

The first draft genome of the Amazonian manatee (*Trichechus inunguis*) provides insights into genome evolution and freshwater adaptation in Sirenia

Adauto Lima Cardoso

São Paulo State University

Flávia dos Santos Tavares

Universidade Federal do Pará

Camila do Nascimento Moreira

São Paulo State University

Victor Ramos Pap

São Paulo State University

Luís Adriano Santos Nascimento

Guamá Science and Technology Park

Angélica Lúcia Figueiredo Rodrigues

Instituto Biologia e Conservação dos Mamíferos Aquáticos da Amazônia (BioMA)

Leonardo Santos Sena

Universidade Federal do Pará

Cesar Martins

São Paulo State University

Renata Coelho Rodrigues Noronha

renatacrn@gmail.com

Universidade Federal do Pará

Research Article

Keywords: Amazon basin, endangered species, aquatic mammal, genome evolution

Posted Date: August 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10542503/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

This study presents the first draft genome for the Amazonian manatee (*Trichechus inunguis*), the only exclusively freshwater sirenian species, contributing to the understanding of genomic evolution, adaptation to the freshwater environment and species conservation. The genome was sequenced using Illumina technology, employing 151 bp paired-end reads and achieving approximately 37× coverage. Initial assemblies were highly fragmented; however, reference-guided scaffolding with RagTag using the chromosome-level genome of *T. senegalensis* increased BUSCO completeness to 98.8% for the female and 96.7% for the male. The genome sizes were 2.89 Gb (female) and 2.75 Gb (male), smaller than those of *T. manatus* (3.10 Gb) and *T. senegalensis* (3.19 Gb). A total of 19,001 protein-coding genes were identified in *T. inunguis*, with 89.1% of BUSCO genes annotated as complete. Repetitive elements accounted for 48.52% of the female genome and 46.54% of the male genome, with a predominance of retroelements and LINE-1/L1-CIN4, which corresponded to 13.46% of the genome. Comparative analysis revealed an expansion of gene families associated with G-protein-coupled receptor signaling, ion channels, oxidative metabolism and transport, potentially linked to adaptation to the Amazonian environment. This genome serves as a fundamental resource for evolutionary, population and conservation studies.

INTRODUCTION

Sirenians (Order Sirenia) constitute a unique group of herbivorous aquatic mammals comprising only four living species: the dugong (*Dugong dugon*), the African manatee (*Trichechus senegalensis*), the West Indian manatee (*Trichechus manatus*) and the Amazonian manatee (*Trichechus inunguis*). The evolutionary history of this group is marked by a successful transition to aquatic environment, involving profound morphological, physiological and behavioral modifications associated with locomotion, feeding and thermoregulation (Marsh et al. 2011).

Among the current species of sirenians, the Amazonian manatee occupies a unique ecological position. This is the only exclusively freshwater species in the group, distributed throughout the Amazon basin, where it inhabits rivers, floodplain lakes and other seasonally flooded environments (Best 1984; Marmontel et al. 2016). Unlike its congeners, which exploit marine and estuarine environments, the Amazonian manatee is subject to intense hydrological fluctuations resulting from the annual Amazonian flood pulse, one of the most striking ecological phenomena in tropical ecosystems (Junk et al. 1989). These variations directly affect food availability, the physicochemical conditions of the water, and the connectivity between habitats, imposing particular selective pressures on the populations of this species.

In addition to its ecological uniqueness, *T. inunguis* has high relevance for conservation. The species is classified as vulnerable by the International Union for Conservation of Nature (IUCN) due to historical hunting, incidental capture in fishing activities, habitat degradation and increasing anthropogenic pressure on Amazonian ecosystems (Marmontel et al. 2016). Additionally, molecular and cytogenetic

studies have revealed the occurrence of natural hybridization between *T. inunguis* and *T. manatus* in the Amazonian estuary region, resulting in the formation of fertile hybrid populations and raising important questions related to evolution, species delimitation and conservation management (Vianna et al. 2006; Lima et al. 2019; De Oliveira et al. 2022; Noronha et al. 2022).

Although significant advances have been made in understanding the biology and ecology of the Amazonian manatee, the absence of a reference genome for this species has limited the understanding of its genomic architecture, evolutionary history, and the molecular basis of its adaptations to the freshwater environment. High-quality genomes for other species of the genus, such as *T. manatus* and *T. senegalensis*, have already been made available, allowing comparative studies that reveal patterns of genomic evolution in sirenians. However, the Amazonian manatee, due to its unique ecological characteristics, remained an important gap in the genomic knowledge of the group.

This study presents the first reference genome of the Amazonian manatee and performs a comprehensive comparative analysis among the three current species of the genus *Trichechus*. Specifically, we investigated the composition and evolutionary dynamics of repetitive DNA, characterized the gene repertoire and assessed patterns of gene family expansion potentially associated with environmental adaptation. The results provide new insights into the genomic evolution of sirenians, the mechanisms associated with adaptation to the Amazonian freshwater environment, and the development of genomic tools for the conservation of one of the most emblematic and threatened species of South American aquatic megafauna.

