they share. For instance, *A. sinensis* almost does not share satDNAs with the other species.

Expanding our analyses, as we found a significant RUL variability in each of those groups, we generated a global dot-plot with sequences from the abovementioned first three groups (groups 1, 2, and 3) (Fig. 2). As expected, sequences belonging to a same group showed similarities as revealed by the dotplots, which also indicated that

most longer satDNAs within groups probably emerged from the diversification of pre-existing shorter satDNAs.

We generated dotplots for each satDNA monomer supporting this view (Fig. 2). Thus, for example, among group 1 satDNAs we observed that the 41-bp satellites are made up of a structure composed of two subrepeats (21+20 bp) (Fig. 2). Comparisons between these two subrepeats suggest that a satDNA of about 20 bp in length must have existed and that a new satDNA composed of 41-bp repeats emerged through a process of duplication and subsequent divergence (in fact, one of the satellites in this group, PtrSat11-23, is 23 bp long). Therefore, when comparing the mean divergence that exists between 41-bp repeat units (inter-repeat divergence) with the mean divergence that exists between 20/21 bp subunits that compose each repeat (intra-repeat divergence), we always find that the former is smaller than the latter (Additional file 2: Table S3 and Additional file 1: Fig. S2). Satellites of this group with lengths longer than 41 bp show a complex pattern of several cycles of duplication and divergence of subunits of about 40 bp or more. For example, the analysis of ClaSat06-1063 demonstrates a complex evolutionary pattern based on different cycles of duplication and divergence of subrepeats of approximately 40/80 bp (including intervening sequences) (Additional file 1: Fig. S3). Similarly, group 2 satellites that are composed of 60 bp repeat units have a pattern of 20 bp subunits that again point to a formation of 60-bp satellites from smaller satellites (in fact, one of the satellites in this group, ClaSat12-24, is 24 bp long) (Fig. 2). Also in this case, mean inter-repeat divergence is smaller than mean intra-repeat divergence (Additional file 2: Table S4). Finally, group 3 is the only example of shared satDNAs between all the analyzed species here. Repeat monomers of 40 bp are predominant among these satDNAs, and dot-plot analysis revealed a heterogeneous structure based on two different subrepeats (29 bp+11 bp, Fig. 2). In this case again, we show

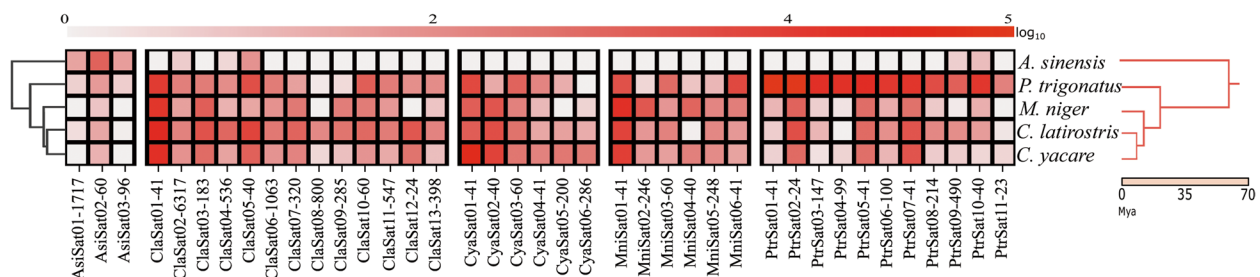
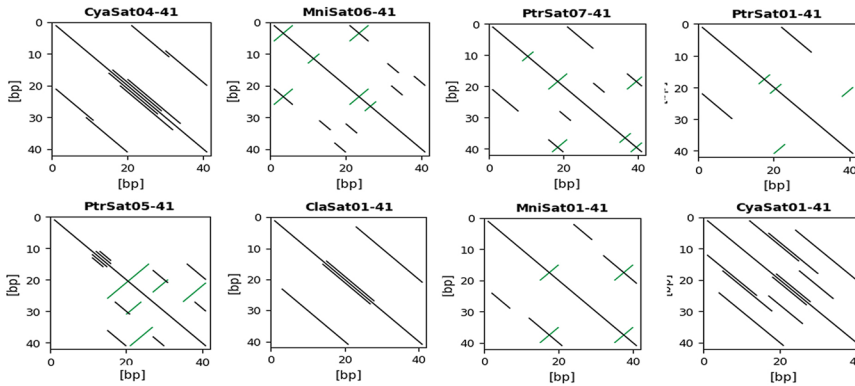
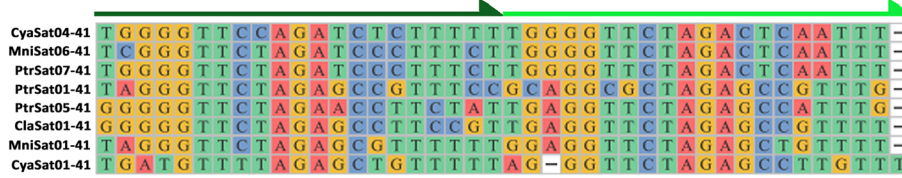
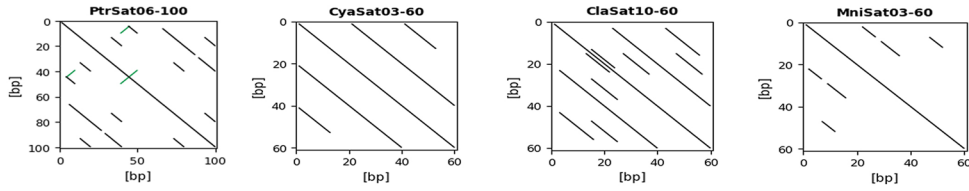


Fig. 1 Clustermap evidencing presence/absence and abundance of the satDNAs across Alligatoridae species. Abundances were calculated as the \log_{10} of the proportion of short reads masked as satDNAs with a maximum of 20% of divergence and normalized by single-copy genes. On the left, hierarchical calculated clusters; on the right, species name and the proposed phylogeny for the group with their respective divergence times, based on data generated by [65]

Group 1



Group 2



Group 3

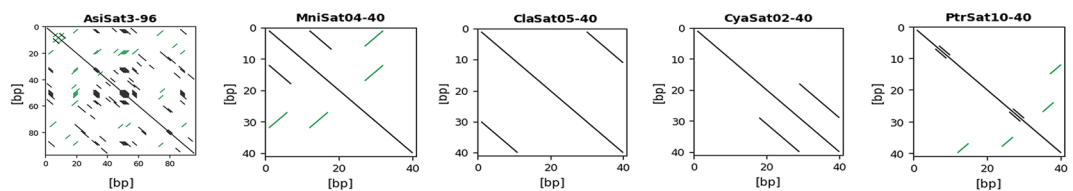
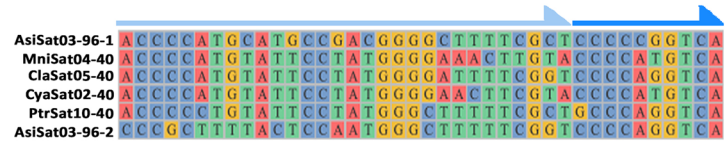
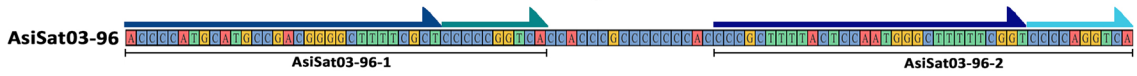


Fig. 2 Sequence alignment of satellite repeats in groups 1, 2, and 3, demarcated by arrows in the alignments, representing the subunits from which the sequences in each group originated. In addition, dotplots of each satDNA, presenting the internal repetitions among all monomers, are also indicated

that mean intra-repeat divergence is greater than mean inter-repeat divergence (Additional file 2: Table S5 and Additional file 1: Fig. S4). Remarkably, in *A. sinensis*, a 96-bp-long satDNA was characterized in this group, and its monomer sequence reveal a complex structure composed of two 40-bp subrepeats and an intervening sequence (40 bp + 16 bp + 40 bp, Fig. 2).

BLAST searches against the genome of *A. sinensis* revealed that satDNAs classified as groups 1 and 2 were not found in this species, while matches were observed for sequences belonging to group 3 (CyaSat02-40, ClaSat05-40, MniSat04-40, PtrSta10-40, and AsiSat03-96) and group 4 (ClaSat04-536 and PtrSat09-490). BLAST searches also resulted in matches against ClaSat02-6317 and ClaSat13-398, satDNAs that are not classified in any group (results are summarized in Additional file 2: Table S6). These results suggest that groups 1 and 2 of sequences emerged after the split of Caimaninae and Alligatorinae, while groups 3 and 4 are shared among the representatives of both subfamilies. In addition, Alligatorinae-specific AsiSat01-1717 and AsiSat02-60 satDNAs returned abundant significant matches, as expected.

ClaSat02-6317 and ClaSat13-398 produced multiple hits against the *A. sinensis* genome ($n=6264$ and 1712, respectively). Remarkably, the obtained TSI for ClaSat13-398 was low (TSI=0.21), suggesting that this sequence is dispersed along the genome. While ClaSat02-6317 exhibited a higher TSI (TSI=0.74), we hypothesize that this is due to its larger monomer size. Since the fragments of paired-end sequencing are usually around 300–400 bp and the monomer of this satDNA is >6000 bp, the obtained TSI is most likely due to mapping in the same monomer, not mapping in adjacent monomers. In fact, both ClaSat02-6317 and

ClaSat13-398 do not show FISH hybridization signals in this species supporting their scattering as short tandems throughout the genome. Interestingly, a RepeatMasker search on the vertebrate database of Repbase revealed that the former is homologous to endogenous retroviruses (ERVs) (62% identity; 70% of the element) and the latter to LINE sequences (70% identity; 53% of the element) (Supplemental Table S7). On the other hand, this search also revealed that AsiSat1-1717 shared a 54 and 82% of its sequence with two satellite DNAs previously found in the Nile crocodile (NCBI accession numbers: OP480175 and OP480176).

