Universidade Federal de São Carlos

Centro de Ciências Biológicas e da Saúde

Iago Santos de Oliveira

Genomic organization and evolutionary dynamics of repetitive sequences in Cathartidae (Aves, Cathartiformes)

São Carlos

2026

Universidade Federal de São Carlos

Centro de Ciências Biológicas e da Saúde

Iago Santos de Oliveira

Genomic organization and evolutionary dynamics of repetitive sequences in Cathartidae (Aves, Cathartiformes)

Research project submitted as one of the requirements for a bachelor's degree in Biotechnology at Universidade Federal de São Carlos (UFSCar)

Supervisor: Prof. Dr. Marcelo de Bello Cioffi

São Carlos

2025

I dedicate this work to my parents Maristela and Luiz Clesio, and my brother Nicholas. I would be nothing without your love, and support. I'm thankful every single day of my life for having you by my side. I love you more than anything.

Agradecimentos

Em primeiro lugar agradeço a minha família e meus amigos. Eu não seria absolutamente nada sem vocês ao meu lado, tanto nos momentos bons quanto nos momentos ruins. Pai e Mãe, vocês são quem me puxam para o chão quando minha mente se perde nas nuvens da ansiedade e da preocupação, e também quem me propulsionam para frente para ser sempre a melhor versão de mim. Admiro vocês todos os dias por terem chegado aonde estão hoje através de muito esforço e amor por mim e pelo Nicholas. Se um dia tiver 10% da bravura e dedicação que vocês tiveram, ficarei em paz sabendo que já é muito. Eu amo vocês. Nicholas você é tudo que eu poderia pedir de um irmão. Agradeço profundamente por termos ficado próximos e sinto que cada dia que passa você representa mais para mim, e se torna cada vez mais meu amigo, e meu irmão. Daiane, Gilda, Liam, Otto, Paulo, Rafael, Raissa, Sandra... vocês me fazem mais bem do que podem imaginar, e um simples almoço em família com vocês pode alegrar toda minha semana. Amo revisitar histórias da família com vocês, falar sobre a vida e principalmente viajar com vocês. Eu amo muito mesmo cada um de vocês. Ilda, você representa muito mais do que pode imaginar. Seu carinho seu sorriso e seu cuidado durante minha vida toda me fizeram ser quem eu sou hoje. Aos meus amigos Bruno, Julia, Pimenta, Dantas, Thales, Biel, Marina e Juan. Vocês são quem tornam meus dias serem mais fáceis, e espero que estejam neles por muito mais tempo. Obrigado pelos lanches, pelos filmes, pelas séries, pelas risadas, pelas festas, pelas mesas de RPG pelas risadas (são muitas) e pelas conversas. As minhas amigas de curso Julia Franciele e Isabela, que foram fundamentais para que eu me graduasse ou que pelo menos que eu me graduasse com algum resquício de sanidade mental, e agora são amigas para a vida.

Agradeço também ao meu orientador e professor Marcelo de Bello Cioffi que me conheceu em minha primeira matéria da graduação, e que agora me recebe ao final dela para me guiar em minha trajetória no mundo acadêmico. Obrigado pelas oportunidades e por me incentivar a ser a melhor versão profissional de mim que eu possa ser.

Agradeço aos meus companheiros de laboratório Guilherme, Jhon, Príncia, Geize Gustavo, Alan, e Pablo, pela recepção e pelos incontáveis momentos em que me ajudaram nos últimos meses. Seja com minhas inúmeras dúvidas em experimentos, com protocolos, com a escrita, com papos sobre ciência ou crises existenciais, vocês sempre estiveram lá. Este projeto seria impossível sem vocês.

"Amaze! Amaze! Amaze!"

- Rocky

Abstract

In this study, we examined how repetitive DNAs influence chromosomal diversification and genome architecture in New World vultures (Cathartidae), a lineage distinguished by exceptional karyotypic stability (2n = 80). We examined five species and three distinct populations of *Coragyps atratus* using a combination of low-coverage genome sequencing, satellite DNA and repetitive DNA profiling, and *in situ* and *in silico* mapping. We found significant diversity in repeat composition between species and populations despite consistent chromosome structure, primarily due to lineage-specific amplification of transposable elements and satellite DNAs. The repeatome was primarily composed of retroelements, particularly LINEs and LTRs, whereas satellitomes exhibited both conserved and rapidly evolving families. Some satellite DNA groups exhibited population-specific divergence, indicating rapid turnover, whereas others were common across all species. The predominant occurrence of satellite DNAs, evidenced by *in situ* and *in silico* mapping, in centromeric and pericentromeric regions suggests their structural role in chromosome organization. However, changes in repeat distribution and abundance occurred without notable chromosomal rearrangements, indicating a disjunction between karyotypic evolution and repeat dynamics. Our findings emphasize the necessity of integrating genomic and cytogenetic methodologies to understand genome evolution, illustrating that repetitive DNA serves as a principal catalyst for genomic diversity in birds and demonstrating that substantial genomic innovation can transpire within conserved chromosomal structures.

Key words: Cytogenetics; Comparative cytogenomics; Repeatome; Satellite DNA; Transposable elements.

Resumo

Neste estudo, analisamos como os DNAs repetitivos influenciam a diversificação cromossômica e a arquitetura genômica em abutres do Novo Mundo (Cathartidae), uma linhagem que se distingue por uma estabilidade cariotípica excepcional (2n = 80). Examinamos cinco espécies e três populações distintas de *Coragyps atratus* utilizando sequenciamento genômico de baixa cobertura para realizar a caracterização de seus DNAs repetitivos, e executamos mapeamento *in situ* e *in silico*. Encontramos uma diversidade significativa na composição de repetições entre espécies e populações, apesar da estrutura cromossômica consistente, principalmente devido à amplificação específica de elementos transponíveis e DNAs satélites. O repeatoma foi composto principalmente por retroelementos, particularmente LINEs e LTRs, enquanto os satelitomas exibiram famílias conservadas, mas também de rápida evolução. Alguns grupos de DNA satélite apresentaram divergência específica da população, indicando rápido turnover, enquanto outros foram comuns a todas as espécies. A ocorrência predominante de DNAs satélites, evidenciada pelo mapeamento *in situ* e *in silico*, em regiões centroméricas e pericentroméricas sugere seu papel estrutural na organização cromossômica. Contudo, as alterações na distribuição e abundância de repetições ocorreram sem rearranjos cromossômicos notáveis, indicando uma disjunção entre a evolução cariotípica e a dinâmica das repetições. Nossos resultados enfatizam a necessidade de integrar metodologias genômicas e citogenéticas para compreender a evolução do genoma, ilustrando que o DNA repetitivo atua como um catalisador principal para a diversidade genômica em aves e demonstrando que alterações genômicas substanciais podem ocorrer dentro de estruturas cromossômicas conservadas.

Palavras-chave: Citogenética; Citogenômica comparativa; DNA satélite; Elementos transponíveis; Repeatoma.