MATERIAL AND METHODS

Samples

Blood of a male named “Buriti” and a female named “Neguinha” of *T. inunguis* that live in captivity at ZooUnama (Santarém, PA, Brazil) and Bioparque da Amazônia (Macapá, AP, Brazil), respectively, were collected with authorization of Chico Mendes Institute for Biodiversity Conservation (SISBIO 44915-1, 92853-1). The procedures were conducted in accordance with the Research Ethics Committee of the Universidade Federal do Pará (Protocol 8803211223). These samples were used for whole genome sequencing.

Assembled genomes of a *T. manatus* female named “Lorelei” (GCA_000243295.1; Foote et al. 2015) and of a *T. senegalensis* female named “Xixi” (GWHERCE000000000; Huang et al. 2024) were downloaded from NCBI/GenBank (USA) and National Genomics Data Center (China), respectively.

Karyotyping

Metaphase chromosome preparations from *T. inunguis* samples were obtained with temporary lymphocyte culture, according to the protocol developed by Moorhead et al. (1960) with modifications.

Whole blood (0,5 mL) was seeded in a flask with 5 mL of RPMI-1640 medium supplemented with phytohemagglutinin (2%) and fetal bovine serum (10%) and maintained for 72 hours at 37°C. After, N-desacetyl-N-methylcolchicine (0.01 µg/mL; KaryoMAX™ Colcemid™ Solution in PBS, Gibco™) was added and treated for 1h. The samples were submitted to hypotonic treatment with 0.075 M potassium chloride (KCl) for 20 minutes and fixed in Carnoy solution (3 methanol: 1 acetic acid). Cell suspensions were dropped onto slides to chromosome analysis. Chromosomes were stained with Giemsa and visualized in a microscopy.

Sequencing, assembly and annotation of nuclear genome

DNA from *T. inunguis* samples was purified with the PureLink Genomic DNA Mini Kit (Invitrogen, Waltham, MA, United States of America), following the manufacturer's protocol. DNA quality and size was checked by electrophoresis on 1% agarose gel and 2100 Bioanalyzer (Santa Clara, CA, United States of America). DNA quantification was done by picogreen method using Victor 3 fluorometry (Norwalk, CT, United States of America). Libraries were constructed with TruSeq Nano DNA Kit (Illumina, San Diego, CA, United States of America). Sequencing processes were performed in NovaSeq 6000 Sequencing System (Illumina, San Diego, CA, United States of America). For each sample were generated fragments of 151 base pairs which cover about 37x the genome. The read quality of raw data was checked by the software FastQC (Andrews, 2010). Reads with low quality were discarded and the Illumina adapters were trimmed using Trimmomatic-0.39 (Bolger et al., 2014) and BBDuk-38.49 (Bushnell et al. 2017) tools. The first 10 bases of the read were cut off, and reads shorter than 95 base pairs were dropped off.

De novo nuclear genome assemblies were performed using Meraculous-v2.2.6 (Chapman et al. 2011), which performed several trial runs of assemblies with Kmer values in a range from 31 to 121. The final assembly was chosen based on the metrics values of the software QUAST-v5.0.2 (Gurevich et al. 2013) and Assembly-Stats (<https://github.com/rjchallis/assembly-stats>). *T. inunguis* and *T. manatus* (GCA_000243295.1) assemblies were scaffolded with RagTag-v2.1.0 (Alonge et al. 2022) using the assembled genome of *T. senegalensis* (Huang et al. 2024) as a reference. The results were evaluated with QUAST-v5.0.2 and BUSCO-v6.0.0 (Simão et al. 2015). Protein coding genes were annotated in *T. inunguis*, *T. manatus* and *T. senegalensis* assemblies with Helixer-v0.3.5 (Holst et al. 2026) and the annotation completeness was assessed with BUSCO-v6.0.0 with mammalia_odb10 gene sets. For repeat annotation, first custom repeat libraries were created with RepeatModeler for each genome (Smit and Hubley 2015), using the default parameters. Then, we use these libraries as input for the RepeatMasker-4.1.1 (Smit et al. 2013).

Expanded gene families

Orthologous were identified with OrthoFinder-v.2.5.5 (Emms & Kelly 2015, 2019). Gene family expansion analyses were conducted with CAFE-v5.1.0 (Mendes et al. 2021) based on the results from OrthoFinder and only gene families with a p-value < 0.05 were considered as having undergone significant change.

Gene Ontology (GO) enrichment analyses were performed on the expanded gene families using clusterProfiler package (Yu et al. 2012). A statistical significance threshold of a p-value < 0.05 was considered to identify significantly enriched GO terms.

RESULTS

Genome assembly of Trichechus species

The female and male samples of *T. inunguis* showed a karyotype with diploid number (2n) of 56 chromosomes and autosomal fundamental number (NFa) of 92. A homomorphic sex chromosome pair (XX) was observed in the female, while the male showed a heteromorphic sex chromosome pair (XY) (Supplementary Fig. 1).

After Illumina sequencing, a total of 756,028,680 and 916,133,222 of paired-end reads (151 bp) were generated for the female and male of *T. inunguis*, respectively. After filtering, 755,885,870 and 915,905,042 pair-end reads with length of 95–141 bp were retained for female and male, respectively. The assemblies showed very similar metrics but remained highly fragmented, with 243,856 and 241,005 scaffolds for female and male, respectively, and recovering 60.2% and 59.8% of the conserved mammalian genes in the mammalia_odb10 dataset, respectively (Supplementary Table 1).

