



Article

The Cytogenetic Map of the Nile Crocodile (*Crocodylus niloticus*, Crocodylidae, Reptilia) with Fluorescence In Situ Localization of Major Repetitive DNAs

Svetlana A. Romanenko ^{1,*} , Dmitry Yu. Prokopov ¹ , Anastasia A. Proskuryakova ¹ , Guzel I. Davletshina ^{1,2}, Alexey E. Tupikin ³, Fumio Kasai ^{4,5}, Malcolm A. Ferguson-Smith ⁵ and Vladimir A. Trifonov ^{1,6}

¹ Institute of Molecular and Cellular Biology, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia

² Institute of Cytology and Genetics, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia

³ Institute of Chemical Biology and Fundamental Medicine, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia

⁴ Japanese Collection of Research Bioresources (JCRB) Cell Bank, Laboratory of Cell Cultures, The National Institute of Biomedical Innovation, Health and Nutrition, Saito-Asagi, Ibaraki 567-0085, Osaka, Japan

⁵ Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, UK

⁶ Department of Natural Science, Novosibirsk State University, 630090 Novosibirsk, Russia

* Correspondence: rosa@mcb.nsc.ru; Tel.: +7-383-363-90-63



Citation: Romanenko, S.A.; Prokopov, D.Y.; Proskuryakova, A.A.; Davletshina, G.I.; Tupikin, A.E.; Kasai, F.; Ferguson-Smith, M.A.; Trifonov, V.A. 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**, *23*, 13063. <https://doi.org/10.3390/ijms232113063>

Academic Editor: Elena Giulotto

Received: 31 August 2022

Accepted: 24 October 2022

Published: 27 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Tandemly arranged and dispersed repetitive DNA sequences are important structural and functional elements that make up a significant portion of vertebrate genomes. Using high throughput, low coverage whole genome sequencing followed by bioinformatics analysis, we have identified seven major tandem repetitive DNAs and two fragments of LTR retrotransposons in the genome of the Nile crocodile (*Crocodylus niloticus*, $2n = 32$). The repeats showed great variability in structure, genomic organization, and chromosomal distribution as revealed by fluorescence in situ hybridization (FISH). We found that centromeric and pericentromeric heterochromatin of *C. niloticus* is composed of previously described in *Crocodylus siamensis* CSI-HindIII and CSI-DraI repetitive sequence families, a satellite revealed in *Crocodylus porosus*, and additionally contains at least three previously unannotated tandem repeats. Both LTR sequences identified here belong to the ERV1 family of endogenous retroviruses. Each pericentromeric region was characterized by a diverse set of repeats, with the exception of chromosome pair 4, in which we found only one type of satellite. Only a few repeats showed non-centromeric signals in addition to their centromeric localization. Mapping of 18S–28S ribosomal RNA genes and telomeric sequences (TTAGGG)_n did not demonstrate any co-localization of these sequences with revealed centromeric and pericentromeric heterochromatic blocks.

Keywords: C-banding; CDAG-banding; FISH; GTG-banding; karyotype; heterochromatin; high-throughput sequencing; repeats; reptile; ribosomal DNA; tandem repeats; telomeric DNA

1. Introduction

Reptiles are a large and paraphyletic class of predominantly terrestrial vertebrates. The taxonomy of the group is complex and has changed repeatedly. Living reptiles comprise turtles, crocodylians, squamates (lizards and snakes), and rhynchocephalians (tuatara). Reptiles show a vast diversity in diploid chromosome number ($2n$), and karyotype morphology, with various combinations of macro- and microchromosomes, and sex-determination systems [1–3]. The genome sequence databases contain only a limited amount of reptile data and the complete analysis of chromosomes and genomes is still lacking for many species.

Crocodylians first appeared 95 million years ago in the Late Cretaceous period, and are the closest living relatives of birds. Together they are combined into a single clade of diapsid reptiles—Archosauria [4,5]. The order Crocodylia is divided into three families

(Alligatoridae, Crocodylidae, and Gavialidae), and the number of currently described species ranges from 23 to 27 [6–13]. Cytogenetic data for Crocodylia are usually restricted to the description of chromosome numbers, karyotype composition, and some conventional banding; only a few studies have used molecular cytogenetic tools [14–19]. Unlike birds and turtles, crocodylian karyotypes do not contain typical reptile dot-shaped microchromosomes. It has been proposed that all microchromosomes disappeared by fusion events after crocodylian divergence from their ancestors [20].

Repetitive DNA sequences make up a large proportion of every eukaryotic genome, and among these, the majority is satellite DNA, organized in large head-to-tail blocks of tandem repeats [21–23]. Satellite DNA can differ in the nucleotide sequence, repeating unit length, and copy number both within the genome and between genomes of different species [24]. Due to the fact that satellite DNA forms long arrays of almost identical repeating units, this part of the genome is still difficult for modern sequencing, assembly, and mapping methods [25,26].

Satellite sequences are a major component of centromeric and pericentromeric heterochromatin (see, for example, [27–30] for reviews). Despite satellite DNA being earlier considered as “junk”, it has now become clear that these DNA arrays are integral to centromere function and stability and they are functionally significant at the level of the spatial organization of chromosomes, their pairing, and segregation during meiosis [24,30–32]. Satellite DNA transcripts are also involved in the formation and maintenance of the heterochromatin structure [33,34]. Very often satellite DNAs are co-localized with other tandemly arranged sequences, such as ribosomal RNA genes (rDNA) and telomeric sequences [35,36].

To date, only a few studies have addressed the characterization of repetitive sequences in crocodylian genomes, and they are mostly focused on such common repeats as telomeric and ribosomal DNA [17,18]. When localizing the telomeric (TTAGGG)_n sequences, no interstitial signals were found. Single rDNA sites have been found in the Siamese crocodile (*C. siamensis*, CSI) and all studied Alligatoridae species, except both examined *Paleosuchus* species, in which three chromosome pairs bear these sequences [17,18].

Using high-throughput sequencing, it was found that the frequency and distribution of microsatellites differ between the Indian gharial (*Gavialis gangeticus*), the American alligator (*Alligator mississippiensis*), and saltwater crocodile (*C. porosus*), and their genomic sequences were dominated by dinucleotide repeats [37]. Physical mapping on chromosomes was performed for (CGG)₁₀ microsatellites only, which marked from four to eight chromosomes of Alligatoridae species [18]. Analysis of the molecular structure of centromeric heterochromatin in the Siamese crocodile showed the centromeric heterochromatin was composed mainly of two repetitive sequence families—two types of GC-rich CSI-*Hind*III family sequences (305 bp and 405 bp) and the 94 bp CSI-*Dra*I [17]. It was proposed that the CSI-*Dra*I-01 fragment is Crocodylidae-specific and present in the *C. niloticus* (CNI) genome, as well as that both types of CSI-*Hind*III family sequences are characteristic for all three crocodylian families (the Crocodylidae, Gavialidae, and Alligatoridae) [17]. However, chromosomal localization of both types of the CSI-*Hind*III family sequences and the CSI-*Dra*I-01 fragment has not been performed on any other species except for *C. siamensis* [17].

The main aim of this study was the identification, initial description, and cytogenetic mapping of major fractions of repetitive DNAs in *C. niloticus*. Using a combination of low-coverage whole-genome high-throughput sequencing, bioinformatics analysis, and molecular cytogenetics we estimated the genomic abundance of repeat families and physically mapped them on chromosomes by FISH. We tested the distribution of identified satellite sequences in the available reference genomic assemblies of various crocodylian species. In order to improve the Nile crocodile karyotype description, we also performed chromosome mapping of the 18S–28S ribosomal RNA genes and telomeric (TTAGGG)_n sequences and made C-banding and chromomycin a3 (CMA₃)/DAPI staining of the Nile crocodile chromosomes.

2. Results

2.1. *C. niloticus* Karyotype

The *C. niloticus* karyotype with $2n = 32$ has been described previously [14]. GTG-banding made it possible to unambiguously identify homologous chromosomes in the karyotype of the species (Figure 1). CBG-banding revealed blocks of pericentromeric heterochromatin on all pairs of chromosomes. Less intensely stained heterochromatin blocks were found in the distal regions of chromosome 1 q-arm and interstitial regions of all chromosomes, except for pairs 4, 7, 9, and 15 (Figure 1).