List of Figures

List of Tables

Supplementary Tables

SUMÁRIO

INTRODUCTION

1.1 Repetitive DNAs

Population geneticists such as Hermann J. Muller had already inferred around 1966 that most of DNA was non-coding (Muller., 1966) based on the idea that too many genes could impose a burden of deleterious mutations. However, it was not until the development of Sanger sequencing (Sanger et al., 1977) that these predictions were gradually verified throughout the next decades in large-scale studies such as the human genome project (Collins; Fink, 1995). Initially, those sequences were termed “junk DNA” (Ohno et al., 1972), and the view that they lack utility was maintained for the following years. However, evidence started to pile up on multiple non-coding regions serving biological functions. These include regulatory elements such as TATA box (Lifton et al., 1978), enhancers (Mulligan; Berg., 1983), as well as RNAs capable of relevant functions like gene silencing through siRNAs e miRNAs (Fire et al., 1998).

Simultaneously, cytogenetic studies highlighted the structural importance of non-coding DNA in key chromosomal components like telomeric and centromeric regions (Mitchel et al., 1992) which are largely composed of repetitive DNA, more specifically satellite DNA (satDNA) (Heller et al., 1996). satDNAs are highly repetitive sequences arranged in tandem, whose name derived from their distinct banding in density gradient centrifugation (Kit., 1961). It can be one of the most abundant components of a species repertoire of repetitive DNA (repeatome) not only playing a structural role (Dudka et al., 2025) but also contributing to processes that are not yet fully elucidated such as its capacity to act selfishly during meiosis, biasing its transmission to daughter cells (Courret et al., 2024). Despite their widespread occurrence, the evolutionary mechanisms governing repeatomes and their functional roles in genome architecture remain poorly understood, particularly in lineages marked by limited cytogenomic investigation.

Currently, the term “junk DNA” is considered outdated and misleading, as it fails to reflect the diverse functions of what is now more accurately referred to as non-coding DNA, while an increasing body of research continues to refine our understanding of its biological significance.

1.2 Role of cytogenetics and genomics in understanding repetitive DNAs

Before the genomic era, cytogenetic analysis of satellite DNA and other repetitive elements relied heavily on restriction fragment length polymorphism (RFLP) analysis, followed

by Southern blot hybridization (Keyser et al., 1996). This process allowed researchers to identify some satellite DNA families and other tandem repeats, revealing their strong association with constitutive heterochromatin, a chromosomal component typically enriched in repetitive DNAs and commonly located in centromeric, pericentromeric, and telomeric regions (Charlesworth et al., 1994; Garrido-Ramos, 2017). In addition, repetitive sequences have been shown to play an important role in the differentiation and evolution of sex chromosomes, where the accumulation and diversification of satellite DNAs and other repetitive elements are frequently associated with recombination suppression and structural divergence between sex-specific chromosome pairs (Bachtrog, 2013).

it was fundamentally constrained by its dependence on prior knowledge: researchers had to predict which restriction enzymes would be informative and design specific probes for the sequences of interest, or attempt ones that had been previously tested in closely related species, which yielded unpredictable or limited results. This limitation meant that novel repetitive sequences often went unnoticed, and the method provided only partial views of the total repetitive DNA landscape within a genome. Therefore, by identifying and quantifying only part of these repetitive DNA classes through these restriction-based methods, researchers could make limited cross-species comparisons, and only for those sequence families for which probes had been designed.

Over the following decades, advances in molecular biology, DNA sequencing technologies, and bioinformatics progressively expanded researchers' ability to investigate repetitive genomic regions at larger scales and with greater resolution. The development of high-throughput sequencing platforms paired with the sharp reduction in sequencing costs and the emergence of computational tools capable of clustering and annotating repetitive reads, took repeatome analysis from targeted to genome-wide. These developments ultimately culminated in the emergence of high-throughput sequencing approaches specifically suited for repeatome analysis. Subsequent advances in next-generation sequencing and computational pipelines have enabled the full characterization of the repeatome using low-coverage genomic data (Novák et al., 2020). Short-read sequencing revolutionized the field by allowing massive parallel sequencing of millions of DNA fragments simultaneously, significantly reducing both sequencing costs and time-to-results from months to days. This increased throughput allowed researchers to generate sufficient coverage of repetitive sequences even at low genomic coverage, while computational algorithms improved accuracy in assembling and classifying

reads derived from highly similar repetitive families. By identifying and quantifying the major repetitive DNA classes, these methods enabled cross-species comparisons and the identification of amplification patterns unique to particular lineages. The research on the collection of satDNAs withing a genome (satellitomes), specifically has yielded novel insights on tandem repeat evolution and chromosomal function (Ruiz-Ruano et al., 2016). The library hypothesis proposes that closely related species might share a suite of satellite DNA families from a common ancestor (Fry and Salser., 1977; Camacho et al., 2022), reflecting phylogenetic relationships. On the other hand, comparative studies have revealed that repeat landscapes may differ significantly between closely related species, suggesting that lineage-specific repeat turnover (the gain, loss, and replacement of repetitive DNA sequences over time) is a major driver of genome diversification (Garrido-Ramos et al., 2017; Camacho et al., 2022).

Although a powerful tool, sequence-based analysis alone is not capable of fully resolving the functional roles of repetitive DNAs in genome organization. Even with chromosome-level genome assemblies, repetitive DNA remains a major challenge for complete and accurate assemblies of sequencing reads (Peona et al., 2018). In this scenario, cytogenetics provides a powerful framework to fullfil the gaps in novel technologies. I*n silico* methods such as RIdeogram and karyoploteR enable the mapping and visualization of repetitive DNA reads across assembled genomes, but an accurate, high-resolution assessment of the physical distribution of these sequences along the chromosome can only be achieved through *in situ* mapping of repeats, thereby linking sequence composition to genome architecture (Hao et al., 2020; Gel and Serra, 2017; Peona et al., 2018; Veseljak et al., 2026). The integration of repeatome profiling and fluorescent *in situ* hybridization (FISH) facilitates a multi-scale comprehension of genome evolution, elucidating the spatial arrangement of repetitive elements and their contributions to structural and functional genomic features, such as centromeres, telomeres, and sex chromosomes (Deakin et al., 2019). Overall, cytogenetic analyses contribute substantially to understanding evolutionary relationships in several vertebrate groups, including crocodilians and bats, particularly through the chromosomal mapping of repetitive DNAs and the identification of conserved and lineage-specific genomic patterns (Sales-Oliveira et al., 2024; Deon et al., 2025). Other techniques such as Chromosome Painting (WCP) for example, uses fluorescently labeled DNA probes across entire chromosomes (Pety et al., 2025) which can be useful to detect chromosomal rearrangements. This technique can also be used to identify avian sex chromosomes using glass-microdissected Z chromosomes derived even from phylogenetically distant species such as the red-legged Seriema *Cariama cristata* (Souza et al.,

2025). Besides being a useful tool, this also provides additional evidence for the high level of chromosomal conservation within Aves (Kretschmer; Ferguson-Smith; De Oliveira., 2018)

1.3 Bird Genomes and the Cathartidae family

Bird genomes are generally characterized by long-term conservation of genome organization, punctuated by episodic structural changes (revised in Degrandi et al., 2020), and relatively low proportions of repetitive DNA compared to other vertebrate groups (Ellegren., 2010; Zhang et al., 2014). Yet despite this genomic stability, tandem repeat clusters, such as satellite DNAs (satDNAs) and transposable elements (TEs), serve as key mediators of structural variation, promoting insertions, deletions, and chromosomal rearrangements that fuel adaptive evolution (Schrader & Schmitz., 2019). Beyond structural dynamics, they also shape centromere and heterochromatin evolution via satellite repeat turnover and drive speciation through rapid population divergence (Raskina et al., 2008; Garrido-Ramos et al., 2025).