Due to the high number of scaffolds in the genome assemblies of *T. inunguis* (generated in this study) and *T. manatus* (generated by Foote et al. 2015), rescaffolding and patching were performed using the software RagTag with *T. senegalensis* chromosome-level genome (Huang et al. 2024) as a reference. The benchmarking universal single-copy orthologs (BUSCO) analysis showed that 98.8%, 96.7% and 97.8% of the conserved mammalian genes dataset were identified as complete in the new versions of *T. inunguis* female, *T. inunguis* male and *T. manatus*, respectively. Overall, RagTag substantially improved the assembly metrics and the assessment results showed a high completeness of the assemblies (Supplementary Table 1). The genome size of *T. inunguis* female is 2,889.21 Mb, while the male is 2,752.79 Mb; *T. manatus* is 3,103.92 Mb, and *T. senegalensis* is 3,185.65 Mb (Table 1; Fig. 1a-c).

Table 1
Statistics of the *Trichechus* genomes.

Features	<i>T. inunguis</i>	<i>T. manatus</i>	<i>T. senegalensis</i>
Genome size (Mb)	2,889.21	3,103.92	3,185.65
Contig N50 (Mb)	59	37	118
Scaffold N50 (Mb)	121	129	136
Number of contigs	622,658	166,490	47
Number of scaffolds	5,850	5,186	39
GC content (%)	40.68	40.73	40.74
Repeat content (%)	48.52	44.94	55.01
Annotated protein-coding genes	19,001	19,129	19,418
BUSCO results for assembly	C: 98.8% [S: 98.3%, D: 0.5%]	C: 97.8% [S: 97.2%, D: 0.6%]	C: 99.4% [S: 98.7%, D: 0.7%]
	F: 0.3%	F: 0.7%	F: 0.1%
	M: 0.9%	M: 1.5%	M: 0.5%
	n: 9226	n: 9226	n: 9226
BUSCO results for annotation	C: 89.1% [S: 88.8%, D: 0.3%]	C: 87.1% [S: 86.8%, D: 0.3%]	C: 90.1% [S: 89.7%, D: 0.4%]
	F: 4.3%	F: 5.5%	F: 4.0%
	M: 6.6%	M: 7.4%	M: 5.9%
	n: 9226	n: 9226	n: 9226

Genome annotation of *Trichechus* species

To comprehensively understand the genome composition of the genus *Trichechus*, genome annotations were performed across all three species (Fig. 1). A total of 1,401.82 Mb (48.52%), 1,281.20 Mb (46.54%), 1,394.98 Mb (44.94%), and 1,752.57 Mb (55.01%) of repetitive sequences were predicted in *T. inunguis* female, *T. inunguis* male, *T. manatus* and *T. senegalensis*, respectively. Retroelements represent the most abundant class of repetitive sequences in all genomes analyzed. Specifically, the element L1/CIN4 is the most abundant sequence representing 13.46% of *T. inunguis*, 13.84% of *T. manatus* and 22.86% of *T. senegalensis* genomes. DNA transposons constitute 0.55%, 0.5% and 0.66% of the genome of these species, respectively. Unclassified elements account for 23.51%, 17.65% and 21.37% of the genomes, respectively (Fig. 1d-f). The genomes of *Trichechus* species showed a widely shared evolutionary pattern, marked by a relatively recent expansion of transposable elements of the LINE class, with

emphasis on the L1 family. However, the chronology of this expansion differed between species: while the expansion of LINEs appears to have expired earlier in *T. inunguis* and *T. manatus*, in *T. senegalensis* it persisted longer (Fig. 1g-i).

A total of 19,001, 19,129 and 19,418 protein coding genes were annotated in *T. inunguis*, *T. manatus* and *T. senegalensis*, respectively. The BUSCO analysis showed that 89.1%, 87.1% and 90.1% of the conserved mammalian genes were annotated to be complete in these genomes, respectively (Table 1).

Environmental adaptation in *Trichechus* genomes

Functional enrichment analyses of expanded gene families revealed both shared and species-specific patterns among the three *Trichechus* species. Gene Ontology (GO) enrichment analyses identified distinct functional categories associated with the expanded gene repertoires of each species (Fig. 2a-c). In *T. inunguis*, the most significantly enriched GO terms included G protein-coupled receptor (GPCR) signaling pathway, proteolysis, olfactory receptor activity, ion channel activity, monooxygenase activity, oxidoreductase activity, iron ion binding, serine-type exopeptidase activity, vitamin transport, and voltage-gated sodium channel activity (Fig. 2a). In *T. manatus*, enriched categories were mainly related to olfactory receptor activity, nucleosome organization, transcription regulator complex, protein heterodimerization activity, proteolysis, serine-type exopeptidase activity, protein tyrosine kinase activity, antigen processing and presentation, protein glycosylation, and MHC class II protein complex (Fig. 2b). Similarly, *T. senegalensis* showed significant enrichment for olfactory receptor activity, serine-type exopeptidase activity, transcription regulator complex, nucleosome organization, protein heterodimerization activity, transferase activity transferring hexosyl groups, carbohydrate metabolic process, motor activity, mRNA export from nucleus, and ferric iron binding (Fig. 2c).