Chromosomal location of satDNAs with differential abundance between species

We analyzed the chromosomal location of satDNAs that were successfully amplified by PCR belonging to group 1 (ClaSat01-41; ClaSat06-1063; ClaSat07-320; ClaSat08-800 and ClaSat11-547), group 2 (ClaSat10-60), group 3 (ClaSat05-40), and group 4 (ClaSat04-536) in addition to the two exclusively ones found in *A. sinensis* (AsiSat01-1717 and AsiSat02-60) in all Alligatoridae species to check their chromosomal distribution. Additionally, the ungrouped satellites ClaSat02-6317 and ClaSat13-398 were tested but none of them yielded visible FISH signals in any species (data not shown).

Concerning the ClatSatDNAs, except for the satellite ClaSat04-536 (group 4), which showed no hybridization signals in any species, all the other satDNA sequences were found in (peri-) centromeric heterochromatin regions in all Caimaninae species (Figs. 3, 4, and 5). Both alligators (*A. sinensis* and *A. mississippiensis*) showed no hybridization signal for any of the ClaSatDNAs investigated (data not shown), which is in

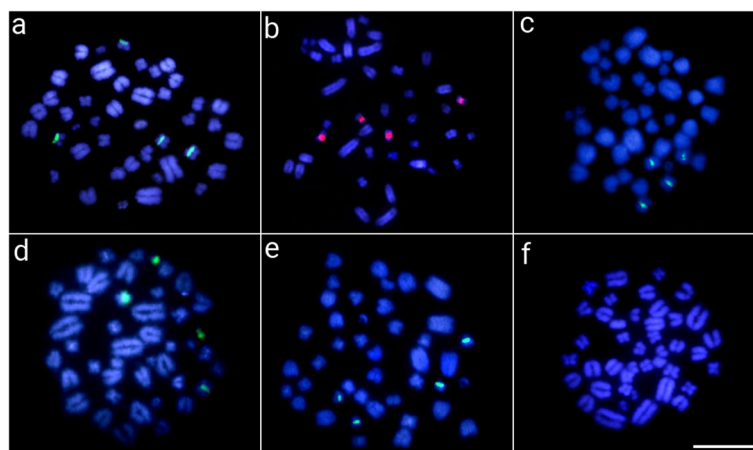


Fig. 3 Metaphase chromosomes from *C. crocodilus* (a), *C. latirostris* (b), *C. yacare* (c), *M. niger* (d), *P. palpebrosus* (e), and *P. trigonatus* (f) after in situ mapping of ClaSat01-41 (group 1). The satDNA FISH signals are highlighted in green (ATTO488 labeled) or red (ATTO550 labeled) and the chromosomes were counterstained with DAPI (blue). Scale bar = 20 μ m

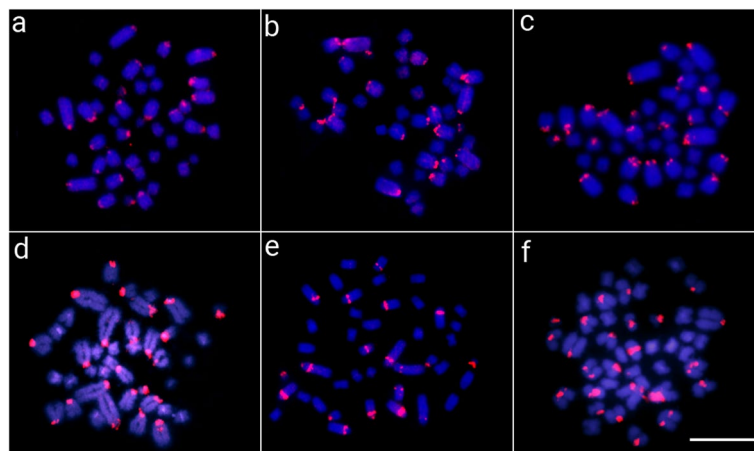


Fig. 4 Metaphase chromosomes from *C. crocodilus* (a), *C. latirostris* (b), *C. yacare* (c), *M. niger* (d), *P. palpebrosus* (e), and *P. trigonatus* (f) after in situ mapping of ClaSat10-60 (group 2). The satDNA FISH signals are highlighted in red (ATTO550 labeled) and the chromosomes were counterstained with DAPI (blue). Scale bar = 20 μ m

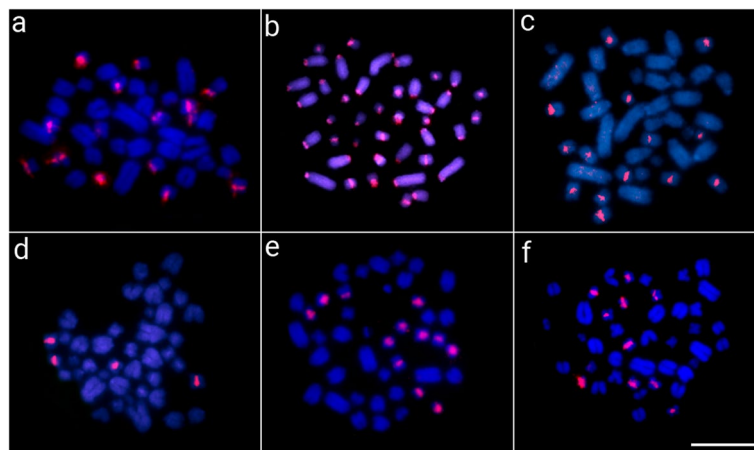


Fig. 5 Metaphase chromosomes from *C. crocodilus* (a), *C. latirostris* (b), *C. yacare* (c), *M. niger* (d), *P. palpebrosus* (e), and *P. trigonatus* (f) after in situ mapping of ClaSat05-40 (group 3). The satDNA FISH signals are highlighted in red (ATTO550 labeled) and the chromosomes were counterstained with DAPI (blue). Scale bar = 20 μ m

accordance with the clustermap analysis. Here, to illustrate, we present the results for representative selected ClaSatDNAs from each of the major groups identified (Figs. 3, 4, and 5). The satDNA ClaSat01-41, belonging to group 1 (the most frequent group present in each species), were mapped in two chromosomal pairs in all Caimaninae species except *P. trigonatus*, which did not display any hybridization signal (Fig. 3). However, despite sharing the same motifs, some divergent and species-specific chromosomal location patterns were observed among ClaSatDNAs from group 1 among species (Additional file 1: Figs. S5–S6). For satellites in groups 2 and 3, numerous chromosomal pairs containing these sequences were found in nearly all species (Figs. 4 and 5). *M. niger* is distinctive for displaying

hybridization signals on only two chromosomal pairs for group 3 satellites (Fig. 5d).

Besides, we also mapped the two exclusive satDNAs presented in *A. sinensis* genome (AsiSat01-1717 and AsiSat02-60) in all Alligatoridae species. Both AsiSatDNAs showed hybridization signals only in *Alligator* species. While AsiSat01-1717 was exclusively mapped in several chromosomes, AsiSat02-60 was mapped in all centromeres of both species (Fig. 6). Collectively, our analyses revealed, for *A. sinensis*, that (i) although group 3 and group 4 satellites are present in the *A. sinensis* genome, these satellites are poorly represented and possibly organized in short tandems scattered throughout the genome as can be deduced from TSI values, BLAST search and FISH: hybridization signals were not visible,

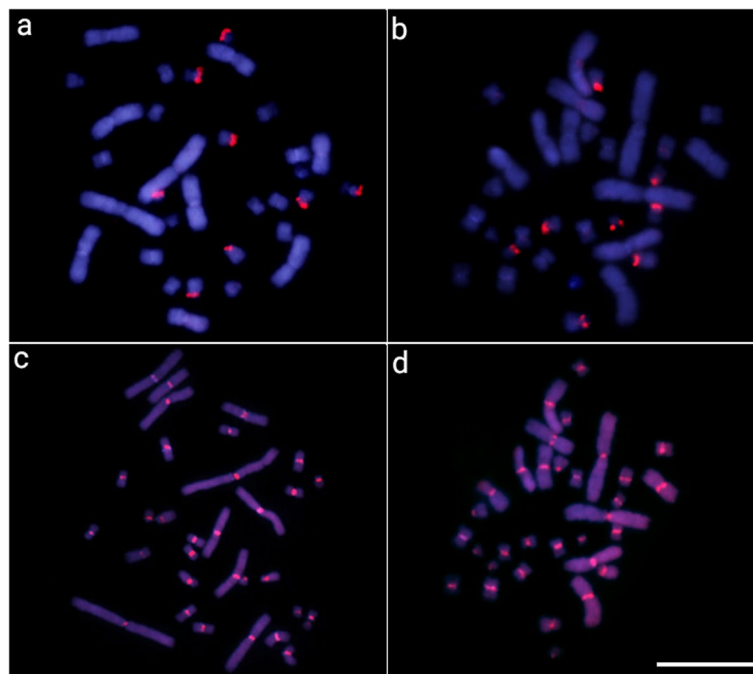


Fig. 6 Metaphase chromosomes from *A. sinensis* (a and c) and *A. mississippiensis* (b and d) after in situ mapping with AsiSat01-1717 (a and b) and AsiSat02-60 (c and d) probes. The satDNA FISH signals are highlighted in red (ATTO550 labeled) and the chromosomes were counterstained with DAPI (blue). Scale bar = 20 μ m

and satellites exhibited high TSI, but low number of alignments in BLAST; and (ii) this is in contrast to the alligator-specific satellites that appear clustered at loci on long arrays, consistent with results obtained in FISH experiments in which these satellites give conspicuous FISH bands, high TSI values, and a large number of alignments in BLAST.