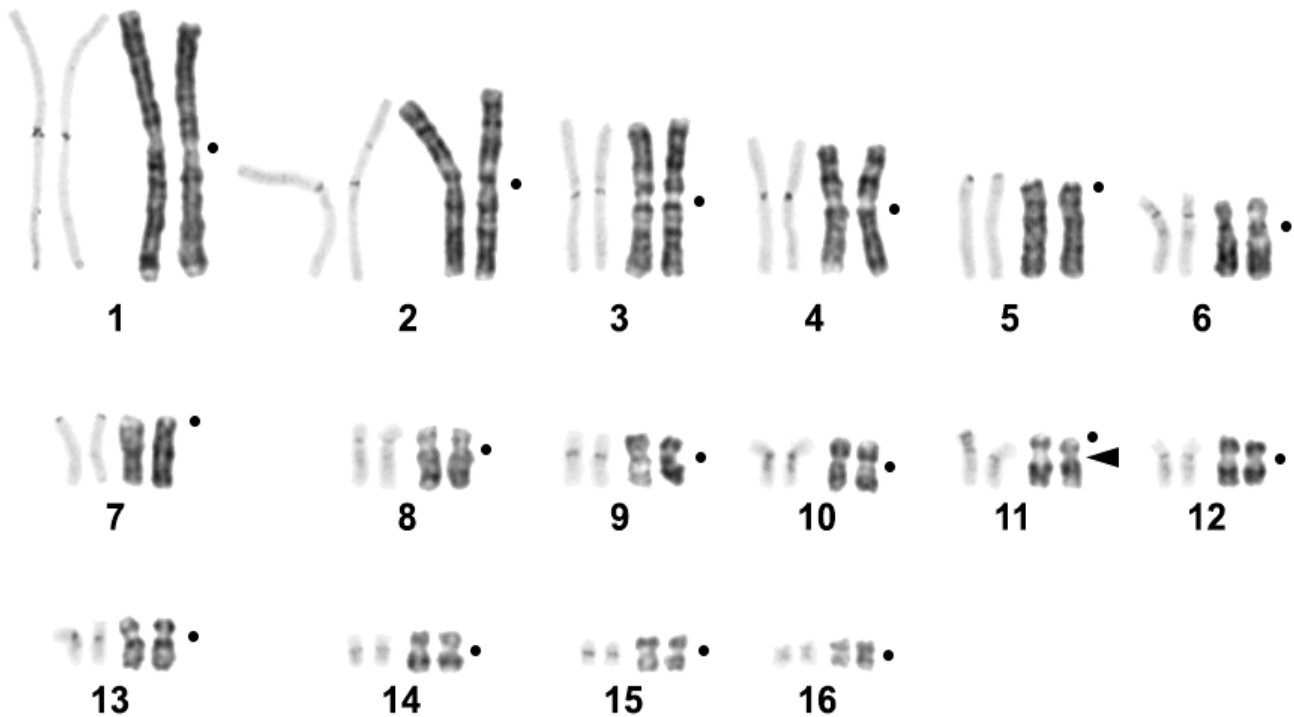


Figure 1. CBG- and GTG-banded *Crocodylus niloticus* karyotype. The black arrow marks the region of 18S/28S-rDNA probe localization. The black dots mark the positions of centromeres.

CDAG-staining (Chromomycin A3-DAPI-after G-banding) makes it possible to reveal the AT-/GC-composition and heterochromatin position on differentially stained chromosomes [38]. Intensive CMA₃-positive pericentromeric blocks were seen in all *C. niloticus* chromosomes except pair 2. Chromosomes 6 and 11 carry large bright CMA₃-positive blocks in p-arms. Weak 4',6-diamidino-2-phenylindole (DAPI)-positive blocks were revealed in pericentromeric regions of chromosomes 1 and 2 (Figure 2a). The 18S/28S-rDNA probe gave the only interstitial signal on the pair 11 (Figures 1 and 2b). Telomeric repeat (TTTAGG)_n marked the distal parts of all chromosomal arms and did not demonstrate any interstitial signals (Figure 2c).

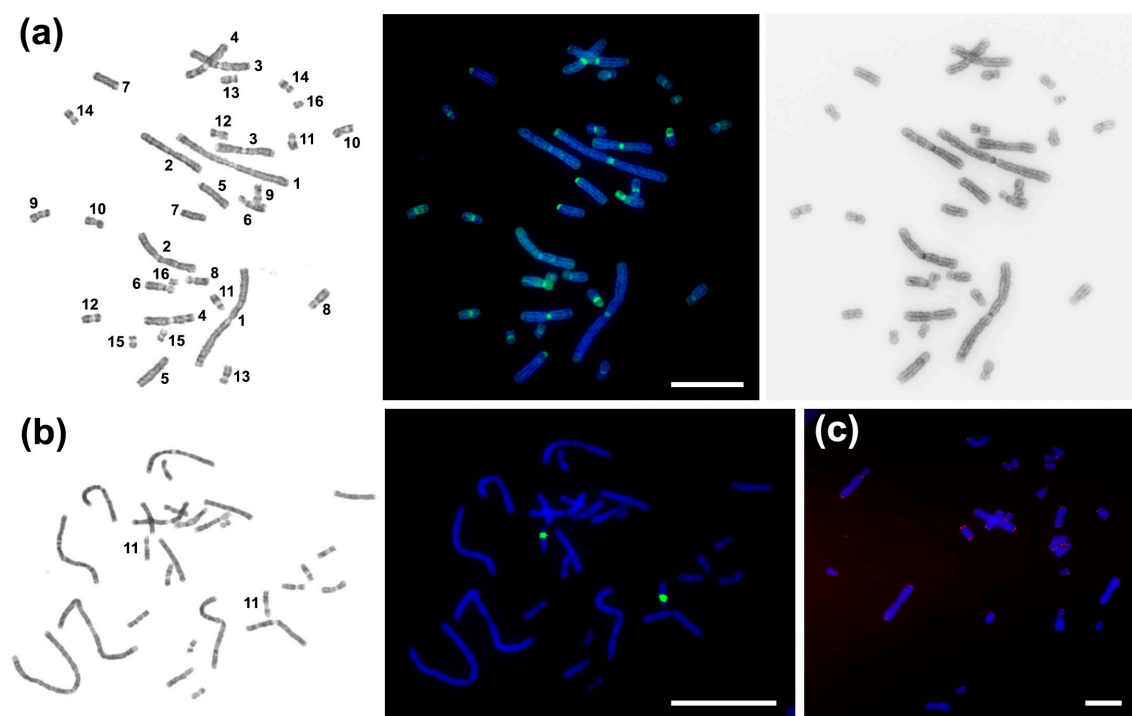


Figure 2. CDAG-staining and localization of 18S/28S-rDNA and telomeric probes on *Crocodylus niloticus* metaphase chromosomes. (a) CDAG-staining. From left to right: GTG-banded chromosomes, CDAG-staining, inverted DAPI-banding. Bright blue blocks contain AT-rich DAPI-positive heterochromatin. Bright green blocks consist of GC-rich CMA₃-positive heterochromatin; (b) Localization of 18S/28S-rDNA probe: GTG-banding on the left; (c) Localization of telomeric repeat (TTTAGG)_n. Scale bar 10 μm.

2.2. Bioinformatics Analysis of *C. niloticus* Repetitive DNA

The genome of *C. niloticus* was sequenced for the first time in this study. BGI MGISEQ-2000 sequencing produced 5,118,575 read pairs (≈ 1 Gb). After pre-processing, only 3,439,364 read pairs (≈ 516 Mb) were analyzed with TAREAN for tandem repeats and putative LTR consensus reconstruction. The nine most represented clusters of satellites (Sat) and LTR were used in this work. A list of identified and characterized sequences is given in Table 1.

Seven sequences (CNI-Sat-4, CNI-Sat-19, CNI-Sat-36, CNI-Sat-58, CNI-Sat-67, CNI-Sat-93, and CNI-Sat-96) were identified as satellites. Their genome proportion ranged from 0.046% to 0.59%, and the length of the consensus sequence varied from 31 to 162 bp. GC content varied from 37.65% in CNI-Sat-96 to 65% in CNI-Sat-4. Three sequences (CNI-Sat-4, CNI-Sat-19, and CNI-Sat-67) shared homology with known satDNA previously described for *C. siamensis* [17] (Table 1). CNI-Sat-58 has about 98% identity with the tandem satellite SAT-2_Crp found in *C. porosus*. For CNI-Sat-36 and CNI-Sat-93 no significant homology to previously described repetitive DNAs was found. CNI-Sat-96 has a distant similarity to a fragment of Gypsy-9_OD-I of *Oikopleura dioica*.

The sequences CNI-LTR-48 and CNI-LTR-68 were identified as fragments of putative LTR retrotransposons. Their genome proportion was 0.17% and 0.097%, respectively. CNI-LTR-48 had a 576 bp consensus length and its GC content was 40.97%. The consensus length for CNI-LTR-68 was 5414 bp and GC content was 50.68% (Table 1). Both sequences obviously belong to the ERV1 family of endogenous retroviruses, whereas CNI-LTR-48 represents a reduced fragment with no coding sequences, while CNI-LTR-68 contains all essential ORFs (gag, pol, and env). Different parts of CNI-LTR-68 (from 468 to 2707 bp in size) demonstrate homology to different EVR1 elements from crocodiles and turtles, with the largest percentage of identity with ERV1-2_PSi-I of *Pelodiscus sinensis*.