The New World vultures (Cathartidae) represent a notably appealing yet poorly analyzed model for the study of genomic evolution in avian species. Vulture taxonomy has undergone significant reclassifications over two centuries and is still somewhat debated. Originally, Old World and New World vultures were grouped together in a single genus due to morphological and behavioral similarities between them. Soon, genetic evidence led to the realization that these resemblances were due to evolutionary convergence rather than shared phylogenetic ancestry, and the Cathartidae family was created to allocate the New World vultures while the Old-World vultures were kept in the family Accipitridae (Wink., 1995). Simultaneously, Sibley and Ahlquist's DNA-DNA hybridization studies suggested that New World vultures were more closely related to storks, (Avise et al., 1994; Sibley and Ahlquist., 1990) possibly belonging to Ciconiiformes, which led to the reallocation of vultures to this group. Their conclusion, however, was misleading: the fact that Storks and Vultures' genomes hybridizing is insufficient evidence for close genetic relatedness, and the authors overestimated the resolution of DNA-DNA hybridization. A more plausible explanation involves the conserved karyotypes of New World vultures probably accumulated fewer mutations over time, creating false genetic similarity to equally conserved stork lineages despite their substantial evolutionary separation. Their phylogenetic distance was then clarified through genomic (Jarvis et al., 2014) and cytogenetic (Tagliarini et al., 2010) studies, which provided enough evidence to leave New World vultures within Accipitriformes. More recently however, given the divergence time of Cathartidae from Accipitriformes, Cathartidae has been progressively treated as its own order:

the Cathartiformes specially by ornithology societies such as the American Ornithological Society (AOS).

This small group of Afroaves, which includes seven extant species found in the Americas (being widely distributed across the continent, from temperate North America to the southern regions of South America) exhibits unique ecological and evolutionary features (**Figure 1).**

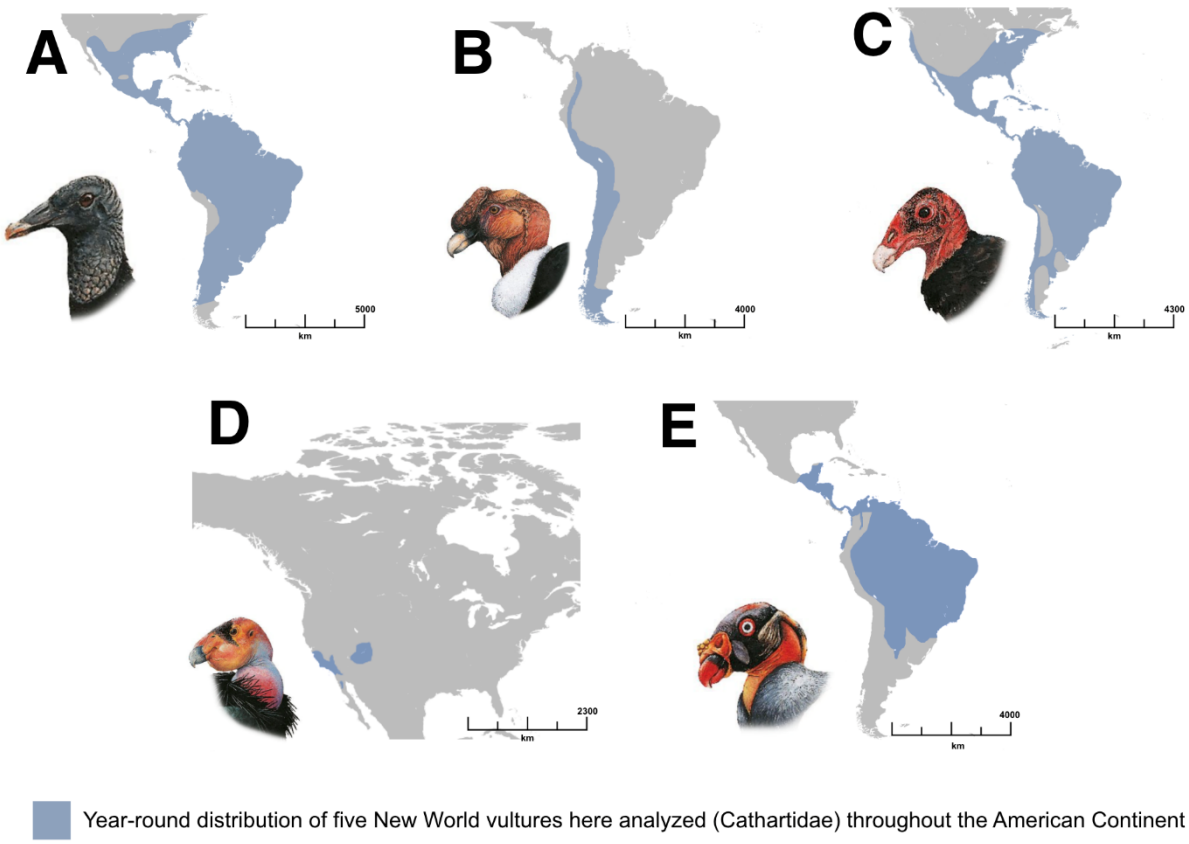

Figure 1: Year-round geographic distribution of the five New World vulture species here analyzed across the Americas: (A) *Coragyps atratus*, (B) *Vultur gryphus*, (C) *Cathartes aura*, (D) *Gymnogyps californianus*, and (E) *Sarcoramphus papa*.

These species present several unique physiological adaptations, such as being obligate scavengers, enhanced visual and olfactory capabilities, and remarkable tolerance to pathogens associated with carrion consumption (Lisney et al., 2013). Together, these traits enable them to locate and rapidly eliminate carcasses, accelerating the flow of energy and nutrients through the food web and also reducing the proliferation of pathogenic microorganisms in the environment. These features make them stand out from other birds. Cytogenetic studies consistently demonstrate conserved 2n = 80 chromosomes for all analyzed Cathartidae species (revised in Degrandi et al., 2020), closely reflecting the hypothesized ancestral bird karyotype, first proposed by Borgaonkar (1969) and subsequently supported by comparative cytogenetic

studies (revised in Kretschmer et al., 2018). Despite this apparent karyotypic stability, very little is known about the organization, diversity, and evolutionary dynamics of repetitive DNA in these species. This knowledge gap is intensified in the black vulture (Coragyps atratus), a species whose populations are genetically structured across North, Central, and South America (Erikson et al., 2022). Using genome-wide population analyses based on reduced-representation sequencing data, the authors identified three major genetic clusters (K = 3), indicating restricted gene flow and partial evolutionary differentiation among geographic regions. This population structuring is likely associated with the species' predominantly non-migratory behavior, which limits long-distance dispersal and promotes regional genetic divergence over time (Winker, 1995). Such structuring has important evolutionary implications, as geographically differentiated populations may accumulate distinct genomic features, including variation in repetitive DNA content and chromosomal organization, potentially contributing to lineage-specific patterns of genome evolution. Since satellite DNAs evolve rapidly, the characterization of satellite DNA families in geographically distinct populations of Coragyps atratus may provide valuable insights into patterns of genomic and chromosomal organization across the Americas.