Discussion

Despite the fact that both alligators' (*A. sinensis* and *A. mississippiensis*) complete genomes were characterized some years ago [63, 67, 68], genome-wide investigations of satDNAs in this group of organisms were never undertaken. SatDNAs are well known to be underrepresented in genome assemblies [4], particularly those genomes assembled using short-read sequencing technology, as is the case with alligators. In this context, knowledge about satDNAs in crocodylians was limited to just a few works [69, 70]. Given that high-throughput satellitome analysis has been very enlightening for understanding the satDNA evolution in various organisms, we used a chromosome- and genomic-based approach to try to describe the satellitome from members of all current Alligatoridae genera for the first time. In a period of around ~70 Myr (million years), many satDNA sequences are shared among the species, assisting in the hypothesis that they are derived from small sequences, as shown in Fig. 2.

Furthermore, in following fluorescence in situ tests the distinct hybridization patterns for the identical ortholog satDNAs were observed.

After mining satellite DNAs using well-established bioinformatic pipelines [6, 36], we found that alligators' satellitomes are among the smallest catalogs described until now, varying between 3 and 13 satDNAs, in *A. sinensis* and *C. latirostris*, respectively. In recent years, several satellitomes from a wide range of species, including plants and animals, have been identified [6, 42–44, 71–73]. These investigations showed that satellite DNA profiles are very dynamic. For example, characiform fish satellitomes display a significant quantitative and qualitative variation, with some species exhibiting a few dozen [44], while others can show more than one hundred satDNAs [74]. Here, we found that all alligators are similarly satDNA-poor constituting a common trend in this group.

Novel satDNA families can emerge by variable mechanisms and from multiple genomic regions, like introns, transposable elements, and/or existing satDNA families [38, 75, 76]. Our findings indicated that there was little intraspecific variation in satellite DNA, indicating that most new satellite sequences evolved from pre-existing ones. For instance, *C. latirostris* exhibited 13 satDNAs, but six and three of them were grouped as superfamilies (sequences showing more than 50% of similarity and less than 80%), named here as groups 1 and 2, respectively

(Table 1). Interestingly, this limited diversity is also apparent at the interspecific level, where over 90% of the 39 satDNAs described for Alligatoridae can be categorized into 4 main groups of sequences. After their origin, new longer satellites derived from the complex diversification of shorter ones already existing in the genome throughout different and successive cycles of duplication and divergence, which has been extensively documented in other species [46, 52, 72].

The long-term evolution of satellite DNA catalogs in related species can be explained by the library hypothesis. Fundamentally, it states that changes in the profiles of satDNAs among species are mostly quantitative in the “library,” rather than multiple de novo origins [77]. Here, we could track the origin of the ancestral forms of satDNAs belonging to groups 1–4 to, at least, the common ancestor of Caimaninae (groups 1, 2, and 4) and Alligatoridae (group 3). We observed a substantial degree of similarity in satDNAs among species, with only four being species-specific. The long-term maintenance of satDNAs is notable. In this context, the conservation could be related to the acquisition of cellular function [42, 47–51, 77], particular genomic organization [32], or slow rates of evolution [52]. Previous studies found slow rates of molecular evolution within crocodylians [63]; thus, we hypothesize that satDNAs also evolved slowly in this group (as discussed below). In squamate reptiles, while the great majority of sequences are of recent origin and only observed in closely related species [78–83], several (and most common ones) are largely conserved in unrelated species [84].

The chromosomal mapping analysis revealed that all characterized satellites showed the general same chromosomal location (i.e., large peri- and centromeric blocks) among species, showing specific patterns for each one (Figs. 3, 4, 5, and 6 and Additional file 1: Figs. S5–S6). On the other hand, it is interesting to see that group 1 satellites, even being the most abundant in the Caimaninae genome, show a visible block of FISH signal in only two chromosomal pairs. When using the FISH technique, as a specific satDNA sequence can actually display a variety of array structures (dispersed and/or clustered into long and nonrandom arrangements) among species, it results in a range of labeling patterns at the chromosomal level. This is particularly true, for example, for the group 3 ClatSat05-40 because, although being abundant in the genome of *A. sinensis* (as indicated by our BLAST results in Additional file 2: Table S6 and clustermap using RepeatMasker data), it exhibits a non-cluster organization, which hindered in situ experiments from producing any detectable hybridization signals at the chromosomal level. We hypothesize that this could well explain the FISH

patterns observed in Caimaninae for group 1 satDNAs, although we cannot verify this as we do not have the complete sequence of their genomes nor are these satDNAs present in the *A. sinensis* genome for comparisons. In this context, different satellites of groups 1 and 2 show TSI values that are compatible with a dual organization, both forming loci visible by FISH and forming short arrays scattered throughout the genome not detectable by FISH (Table 1).

On the other hand, it is remarkable that two of the satellites studied in this paper (ClaSat02-6317 and ClaSat13-398), which appear to be dispersed according to BLAST and FISH results, are related to mobile elements and show homology of an important part of their sequences with such elements, which suggests that these satellites have evolved from this type of elements. Specifically, ClaSat02-6317 is related to ERVs, while ClaSat13-398 is related to LINES. There is increasing evidence that TEs are a major source of satellites (Šatović-Vukšić and Plohl, 2023) and these results support this proposal. Interestingly, it has been shown that the majority of within-crocodylian TE activity is derived from ERVs (Chong et al. 2014; Sotero-Caio et al. 2017). Our results therefore also support that these elements can constitute a source for satellites in Crocodylia.

Our current findings are in line with the karyotype patterns described for the family, which show a stable dichotomy between the genera *Alligator* ($2n=32$) and *Caiman*, *Melanosuchus*, and *Paleosuchus* ($2n=42$), with $2n=32$ representing the most likely ancestral state [revised in 66]. The two main divergent karyotype groups to which these reptiles belong are reflected in both the specificities of their respective satDNA libraries in terms of their sequence composition and chromosomal locations. However, all the satDNAs were mapped in the constitutive heterochromatin that is limited to the pericentromeric areas in all Alligatoridae species [66]. It is reasonable to consider that some of these satellites would be a component of the centromeric chromatin, much like in other species [4, 5]. Although the presence of multiple dispersed loci composed of a single copy or a few tandem copies of a satDNA is a fact today [23], the accumulation of satDNAs (as well as other repetitive DNA families) in centromeres and in heterochromatic regions is characteristic, as observed in many other groups [4, 23, 85–87]. Such colocalization (i.e., the tendency to occupy similar locations on non-homologous chromosomes) might have been facilitated by the reunion of centromeres at the first meiotic prophase bouquet [6, 88]. This is especially true in Caimaninae since the karyotypes of all species are dominated by acrocentric chromosomes. In this context, the existence of large and small chromosomes in Caimaninae could be favoring the structural differences

at the (peri)centromeric level between different chromosomes [89].

Both alligator species, *A. sinensis* and *A. mississippiensis*, displayed hybridization signals only for two (AsiSat01-1717 and AsiSat02-60) among all the investigated satDNAs (Fig. 6). Furthermore, AsiSat02-60 was exclusively mapped in all centromeres of both *Alligator* species. That is, these two species have conserved the same (peri)centromeric satDNA in all their chromosomes underscoring its possible important role in the centromeric and pericentromeric organization, a role that it may be shared with AsiSat01-1717 in some chromosomes. Alligatorinae long diverged (~70 Myr) from all the other Caimaninae and have highly rearranged karyotypes ($2n=32$) that are predominantly metacentric, in contrast with all Caimaninae species that have $2n=42$ chromosomes and karyotypes dominated by acrocentric chromosomes [66]. We have proposed that $2n=32$ represents the likely ancestral state and that the karyotype diversification in Caimaninae was followed by a series of Robertsonian rearrangements in which centric fissions played a key role [66]. Accordingly, alligators' satellitomes are among the smallest catalogs described until now for any species, with only 3 satDNAs identified.

Taking together the data obtained in this work, we can conclude that this group of ancient species that have survived on Earth for more than 100 Myr, has a very small common catalog of satDNA families. Nevertheless, each of the two lineages analyzed (Caimaninae and Alligatorinae), which have diverged for more than 70 Myr, is differentiated by the satDNAs that have been amplified in each group at the centromeric level. What stands out in this study is that these satellites have been conserved during all this time and persist for reasons that we have to analyze below. While the same satellite has been conserved in centromeres of Alligatorinae species for about 70 Myr, the chromosomal rearrangements that have taken place in the Caimaninae lineage would have caused the emergence and diversification of new satellite DNAs that have replaced them in the (peri)centromeric regions. Some of them, such as those of Group 3, were already present in a dispersed form in the ancestral genome of Alligatoridae, as was possibly the case with the satDNAs of Group 4 and the ungrouped satDNAs ClaSat02-6317 and ClaSat13-398 (Additional file 2: Table S2), still dispersed in all Alligatoridae species. In fact, the replacement of some satDNAs by others is common at the centromeric level even among closely related species in both animals and plants (reviewed in [4, 5, 90], see also the “Background” section). In the case of Alligatoridae, the slow evolution of their genomes may also be affecting in turn, as it was suggested for satDNAs from sturgeons [52, 53]. Extant crocodiles have limited rates of morphological [91, 92],

molecular [63], and karyotype diversification [66, 93, 94]. Likewise, the present-day satellitome (particularly the Caimaninae species) shares common satDNA libraries among its species, despite their long time of divergence. Therefore, the following questions arise: (a) why have they also changed so little in such a highly variable genome fraction over such an enormous span of time?; (b) would such low genetic, karyotype, and morphological variability be related to the low number of extant crocodylian species?