Table 1. Putative satellites (Sat) and putative LTR elements (LTR) revealed in the *Crocodylus niloticus* genome.

| Repetitive Sequence Name | Genome Proportion | Consensus Length | GC-Content | Accession Number | Homology to Known DNA (Best Hit) | Query Coverage | Percent Identity |
|--------------------------|-------------------|------------------|------------|------------------|--|----------------|------------------|
| CNI-Sat-4 | 0.59% | 40 | 65% | OP480173 | <i>Crocodylus siamensis</i> DNA, centromere-specific GC-rich repetitive sequences: clones CSI-HindIII-M02, CSI-HindIII-S40, CSI-HindIII-S07 (GenBank AB434504.1; AB434503.1; AB434500.1) | 100% | 97.5% |
| CNI-Sat-19 | 0.3% | 93 | 40.86% | OP480174 | <i>Crocodylus siamensis</i> DNA, centromere-specific repetitive sequences: clones CSI-DraI-01, CSI-DraI-05 (GenBank AB434505.1; AB434506.1) | 95% | 85.56% |
| CNI-Sat-36 | 0.24% | 101 | 60.4% | OP480175 | - | - | - |
| CNI-Sat-58 | 0.12% | 112 | 39.29% | OP480176 | <i>Crocodylus porosus</i> , satellite repeat, SAT-2_Crp | - | ~98% |
| CNI-Sat-67 | 0.1% | 94 | 44.68% | OP480177 | <i>Crocodylus siamensis</i> DNA, centromere-specific repetitive sequences: clone CSI-DraI-05 (GenBank AB434506.1) | 79% | 90.67% |
| CNI-Sat-93 | 0.05% | 31 | 67.74% | OP480178 | - | - | - |
| CNI-Sat-96 | 0.046% | 162 | 37.65% | OP480179 | <i>Oikopleura dioica</i> , part of Gypsy-9_OD-I | - | 71.7% |
| CNI-LTR-48 | 0.17% | 576 | 40.97% | OP480180 | <i>Crocodylus porosus</i> , ERV1-2B_Crp-LTR | - | 85.64% |
| CNI-LTR-68 | 0.097% | 5414 | 50.68% | OP480181 | <i>Pelodiscus sinensis</i> , ERV1-2_Psi-I | - | 80.68% |

2.3. Comparative Analysis of Satellite Sequences

We carried out a search for similar satellite sequences in the available reference genomic assemblies of various crocodylian species (Table 2). Six sequences (CNI-Sat-4, CNI-Sat-19, CNI-Sat-36, CNI-Sat-58, CNI-Sat-67, and CNI-Sat-96) were revealed in the *C. porosus* genome. CNI-Sat-67 is probably a *Crocodylus*-specific satellite, as it is missing from other families. CNI-Sat-19 has not been identified in the genomes of the Alligatoridae families, but it does occur in Gavialidae. CNI-Sat-93 was specific to the *C. niloticus* genome.

Table 2. Representation of satellite sequences in genome assemblies of four Crocodylia species. Percentage of identity for the best hit (obtained using blastn) shown for each satellite.

| Satellites Revealed in <i>Crocodylus niloticus</i> Genome | <i>Alligator sinensis</i> (GCA_000455745.1) | <i>Alligator mississippiensis</i> (GCA_000281125.4) | <i>Gavialis gangeticus</i> (GCA_001723915.1) | <i>Crocodylus porosus</i> (GCA_000768395.2) |
|---|---|---|--|---|
| CNI-Sat-4 | 84.62% | 89.47% | 87.18% | 100% |
| CNI-Sat-19 | - | - | 74.68% | 94.62% |
| CNI-Sat-36 | 81.82% | 78.08% | 79.73% | 96.04% |
| CNI-Sat-58 | 83.81% | 84.26% | 88.89% | 97.27% |
| CNI-Sat-67 | - | - | - | 89.77% |
| CNI-Sat-93 | - | - | - | - |
| CNI-Sat-96 | 73.38% | 72.66% | 77.03% | 82.53% |

2.4. Chromosomal Distribution of Repetitive Sequences

All analyzed repeats, with the exception of CNI-LTR-48, demonstrated a cluster organization and were localized mainly in the pericentromeric regions of chromosomes (Figure 3). The repeats showed specific chromosomal localization.

CNI-Sat-4 was localized in the pericentromeric regions of chromosomes CNI1, 3–16. The signal on CNI16 was very weak (Figure 3a,b). The localization of repeat CNI-Sat-36 was similar to CNI-Sat-4. Although CNI-Sat-36 did not hybridize on CNI1, it labeled the pericentromeric regions of chromosomes 3–16 (Figure 3a,e,f). At the same time, the signal on CNI4 was very weak. The intensity of the signals between the repeats CNI-Sat-4 and CNI-Sat-36 differed greatly.

Localization of CNI-Sat-19 was mainly restricted to pericentromeric regions of CNI chromosomes 1–3, and 16 (Figure 3c,d). CNI-LTR-48 is a dispersed repeat, which forms fairly clear blocks during hybridization in the subtelomeric region of the q-arm of CNI1 and in the pericentromeric regions of CNI6, 9, and 11 (Figure 3e).

CNI-Sat-58 marked the p-arms of chromosomes 13, 14, and 16. At some metaphase plates, it slightly marked the distal part of the p-arm of CNI8 (Figure 3g). CNI-Sat-67 hybridized with pericentromeric regions of all chromosomes, except CNI4. The signal on CNI8 was very weak (Figure 3b,f,h,i).

A low signal intensity was observed during the hybridization of CNI-LTR-68 (Figure 3h). It gave weak signals in the pericentromeric regions of chromosomes 1 and 2, and also weakly marked distal regions of all chromosome arms, similarly to the telomeric repeat.

CNI-Sat-93 hybridized with the pericentromeric regions of chromosomes 5, 7, 10–12, 14, and 15; moreover, it gave an interstitial signal on the CNI9 q-arm (Figure 3g,j). On a pair of chromosomes 12, the signal was observed only on one of the homologs. CNI-Sat-96 marked pericentromeric regions of CNI1-3, 12, and 16 (Figure 3d,i,j). On a pair of chromosomes 12, the signal was observed only on one of the homologs (the same homolog contained CNI-Sat-93).

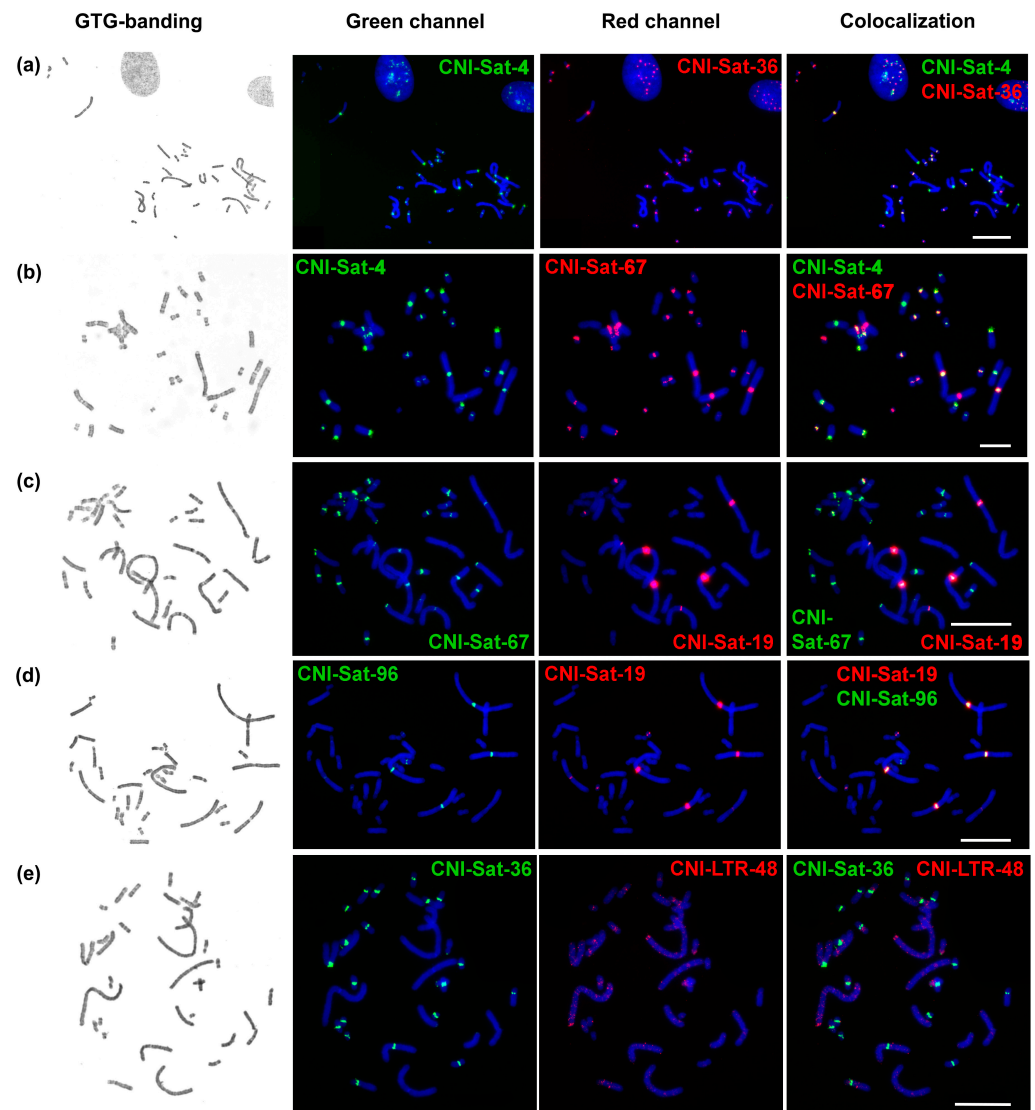


Figure 3. Cont.