DISCUSSION

In the present work we sequenced, assembled and annotated the first draft genome of the Amazonian manatee (*Trichechus inunguis*), providing a critical genome resource for the only extant sirenian species that inhabits exclusively freshwater ecosystems. Like other manatee species, *T. inunguis* is threatened with extinction (International Union for Conservation of Nature's Red List of Threatened Species) and this first genome description represents a fundamental tool for understanding its biology, adaptive evolution and for supporting management and conservation strategies.

Previous molecular and cytogenetic studies demonstrated that the Amazonian manatee (*T. inunguis*) and the marine West Indian manatee (*T. manatus*) hybridize in the Amazonian estuary, generating naturally fertile hybrid individuals (Vianna et al. 2006; Lima et al. 2019; De Oliveira et al. 2022; Noronha et al. 2022). Therefore, karyotype analysis is essential to accurate species identification and for distinguishing pure individuals from hybrids. The two individuals analysed in the present study exhibited the chromosomal characteristics previously described for *T. inunguis*, confirming their taxonomic identity and excluding the possibility of hybrid origin.

The genome assemblies revealed sex-associated variation in genome size, with the female genome (2.88 Gb) being slightly larger than the male genome (2.75 Gb), consistent with the presence of two X chromosomes in females versus one X chromosome and a small Y chromosome in males (Assis et al., 1988; De Oliveira et al., 2022; Noronha et al., 2022). Interspecifically, *T. inunguis* had the smallest genome among the three congeners, followed by *T. manatus* (3.10 Gb) (Foote et al. 2015) and *senegalensis* (3.18 Gb) (Huang et al. 2024). This smaller genome size of *T. inunguis* correlates with a lower proportion of repetitive DNA (46.54–48.52%) compared to *T. senegalensis* (55.01%), suggesting that lineage-specific differences in transposable element dynamics contributed to the genomic contraction in the Amazonian lineage. This pattern is evolutionarily significant, given that *T. inunguis* diverged from *T. manatus* relatively recently (approximately 1.34 million years ago) during the Plio-Pleistocene, a period of intense geological and climatic changes in the Amazon Basin (De Souza et al. 2019). Thus, the smaller genome and reduced repetitive content in *T. inunguis* may reflect more recent genomic compaction associated with its adaptation to the peculiarities of the Amazonian freshwater environment.

The predominance of retroelements, particularly LINE-1 (L1/CIN4), in the three species is consistent with the general pattern observed in mammals (Kazazian 2004; Kapusta et al. 2017). However, a striking difference was observed: *T. senegalensis* presented a substantially higher L1/CIN4 content (22.86%) compared to *T. inunguis* (13.46%) and *T. manatus* (13.84%). Divergence analyses indicated that LINE expansion persisted more recently in *T. senegalensis*, while its activity ceased earlier in the American species. This earlier interruption in *T. inunguis* may have limited copy accumulation, contributing to its smaller genome size. The differential dynamics of LINE elements among these closely related species may result from multiple factors, including epigenetic silencing, demographic history and the efficiency of purifying selection (Lynch & Conery 2003; Bourque et al. 2018). For *T. inunguis*, which inhabits a highly seasonal and relatively isolated freshwater system, population bottlenecks or founder effects during the Pleistocene may have accelerated the removal or reduced the fixation of new transposable element insertions. Another relevant aspect is the high proportion of unclassified repetitive elements – 23.51% in *T. inunguis*, 17.65% in *T. manatus* and 21.37% in *T. senegalensis*. This suggests the existence of specific or highly divergent Sirenia repeat families not represented in current databases, emphasizing the need for further characterization.

Annotation of protein-coding genes revealed remarkably similar numbers among the three species of the genus *Trichechus*: 19,001 (*T. inunguis*), 19,129 (*T. manatus*) and 19,418 (*T. senegalensis*). These values fall within the expected range for placental mammals, whose typical gene content varies between 19,000 and 22,000 genes (Frankish et al. 2019; Amaral et al. 2023). The numerical proximity suggests that, despite the ecological and geographic differences between lineages, the core gene repertoire has remained stable throughout the diversification of the genus, consistent with the pattern of gene conservation observed in other groups of aquatic mammals, such as cetaceans (Yim et al. 2014; Foote et al. 2015).

Completeness analyses using BUSCO (Benchmarking Universal Single-Copy Orthologs) indicated that 89.1% (*T. inunguis*), 87.1% (*T. manatus*) and 90.1% (*T. senegalensis*) of the conserved genes in the

mammalia_odb10 set were annotated as complete. These percentages are substantially higher than the completeness of the initial version of assemblies, which recovered only 59.8–60.2% of the same genes, demonstrating that the annotation process, combined with rescaffolding using RagTag, was effective in recovering coding regions from fragmented assemblies (Alonge et al. 2022). These values are comparable to other non-model mammalian draft genomes such as Andean Bear (*Tremarctos ornatus*) (Seremi et al. 2021), indicating that the annotation quality is compatible with the state-of-the-art for reference genomes of threatened mammals.