Crocodylomorpha (a clade that comprises living and extinct crocodylians) first appeared roughly 250 million years ago, and its 28 existing species are among the biggest living ectothermic animals. As a result, their survival over such a long geological period is of great evolutionary importance. They do, however, have a rich fossil history that includes hundreds of extinct species, revealing a hidden past of incredible variety and complexity [95, 96]. Oaks [65] has questioned the traditional notion of crocodiles as old “living fossils,” arguing that most extant crocodylians are remnants of formerly successful lineages in terms of diversity and range. Crocodylomorpha is the only pseudosuchians to have survived the Triassic-Jurassic (TJ) extinction event, which happened around 200 million years ago [97, 98]. Furthermore, after the mid-Miocene climatic optimum, there was a huge drop in crocodylian diversity, which coincided with global cooling and glacial advancement. During the Pliocene, the number of taxa is believed to have decreased from around 26 to 8, representing the greatest extinction rate over the previous 100 million years [99]. As a result, the selection of an “evolutionary package” with similar genomic, chromosomal, morphology, and physiology to what is currently observed among extant species most likely resulted from drastic demographic declines or founder events and represented evolutionary responses to a long-term bottleneck history.

Conclusions

This study is the first to offer a comparative mapping of the satDNA families in Alligatoridae. We observe some level of interspecific divergence even with so strong sequence conservatism through Caimaninae. With the results, we learn that satDNA orthologs indicate their evolutionary process according to different hybridization patterns. After rounds of mining, we discover the alligators' satellitomes are one of the smallest satDNA libraries described so far, with just four groups of satDNAs and four sequences species-specific between all species, possibly showing as ancestral features for the group, conserved throughout the crocodylians for a long time. With additional studies about repetitive DNAs in the other families of Crocodylia, it is important to demonstrate the

evolution of these sequences and provide more information about the chromosomal evolution in reptiles.

Methods

Samples, DNA extraction, and chromosomal preparation

Table 2 summarizes the collecting sites, number, and sex of individuals used in this investigation. The sampling is similar to that previously examined by [66]. In vitro blood cultures were used to obtain chromosomal preparations [100, 101]. The usual phenol–chloroform–isoamyl alcohol procedure was used to extract genomic DNA (gDNA) from blood stored in 100% ethanol [102].

Sequencing data

Two broad-snouted caimans *C. latirostris* and the Schneider's smooth-fronted caiman *P. trigonatus* were selected for low-pass shotgun sequencing on the BGISEQ-500 platform at BGI (BGI Shenzhen Corporation, Shenzhen, China), yielding 2.76, 2.76, and 2.67 Gb, respectively (Additional file 2: Table S7). Raw reads are available in the Sequence Read Archive from the NCBI (SRA-NCBI) under the accession numbers: SRR19901397 (*C. latirostris* male), SRR19901398 (*C. latirostris* female), SRR19901554 (*P. trigonatus* female). To search and compare satDNAs in other Alligatoridae species, we also collected genomic data available in the SRA-NCBI for the Yacare caiman *Caiman yacare* (SRR1609243), the black caiman *Melanosuchus niger* (SRR1609245) and for the Chinese alligator *Alligator sinensis* (SRR953089), thus encompassing all the extant Alligatoridae genera. The general features of sequencing data are summarized in Additional file 2: Table S8.

Satellite DNA characterization and comparative analyses

After gathering sequencing data for all the species as mentioned earlier, we performed a quality ($Q > 30$) and adapter trimming with Trimmomatic [103] for each library separately. After that, we proceeded to the characterization of satDNAs in each species. We performed several iterations of RepeatExplorer2 [36] and filtered the identified satDNAs with DeconSeq [104] following the protocol of [6]. We analyzed $2 \times 500,000$ reads in each iteration until no low- or high-confidence satellite DNA was found. After multiple iterations, we filtered and removed multigene families (5S rDNA and/or U snDNA) from the catalog. Then, we performed a similarity search among the remaining sequences with RepeatMasker using a custom python script (https://github.com/fjuizruano/ngs-protocols/blob/master/rm_homology.py), grouping them as the same sequence variant ($\geq 95\%$ of similarity), variant ($\geq 80\%$ of similarity) or different satDNA sharing a same superfamily ($\geq 50\%$ of similarity) in each species [6]. After that, we estimated Kimura's divergence, using Kimura 2-parameter model from the script calcDivergenceFromAlign.pl of RepeatMasker suite and abundance values for all satDNAs families with the "cross_match" option in RepeatMasker software [105], using $2 \times 5,000,000$ reads for each library, except for *Melanosuchus niger* and *Caiman yacare*, because their libraries had fewer reads, performing the analysis with $2 \times 1,213,376$ and $2 \times 1,608,245$, respectively (Table 1; Additional file 1: Fig. S1). Genomic abundance of each satDNA was given as the number of mapped reads in each satDNA divided by the number of analyzed nucleotides. Finally, we classified each satellite based on decreasing abundance order, as Ruiz-Ruano et al.

Table 2 Species, sample size (N), sex, and locality of the analyzed individuals. The species whose satellitomes were studied are highlighted in bold

Species	N	Locality/origin of samples
<i>Caiman crocodilus</i> (Spectacled caiman)	2♀, 2♂	Amazonas (BR) (Amazon Basin) 3° 22' 34.7" S 60° 19' 20.7" W
<i>Caiman latirostris</i> (Broad-snouted caiman)	4♀, 6♂	São Paulo (BR) (Cerrado) 22° 33' 53.1" S 48° 00' 35.2" W
<i>Caiman yacare</i> (Yacare caiman)	2♀, 8♂	Mato Grosso (BR) (Pantanal) 16° 19' 32.0" S 57° 46' 35.7" W
<i>Melanosuchus niger</i> (Black caiman)	2♀, 2♂	Amazonas (BR) (Amazon Basin) 3° 25' 50.4" S 66° 02' 35.0" W
<i>Paleosuchus palpebrosus</i> (Cuvier's dwarf caiman)	3♀, 3♂	Pará (BR) (Amazon Basin) 1° 18' 19.7" S 48° 19' 05.0" W
<i>Paleosuchus trigonatus</i> (Schneider's smooth-fronted caiman)	3♀, 4♂	Amazonas (BR) (Amazon Basin) 3° 06' 52.0" S 60° 01' 58.0" W
<i>Alligator mississippiensis</i> (American alligator)	2♀, 2♂	Canberra University collection (Australia)
<i>Alligator sinensis</i> (Chinese alligator)	4♀, 2♂	Private collections (Germany)

[6] suggested. The specific features of each satDNA are observed in Table 1. Each catalog of satDNAs was deposited on the GenBank with accession numbers OP169024–OP169026 (*A. sinensis*), OP169027–OP169032 (*C. yacare*), OP169033–OP169038 (*M. niger*), OP169039–OP169049 (*P. trigonatus*), and OP169050–OP169062 (*C. latirostris*). One additional and independent RepeatExplorer2 run was performed with a concatenated genomic library containing 150,000 reads from each species, using the “Perform comparative analysis” option.

To compare the satellitomes of multiple species, we performed a similarity search with RepeatMasker (https://github.com/fjruirozano/ngs-protocols/blob/master/rm_homology.py) considering all the de novo-characterized satDNA sequences. Then, we aligned the monomers of all satDNAs showing at least 50% similarity with MUSCLE [106]. In addition, we generated individual self-dotplots of the satDNA sequences and a general one with Flexidot [107].

For a general visualization of abundance and presence/absence of each satDNA in the different species, we ran RepeatMasker [105] against the complete catalog of Alligatoridae using each of the genomic libraries. After that, we normalized read coverage of the samples relative to single-copy genes. For this, we retrieved three single-copy genes in Sauropsida (options: Present in all species; Single-copy in all species) in the OrthoDB (<https://www.orthodb.org/>; accessed in July 30th) and mapped the genomic libraries against the genes using bowtie2 [108] with the preset values –sensitive and –local. Then, a normalization factor was calculated as: [(number of mapped reads × read sizes × gene sizes)/number of analyzed reads] (Additional file 2: Table S9). A final step of summing up the log₁₀ of normalized read counts from RepeatMasker (0 to 20% of Kimura divergence) was performed. With the final matrix (Additional file 2: Table S10), we generated a Clustermap (Fig. 1) with Seaborn using the seaborn.clustermap function (<https://seaborn.pydata.org/generated/seaborn.clustermap.html>).

Taking advantage of the fact that the genome of *Alligator sinensis* is available in the NCBI (GCA 000455745.1), we also conducted a BLAST (blastn, word size = 11, e-value = 1e-6) to search the entire list of satDNAs against this genome that was assembled using Illumina HiSeq2000 [67]. We did not perform any structural or quantitative analysis on array sizes and/or organization because only short reads were employed for this assembly [67]. As a result, BLAST searches provided more useful information on the presence or absence of satDNAs in the genome of *A. sinensis*. To get an estimation of the degree of tandem structure for the satDNAs in this species, we calculated the Tandem Structure Index (TSI), as demonstrated in [71]. This value is calculated as the

quotient of the number of paired reads mapped against a satDNA and the total number of reads (<https://github.com/fjruirozano/SatIntExt>). Thus, higher TSI values indicate the occurrence of longer arrays in the analyzed species. One must note that once the FISH probes were labeled and hybridized in groups, the TSI values are not completely suitable for comparison with the FISH results of satDNAs within groups 1 to 4.