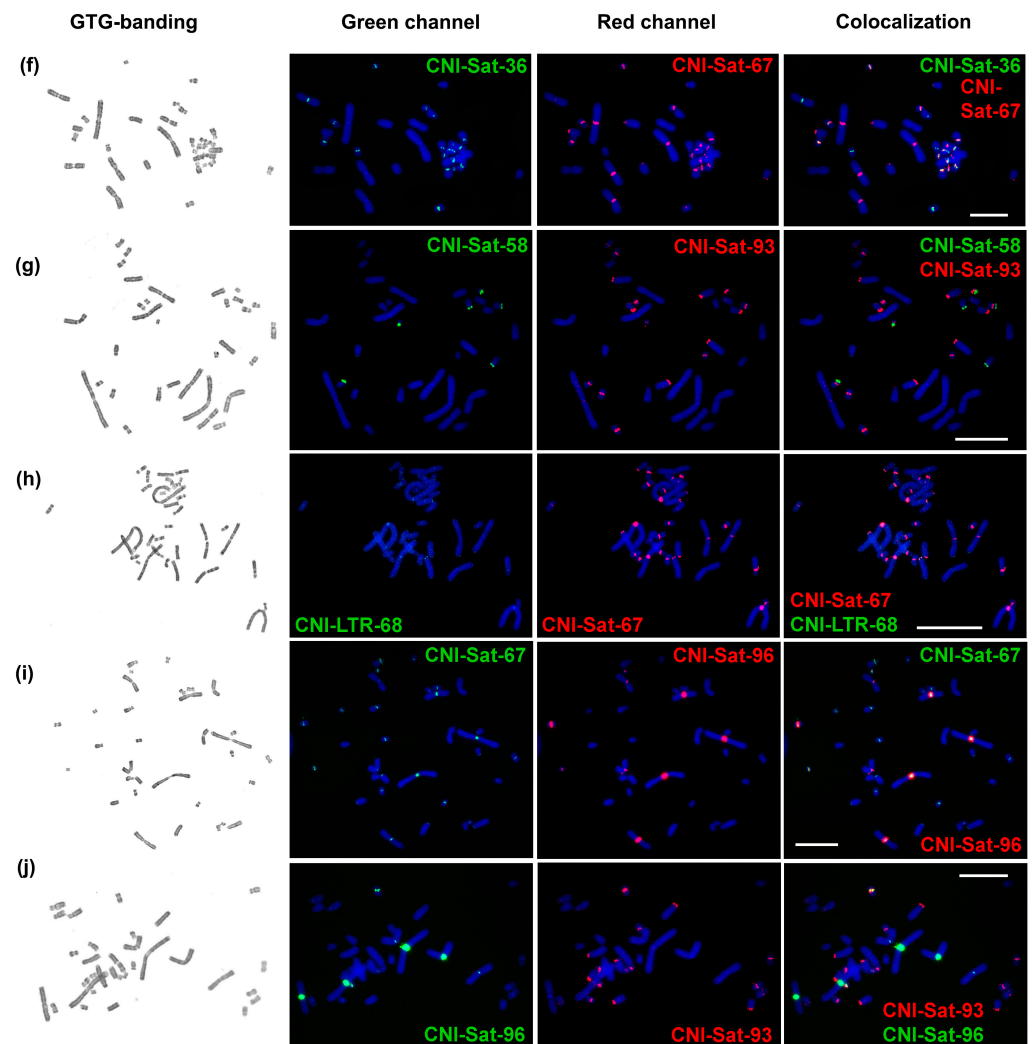


Figure 3. FISH of repetitive DNA probes on metaphase plates of *Crocodylus niloticus*. For each pair of probes from left to right, GTG-banded metaphase, green channel, red channel, and the colocalization of probes are shown: (a) CNI-Sat-4 (green) and CNI-Sat-36 (red); (b) CNI-Sat-4 (green) and CNI-Sat-67 (red); (c) CNI-Sat-19 (red) and CNI-Sat-67 (green); (d) CNI-Sat-19 (red) and CNI-Sat-96 (green); (e) CNI-Sat-39 (green) and CNI-LTR-48 (red); (f) CNI-Sat-36 (green) and CNI-Sat-67 (red); (g) CNI-Sat-58 (green) and CNI-Sat-93 (red); (h) CNI-Sat-67 (red) and CNI-LTR-68 (green); (i) CNI-Sat-67 (green) and CNI-Sat-96 (red); (j) CNI-Sat-93 (red) and CNI-Sat-96 (green). Scale bar 10 μm .

3. Discussion

Tandem repeats content is well investigated in the genomes of only such vertebrates as humans, mice, and some birds [30,39–44]. The number of works on other species is steadily increasing (e.g., [45–50]). Tandem repeats are highly prevalent at centromeres of both animal and plant genomes, but the repeat monomers vary greatly in composition and sequence length [51]. The size of repeats ranges from a few base pairs (microsatellites) up to several kilobases. Due to the very fast evolution and thus a high propensity for block size polymorphism, repetitive sequences have proven to be good genetic [39,43,52,53] and molecular cytogenetic markers demonstrating size-dependent chromosomal localization in species with clearly fragmented genomes (birds and most reptiles) [54,55], as well as making it possible to distinguish between paralogous chromosomes in the genomes of paleopolyploid species [45,46].

Bird macro- and microchromosomes are very conserved in evolution and often syntenic to turtle macro- and microchromosomes, as shown by comparative mapping of protein-coding genes [56]. Using the example of birds and turtles, it was shown that mi-

crochromosomes are enriched in GC-rich genes [57]. Comparative painting studies have demonstrated that rearrangements between macro- and microchromosomes in birds occurred very rarely [58,59]. Homogenization of the centromeric repetitive sequences did not take place between macro- and microchromosomes owing to their structural differences [17]. These structural differences between macro- and microchromosomes are believed to have been highly conserved during the evolution of the Archosauromorpha over 260 million years since this lineage diverged from other diapsids (Lepidosauria) [60–63]. However, in some birds the phenomenon of repetitive DNA homogenization between macro- and microchromosomes has been detected [64].

Crocodylians lack genome compartmentalization depending on chromosome size, which seems to be a consequence of multiple microchromosome fusions after crocodylian divergence from their common ancestors with birds [20]. Analysis of centromeric repeat sequences in the Siamese crocodile showed that the separation of centromeric repeat sequences into macro- and microchromosome-specific, was lost in Crocodylia, and homogenization of centromeric repeat sequences between all chromosomes, except chromosome 2, took place [17].

Repeated sequences identified here in the Nile crocodile genome (Table 1), showed significant differences in structure, genomic organization, and distribution on chromosomes (Figure 4). Seven sequences (CNI-Sat-4, CNI-Sat-19, CNI-Sat-36, CNI-Sat-67, CNI-Sat-93, CNI-Sat-96, and CNI-LTR-68) were located in the pericentromeric regions of the Nile crocodile chromosomes (Figure 4). In view of the fact that constitutive heterochromatin is composed mainly of repetitive elements, the chromosomal localization of the repeats was perfectly correlated with the location of constitutive heterochromatin blocks revealed by C-banding (Figures 1 and 4).