2. OBJECTIVES

This work aimed to investigate the evolution and genomic organization of repetitive DNA in New World vultures (Cathartidae) through an integrative approach combining genomic and cytogenetic analyses. Specifically, it seeks to (1) characterize the repeatome and satellitome of five representative species, (2) identify patterns of repeat composition, diversification, and lineage-specific amplification, (3) evaluate repeat dynamics across geographically distinct populations of *Coragyps atratus,* and (4) map selected repetitive elements onto chromosomes using fluorescence *in situ* hybridization (FISH). By integrating sequence-based and chromosomal data, this study intends to shed light on the role of repetitive DNA in shaping genome architecture and evolution in a group with conserved karyotypes.

3. MATERIAL AND METHODS

3.1 Sampling

Two black vulture (*Coragyps atratus*) individuals were sampled: one male at the Municipal Zoo of Guarulhos, Guarulhos, São Paulo, Brazil (23°26′35.533″S, 46°33′04.036″W), and one female on the campus of the Universidade Federal de São Carlos (UFSCar), São Carlos, São Paulo, Brazil (21°58′S, 47°52′W). All sampling procedures were conducted under a permit issued by the Brazilian Environmental Agency (ICMBIO/SISBIO) (license 100206-1) and followed the ethical guidelines approved by the Animal Experimentation Ethics Committee of the Universidade Federal de São Carlos (CEUA-UFSCar - 7994170423). To enable comparative genomic analyses across geographic and phylogenetic scales, additional genomic data were retrieved from publicly available NCBI`s sequence read archive (SRA), including two other *C. atratus* populations from Central America (CA) and North America (NA), as well as four additional Cathartidae species with reference genomes assembled at the chromosomal level currently available (**Table 1**).

Table 1: Species included in this study, with their diploid chromosome number (2n) and corresponding NCBI accession codes. Populations of *C. atratus* are distinguished by geographic location: South America (SA), Central America (CA), and North America (NA).

Species	2n	Reference	Short-read accession	NCBI Assembly accession
Cathartes aura	80	Willians & Benirschke (1976)	SRR954276	GCA_000699945.1
Coragyps atratus	80	-	-	GCA_055772195.1
Coragyps atratus (CA)	80	-	SRR19050532	-
Coragyps atratus (NA)	80	-	SRR19050562	-
Coragyps atratus (SA)	80	Tagliarini et al., (2007); present study	present study	-
Gymnogyps californianus	80	Raudsepp et al., (2002)	SRR14067634	GCA_018139145.2
Sarcoramphus papa	80	Takagi & Sasaki (1974)	SRR19167646	GCA_037962945.1
Vultur gryphus	80	Takagi & Sasaki (1974)	SRR33697495	GCA_039700855.2

3.2 Conventional cytogenetics, DNA extraction, and Genome sequencing of *C. atratus*

Fibroblast cell cultures were established from feather pulp samples obtained from a male *C. atratus* and from an embryo of a female *C. atratus*, both from the South American (SA) population. Mitotic chromosomes were then obtained from those cell cultures following the protocol described by Sasaki et al., (1968) and Cioffi et al., (2026). Genomic DNA (gDNA) was extracted using a commercial kit from Cellco (Brazil, São Carlos – SP) based on silica column purification, following the manufacturer's instructions. Detection of constitutive heterochromatin was performed using the C-banding protocol described by Sumner (1972) with adaptations from Lui et al (2012) , which involves acid treatment, alkaline denaturation, and saline incubation of chromosome preparations followed by staining with propidium iodide, allowing the visualization of highly condensed heterochromatic regions enriched in repetitive DNA. From the female sample, paired-end libraries with a 150 bp read length were constructed and subjected to low-coverage sequencing on the BGISEQ-500 platform at BGI (BGI Shenzhen Corporation, Shenzhen, China). The sequencing produced a total of 18.7 GB of raw reads, currently deposited in the Sequence Read Archive (SRA) under accession number SRR38244970.

3.3 Repeatome characterization

Repeatome analysis integrated the newly generated data for *C. atratus* (see above) with short-read datasets from two other populations (Panama, Central America - CA, and United States, North America - NA), as well as from four other Cathartidae species (**Table 1**). Characterization of satellite DNA was performed via graph-based clustering of repetitive elements using RepeatExplorer and TAREAN pipelines (Novák et al., 2020), implemented on the Galaxy server (https://repeatexplorer-elixir.cerit-sc.cz/galaxy/). Quality assessment ($Q < 20$) and trimming of sequencing reads were performed with Trimmomatic v.3.0 (Bolger et al., 2011) using the parameters LEADING:3, TRAILING:3, SLIDINGWINDOW:4:20, MINLEN:100, and CROP. A subset of $2 \times 500{,}000$ paired reads from the three populations of *C. atratus* and the remaining Cathartidae species was produced and individually loaded into TAREAN. SatDNAs were selected based on their circular topography graph and filtered from trimmed reads using Deconseq v0.4.3 (Schmieder and Edwards., 2011). For each library, $2 \times 500{,}000$ paired reads were subsampled and iteratively run through TAREAN and Deconseq until no new satDNAs emerged. This process is known as the satminer protocol (Ruiz-Ruano

et al., 2016), which is a toolkit used to characterize underrepresented satDNAs in the original raw reads, but detectable after the filtrage of found satellite sequences (**Figure 2**). Finally, each satellitome dataset was screened for non-satellite repetitive DNAs (e.g., multigene families and transposable elements). Satellite DNA catalogs are available in GenBank under accession numbers PZ325739-PZ325831, PZ244963-PZ244975 and PZ407887-PZ407897.

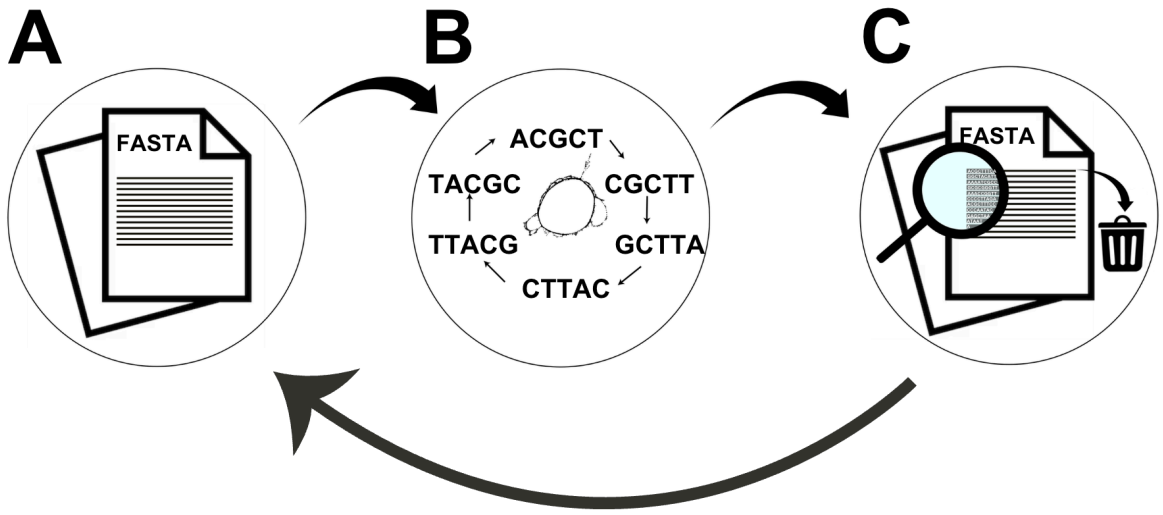

Figure 2: SatMiner protocol workflow. Genome short reads (A) are submitted to TAREAN for putative satDNA identification through graph-based sequence clustering (B), generating consensus sequences. Identified satellite sequences are then filtered out from the original FASTA reads (C), and the remaining reads are resubmitted for iterative detection of additional satDNA families until no new satellites are found.