Primer design and polymerase chain reaction (PCR)

We designed primer pairs for 12 satellite DNA families characterized from *C. latirostris* and two satellite DNA families characterized for *A. sinensis*, creating convergent primers for satellites larger than 1000 bp and divergent primers for satellites smaller than 1000 bp. We verified if those primer pairs anchors in conserved regions of monomers and used them to PCR-amplify in all Alligatoridae species. The PCRs contained 1 × PCR buffer, 1.5 mM of MgCl₂, 200 μM of each dNTP, 0.5 μL of each primer, 10–100 ng/μL of gDNA, and 0.2 μL of Taq DNA polymerase in a total volume of 25 μL. The PCR program included an initial denaturation at 95 °C for 7 min, followed by 34 cycles at 95 °C for 45 s, 61 °C for 1 min, 72 °C for 1 min, and a final extension at 72 °C for 7 min. The PCR products were checked in 2% agarose gel.

Probe labeling and fluorescence in situ hybridization (FISH)

Except for ClaSat03-183, ClaSat09-285, ClaSat12-24, and AsiSat03-96, all the other satDNAs were successfully amplified and the PCR products were labeled with Atto550-dUTP (red) or Atto488-dUTP (green) according to the manufacturer’s recommendations using the Nick-Translation mix kit (Jena Bioscience, Jena, Germany). The probes were then hybridized in all other Alligatoridae species according to the methodology reported by [109]. To corroborate the FISH results, at least 30 metaphase spreads were examined in each individual. Photos were obtained with CoolSNAP on an Olympus BX50 microscope (Olympus Corporation, Ishikawa, Japan), and the images were processed using Image-Pro Plus 4.1 software (Media Cybernetics, Silver Spring, MD, USA).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-024-01847-8>.

Additional file 1: Figures S1–S6. Fig S1—All-against-all dotplot of Caimaninae satellite DNA, with the division of these sequences among the three groups presented in the analysis in green (group 1), red (group 2), and blue (group 3), and the similarity among each satDNA family, represented by a white to the black color ladder. **Fig S2**—This figure shows the alignment between two consecutive repeats and the next two. (a) The scheme shows the four 41 bp repeats (A, B, C and D) aligned two by two with their corresponding 21/20 bp subunits (alpha and beta) distinguished by being shaded in gray (alpha) or blue (beta). Asterisks

indicate similarity between aligned sequences. Nucleotide positions that are divergent between alpha and beta subunits of each repeat unit have been marked in red. (b) The alignment of the four 41 bp repetitive units (A, B, C and D) is shown on one side and the separate alignments of the alpha (gray) and beta (blue) subunits are shown on the other. Asterisks indicate similarity between aligned sequences. Divergent nucleotides are not shaded. (c) Multiple alignment of all alpha and beta subunits. Asterisks indicate similarity between aligned sequences. Divergent nucleotides are not shaded. Consistent with Table S2, it can be observed in Figure S2 that inter-repeat alignments show more conserved nucleotide positions than intra-repeat ones. In addition, it can be observed that the most divergent part between alpha and beta subunits occurs at the 3' end. **Fig S3**—The figure is an example of the internal organization of long satellites. In this case, ClaSat06-1063. As can be seen, the repetitive unit of this satellite is made up of 12 subrepeats of 41 bp between which the average divergence is 0.30, followed by a short intervening sequence of 23 bp and then 6 repeats of a sequence of 81 bp (41 + 40 bp) with an average divergence between subrepeats of 0.35. Finally, a 62 bp fraction of the latter subrepeats. Therefore, the repetitive unit of this satellite has evolved through different cycles of duplication and divergence first from 40 bp subunits (but not 20 bp) and then from 80 bp subunits through a complex process in which partial sequences of the 40 bp subunit have been interspersed (pointing to unequal crossing over as a molecular tool towards the consolidation of a current 1063 bp unit). In red, the average sequence of 41 bp subunits and 81 bp subunits. **Fig S4**—Phylogenetic tree (minimum spanning tree, mst) comparing alpha (red) and beta (blue) subunits of the ClaSat05-40 satellite. We can observe two distinct clades, one formed by alpha sequences and others by beta sequences. **Fig S5**—Metaphase chromosomes from *C. crocodilus* (a–d), *C. latirostris* (e–h) and *C. yacare* (i–l) after in situ mapping with satDNA probes belonging to group 1 (ClaSat06-1063; ClaSat07-320; ClaSat08-800 and ClaSat11-547). The satDNA FISH signals are highlighted in green (ATTO488 labeled) or red (ATTO550 labeled) and the chromosomes were counterstained with DAPI (blue). Scale bar = 20 µm. **Fig S6**—Metaphase chromosomes from *M. niger* (a–d), *P. palpebrosus* (e–h) and *P. trigonatus* (i–l) after in situ mapping with satDNA probes belonging to group 1 (ClaSat06-1063; ClaSat07-320; ClaSat08-800 and ClaSat11-547). The satDNA FISH signals are highlighted in green (ATTO488 labeled) or red (ATTO550 labeled) and the chromosomes were counterstained with DAPI (blue). Scale bar = 20 µm.

Additional file 2: Tables S1–S10. Table S1—Within-group pairwise similarity of satDNAs. **Table S2**—Summary of the results obtained from the read clustering analysis with multiple species analyzed simultaneously. **Table S3**—We have compared hundreds of repetitive units from each satellite and compared the mean divergence that exists between them (inter-repeat divergence) with the mean divergence that exists between the subunits that compose each repetitive unit (intra-repeat divergence). For the case of the sequences of the Group 1 constituted by repetitive units of 41 bp, we have considered each of its subunits (alpha and beta) of 21 and 20 bp, respectively, obtaining that the intra-repeat divergence was always greater than the inter-repeat divergence. **Table S4**—As for the Group 1 satellites, we have analyzed the 60 bp long satellites of Group 2. Here we verified that their repetitive units were made up of subunits of 20 bp each, which we have named alpha, beta, and gamma. The table shows the analysis of intra-repeats versus inter-repeats divergences, and it can also be seen here that the former is greater than the latter. **Table S5**—While Groups 1 and 2 were more homogeneous, Group 3 sequences are characterized as heterogeneous ones, as their sequences were derived from a repetition of two subunits of 29 bp and 11 bp. This table demonstrates that the divergence between alpha (29 bp) and beta (11 bp) subrepeats is greater than the divergence between complete repeats of these satellites. Thus, looking at their structure, we can affirm that the Group 3 satellites are not formed by repetitive units of 20 bp. In any case, it is possible that the ancestral satellite that gave rise to the group 3 satellites had a length of 29 bp and that later rearrangements, possibly mediated by unequal crossing over (see, for example, Navajas-Pérez et al. 2005 for RAYSI satDNA), generated a 40 bp repetitive unit formed by one unit plus one

third of another, being now the unit of homogenization of these satellites.

Table S6—Summary of the BLAST searches against Alligator sinensis genome. **Table S7**—Main features of the analyzed genomic libraries.

Table S8—Results of RepeatMasking against the Repbase database.

Table S9—Read mapping against single-copy genes for normalization of satDNA abundance. **Table S10**—log10 of normalized abundances of each satDNA, considering RepeatMasker counts from 0 to 20% of divergence.

Acknowledgements

We also would like to thank the Crocodile working group of the German Society for Herpetology and Terrarium Science Association (DGHT AG Krokodile) for providing contact to keepers of alligators. We are grateful to Patrik Ferreira Viana, Hugmar Pains da Silva, Milena, Breno Almeida, Alexander Meurer, and Sebastian Scholz for all the valuable support provided with sampling.

Authors' contributions

M.B.C and R.U conceived the study. V.C.S.O., R.Z.S., C.A.G.G., R.M.C., M.A.G.R., and M.A. performed experiments. V.C.S.O., R.Z.S., C.A.G.G., R.M.C., M.A.G.R., M.A., T.E., T.L., F.P., R.U., and M.B.C. analyzed the data. V.C.S.O., R.Z.S., R.U., and M.B.C wrote the manuscript. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. V.C.S.O was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (401036/2022–7). M.A. was supported by the Charles University Research Centre program No. UNCE/24/SCI/006 and by Czech Science Foundation Project No. 20-27236 J. M.d.B.C. was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (302928/2021–9) and by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) (2023/00955–2). M.d.B.C. and T.L. were supported by Alexander von Humboldt Foundation (Research Group Linkage Programme). We acknowledge support by the German Research Foundation Projekt-Nr. 512648189 (T.L.) and the Open Access Publication Fund of the Thuringer Universitaets- und Landesbibliothek Jena. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Finance Code 001.

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available in the NCBI database (<https://www.ncbi.nlm.nih.gov/bioproject/>) under accession numbers OP169024–OP169026 (*A. sinensis*), OP169027–OP169032 (*C. yacare*), OP169033–OP169038 (*M. niger*), OP169039–OP169049 (*P. trigonatus*), and OP169050–OP169062 (*C. latirostris*) and, in the SRA-NCBI for the Yacare caiman *Caiman yacare* (SRR1609243), the black caiman *Melanosuchus niger* (SRR1609245), and for the Chinese alligator *Alligator sinensis* (SRR953089).

All other data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Blood samples were collected from free-living South American animals with the permission of the environmental agencies ICMBIO/SISBIO (License no 71857–7) and SISGEN (ABFF266). *Alligator sinensis* blood samples were obtained from lawfully housed animals in Europe (CITES certificate numbers EU 0228–1057/14, ES-CC-00041/07C, ES-CC-00036/07C, 50721–18, DE-DA190814-5, DE-DA190814-6). There were no major injuries to the animals, and all free-living individuals were returned to their respective collecting sites.