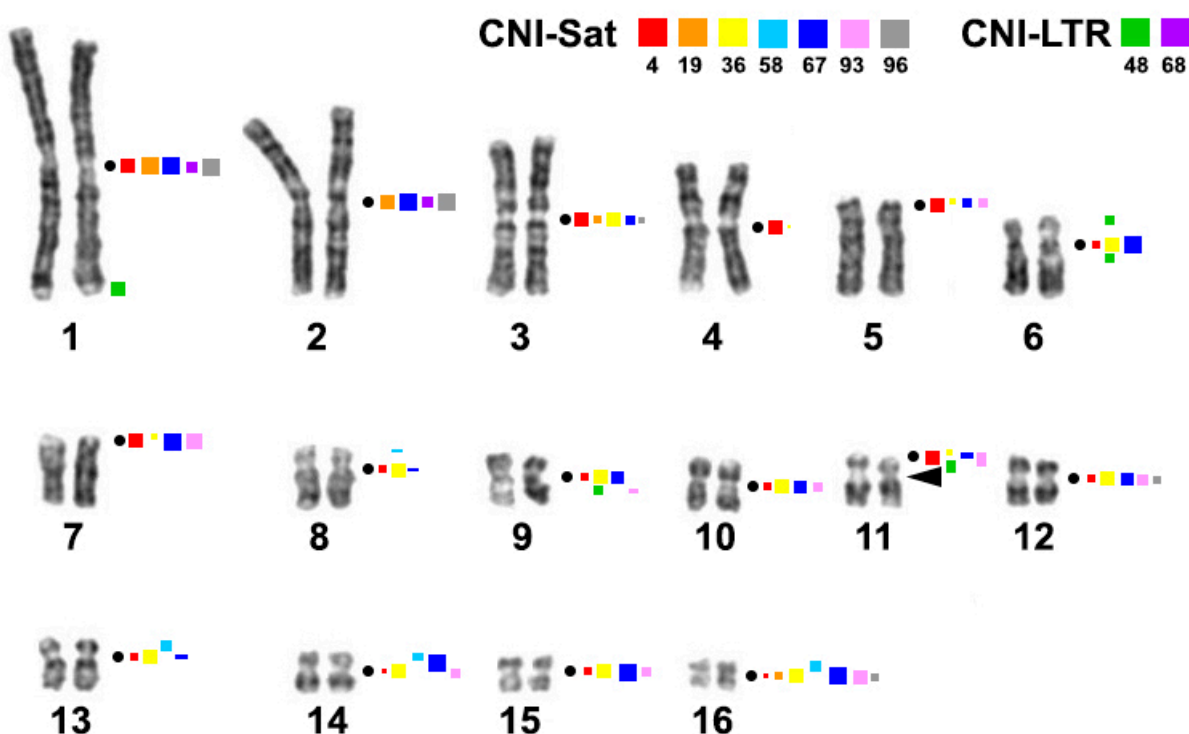


Figure 4. Repetitive DNA probe location on the GTG-banded *Crocodylus niloticus* karyotype. Each cluster of repeated sequences has its own color code (shown in the upper figure right corner). The sizes of the squares roughly show signal intensity. The black arrow marks the region of 18S/28S-rDNA probe localization. The black dots mark the positions of centromeres.

The GC-rich CNI-Sat-4 repeat with a sequence length of 40 bp (Table 1), showing a high degree of similarity to the CSI-*Hind*III family of sequences, was found in the Siamese

crocodile genome [17]. These repeated sequences showed a roughly similar localization pattern in the karyotypes of the two crocodylian species. Clone CSI-*HindIII*-S01 did not label CSI2, and four small CSI chromosomes, and, similarly, clone CSI-*HindIII*-M02 did not label CSI2 but labeled only one of the homologues of one small pair of CSI chromosomes [17], CNI-Sat-4 did not produce a signal on CNI2 and 14 (Figure 4). Interestingly, CSI-*HindIII* showed incomplete tandem arrays in the Siamese crocodile, with this family of repeats conserved in three crocodylian families, Crocodylidae, Gavialidae, and Alligatoridae [17]. Our data showing the presence of CNI-Sat-4 in genome assemblies of the three crocodylian families further supports a possible ancestral status of the CSI-*HindIII* family of sequences in crocodylians (Table 2).

CNI-Sat-19 and CNI-Sat-67 have a similar consensus length (93 and 94 bp, respectively (Table 1)), and they share a high degree of similarity with the CSI-*DraI* sequence family identified in the *C. siamensis* genome [17]. However, in this case, the pattern of repeat localization was very different. While the CSI-*DraI* repeated family was localized in the pericentromeric regions of CSI2 and four small CSI chromosomes [17], CNI-Sat-19 was localized in the pericentromeric regions of CNI1, 2, 3, and 16, and CNI-Sat-67 labeled all CNI pairs except CNI4 (Figure 4). The CSI-*DraI* repeat has been shown to be organized in long tandem arrays in the Siamese crocodile and categorized as a typical centromeric satellite DNA, with this family of repeats being characteristic only of the genus *Crocodylus* but not of higher-ranking taxa [17]. CNI-Sat-67 may be genus-specific as it is distributed in two *Crocodylus* species, while the detection of CNI-Sat-19 in *Gavialis gangeticus* indicates that the CSI-*DraI* repeated family is common for at least two crocodylian families (Table 2).

Clusters of three tandem repeats (CNI-Sat-36, CNI-Sat-93, and CNI-Sat-96) found in the Nile crocodile chromosomes bear no homology with the previously described repeat sequences from GenBank, but one of the satellites (CNI-Sat-58) was revealed in *C. porosus* genome by RepBase [65]. Four repeats (CNI-Sat-36, CNI-Sat-93, CNI-Sat-96, and CNI-LTR-68) were localized in the pericentromeric regions of the *C. niloticus* chromosomes, showing chromosome-specific distribution. Like the CSI-*HindIII* and CSI-*DraI* repeat families, there was no obvious attraction of a certain type of repeat to larger or smaller chromosomes in the Nile crocodile karyotype (Figure 4).

Two clusters of repeats (CNI-LTR-48 and CNI-Sat-58) localized outside the centromeric regions (Figure 4). In addition to separate blocks on chromosomes 1, 6, 9, and 11 of the Nile crocodile, CNI-LTR-48 also gave dispersed signals mainly on large chromosomes (Figure 3e). Despite the fact that CNI-Sat-96 gave weak signals in the subtelomeric regions of chromosomes, we did not find homology of this repeat to telomeric sequences.

We noticed that some satellites co-localized on many chromosomes. Thus, CNI-Sat-4, CNI-sat-36, and CNI-Sat-67 are located in pericentromeric regions of most Nile crocodile chromosomes (CNI3, 5–16) (Figure 4). These three satellites together with CNI-Sat-93 were found in the pericentromeric regions of CNI5, 7, 9, and 11–16 (Figure 4). Satellites CNI-Sat-19, CNI-Sat-67, CNI-Sat-96, and CNI-LTR-68 were found in the two largest chromosomes (CNI1 and 2) (Figure 4). This pattern of localization indicates the evolutionary processes driving the linked spreading of satellites across different chromosomes.

Of the sequences identified in this work, CNI-Sat-4, 36, and 93 were GC-rich (Table 1). These repeated sequences were not detected on chromosome pair 2 (Figures 3 and 4), which is in excellent agreement with the results of CDAG-staining showing the enrichment of AT-rich heterochromatin in the pericentromeric region of chromosome 2 (Figure 2a). The GC-rich interstitial heterochromatic block on chromosome 11 corresponds to the ribosomal gene cluster (Figure 2a,b).

CNI-Sat-93 seems to be the only species-specific satellite of *C. niloticus*. All other tandem repeats, except genus-specific CNI-Sat-67, are shared between different crocodylian families. Despite the 80–100 mya radiation [66] no significant divergence of tandemly arranged repetitive elements is revealed, which indicates quite slow rates of molecular evolution in the taxon.

The presence of endogenous retroelements is an expected false-positive result of satellite DNA identification using TAREAN due to the presence of direct terminal repeats [67]. Two LTR retrotransposons discovered here belong to the ERV1 class. It was well established that ERVs represent less than 2% of crocodile genomes and many of the elements are species-specific [68]. The high similarity to ERV elements from the turtle demonstrated here is not surprising, as shared endogenous retroviruses have been found previously in the same study. The dispersed nature of the retrotransposons was confirmed by FISH analysis (Figure 3e,h). Some enrichment in fluorescent signal density was observed in pericentromeric and subtelomeric regions, which indicates an uneven distribution of retroviruses across the *C. niloticus* genome.

4. Materials and Methods

4.1. Cell Line Establishment and Karyotype Analysis

The *C. niloticus* cells were grown from embryonic tissues obtained from La Ferme aux Crocodiles [16,69] and deposited in the Cambridge Resource Center for Comparative Genomics, Department of Veterinary Medicine, UK. The cell culture was provided to the Institute of Molecular and Cellular Biology, SB RAS, Russia for joint research. The cell line of *C. niloticus* was deposited in the IMCB SB RAS cell bank ("The general collection of cell cultures", 0310-2016-0002). Chromosome suspensions from the cell culture were obtained in the Laboratory of Animal Cytogenetics, IMCB SB RAS, Novosibirsk, Russia, as described previously [70,71].

CBG-banding was made as described by Gladkikh et al. [72]. GTG-banding was performed prior to FISH using the standard trypsin/Giemsa treatment procedure [73]. CDAG-banding was conducted as described before [38].