Among the adaptations distinguishing *T. inunguis*, a notable feature is the significant expansion of gene families linked to G-protein-coupled receptor (GPCR) signaling and voltage-gated ion channels (particularly sodium channels). These neuromuscular and electrophysiological regulatory pathways are crucial for maintaining osmotic homeostasis and water-salt balance in an exclusively freshwater ecosystem, where osmotic challenges differ drastically from those faced by marine or estuarine species (Marsh et al. 2011). Furthermore, the genomic profile of *T. inunguis* reflects marked specialization in redox physiology and substance transport, evidenced by the enrichment of genes associated with monooxygenase and oxidoreductase activity, iron-ion binding, and vitamin transport. These pathways suggest refined metabolic adaptations for coping with oxidative stress and fluctuations in partial oxygen pressure during diving within the dynamic and variable waters of the Amazon River system (Junk et al. 1989; Yim et al. 2014).

Although *T. inunguis* exhibits this unique set of metabolic and ionic specializations, the species shares an expansion of olfactory receptors and proteolytic enzymes (such as serine-type exopeptidases) with its congeners. This chemosensory expansion underscores the importance of olfaction for navigation and foraging by *T. inunguis* in the low-visibility waters of the Amazon Basin, while the proteolytic machinery meets the demands of a strictly herbivorous diet based on aquatic macrophytes (Best 1984; Marsh et al. 2011). In contrast to the osmotic and redox adaptations dominant in the Amazonian manatee, marine and estuarine species expanded their gene repertoires in other directions: *T. manatus* invested heavily in adaptive immunity (via MHC class II and antigen processing) to combat the diversity of marine pathogens, whereas *T. senegalensis* prioritized carbohydrate metabolism and cellular transport dynamics. Thus, the Amazonian manatee genome stands as an exemplary model of evolutionary divergence centered on survival, osmotic regulation and metabolic efficiency within the planet's largest river system.

From a conservation standpoint, the first *T. inunguis* genome provides a powerful tool for addressing pressing issues. Given the occurrence of natural hybridization with *T. manatus* in the Amazon estuary (Vianna et al. 2006; De Oliveira et al. 2022. Noronha et al. 2022), this genome will allow the development of high-resolution molecular markers for the identification of pure and hybrid individuals, which is essential for defining conservation units and management strategies. Furthermore, the genome serves as a basis for population genomics studies, allowing the assessment of genetic diversity, inbreeding levels and population structure throughout the species' fragmented distribution in the Amazon Basin

(Marmontel et al. 2016). This is particularly urgent considering the ongoing threats of habitat loss, waterway pollution and incidental capture in fishing gear (Cantanhede et al. 2005; White et al. 2026).

In conclusion, this first draft of the *Trichechus inunguis* genome not only provides a high-quality genomic resource for comparative studies within the genus *Trichechys*, but, more importantly, establishes a basis for understanding freshwater adaptation in an exclusively riverine sirenian. The observed genomic signatures (smaller genome size, reduced repetitive content, early cessation of LINE activity and expansions of gene families related to metabolism, sensory perception, and environmental response) collectively suggest that *T. inunguis* has undergone unique evolutionary trajectories shaped by the Amazonian freshwater ecosystem. This genome represents a crucial step for the conservation genomics of this vulnerable species and opens new avenues for research on the evolution of aquatic mammals in extreme and seasonal environments.

DECLARATIONS

FUNDING

This work was supported by grants from the National Council for Scientific and Technological Development (CNPq) (grant numbers 313798/2022-2, 309711/2023-1, 306665/2025-5 and 444448/2024-0). FST received a PhD scholarship from CAPES (PPGNBC/UFGA). LASN and RCRN were funded by CNPq and the São Paulo Research Foundation (FAPESP) (grant numbers 2022/1437745, 167/2024, and 073/2023). CM were supported by the São Paulo Research Foundation (grant number 2015/16661-1).

Author Contribution

A.L.C. wrote and contributed to the entire manuscript. F.S.T., C.N.M., and V.R.P. contributed to the methodology. L.S.S. and A.L.F.R. provided biological samples. L.A.S.N. and C.M. contributed resources. R.C.R.N. contributed funding, resources, overall manuscript revision, and structural organization. All authors reviewed and approved the final manuscript.

Acknowledgement

The authors are grateful to the Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) for the collection permit; the support by Zoológico da Universidade da Amazônia - ZooUnama (Santarém, PA, Brazil) and Bioparque da Amazônia (Macapá, AP, Brazil) responsible for the management of animals and collection of samples; and National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES) and São Paulo Research Foundation for the financial support.

Data Availability

The raw data and assembled genomes generated by the present work has been submitted to the NCBI and will be released to the public after publication of this article. NCBI Bioproject numbers are PRJNA1131814 and PRJNA1131815 for female and male, respectively. Raw data accession numbers are SRR29792334 and SRR29834295 for female and male, respectively. Genome assemblies accession numbers are SRR38003954 and SRR38019779 for female and male, respectively.