Institutional review board statement

The study was conducted according to the guidelines of the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos (Process number CEUA 4617090919). Collections were done under the authorization of the Chico Mendes Institute for Biodiversity Conservation (ICMBIO), System of Authorization and Information about Biodiversity (SISBIO-License No. 71857–7), and National System of Genetic Resource Management

and Associated Traditional Knowledge (SISGEN-ABFF266). Blood samples from *Alligator sinensis* came from animals legally kept in Europe (CITES certificate number EU 0228–1057/14, ES-CC-00041/07C, ES-CC-00036/07C, 50721–18, DE-DA190814-5, DE-DA190814-6).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 11 August 2023 Accepted: 15 February 2024

Published online: 27 February 2024

References

- Biémont C, Vieira C. Junk DNA as an evolutionary force. *Nature*. 2006. <https://doi.org/10.1038/443521a>.
- López-Flores I, Garrido-Ramos MA. The repetitive DNA content of eukaryotic genomes. In: Garrido-Ramos MA, editor. *Repetitive DNA*. Basel: Karger; 2012. p. 1–28.
- Plohl M, Meštrović N, Mravinac B. Centromere identity from the DNA point of view. *Chromosoma*. 2014. <https://doi.org/10.1007/s00412-014-0462-0>.
- Garrido-Ramos MA. Satellite DNA: An evolving topic. *Genes*. 2017. <https://doi.org/10.3390/genes8090230>.
- Garrido-Ramos MA. The genomics of plant satellite DNA. In: Ugarković Đ, editor. *Satellite DNAs in Physiology and Evolution*, Progress in Molecular and Subcellular Biology. Springer: Cham; 2021. p. 103–44.
- Ruiz-Ruano F, López-Leon MD, Cabrero J, Camacho JPM. High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci Rep*. 2016. <https://doi.org/10.1038/srep28333>.
- Csink AK, Henikoff S. Something from nothing: the evolution and utility of satellite repeats. *Trends Genet*. 1998. [https://doi.org/10.1016/S0168-9525\(98\)01444-9](https://doi.org/10.1016/S0168-9525(98)01444-9).
- Kuhn GCS, Sene FM, Moreira-Filho O, Schwarzacher T, Heslop-Harrison JS. Sequence analysis, chromosomal distribution and long-range organization show that rapid turnover of new and old pBuM satellite DNA repeats leads to different patterns of variation in seven species of the *Drosophila buzzatii* cluster. *Chromosome Res*. 2008. <https://doi.org/10.1007/s10577-007-1195-1>.
- Kuhn GCS, Schwarzacher T, Heslop-Harrison JS. The non-regular orbit: three satellite DNAs in *Drosophila martensis* (*buzzatii* complex, *repleta* group) followed three different evolutionary pathways. *Mol Genet Genom*. 2010. <https://doi.org/10.1007/s00438-010-0564-1>.
- Plohl M, Meštrović N, Mravinac B. Satellite DNA evolution. In: Garrido-Ramos MA, editor. *Repetitive DNA*. Basel: Karger; 2012. p. 126–52.
- Feliciello I, Akrap I, Ugarković Đ. Satellite DNA modulates gene expression in the beetle *Tribolium castaneum* after heat stress. *PLoS Genet*. 2015. <https://doi.org/10.1371/journal.pgen.1005466>.
- Prakhongcheep O, Thapana W, Suntronpong A, Singchat W, Pattanatanang K, Phatcharakullawarawat R, Muangmai N, Peyachoknagul S, Matsubara K, Ezaz T, Srikulnath K. Lack of satellite DNA species-specific homogenization and relationship to chromosomal rearrangements in monitor lizards (Varanidae, Squamata). *BMC Evol Biol*. 2017. <https://doi.org/10.1186/s12862-017-1044-6>.
- Palacios-Gimenez OM, Dias GB, de Lima LG, Campos G, Kuhn S, Ramos E, Martins C, Cabral-de-Mello DC. High-throughput analysis of the satellitome revealed enormous diversity of satellite DNAs in the neo-Y chromosome of the cricket *Eneoptera surinamensis*. *Sci Rep*. 2017. <https://doi.org/10.1038/s41598-017-06822-8>.
- Feliciello I, Chinali G, Ugarković Đ. Structure and population dynamics of the major satellite DNA in the red flour beetle *Tribolium castaneum*. *Genetica*. 2011. <https://doi.org/10.1007/s10709-011-9601-1>.
- Kuhn GCS, Küttler H, Moreira-Filho O, Heslop-Harrison JS. The 1.668 repetitive DNA of *Drosophila*: Concerted evolution at different genomic scales and association with genes. *Mol Biol Evol*. 2012. <https://doi.org/10.1093/molbev/msr173>.
- Brajković J, Feliciello I, Bruvo-Madarić B, Ugarković Đ. Satellite DNA-Like elements associated with genes within euchromatin of the beetle *Tribolium castaneum*. *G3: Genes, Genomes, Genetics* 2012. <https://doi.org/10.1534/g3.112.003467>.
- Larracuent AM. The organization and evolution of the *Responder* satellite in species of the *Drosophila melanogaster* group: dynamic evolution of a target of meiotic drive. *BMC Ecol Evol*. 2014. <https://doi.org/10.1186/s12862-014-0233-9>.
- Pavlek M, Gelfand Y, Plohl M, Meštrović N. Genome-wide analysis of tandem repeats in *Tribolium castaneum* genome reveals abundant and highly dynamic tandem repeat families with satellite DNA features in euchromatic chromosomal arms. *DNA Res*. 2015. <https://doi.org/10.1093/dnares/dsv021>.
- Pita S, Panzera F, Mora P, Vela J, Cuadrado A, Sánchez A, Palomeque T, Lorite P. Comparative repeatome analysis on *Triatoma infestans* Andean and Non-Andean lineages, main vector of Chagas disease. *PLoS ONE*. 2017. <https://doi.org/10.1371/journal.pone.0181635>.
- de Lima LG, Svartman M, Kuhn GCS. Dissecting the satellite DNA landscape in three cactophilic *Drosophila* sequenced genomes. *G3*. 2017. <https://doi.org/10.1534/g3.117.042093>.
- Robledillo LÁ, Neumann P, Kobližková A, Novák P, Vrbová I, Macas J. Extraordinary sequence diversity and promiscuity of centromeric satellites in the legume tribe *Fabeae*. *Mol Biol Evol*. 2020. <https://doi.org/10.1093/molbev/msaa090>.
- Milani D, Ruiz-Ruano FJ, Camacho JPM, Cabral-de-Mello DC. Out of patterns, the euchromatic B chromosome of the grasshopper *Abacris flavolineata* is not enriched in high-copy repeats. *Heredity*. 2021. <https://doi.org/10.1038/s41437-021-00470-5>.
- Šatović-Vukšić E, Plohl M. Satellite DNAs-From localized to highly dispersed genome components. *Genes*. 2023. <https://doi.org/10.3390/genes14030742>.
- Joshi SS, Meller VH. Satellite repeats identify X chromatin for dosage compensation in *Drosophila melanogaster* males. *Curr Biol*. 2017. <https://doi.org/10.1016/j.cub.2017.03.078>.
- Rošić S, Köhler F, Erhardt S. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J Cell Biol*. 2014. <https://doi.org/10.1083/jcb.201404097>.
- Ugarković Đ. Functional elements residing within satellite DNAs. *EMBO Rep*. 2005. <https://doi.org/10.1083/jcb.201404097>.
- Pathak R, Mamillapalli A, Rangaraj N, Kumar R, Vasanthi D, Mishra K, Mishra R. AAGAG repeat RNA is an essential component of nuclear matrix in *Drosophila*. *RNA Biol*. 2013. <https://doi.org/10.4161/ra.24326>.
- Jagannathan M, Cummings R, Yamashita YM. A conserved function for pericentromeric satellite DNA. *eLife* 2018. <https://doi.org/10.7554/eLife.34122>.
- Jagannathan M, Cummings R, Yamashita YM. The modular mechanism of chromocenter formation in *Drosophila*. *eLife* 2019. <https://doi.org/10.7554/eLife.43938>.
- Graphodatsky AS. Comparative chromosomics. *Mol Biol*. 2007;41:361–75.
- Deakin JE, Potter S, O'Neill R, Ruiz-Herrera A, Cioffi MB, Eldridge MDB, Fukui K, Graves JAM, Griffin D, Grutzner F, Kratochvíl L, Miura I, Rovatsos M, Srikulnath K, Wapstra E, Ezaz T. Chromosomics: Bridging the gap between genomes and chromosomes. *Genes*. 2019. <https://doi.org/10.3390/genes10080627>.
- dos Santos RZ, Calegari RM, Silva DMZA, Ruiz-Ruano FJ, Melo S, Oliveira C, Foresti F, Uliano-Silva M, Porto-Foresti F, Utsunomia R. A long-term conserved satellite DNA that remains unexpanded in several genomes of Characiformes fish is actively transcribed. *Genome Biol Evol*. 2021. <https://doi.org/10.1093/gbe/evab002>.
- Rovatsos M, Kratochvíl L, Altmanová M, Pokorná MJ. Interstitial telomeric motifs in squamates reptiles: When the exceptions outnumber the rule. *PLoS ONE*. 2015. <https://doi.org/10.1371/journal.pone.0134985>.
- Kretschmer R, Goes CAG, Bertollo LAC, Ezaz T, Porto-Foresti F, Toma GA, Utsunomia R, Cioffi MB. Satellitome analysis illuminates the evolution of ZW sex chromosomes of Triportheidae fishes (Teleostei: Characiformes). *Chromosoma*. 2022. <https://doi.org/10.1007/s00412-022-00768-1>.
- Novák P, Neumann JP, Steinhaisl J, Macas J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic

- repetitive elements from next-generation sequence reads. *Bioinform.* 2013. <https://doi.org/10.1093/bioinformatics/btt054>.
36. Novák P, Robledillo LA, Koblížková A, Vrbová I, Neumann P, Macas J. TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res.* 2017. <https://doi.org/10.1093/nar/gkx257>.
 37. Harris RS, Cechova M, Makova KD. Noise-cancelling repeat finder: uncovering tandem repeats in error-prone long-read sequencing data. *Bioinform.* 2019. <https://doi.org/10.1093/bioinformatics/btz484>.
 38. Vondrak T, Robledillo LA, Novák P, Koblížková A, Neumann P, Macas J. Characterization of repeat arrays in ultra-long nanopore reads reveals frequent origin of satellite DNA from retrotransposon-derived tandem repeats. *Plant J.* 2020. <https://doi.org/10.1111/tj.14546>.
 39. Silva DMZ de A, Utsunomia R, Ruiz-Ruano FJ, Daniel SN, Porto-Foresti F, Hashimoto DT, Oliveira C, Camacho JPM, Foresti F. High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*. *Sci Rep.* 2017. <https://doi.org/10.1038/s41598-017-12939-7>.
 40. Kirov I, Gilyok M, Knyazev A, Fesenko I. Pilot satellitome analysis of the model plant, *Physcomitrella patens*, revealed a transcribed and high-copy IGS related tandem repeat. *Comp Cytogenet.* 2018. <https://doi.org/10.3897/CompCytogen.v12i4.31015>.
 41. Suárez-Santiago VN, Blanca G, Ruiz-Rejón M, Garrido-Ramos MA. Satellite-DNA evolutionary patterns under a complex evolutionary scenario: The case of Acrolophus subgroup (Centaurea L., Compositae) from the western Mediterranean. *Gene.* 2007. <https://doi.org/10.1016/j.gene.2007.09.001>
 42. Chaves R, Ferreira D, Mendes-da-Silva A, Meles S, Adegas F. FA-SAT Is an Old Satellite DNA Frozen in Several Bilateria Genomes. *Genome Biol Evol.* 2017. <https://doi.org/10.1093/gbe/evx212>.
 43. Peona V, Kutschera VE, Blom MPK, Irestedt M, Suh A. Satellite DNA evolution in Corvoidea inferred from short and long reads. *Mol Ecol.* 2022. <https://doi.org/10.1111/mec.16484>.
 44. Goes CAG, dos Santos RZ, Aguiar WRC, Alves DCV, Silva DMZA, Foresti F, Oliveira C, Utsunomia R, Porto-Foresti F. Revealing the satellite DNA history in *Psalidodon* and *Astyanax* Characid fish by comparative satellitomics. *Front Genet.* 2022. <https://doi.org/10.3389/fgene.2022.884072>.
 45. Henikoff S, Ahmad K, Malik HS. The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science.* 2001. <https://doi.org/10.1126/science.1062939>.
 46. Navajas-Pérez R, de la Herrán R, Jamilena M, Lozano R, Rejón CR, Rejón MR, Garrido-Ramos MA. Reduced rates of sequence evolution of Y-linked satellite DNA in *Rumex* (Polygonaceae). *J Mol Evol.* 2005. <https://doi.org/10.1007/s00239-004-0199-0>.
 47. Mravinac B, Plohl M, Meštrović N, Ugarković Đ. Sequence of PRAT satellite DNA "frozen" in some Coleopteran species. *J Mol Evol.* 2002. <https://doi.org/10.1007/s00239-001-0079-9>.
 48. Garrido-Ramos MA, Jamilena M, Lozano R, Rejón CR, Rejón MR. The *EcoRI* centromeric satellite DNA of the Sparidae family (Pisces, Perciformes) contains a sequence motive common to other vertebrate centromeric satellite DNAs. *Cytogenet Genome Res.* 1995. <https://doi.org/10.1159/000134137>.
 49. Mravinac B, Plohl M, Ugarković Đ. Preservation and high sequence conservation of satellite DNAs suggest functional constraints. *J Mol Evol.* 2005. <https://doi.org/10.1007/s00239-004-0342-y>.
 50. Schueler MG, Swanson W, Thomas PJ, NISC Comparative Sequencing Program, Green ED. Adaptive evolution of foundation kinetochore proteins in primates. *Mol Biol Evol.* 2010. <https://doi.org/10.1093/molbev/msq043>.
 51. Fachinetti D, Han JS, McMahon MA, Ly P, Abdullah A, Wong AJ. DNA Sequence-specific binding of CENP-B enhances the fidelity of human centromere function. *Dev Cell.* 2015. <https://doi.org/10.1016/j.devcel.2015.03.020>.
 52. de la Herrán R, Fontana F, Lanfredi M, Congiu L, Leis M, Rossi R, Ruiz Rejón C, Ruiz Rejón M, Garrido-Ramos MA. Slow Rates of Evolution and Sequence Homogenization in an Ancient Satellite DNA Family of Sturgeons. *Mol Biol Evol.* 2001;18(3):432–6. <https://doi.org/10.1093/oxfordjournals.molbev.a003820>.
 53. Robles F, de la Herrán R, Ludwig A, Rejón CR, Rejón MR, Garrido-Ramos MA. Evolution of ancient satellite DNAs in sturgeon genomes. *Gene.* 2004. <https://doi.org/10.1016/j.gene.2004.06.001>.
 54. Plohl M, Petrović V, Luchetti A, Ricci A, Šatović E, Passamonti M, Mantovani B. Long-term conservation vs high sequence divergence: the case of an extraordinarily old satellite DNA in bivalve mollusk. *Heredity.* 2010. <https://doi.org/10.1038/hdy.2009.141>.
 55. Lorite P, Muñoz-López M, Carrillo JA, Sanllorente O, Mora P, Tinaut A, Torres MI, Palomeque T. Concerted evolution, a slow process for ant satellite DNA: study of the satellite DNA in the *Aphaenogaster* genus (Hymenoptera, Formicidae). *Org Divers Evol.* 2017. <https://doi.org/10.1038/hdy.2009.141>.
 56. Halbach R, Miesen P, Joosten J, Taşköprü E, Rondeel I, Pennings B, Vogels CBF, Merklings SH, Joenraadt CJ, Lambrechts L, van Riji RP. A satellite repeat-derived piRNA controls embryonic development of *Aedes*. *Nature.* 2020. <https://doi.org/10.1038/s41586-020-2159-2>.
 57. Brochu CA. Phylogenetic approaches toward Crocodylian history. *Annu Rev Earth Planet Sci.* 2003. <https://doi.org/10.1146/annurev.earth.31.100901.141308>.
 58. Grigg G, Seebacher F, Franklin CE. *Crocodylian Biology and Evolution*. 1st ed. Surrey Beatty: Chipping Norton; 2001.
 59. Brochu CA, Sumrall CD, Theodor JM. When clocks (and communities) collide: Estimating divergence time from molecules and the fossil record. *J Paleontol.* 2004. [https://doi.org/10.1666/0022-3360\(2004\)078%3c0001:WCACCE%3e2.0.CO;2](https://doi.org/10.1666/0022-3360(2004)078%3c0001:WCACCE%3e2.0.CO;2).
 60. Bronzati M, Montefeltro FC, Langer MC. Diversification events and the effects of mass extinction on Crocodyliformes evolutionary history. *R Soc Open Sci.* 2015. <https://doi.org/10.1098/rsos.140385>.
 61. Janke A, Arnason U. The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent Archosauria (birds and crocodiles). *Mol Biol Evol.* 1997. <https://doi.org/10.1093/oxfordjournals.molbev.a025736>.
 62. Iwabe N, Hara Y, Kumazawa Y, Shibamoto K, Saito Y, Miyata T, Katoh K. Sister group relationship of turtles to the bird-crocodylian clade revealed by nuclear DNA-coded proteins. *Mol Biol Evol.* 2005. <https://doi.org/10.1093/molbev/msi075>.
 63. Green RE, Braun EL, Armstrong J, Earl D, Nguyen N, Hickey G, Vandewege MW, St John JA, Capella-Gutiérrez S, Castoe TA. Three crocodylian genomes reveal ancestral patterns of evolution among archosaurs. *Science.* 2014. <https://doi.org/10.1126/science.1254449>.
 64. The Reptile Database. Richmond, USA. 1995. <http://www.reptile-database.org> Accessed 29 March 2023.
 65. Oaks JR. A time-calibrated species tree of Crocodylia reveals a recent radiation of the true crocodiles. *Evolution.* 2011. <https://doi.org/10.1111/j.1558-5646.2011.01373.x>.
 66. Oliveira VCS, Altmanová M, Viana PF, Ezaz T, Bertollo LAC, Ráb P, Liehr T, Al-Rikabi A, Feldberg E, Hatanaka T, et al. Revisiting the karyotypes of Alligators and Caimans (Crocodylia, Alligatoridae) after a half-century delay: Bridging the gap in the chromosomal evolution of Reptiles. *Cells.* 2021. <https://doi.org/10.3390/cells10061397>.
 67. Wan Q-H, Pan S-K, Hu L, Zhu Y, Xu P-W, Xia J-Q, Chen H, He G-Y, He J, Ni X-W, et al. Genome analysis and signature discovery for diving and sensory properties of the endangered Chinese alligator. *Cell Res.* 2013. <https://doi.org/10.1038/cr.2013.104>.
 68. Rice ES, Kohno S, John JS, Pham S, Howard J, Lareau LF, O'Connell BL, Hickey G, Armstrong J, Deran A, et al. Improved genome assembly of American alligator genome reveals conserved architecture of estrogen signaling. *Genome Res.* 2017. <https://doi.org/10.1101/gr.213595.116>.
 69. Romanenko SA, Prokopov DY, Proskuryakova AA, Davletshina GI, Tupikin AE, Kasai F, Ferguson-Smith MA, Trifonov VA. The Cytogenetic Map of the Nile Crocodile (*Crocodylus niloticus*, Crocodylidae, Reptilia) with Fluorescence In Situ Localization of Major Repetitive DNAs. *Int J Mol Sci.* 2022. <https://doi.org/10.3390/ijms23113063>.
 70. Kawagoshi T, Nishida C, Ota H, Kumazawa Y, Endo H, Matsuda Y. Molecular structures of centromeric heterochromatin and karyotypic evolution in the Siamese crocodile (*Crocodylus siamensis*) (Crocodylidae, Crocodylia). *Chromosome Res.* 2008. <https://doi.org/10.1007/s10577-008-1263-1>.
 71. Camacho JPM, Cabrero J, López-León MD, Martín-Peciña M, Perfectti F, Garrido-Ramos MA, Ruiz-Ruano FJ. Satellitome comparison of two oedipode grasshoppers highlights the contingent nature of satellite DNA evolution. *BMC Biology* 2022. <https://doi.org/10.1186/s12915-021-01216-9>.