4.2. DNA Extraction, Library Preparation, and Whole-Genome DNA Sequencing

Whole-genome DNA of *C. niloticus* was extracted from cell culture using Thermo Scientific GeneJET Genomic DNA Purification (Thermo Fisher Scientific Baltics, Vilnius, Lithuania) according to the manufacturer's protocol. Genomic DNA was fragmented in microTUBE AFA Fiber Pre-Split tubes on a Covaris S2 device (Covaris, Woburn, MA, USA) with a median fragment length of 400 bp according to the manufacturer's protocol. Libraries were prepared from fragmented DNA using the MGIEasy Universal DNA Library Prep Set (MGI, Shenzhen, China) according to the manufacturer's protocol. Library sequencing was performed on an MGISEQ-2000 using the DNBSEQ-G400RS High-throughput Sequencing Set (FCL PE100) in the SB RAS Genomics Core Facility (ICBFM SB RAS, Novosibirsk, Russia). The data were deposited in the NCBI SRA database (PRJNA861923).

4.3. Repetitive DNA Identification and Characterization

Preprocessing of raw reads was carried out using fastp 0.23.2 [74]. A total of 4,321,634 pairs of high-quality reads with a length of 75 bp were randomly selected and 3,439,364 read pairs (39.1% of genome size [69]) used in the analysis with the TAREAN 2.3.7 tool [67], which identified clusters of the most abundant tandemly arranged repeats. NCBI BLAST [75] and RepBase [65] databases were used to compare consensus tandem repeat sequences with previously described sequences. Putative LTR retrotransposons were analyzed using the Conserved Domain Database [76].

4.4. Fluorescence In Situ Hybridization (FISH) with Repeats

Primers for PCR amplification and labeling of nine probes (for repeats CNI-Sat-4, 19, 36, 58, 67, 93, 96, CNI-LTR-48, and CNI-LTR-68) were designed with Gene Runner (version 6.5.52) (Table S1). In the case of CNI-LTR-48 and CNI-LTR-68, the primers were matched to a region of the sequences of these putative LTR elements. PCR amplification was performed as described earlier [45]. The telomeric DNA probe was generated by PCR with oligonucleotides (TTAGGG)₅ and (CCCTAA)₅ [77]. Clones of human ribosomal DNA (rDNA) containing a partial 18S ribosomal gene, the full 5.8S gene, a part of the 28S gene,

and two internal transcribed spacers were obtained as described elsewhere [78]. Labeling was performed using PCR by incorporation of biotin-dUTP and digoxigenin-dUTP (Sigma). FISH was performed in accordance with previously published protocols [79]. Images were captured using the VideoTest-FISH software (Imicrotec, New York, NY, USA) with a JenOptic charge-coupled device (CCD) camera mounted on an Olympus BX53 microscope. Hybridization signals were assigned to specific chromosome regions identified by means of GTG-banding patterns photographed by the CCD camera. All images were processed in Adobe PhotoShop 2021 (Adobe, San Jose, CA, USA).

5. Conclusions

Despite the important functional role in centromere organization, tandemly arranged repetitive elements remain poorly studied and are mostly missing from genome assemblies. Here we described the major repetitive DNAs of the Nile crocodile (*C. niloticus*) for the first time and mapped the discovered elements to chromosomes. We confirmed four (CNI-Sat-4, CNI-Sat-19, CNI-Sat-58, and CNI-Sat-67) previously described annotated tandemly arranged elements and revealed three (CNI-Sat-36, CNI-Sat-93, and CNI-Sat-96) novel tandem repeats. We demonstrated that almost all major clusters have centromeric and pericentromeric localization, which may assume the role of some of them in centromere assembly. All pericentromeric regions are characterized by a diverse set of repeated sequences, except for chromosome pair 4, in which we found only one type of repeat. Very low rates of molecular evolution of most discovered repetitive elements may suggest their functional significance. The tandemly arranged repeated elements discovered here may be involved in centromere organization, but additional experiments including chromatin anti-CENP-B immunoprecipitation are necessary to resolve this issue. This work complements available genome assemblies and opens new perspectives in repetitive DNA studies of other crocodiles and reptiles in general.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232113063/s1>.

Author Contributions: Conceptualization, S.A.R., D.Y.P. and V.A.T.; methodology, S.A.R., A.A.P., G.I.D., A.E.T. and F.K.; software, D.Y.P.; validation, S.A.R.; formal analysis, S.A.R. and D.Y.P.; investigation, S.A.R. and D.Y.P.; resources, F.K. and M.A.F.-S.; data curation, S.A.R. and D.Y.P.; writing—original draft preparation, S.A.R. and D.Y.P.; writing—review and editing, M.A.F.-S. and V.A.T.; visualization, S.A.R.; supervision, V.A.T.; project administration, V.A.T.; funding acquisition, V.A.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by RFBR grant No 19-54-26017 (V.A.T.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The sequenced genomic reads were deposited in the NCBI SRA database under the accession number PRJNA861923. The obtained repetitive element consensus sequences were deposited in the GenBank database (OP480173-OP480181).

Acknowledgments: Computational resources were provided by the ELIXIR-CZ project (LM2015047), part of the international ELIXIR infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Deakin, J.E.; Ezaz, T. E-Mail Understanding the Evolution of Reptile Chromosomes through Applications of Combined Cytogenetics and Genomics Approaches. *Cytogenet. Genome Res.* **2019**, *157*, 7–20. [[CrossRef](#)] [[PubMed](#)]
2. Olmo, E. Trends in the evolution of reptilian chromosomes. *Integr. Comp. Biol.* **2008**, *48*, 486–493. [[CrossRef](#)] [[PubMed](#)]
3. Straková, B.; Rovatsos, M.; Kubička, L.; Kratochvíl, L. Evolution of Sex Determination in Amniotes: Did Stress and Sequential Hermaphroditism Produce Environmental Determination? *BioEssays* **2020**, *42*, 2000050. [[CrossRef](#)] [[PubMed](#)]
4. Chiari, Y.; Cahais, V.; Galtier, N.; Delsuc, F. Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). *BMC Biol.* **2012**, *10*, 65. [[CrossRef](#)]