REFERENCES

1. Alonge M, Lebeigle L, Kirsche M et al (2022) Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol* 23:258. <https://doi.org/10.1186/s13059-022-02823-7>
2. Amaral P, Carbonell-Sala S, De La Vega FM et al (2023) The status of the human gene catalogue. *Nature* 622:41–47. <https://doi.org/10.1038/s41586-023-06490-x>
3. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
4. Assis MFL, Best RC, Barros RMS, Yonenaga-Yassuda Y (1988) Cytogenetic study of *Trichechus inunguis* (Amazonian manatee). *Braz J Genet* 11:41–50
5. Best RC (1984) The aquatic mammals and reptiles of the Amazon. In: Sioli H (ed) *The Amazon. Monographiae Biologicae*, vol 56. Springer, Dordrecht, pp 371–412. https://doi.org/10.1007/978-94-009-6542-3_15
6. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
7. Bourque G, Burns KH, Gehring M et al (2018) Ten things you should know about transposable elements. *Genome Biol* 19:199. <https://doi.org/10.1186/s13059-018-1577-z>
8. Bushnell B, Rood J, Singer E (2017) BBMerge—accurate paired shotgun read merging via overlap. *PLoS ONE* 12:e0185056. <https://doi.org/10.1371/journal.pone.0185056>
9. Cantanhede AM, Da Silva VM, Farias IP et al (2005) Phylogeography and population genetics of the endangered Amazonian manatee, *Trichechus inunguis* Natterer, 1883 (Mammalia, Sirenia). *Mol Ecol* 14:401–413. <https://doi.org/10.1111/j.1365-294X.2004.02413.x>
10. Chapman JA, Ho I, Sunkara S et al (2011) Meraculous: De novo genome assembly with short paired-end reads. *PLoS ONE* 6:e23501. <https://doi.org/10.1371/journal.pone.0023501>
11. De Oliveira EHC, Gomes AJB, Costa AF et al (2022) Karyotypical confirmation of natural hybridization between two manatee species, *Trichechus manatus* and *Trichechus inunguis*. *Life* 12:616. <https://doi.org/10.3390/life12050616>
12. De Souza ÉMS, Freitas L, da Silva Ramos EK et al (2021) The evolutionary history of manatees told by their mitogenomes. *Sci Rep* 11:3564. <https://doi.org/10.1038/s41598-021-82390-2>

13. Emms DM, Kelly S (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16:157.
<https://doi.org/10.1186/s13059-015-0721-2>
14. Emms DM, Kelly S (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 20:238. <https://doi.org/10.1186/s13059-019-1832-y>
15. Foote AD, Liu Y, Thomas GW et al (2015) Convergent evolution of the genomes of marine mammals. *Nat Genet* 47:272–275. <https://doi.org/10.1038/ng.3198>
16. Frankish A, Diekhans M, Ferreira AM et al (2019) GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 47:D766–D773. <https://doi.org/10.1093/nar/gky955>
17. Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
18. Holst F, Bolger AM, Kindel F et al (2026) Helixer: ab initio prediction of primary eukaryotic gene models combining deep learning and a hidden Markov model. *Nat Methods* 23:732–739.
<https://doi.org/10.1038/s41592-025-02939-1>
19. Huang X, Dong G, Fan H et al (2024) The genome of African manatee *Trichechus senegalensis* reveals secondary adaptation to the aquatic environment. *iScience* 27:110394.
<https://doi.org/10.1016/j.isci.2024.110394>
20. Junk WJ, Bayley PB, Sparks RE (1989) The flood pulse concept in river-floodplain systems. *Can Spec Publ Fish Aquat Sci* 106:110–127
21. Kapusta A, Suh A, Feschotte C (2017) Dynamics of genome size evolution in birds and mammals. *Proc Natl Acad Sci USA* 114:E1460–E1469. <https://doi.org/10.1073/pnas.1616702114>
22. Kazazian HH Jr (2004) Mobile elements: drivers of genome evolution. *Science* 303:1626–1632.
<https://doi.org/10.1126/science.1089670>
23. Lima CS, Magalhães RF, Marmontel M et al (2019) A hybrid swarm of manatees along the Guianas coastline, a peculiar environment under the influence of the Amazon River plume. *Acad Bras Cienc* 91:e20190325. <https://doi.org/10.1590/0001-3765201920190325>
24. Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302:1401–1404.
<https://doi.org/10.1126/science.1089370>
25. Marmontel M, de Souza D, Kendall S (2016) *Trichechus inunguis*. The IUCN Red List of Threatened Species 2016. <https://doi.org/10.2305/IUCN.UK.2016-2.RLTS.T22102A43793736.en>.
e.T22102A43793736
26. Marsh H, O'Shea TJ, Reynolds JE III (2011) *Ecology and conservation of the Sirenia*. Cambridge University Press, Cambridge
27. Mendes FK, Vanderpool D, Fulton B, Hahn MW (2020) CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics* 36:5516–5518.
<https://doi.org/10.1093/bioinformatics/btaa1022>