72. Ruiz-Ruano FJ, Navarro-Domínguez B, Camacho JPM, Garrido-Ramos MA. Characterization of the satellitome in lower vascular plants: the case of the endangered fern *Vandenboschia speciosa*. *Ann Bot.* 2019. <https://doi.org/10.1093/aob/mcy192>.
73. Montiel EE, Mora P, Rico-Porras JM, Palomeque T, Lorite P. Satellitome of the red palm weevil, *Rhynchophorus ferrugineus* (Coleoptera: Curculionidae), the most diverse among insects. *Front Ecol Evol.* 2022. <https://doi.org/10.3389/fevo.2022.826808>.
74. Utsunomia R, Silva DMZA, Ruiz-Ruano FJ, Goes CAG, Melo S, Ramos LP, Oliveira C, Porto-Foresti F, Foresti F, Hashimoto DT. Satellitome landscape analysis of *Megaleporinus macrocephalus* (Teleostei, Anostomidae) reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome. *Scientific Rep.* 2019. <https://doi.org/10.1038/s41598-019-42383-8>.
75. Ruiz-Ruano FJ, Cabrero J, López-León MD, Sánchez A, Camacho JPM. Quantitative sequence characterization for repetitive DNA content in the supernumerary chromosome of the migratory locust. *Chromosoma.* 2018. <https://doi.org/10.1007/s00412-017-0644-7>.
76. Valeri MP, Dias GB, Pereira VS, Kuhn GCS, Svartman M. An eutherian intronic sequence gave rise to a major satellite DNA in Platyrrhini. *Biol Lett.* 2018. <https://doi.org/10.1098/rsbl.2017.0686>.
77. Fry K, Salser W. Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell.* 1977. [https://doi.org/10.1016/0092-8674\(77\)90170-2](https://doi.org/10.1016/0092-8674(77)90170-2).
78. Capriglione T. Repetitive DNA as a tool to study the phylogeny of cold-blooded vertebrates. In: Olmo E, Redi CA, editors. *Chromosomes Today*. Switzerland: Birkhäuser; 2000. p. 183–94.
79. Capriglione T, Cardone A, Odierna G, Olmo E. Further data on the occurrence and evolution of satellite DNA families in the lacertid genome. *Chromosome Res.* 1994. <https://doi.org/10.1007/BF01552726>.
80. Rudykh IA, Grechko VV, Kramerov DA, Darevskii IS. Distribution of HindIII-repeats in genomes of Caucasian lizards of the Lacerta species reflect their phylogenetic affiliation. *Dokl Akad Nauk.* 1999;367:563–6.
81. Ciobanu DG, Grechko VV, Darevsky IS. Molecular evolution of satellite DNA CLS at in lizards from the genus *Darevskia* (Sauria: Lacertidae): Correlation with species diversity. *Russ J Genet.* 2003;39:1292–305.
82. Ciobanu DG, Grechko VV, Darevsky IS, Kramerov DA. New satellite DNA in *Lacerta s. str.* lizards (Sauria: Lacertidae): Evolutionary pathways and phylogenetic impact. *J Exp Zool.* 2004 <https://doi.org/10.1002/jez.b.21014>.
83. Grechko VV, Ciobanu DG, Darevsky IS, Kramerov DA. Satellite DNA of lizards of the genus *Lacerta s. str.* (the group *L. agilis*), the family Lacertidae. *Dokl Biochem Biophys.* 2005 <https://doi.org/10.1007/s10628-005-0029-3>.
84. Olmo E, Capriglione T, Odierna G. Different genomic evolutionary rates in the various reptile lineages. *Gene.* 2002. [https://doi.org/10.1016/S0378-1119\(02\)00685-6](https://doi.org/10.1016/S0378-1119(02)00685-6).
85. Stornioli JHF, Goes CAG, Calegari RM, dos Santos RZ, Giglio LM, Foresti F, Oliveira C, Penitente M, Porto-Foresti F, Utsunomia R. The B chromosomes of *Prochilodus lineatus* (Teleostei, Characiformes) are highly enriched in satellite DNAs. *Cells.* 2021. <https://doi.org/10.3390/cells10061527>.
86. Iwata A, Tek AL, Richard MMS, Abernathy B, Fonsêca A, Schmutz J, Chen NWG, Thareau V, Magdelenat G, Li Y, et al. Identification and characterization of functional centromeres of the common bean. *Plant J.* 2013. <https://doi.org/10.1111/tpj.12269>.
87. Shang W-H, Hori T, Toyoda A, Kato J, Popendorf K, Sakakibara A, Fukagawa T. Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. *Genome Res.* 2010. <https://doi.org/10.1101/gr.106245.110>.
88. Mravinac B, Plohl M. Parallelism in evolution of highly repetitive DNAs in sibling species. *Mol Biol Evol.* 2010. <https://doi.org/10.1093/molbev/msq068>.
89. Lanfredi M, Congiu L, Garrido-Ramos MA, de la Herrán R, Leis M, Chicca M, Rossi R, Tagliavini J, Rejón CR, Rejón MR, et al. Chromosomal location and evolution of a satellite DNA family in seven sturgeon species. *Chromosome Res.* 2001. <https://doi.org/10.1023/a:1026739616749>.
90. Garrido-Ramos MA. Satellite DNA in Plants: More than just rubbish. *Cytogenet Genome Res.* 2015. <https://doi.org/10.1159/000437008>.
91. Mook C. The skull characters of *Crocodylus megarhinus* Andrews. *Am Mus Novit.* 1927;289:1–8.
92. Sill WD. The Zoogeography of the Crocodylia. *Copeia.* 1968;1:76–88.
93. Cohen MM, Clark HF. The somatic chromosomes of five crocodylian species. *Cytogenet.* 1967. <https://doi.org/10.1159/000129941>.
94. Cohen MM, Gans C. The chromosomes of the Order Crocodylia. *Cytogenet.* 1970. <https://doi.org/10.1159/000130080>.
95. Mannion PD, Chiarenza AA, Godoy PL, Cheah YN. Spatiotemporal sampling patterns in the 230 million year fossil record of terrestrial crocodylomorphs and their impact on diversity. 2019. *Palaeontol.* <https://doi.org/10.5061/dryad.668950m>.
96. Stubbs TL, Pierce SE, Elsler A, Anderson PSL, Rayfield EJ, Benton MJ. Ecological opportunity and the rise and fall of crocodylomorph evolutionary innovation. *Proc Royal Soc B.* 2021. <https://doi.org/10.1098/rspb.2021.0069>.
97. Brusatte SL, Benton MJ, Ruta M, Lloyd GT. Superiority, competition, and opportunism in the evolutionary radiation of Dinosaurs. *Science.* 2008. <https://doi.org/10.1126/science.1161833>.
98. Toljagic O, Butler R. Triassic-Jurassic mass extinction as trigger for the Mesozoic radiation of crocodylomorphs. *Biol Lett.* 2013. <https://doi.org/10.1098/rsbl.2013.0095>.
99. Markwick PJ. Crocodylian diversity in space and time: the role of climate in paleoecology and its implication for understanding K/T extinctions. *Paleobiology.* 1998;24:470–97.
100. Viana PF, Ribeiro LB, Lima T, de Carvalho VT, Vogt RC, Gross MC, Feldberg E. An optimized protocol for obtaining mitotic chromosomes from cultured reptilian lymphocytes. *Nucleus.* 2016. <https://doi.org/10.1007/s13237-016-0174-3>.
101. Johnson Pokorná M, Altmanová M, Rovatsos M, Velenský P, Vodíčka R, Reháč I, Kratochvíl L. First description of the karyotype and sex chromosomes in the Komodo dragon (*Varanus komodoensis*). *Cytogenet Genome Res.* 2016. <https://doi.org/10.1159/000447340>.
102. Sambrook J, Russell DW. *Molecular Cloning, A Laboratory Manual*. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.
103. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinform.* 2014. <https://doi.org/10.1093/bioinformatics/btu170>.
104. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinform.* 2011. <https://doi.org/10.1093/bioinformatics/btr026>.
105. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0 2013–2015. <http://www.repeatmasker.org>. Accessed at 17/03/2023
106. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004. <https://doi.org/10.1093/nar/gkh340>.
107. Seibt KM, Schmidt T, Heitkam T. FlexiDot: highly customizable, ambiguity-aware dotplots for visual sequence analyses. *Bioinform.* 2018. <https://doi.org/10.1093/bioinformatics/bty395>.
108. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012. <https://doi.org/10.1038/nmeth.1923>.
109. Yano CF, Bertollo LAC, Cioffi MB. Fish-FISH: Molecular Cytogenetics in Fish Species. In: Liehr T, editor. *Fluorescence In Situ Hybridization (FISH)—Application Guide*. Berlin: Springer; 2017. p. 429–44.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.