Transposable elements (TEs) were characterized using DNApipeTE v1.4 (Goubert et al., 2015), a pipeline designed for comprehensive repeat analysis from low-coverage short-read sequencing data. Single-end forward reads from each trimmed library were used as input, and the pipeline was run with default parameters under a 0.1x genome coverage setting. To estimate the proportion of the genome occupied by each repetitive DNA class, we used a reference genome size of 1.3 Gb, based on prior genome assemblies of closely related species (GCA_018139145.2; GCA_037962945.1; GCA_039700855.2; GCA_055772195.1; GCA_000699945.1).

3.4 Estimating the abundance, divergence, and homology of satellite DNAs

The relative abundance of each satellite DNA was calculated as the proportion of satDNA-derived nucleotides relative to the total number of genomic nucleotides analyzed. For this purpose, 2 × 5,000,000 paired-end reads were sampled from each genome. Alignments of the isolated satDNAs against their respective genomes were performed using the "cross-match" tool in RepeatMasker v.4.16 (Smit et al., 2017), together with a custom Python script (https://github.com/Dfam-consortium/RepeatMasker). Genetic distances were estimated using the Kimura-2-parameter model, with calculations being performed through the calcDivergenceFromAlign.py script in RepeatMasker (Smit et al., 2017). This parameter is important to account for substitution rates in DNA bases over time (transversions and transitions) which is a natural occurence. The resulting values estimated by Kimura-2 can then be used to account for repeat changes and homogenitization dynamics within the genome.

Satellite DNAs were numbered according to decreasing genomic abundance and named following the criteria proposed by Ruiz-Ruano et al., (2016). Accordingly, species abbreviations were assigned as follows: *C. aura* (Cau), *C. atratus* (Cat), *G. californianus* (Gca), *S. papa* (Spa), and *V. gryphus* (Vgr). For *C. atratus*, an additional sufix denoting its geographic origin was added for populations from South America (SA), Central America (CA), and North America (NA). To assess the conservation of satDNA sequences across species and track the homology groups of each satellite, similarity searches were conducted using a custom Python script, rm_homology.py (https://github.com/fjruizruano/satminer/blob/master/rm_homology.py) (Ruiz-Ruano et al., 2016), followed by additional inspection through manual sequence alignments. Classification followed the criteria established by Ruiz-Ruano et al., (2016), using identity thresholds of >80% for variants and <80% for superfamilies/Groups.

3.5 Primer Design and Probe labelling

Based on the South American (SA) population of *C. atratus*, eleven primer sets were manually designed for selective satDNA amplification for further *in situ* chromosome mapping (**Supplementary Table 1**). A minimum sequence length threshold of >30 bp was applied to each satDNA family to ensure reliable primer design. Designing primers for repetitive sequences requires special attention to avoid non-specific amplification. This is particularly important for satDNAs, which often share highly similar sequence regions, increasing the risk

that poorly designed primers amplify multiple satellite families. Due to its short repeat unit length, CatSASat06-30 was synthesized directly with biotin molecules incorporated at the 5′ end. The PCR amplification protocol consisted of an initial denaturation step at 95 °C for 5 minutes, followed by 32 cycles of 95 °C for 20 seconds, annealing at 56 °C for 40 seconds, and extension at 72 °C for 50 seconds, with a final extension step at 72 °C for 5 minutes. Amplification was confirmed by agarose gel electrophoresis (2%), and PCR products were quantified using a NanoDrop spectrophotometer (ThermoFisher Scientific, Branchburg, NJ, USA). For probe construction, amplicons were labeled via nick translation using Atto488-dUTP and Atto550-dUTP nucleotides (Jena Biosciences, Jena, Germany) following the manufacturer's instructions. The resulting probes were subsequently applied for Fluorescence *In Situ* Hybridization (FISH) performed following the protocol of Kretschmer et al., (2026).

3.6 *In situ* Mapping of repetitive DNAs

Besides the satDNA families, we also *in situ* mapped other repetitive DNA classes, including the 18S rDNA, which can give us insight into chromossomal rearrangements, and seven microsatellites (A_n, CA_n, CAG_n, CAT_n, GA_n, TAAn, and TAC_n) that can highlight the distribution of repeat regions. While the 18S rDNA fragments were obtained by PCR following the method described in Cioffi et al., (2009), the microsatellites were labeled directly with Cy3 at the 5’ end during synthesis (VBC Biotech, Vienna, Austria) and used for FISH.

3.7 Comparative Genome Hybridization (CGH) and Whole Chromosome Painting (WCP)

To properly distinguish the Z and W sex chromosomes of *C. atratus*, whole chromosome painting experiments (WCP) were performed in addition to intra-specific comparative genome hybridization (CGH). CGH was first utilized to map sex-specific sequences by co-hybridizing differentially labeled male and female genomic DNA onto female metaphase spreads. Male and female gDNAs were labeled using a nick-translation kit (Jena Biosciences) according to manufacturer’s instructions with Atto488-dUTP and Atto550-dUTP, respectively. To block repetitive sequences and reduce nonspecific hybridization, 2 µg of unlabeled species-specific COT-1 DNA (produced according to Zwick et al., 1997) was added to the hybridization mix. COT-1 is commonly used to suppress nonspecific hybridization signals in FISH experiments, and is undispensable when doing CGH due to the high volume of repetitive sequences. This technique consisted of applying male-derived COT-1 DNA to female

metaphase spreads. The final hybridization mixture contained 500 ng each of labeled male and female gDNA, 2 µg of COT-1 DNA, and 20 µl of hybridization buffer (50% formamide, 2× SSC, 10% SDS, 10% dextran sulfate, and Denhardt's buffer, pH 7.0), which was applied to female mitotic chromosome spreads and hybridized for 72 h. The probe-to-COT-1 DNA ratio was selected based on previous investigations (Kretschmer et al., 2024; Oliveira et al., 2024; Souza et al., 2025). Whole chromosome painting experiments were conducted using a Z chromosome-specific probe derived from *Cariama cristata* (Souza et al., 2025), as well as whole chromosome-specific probes derived from *Pseudastur albicollis* (formerly *Leucopternis albicollis*) corresponding to pairs homologous to chicken chromosomes. 1 (LAL 6, LAL 7, LAL 15, and LAL 18), chrm. 2 (LAL 2 and LAL 4), chrm. 3 (LAL 17) and chrm. 5 (LAL 5) obtained from Oliveira et al., (2010) in order to detect rearrangements using the White Hawk as proxy. Using degenerate oligonucleotide-primed PCR (DOP-PCR), the probes were labeled with Spectrum-Orange-dUTP. The labeling reaction consisted of 28 amplification cycles and utilized 1 µL of original DNA template (Yang and Graphodatsky., 2009). All hybridization procedures followed the conditions described in Kretschmer et al., (2026).