28. Moorhead PS, Nowell PC, Mellman WJ et al (1960) Chromosome preparations of leukocytes cultured from human peripheral blood. *Exp Cell Res* 20:613–616. [https://doi.org/10.1016/0014-4827\(60\)90138-8](https://doi.org/10.1016/0014-4827(60)90138-8)
29. Noronha RCR, Almeida BRR, Chagas MCS et al (2022) Karyotypes of manatees: new insights into hybrid formation (*Trichechus inunguis* × *Trichechus m. manatus*) in the Amazon Estuary. *Genes* 13:1263. <https://doi.org/10.3390/genes13071263>
30. Saremi NF, Oppenheimer J, Vollmers C, O'Connell B, Milne SA, Byrne A, Yu L, Ryder OA, Green RE, Shapiro B (2021) An Annotated Draft Genome for the Andean Bear, *Tremarctos ornatus*. *J Hered* 112(4):377–384. <https://doi.org/10.1093/jhered/esab021>
31. Simão FA, Waterhouse RM, Ioannidis P et al (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
32. Smit A, Hubley R (2015) RepeatModeler Open-1.0. <http://www.repeatmasker.org/RepeatModeler/>
33. Smit A, Hubley R, Green P (2013) RepeatMasker Open-4.0. <http://www.repeatmasker.org/>
34. Vianna JA, Bonde RK, Caballero S et al (2006) Phylogeography, phylogeny and hybridization in trichechid sirenians: implications for manatee conservation. *Mol Ecol* 15:433–447. <https://doi.org/10.1111/j.1365-294X.2005.02771.x>
35. White MC, Rivera CJ, Babilonia JS et al (2026) Factors influencing survival of rescued Amazonian manatees (*Trichechus inunguis*) during clinical rehabilitation in Peru. *Conserv Sci Pract* 8:e70235. <https://doi.org/10.1111/csp2.70235>
36. Yim HS, Cho Y, Guang X et al (2014) Minke whale genome and aquatic adaptation in cetaceans. *Nat Genet* 46:88–92. <https://doi.org/10.1038/ng.2835>
37. Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16:284–287. <https://doi.org/10.1089/omi.2011.0118>

Figures

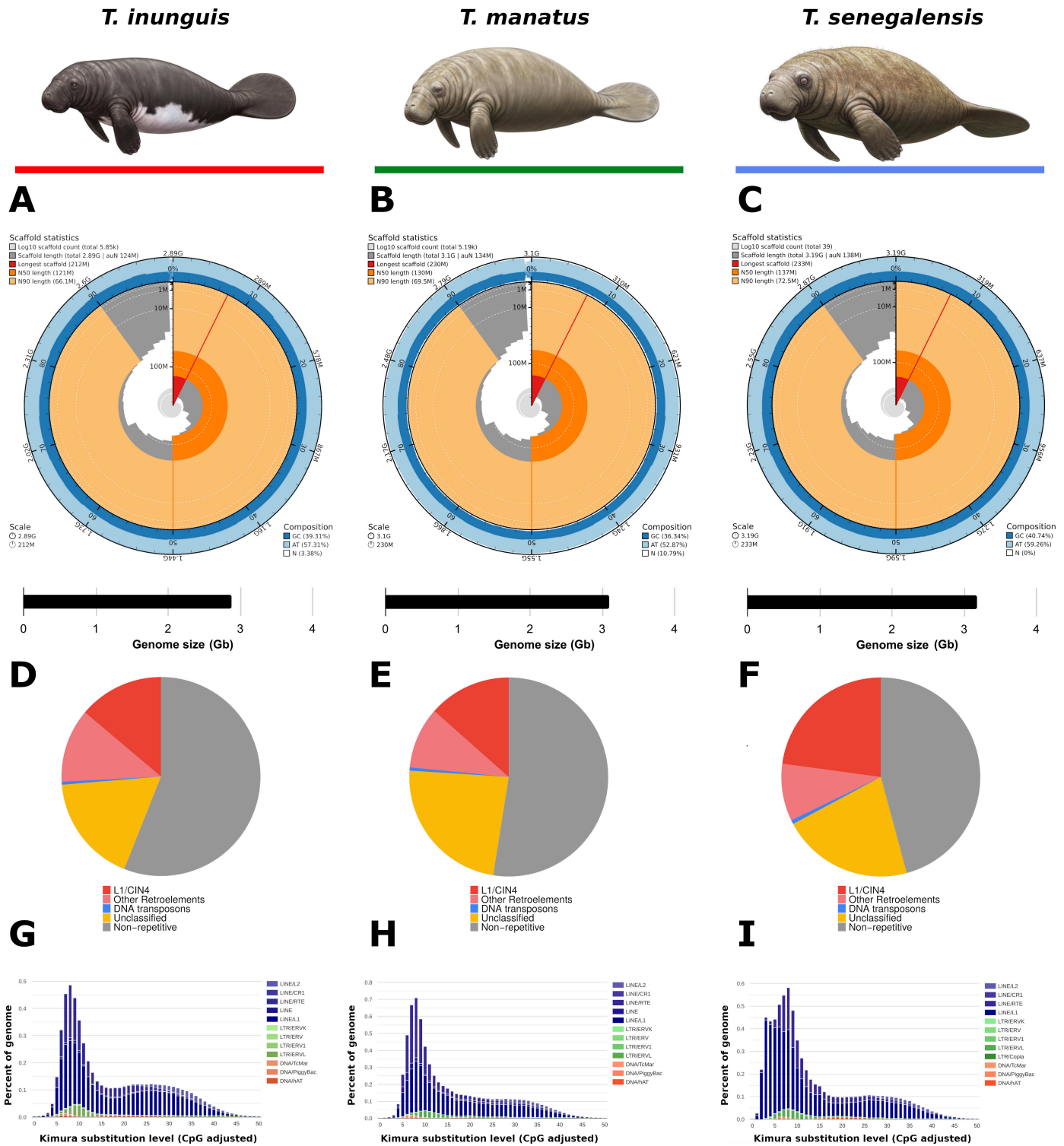


Figure 1

Genome characteristics of *Trichechus* species: genomic metrics of *T. inunguis* (a), *T. manatus* (b) and *T. senegalensis* (c); compositions of repetitive DNAs of *T. inunguis* (d), *T. manatus* (e) and *T. senegalensis* (f); genomics landscape of transposable elements in *T. inunguis* (g), *T. manatus* (h) and *T. senegalensis* (i).