5. Blair Hedges, S.; Poling, L.L. A molecular phylogeny of reptiles. *Science* **1999**, *283*, 998–1001. [[CrossRef](#)]
6. Brochu, C.A. Phylogenetic Approaches Toward Crocodylian History. *Annu. Rev. Earth Planet. Sci.* **2003**, *31*, 357–397. [[CrossRef](#)]
7. Hekkala, E.; Shirley, M.H.; Amato, G.; Austin, J.D.; Charter, S.; Thorbjarnarson, J.; Vliet, K.A.; Houck, M.L.; Desalle, R.; Blum, M.J. An ancient icon reveals new mysteries: Mummy DNA resurrects a cryptic species within the Nile crocodile. *Mol. Ecol.* **2011**, *20*, 4199–4215. [[CrossRef](#)]
8. McAliley, L.R.; Willis, R.E.; Ray, D.A.; White, P.S.; Brochu, C.A.; Densmore, L.D. Are crocodiles really monophyletic?—Evidence for subdivisions from sequence and morphological data. *Mol. Phylogenet. Evol.* **2006**, *39*, 16–32. [[CrossRef](#)]
9. Muniz, F.L.; Ximenes, A.M.; Bittencourt, P.S.; Hernández-Rangel, S.M.; Campos, Z.; Hrbek, T.; Farias, I.P. Detecting population structure of *Paleosuchus trigonatus* (Alligatoridae: Caimaninae) through microsatellites markers developed by next generation sequencing. *Mol. Biol. Rep.* **2019**, *46*, 2472–2484. [[CrossRef](#)]
10. Shirley, M.H.; Vliet, K.A.; Carr, A.N.; Austin, J.D. Rigorous approaches to species delimitation have significant implications for African crocodylian systematics and conservation. *Proc. Biol. Sci.* **2013**, *281*, 20132483. [[CrossRef](#)]
11. Srikulnath, K.; Thapana, W.; Muangmai, N. Role of Chromosome Changes in *Crocodylus* Evolution and Diversity. *Genom. Inform.* **2015**, *13*, 102. [[CrossRef](#)]
12. Martin, S. Global diversity of crocodiles (Crocodylia, Reptilia) in freshwater. *Hydrobiologia* **2008**, *595*, 587–591. [[CrossRef](#)]
13. Uetz, P.; Freed, P.; Aguilar, R.; Hošek, J. The Reptile Database. Available online: <http://www.reptile-database.org> (accessed on 31 August 2022).
14. Cohen, M.M.; Clark, H.F. The somatic chromosomes of five crocodylian species. *Cytogenetics* **1967**, *6*, 193–203. [[CrossRef](#)]
15. Cohen, M.M.; Gans, C. The chromosomes of the order Crocodylia. *Cytogenetics* **1970**, *9*, 81–105. [[CrossRef](#)]
16. Kasai, F.; O'Brien, P.C.M.; Martin, S.; Ferguson-Smith, M.A. Extensive homology of chicken macrochromosomes in the karyotypes of *Trachemys scripta elegans* and *Crocodylus niloticus* revealed by chromosome painting despite long divergence times. *Cytogenet. Genome Res.* **2012**, *136*, 303–307. [[CrossRef](#)]
17. 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). *Chromosom. Res.* **2008**, *16*, 1119–1132. [[CrossRef](#)]
18. Oliveira, V.C.S.; Altmanová, M.; Viana, P.F.; Ezaz, T.; Bertollo, L.A.C.; 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**, *10*, 1397. [[CrossRef](#)]
19. King, M.; Honeycutt, R.; Contreras, N. Chromosomal repatterning in crocodiles: C, G and N-banding and the in situ hybridization of 18S and 26S rRNA cistrons. *Genetica* **1986**, *70*, 191–201. [[CrossRef](#)]
20. Uno, Y.; Nishida, C.; Tarui, H.; Ishishita, S.; Takagi, C. Inference of the Protokaryotypes of Amniotes and Tetrapods and the Evolutionary Processes of Microchromosomes from Comparative Gene Mapping. *PLoS ONE* **2012**, *7*, 53027. [[CrossRef](#)]
21. Charlesworth, B.; Jarne, P.; Assimacopoulos, S. The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. III. Element abundances in heterochromatin. *Genet. Res.* **1994**, *64*, 183–197. [[CrossRef](#)]
22. Elder, J.F.; Turner, B.J. Concerted evolution of repetitive DNA sequences in eukaryotes. *Q. Rev. Biol.* **1995**, *70*, 297–320. [[CrossRef](#)] [[PubMed](#)]
23. Schmidt, T.; Heslop-Harrison, J.S. Genomes, genes and junk: The large-scale organization of plant chromosomes. *Trends Plant Sci.* **1998**, *3*, 195–199. [[CrossRef](#)]
24. Biscotti, M.A.; Canapa, A.; Forconi, M.; Olmo, E.; Barucca, M. Transcription of tandemly repetitive DNA: Functional roles. *Chromosome Res.* **2015**, *23*, 463–477. [[CrossRef](#)] [[PubMed](#)]
25. Miga, K.H. Completing the human genome: The progress and challenge of satellite DNA assembly. *Chromosome Res.* **2015**, *23*, 421–426. [[CrossRef](#)] [[PubMed](#)]
26. Nishibuchi, G.; Déjardin, J. The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome Res.* **2017**, *25*, 77–87. [[CrossRef](#)]
27. Arney, K.L.; Fisher, A.G. Epigenetic aspects of differentiation. *J. Cell Sci.* **2004**, *117*, 4355–4363. [[CrossRef](#)]
28. Hall, A.E.; Keith, K.C.; Hall, S.E.; Copenhaver, G.P.; Preuss, D. The rapidly evolving field of plant centromeres. *Curr. Opin. Plant Biol.* **2004**, *7*, 108–114. [[CrossRef](#)]
29. Bloom, K. Centromere dynamics. *Curr. Opin. Genet. Dev.* **2007**, *17*, 151–156. [[CrossRef](#)]
30. Hartley, G.; O'Neill, R.J. Centromere Repeats: Hidden Gems of the Genome. *Genes* **2019**, *10*, 223. [[CrossRef](#)]
31. Plohl, M.; Meštrović, N.; Mravinac, B. Satellite DNA evolution. *Genome Dyn.* **2012**, *7*, 126–152. [[CrossRef](#)]
32. Ferreira, D.; Meles, S.; Escudeiro, A.; Mendes-da-Silva, A.; Adegá, F.; Chaves, R. Satellite non-coding RNAs: The emerging players in cells, cellular pathways and cancer. *Chromosome Res.* **2015**, *23*, 479–493. [[CrossRef](#)]
33. Grewal, S.I.S.; Elgin, S.C.R. Transcription and RNA interference in the formation of heterochromatin. *Nature* **2007**, *447*, 399–406. [[CrossRef](#)]
34. Trofimova, I.; Krasikova, A. Transcription of highly repetitive tandemly organized DNA in amphibians and birds: A historical overview and modern concepts. *RNA Biol.* **2016**, *13*, 1246–1257. [[CrossRef](#)]
35. Plohl, M.; Luchetti, A.; Meštrović, N.; Mantovani, B. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* **2008**, *409*, 72–82. [[CrossRef](#)]
36. Garrido-Ramos, M.A. Satellite DNA: An Evolving Topic. *Genes* **2017**, *8*, 230. [[CrossRef](#)]

37. Mahfooz, S.; Singh, P.; Akhter, Y. A comparative study of microsatellites among crocodiles and development of genomic resources for the critically endangered Indian gharial. *Genetica* **2022**, *150*, 67–75. [[CrossRef](#)]
38. Lemskaya, N.A.; Kulemzina, A.I.; Beklemisheva, V.R.; Biltueva, L.S.; Proskuryakova, A.A.; Hallenbeck, J.M.; Perelman, P.L.; Graphodatsky, A.S. A combined banding method that allows the reliable identification of chromosomes as well as differentiation of AT- and GC-rich heterochromatin. *Chromosome Res.* **2018**, *26*, 307–315. [[CrossRef](#)]
39. Ames, D.; Murphy, N.; Helentjaris, T.; Sun, N.; Chandler, V. Comparative analyses of human single- and multilocus tandem repeats. *Genetics* **2008**, *179*, 1693–1704. [[CrossRef](#)]
40. Deryusheva, S.; Krasikova, A.; Kulikova, T.; Gaginskaya, E. Tandem 41-bp repeats in chicken and Japanese quail genomes: FISH mapping and transcription analysis on lampbrush chromosomes. *Chromosoma* **2007**, *116*, 519–530. [[CrossRef](#)]
41. Kulak, M.; Komissarov, A.; Fillon, V.; Tsukanova, K.; Saifitdinova, A.; Galkina, S. Genome organization of major tandem repeats and their specificity for heterochromatin of macro- and microchromosomes in Japanese quail. *Genome* **2022**, *65*, 391–403. [[CrossRef](#)]
42. Shang, W.H.; Hori, T.; Toyoda, A.; Kato, J.; Pependorf, K.; Sakakibara, Y.; Fujiyama, A.; Fukagawa, T. Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. *Genome Res.* **2010**, *20*, 1219–1228. [[CrossRef](#)] [[PubMed](#)]
43. Liehr, T. Repetitive Elements in Humans. *Int. J. Mol. Sci.* **2021**, *22*, 2072. [[CrossRef](#)] [[PubMed](#)]
44. Komissarov, A.S.; Gavrilova, E.V.; Demin, S.J.; Ishov, A.M.; Podgornaya, O.I. Tandemly repeated DNA families in the mouse genome. *BMC Genom.* **2011**, *12*, 531. [[CrossRef](#)] [[PubMed](#)]
45. Biltueva, L.S.; Prokopov, D.Y.; Makunin, A.I.; Komissarov, A.S.; Kudryavtseva, A.V.; Lemskaya, N.A.; Vorobieva, N.V.; Serdyukova, N.A.; Romanenko, S.A.; Gladkikh, O.L.; et al. Genomic Organization and Physical Mapping of Tandemly Arranged Repetitive DNAs in Sterlet (*Acipenser ruthenus*). *Cytogenet. Genome Res.* **2017**, *152*, 148–157. [[CrossRef](#)] [[PubMed](#)]
46. Biltueva, L.S.; Prokopov, D.Y.; Romanenko, S.A.; Interesova, E.A.; Schartl, M.; Trifonov, V.A. Chromosome Distribution of Highly Conserved Tandemly Arranged Repetitive DNAs in the Siberian Sturgeon (*Acipenser baerii*). *Genes* **2020**, *11*, 1375. [[CrossRef](#)]
47. Lamelas, L.; Arroyo, M.; Fernández, F.J.; Marchal, J.A.; Sánchez, A. Structural and Evolutionary Relationships in the Giant Sex Chromosomes of Three *Microtus* Species. *Genes* **2018**, *9*, 27. [[CrossRef](#)]
48. Rovatsos, M.; Marchal, J.A.; Giagia-Athanasopoulou, E.; Sánchez, A. Molecular Composition of Heterochromatin and Its Contribution to Chromosome Variation in the *Microtus thomasi*/*Microtus atticus* Species Complex. *Genes* **2021**, *12*, 807. [[CrossRef](#)]
49. Vicari, M.R.; Nogaroto, V.; Noleto, R.B.; Cestari, M.M.; Cioffi, M.B.; Almeida, M.C.; Moreira-Filho, O.; Bertollo, L.A.C.; Artoni, R.F. Satellite DNA and chromosomes in Neotropical fishes: Methods, applications and perspectives. *J. Fish Biol.* **2010**, *76*, 1094–1116. [[CrossRef](#)]
50. Vozdova, M.; Kubickova, S.; Martinková, N.; Galindo, D.J.; Bernegossi, A.M.; Cernohorska, H.; Kadlcikova, D.; Musilová, P.; Duarte, J.M.; Rubes, J. Satellite DNA in Neotropical Deer Species. *Genes* **2021**, *12*, 123. [[CrossRef](#)]
51. Melters, D.P.; Bradnam, K.R.; Young, H.A.; Telis, N.; May, M.R.; Ruby, J.G.; Sebra, R.; Peluso, P.; Eid, J.; Rank, D.; et al. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol.* **2013**, *14*, R10. [[CrossRef](#)]
52. Feliciello, I.; Akrap, I.; Brajkovi, J.; Zlatar, I.; Ugarkovic, D. Satellite DNA as a driver of population divergence in the red flour beetle *Tribolium castaneum*. *Genome Biol. Evol.* **2014**, *7*, 228–239. [[CrossRef](#)] [[PubMed](#)]
53. Warburton, P.E.; Hasson, D.; Guillem, F.; Lescale, C.; Jin, X.; Abrusan, G. Analysis of the largest tandemly repeated DNA families in the human genome. *BMC Genom.* **2008**, *9*, 533. [[CrossRef](#)] [[PubMed](#)]
54. Tanaka, K.; Suzuki, T.; Nojiri, T.; Yamagata, T.; Namikawa, T.; Matsuda, Y. Characterization and chromosomal distribution of a novel satellite DNA sequence of Japanese quail (*Coturnix coturnix japonica*). *J. Hered.* **2000**, *91*, 412–415. [[CrossRef](#)] [[PubMed](#)]
55. Yamada, K.; Nishida-Umehara, C.; Matsuda, Y. Characterization and chromosomal distribution of novel satellite DNA sequences of the lesser rhea (*Pterocnemia pennata*) and the greater rhea (*Rhea americana*). *Chromosome Res.* **2002**, *10*, 513–523. [[CrossRef](#)] [[PubMed](#)]
56. Matsuda, Y.; Nishida-Umehara, C.; Tarui, H.; Kuroiwa, A.; Yamada, K.; Isobe, T.; Ando, J.; Fujiwara, A.; Hirao, Y.; Nishimura, O.; et al. Highly conserved linkage homology between birds and turtles: Bird and turtle chromosomes are precise counterparts of each other. *Chromosome Res.* **2005**, *13*, 601–615. [[CrossRef](#)]
57. Kuraku, S.; Ishijima, J.; Nishida-Umehara, C.; Agata, K.; Kuratani, S.; Matsuda, Y. cDNA-based gene mapping and GC3 profiling in the soft-shelled turtle suggest a chromosomal size-dependent GC bias shared by sauropsids. *Chromosome Res.* **2006**, *14*, 187–202. [[CrossRef](#)]
58. Shibusawa, M.; Nishibori, M.; Nishida-Umehara, C.; Tsudzuki, M.; Masabanda, J.; Griffin, D.K.; Matsuda, Y. Karyotypic evolution in the Galliformes: An examination of the process of karyotypic evolution by comparison of the molecular cytogenetic findings with the molecular phylogeny. *Cytogenet. Genome Res.* **2004**, *106*, 111–119. [[CrossRef](#)]
59. Nishida-Umehara, C.; Tsuda, Y.; Ishijima, J.; Ando, J.; Fujiwara, A.; Matsuda, Y.; Griffin, D.K. The molecular basis of chromosome orthologies and sex chromosomal differentiation in palaeognathous birds. *Chromosome Res.* **2007**, *15*, 721–734. [[CrossRef](#)]
60. Kumazawa, Y.; Nishida, M. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: Statistical evidence for archosaurian affinity of turtles. *Mol. Biol. Evol.* **1999**, *16*, 784–792. [[CrossRef](#)]
61. Janke, A.; Erpenbeck, D.; Nilsson, M.; Arnason, U. The mitochondrial genomes of the iguana (*Iguana iguana*) and the caiman (Caiman crocodylus): Implications for amniote phylogeny. *Proc. Biol. Sci.* **2001**, *268*, 623–631. [[CrossRef](#)]