3.8 Statistics and reproducibility

At least 20 metaphases were examined to confirm the diploid number (2n) and the FISH results. The best metaphases were captured using an Olympus BX50 microscope (Olympus Corporation, Ishikawa, Japan) coupled with the CoolSNAP capture system and Image Pro Plus 4.1 software (Media Cybernetics, Silver Spring, MD, USA).

3.9 *In silico* Mapping of satDNAs

Comparative *in silico* mapping of six satDNA families was performed using chromosome-level assemblies of *G. californianus* (GCA_018139145.2), *S. papa* (GCA_037962945.1), and *V. gryphus* (GCA_039700855.2) (**Table 1**). Selected satellite DNAs that exhibited high sequence similarity and were broadly distributed across nearly all species analyzed. To assess the chromosomal distribution of these repeats, satDNAs were queried against chromosome-level genome assemblies using BLASTn. Resulting alignments were rigorously filtered to retain only high-confidence matches based on strict criteria for e-value, percent identity, and alignment lengths. The coordinates of filtered hits were then converted to BED format to facilitate downstream visualization. Genomic mapping was carried out with the

karyoploteR package in R, which generates chromosomal ideograms and enables precise localization of sequence hits. For each species, a ".genome" file specifying chromosome names and sizes was used alongside the BED files to evaluate the abundance, spatial arrangement, and distribution patterns of satellite DNAs across chromosomes.

4. RESULTS

4.1 Karyotype description and characterization of sex chromosomes in *C. atratus*

All the *C. atratus* specimens analyzed consistently revealed a diploid number of 2n = 80 (**Figure 3**). The karyotype likely resembles the proposed avian ancestral karyotype, comprising ten distinguishable pairs of macrochromosomes and remaining dot-like microchromosomes. This conservatism was further corroborated by whole-chromosome painting (WCP) experiments using *Leucopternis albicollis* (LAL) probes, which revealed homology patterns consistent with the ancestral avian chromosomal organization **(Figure 4).** While the W chromosome is a small metacentric element, the Z chromosome is slightly larger and submetacentric, ranking among the fifth- to sixth-largest macrochromosomes (**Figure 3**). However, despite such morphological divergence, intraspecific comparative genomic hybridization (CGH) did not reveal any accumulation of female-specific sequences on the W chromosome **(Figure 5)** Coragyps atratus are still poorly differentiated at the sequence level. Sequential whole-chromosome painting (WCP) and C-banding further highlighted the heterochromatic content of both sex chromosomes, confirming the moderate enrichment of heterochromatin on the W chromosome alongside its relatively lower abundance on the Z chromosome **(Figure 6).**

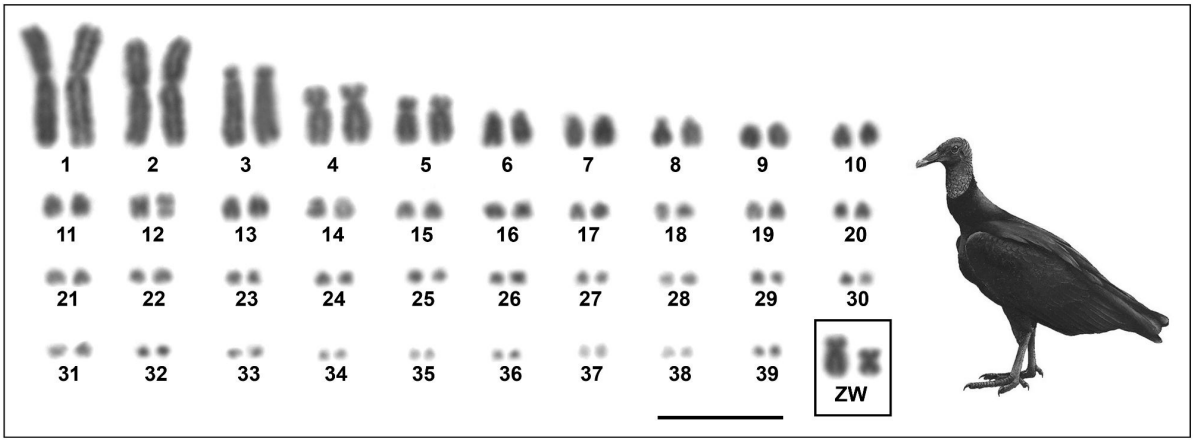

Figure 3: Female Giemsa-stained karyotype of the South American *Coragyps atratus* collected in Brazil. The ZW sex chromosomes are boxed. Scale Bar = 20µm

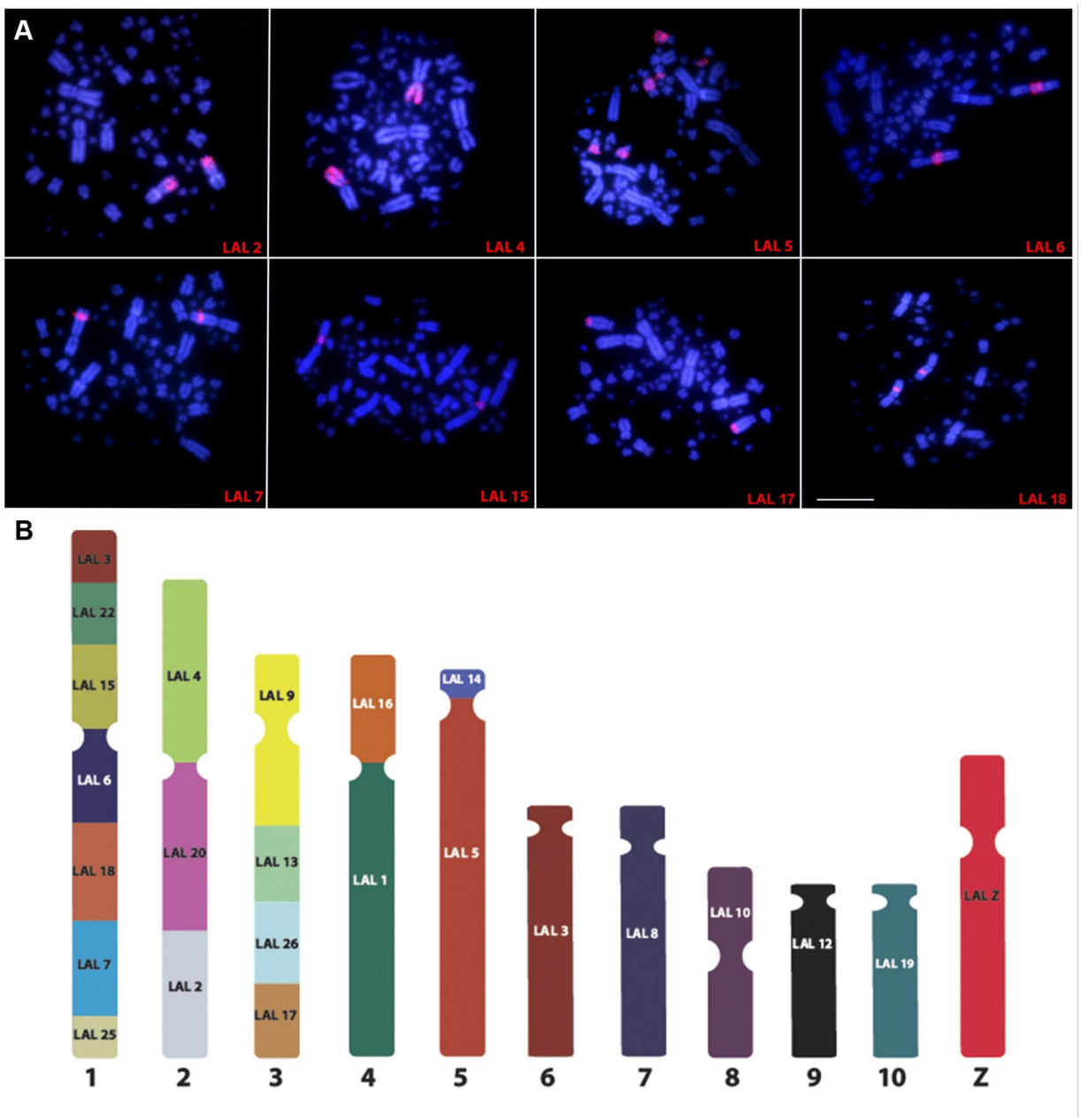

Figure 4: (A) Female metaphase of *Coragyps atratus* hybridized with different LAL WCP-probes **(B)** Homology map between chicken macrochromosomes and white hawk paints based on Furo et al., 2020. Scale bar = 20 µm.