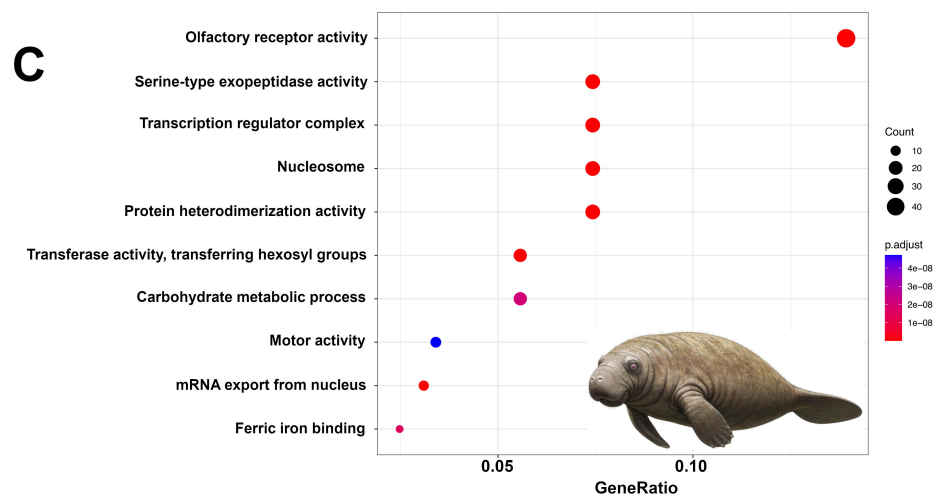
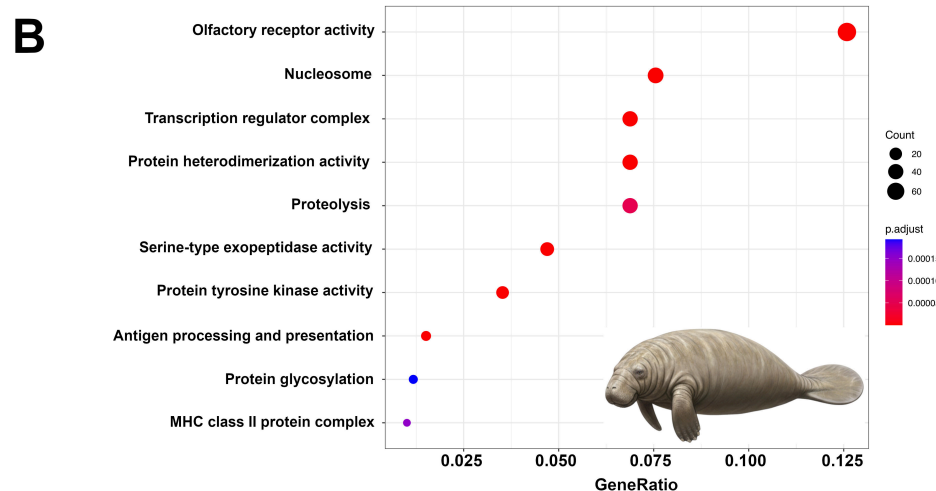
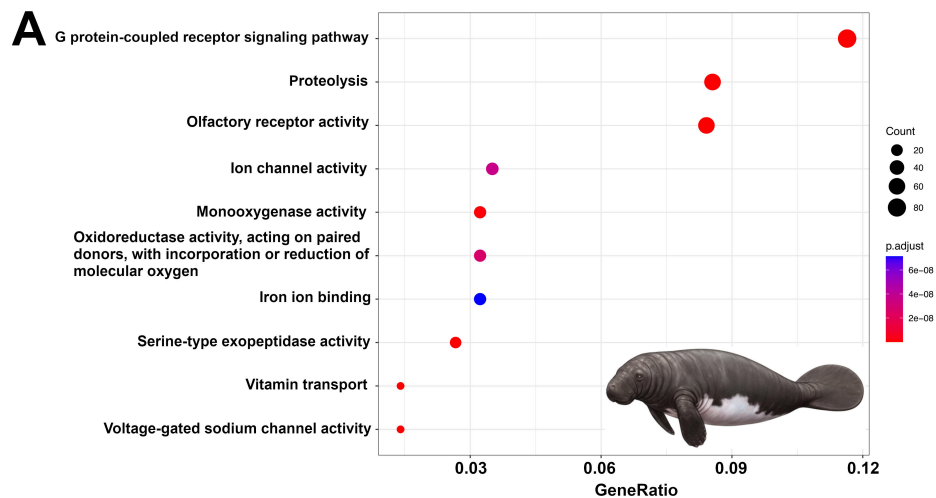


Figure 2

Functional enrichment GO of expanded gene families of *T. inunguis* (a), *T. manatus* (b) and *T. senegalensis* (c).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)