62. Rest, J.S.; Ast, J.C.; Austin, C.C.; Waddell, P.J.; Tibbetts, E.A.; Hay, J.M.; Mindell, D.P. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. *Mol. Phylogenet. Evol.* **2003**, *29*, 289–297. [[CrossRef](#)]
63. Kumazawa, Y. Mitochondrial genomes from major lizard families suggest their phylogenetic relationships and ancient radiations. *Gene* **2007**, *388*, 19–26. [[CrossRef](#)]
64. Yamada, K.; Nishida-Umehara, C.; Matsuda, Y. A new family of satellite DNA sequences as a major component of centromeric heterochromatin in owls (Strigiformes). *Chromosoma* **2004**, *112*, 277–287. [[CrossRef](#)]
65. Kohany, O.; Gentles, A.J.; Hankus, L.; Jurka, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinform.* **2006**, *7*, 474. [[CrossRef](#)]
66. Green, R.E.; Braun, E.L.; Armstrong, J.; Earl, D.; Nguyen, N.; Hickey, G.; Vandewege, M.W.; St John, J.A.; Capella-Gutiérrez, S.; Castoe, T.A.; et al. Three crocodylian genomes reveal ancestral patterns of evolution among archosaurs. *Science* **2014**, *346*, 1254449. [[CrossRef](#)]
67. Novák, P.; Robledillo, L.Á.; 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**, *45*, e111. [[CrossRef](#)]
68. Chong, A.Y.; Kojima, K.K.; Jurka, J.; Ray, D.A.; Smit, A.F.A.; Isberg, S.R.; Gongora, J. Evolution and gene capture in ancient endogenous retroviruses—Insights from the crocodylian genomes. *Retrovirology* **2014**, *11*, 71. [[CrossRef](#)]
69. Kasai, F.; O’Brien, P.C.M.; Ferguson-Smith, M.A. Reassessment of genome size in turtle and crocodile based on chromosome measurement by flow karyotyping: Close similarity to chicken. *Biol. Lett.* **2012**, *8*, 631–635. [[CrossRef](#)]
70. Graphodatsky, A.S.; Yang, F.; Serdukova, N.; Perelman, P.; Zhdanova, N.S.; Ferguson-Smith, M.A. Dog chromosome-specific paints reveal evolutionary inter- and intrachromosomal rearrangements in the American mink and human. *Cytogenet. Cell Genet.* **2000**, *90*, 275–278. [[CrossRef](#)]
71. Yang, F.; O’Brien, P.C.M.; Milne, B.S.; Graphodatsky, A.S.; Solanky, N.; Trifonov, V.; Rens, W.; Sargan, D.; Ferguson-Smith, M.A. A complete comparative chromosome map for the dog, red fox, and human and its integration with canine genetic maps. *Genomics* **1999**, *62*, 189–202. [[CrossRef](#)]
72. Gladkikh, O.L.; Romanenko, S.A.; Lemskaya, N.A.; Serdyukova, N.A.; O’Brien, P.C.M.; Kovalskaya, J.M.; Smorkatcheva, A.V.; Golenishchev, F.N.; Perelman, P.L.; Trifonov, V.A.; et al. Rapid karyotype evolution in *Lasiopodomys* involved at least two autosome—Sex chromosome translocations. *PLoS ONE* **2016**, *11*, e0167653. [[CrossRef](#)]
73. Seabright, M. A rapid banding technique for human chromosomes. *Lancet* **1971**, *2*, 971–972. [[CrossRef](#)]
74. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [[CrossRef](#)] [[PubMed](#)]
75. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [[CrossRef](#)]
76. Marchler-Bauer, A.; Lu, S.; Anderson, J.B.; Chitsaz, F.; Derbyshire, M.K.; DeWeese-Scott, C.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; et al. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **2011**, *39*, D225. [[CrossRef](#)]
77. Ijdo, J.W.; Wells, R.A.; Baldini, A.; Reeders, S.T. Improved telomere detection using a telomere repeat probe (TTAGGG)_n generated by PCR. *Nucleic Acids Res.* **1991**, *19*, 4780. [[CrossRef](#)]
78. Maden, B.E.H.; Dent, C.L.; Farrell, T.E.; Garde, J.; McCallum, F.S.; Wakeman, J.A. Clones of human ribosomal DNA containing the complete 18 S-rRNA and 28 S-rRNA genes. Characterization, a detailed map of the human ribosomal transcription unit and diversity among clones. *Biochem. J.* **1987**, *246*, 519–527. [[CrossRef](#)]
79. Romanenko, S.A.; Biltueva, L.S.; Serdyukova, N.A.; Kulemzina, A.I.; Beklemisheva, V.R.; Gladkikh, O.L.; Lemskaya, N.A.; Interesova, E.A.; Korentovich, M.A.; Vorobieva, N.V.; et al. Segmental paleotetraploidy revealed in sterlet (*Acipenser ruthenus*) genome by chromosome painting. *Mol. Cytogenet.* **2015**, *8*, 90. [[CrossRef](#)]