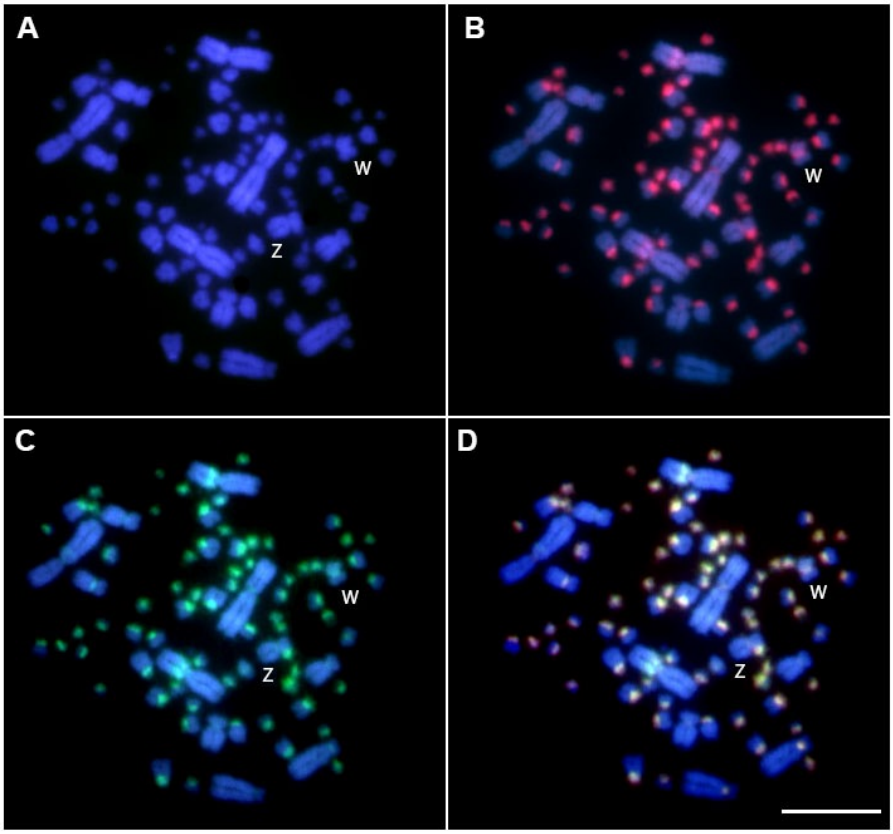

Figure 5: Intraspecific genomic hybridization using male (red) and female (green) genomic DNA probes from *Coragyps atratus* hybridized on female metaphase chromosomes. (A) DAPI-stained metaphase; (B) Hybridization pattern of the male-derived probe (red); (C) hybridization pattern of the female-derived probe (green); and (D) merged images of both genomic probes and DAPI staining. The ZW sex chromosomes are indicated. Bar: 20 μm.

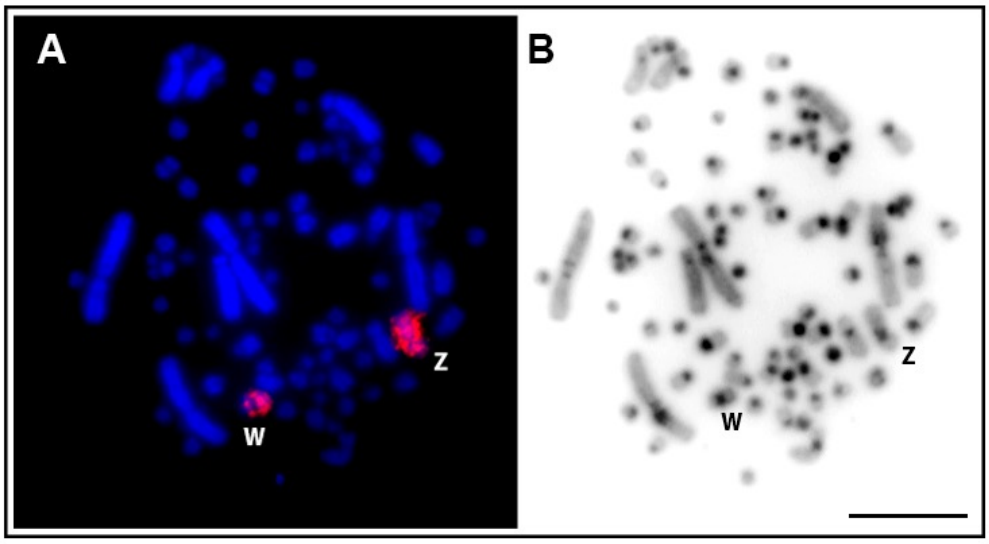

Figure 6: Female chromosomes of *Coragyps atratus* hybridized with WCP-FISH using a Z chromosome-specific probe from *Cariama cristata* (A), sequentially C-banded (B). The ZW sex chromosomes are indicated. Scale bar = 20 μm.

4.2 Characterization and comparison of Cathartidae repeatomes

The repeatome profiling of five Cathartidae species demonstrated different amplifications of the two major TE classes: DNA transposons and retroelements, such as Long Interspersed Nuclear Elements (LINEs) and Long terminal repeats (LTRs) (**Figure 7**, **Supplementary Table 2**). Retrotransposons were the most abundant component across all genomes, ranging from 13% in *C. aurea* to 18% in *V. gryphus*. Consistently, *V. gryphus* exhibited the highest overall repetitive DNA content (23%), whereas the remaining species ranged from 16% to 20.1% **(Figure 7, Supplementary Table 2)**. A detailed survey of retroelements demonstrated that L2/CR1/Rex and retrovirus-related elements are the predominant contributors, accounting for a substantial proportion of the repeatome variation observed among the five species **(Supplementary Table 2**).

Intraspecific comparisons among *C. atratus* lineages revealed pronounced population-specific repeat landscapes (**Figure 8, Supplementary Table 2**). The North American (NA), Central American (CA) and South American (SA) populations showed comparable repeat contents (~16%), dominated by LINEs and LTRs, indicating a more typical retroelement-driven genomic architecture (**Figure 8, Supplementary Table 2**). However, it is important to note that satellite DNA is often underrepresented in DNApipeTE analyses, as a single satDNA can account for up to 5.63% of the genome **(Supplementary Table 3)** when using pipelines specifically designed to characterize satellitomes.

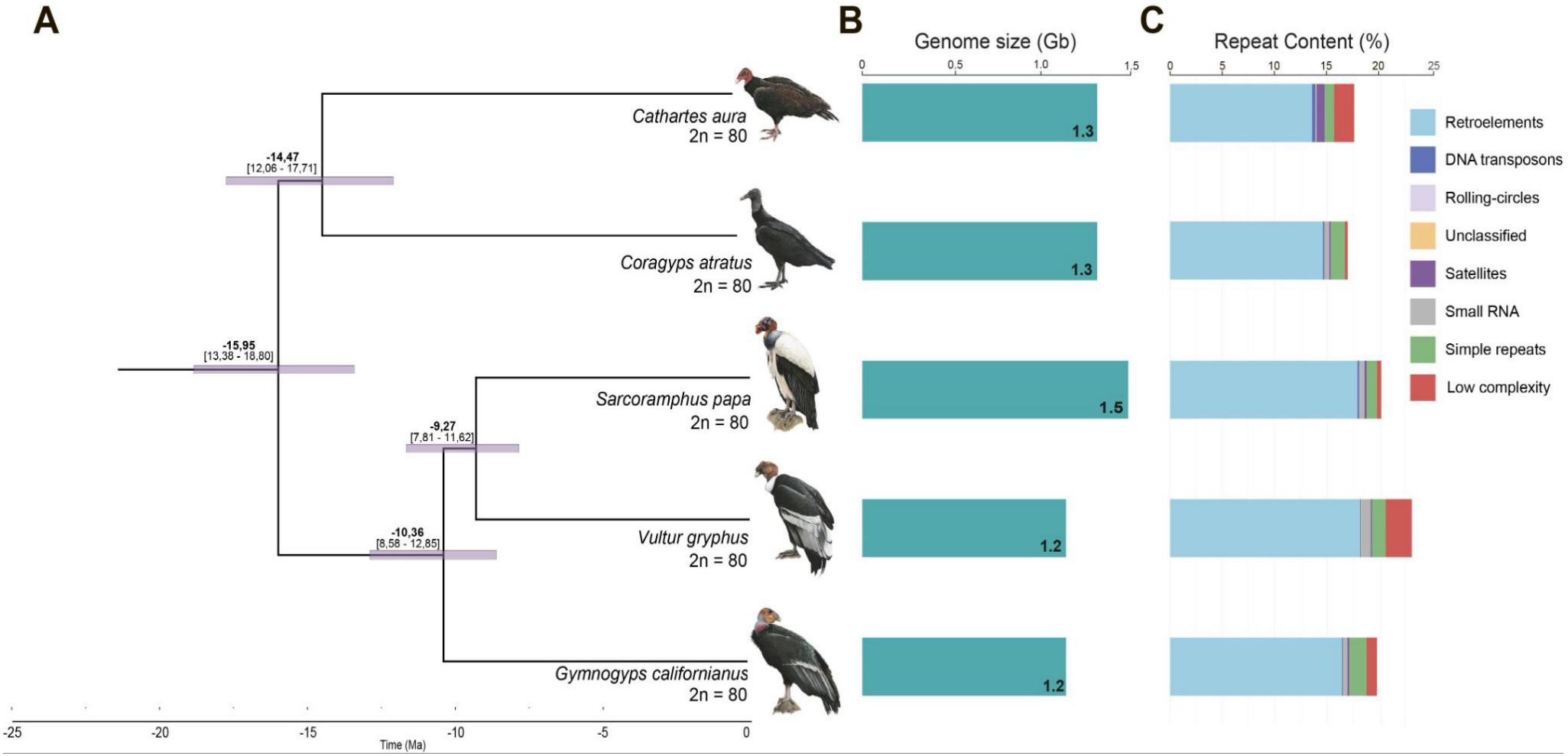

Figure 7: (A) Phylogeny of the five Cathartidae species here analyzed, showing their divergence times (Modified from Johson et al., 2016 (B) The genome sizes and (C) repeat content of each species are also displayed. The mean divergence time is labeled at each node.

4.3 General features and homology of Cathartidae satellitomes

Characterization of satDNA families in *C. atratus* revealed population-level variation in satellitome composition, with 12, 20, and 23 families identified in SA, NA, and CA populations, respectively (**Supplementary Table 3**). A subset of satDNA families (Groups 2, 4, 5, and 8) was shared across all populations, supporting the presence of a conserved satellitome core in *C. atratus*. Within this core, four sets of variants are present in every population (sequences with more than 80% of similarity) (**Table 2, Supplementary Table 4**). In comparison, across other species examined, satellitome sizes were comparable, comprising seven families in *G. californianus*, nine in *C. aura*, 10 in *V. gryphus,* and 11 in *S. papa*. Repeat unit length (RUL) varied from 16 bp to 5021 bp, while GC content was similar among the satDNA families, with most sequences exceeding 50%. However, the *C. atratus* NA population diverged from this pattern, as most satDNA families exhibited higher A+T content, indicating population-level differences in base composition (**Supplementary Table 3**). Kimura-2 substitution (K2P) profiles were largely consistent across species, indicating a recent expansion of satDNAs, as denoted by low divergence values. An exception was CatSASat02-180, which showed markedly elevated divergence, with K2P values reaching ~30% (**Figures 8 and 9**).

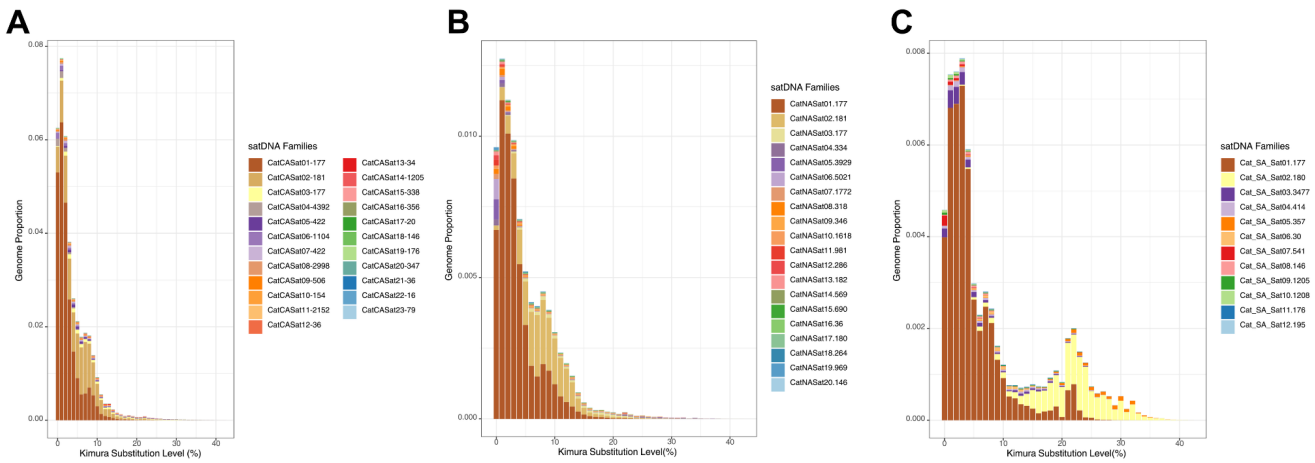

Figure 8: Repeat landscapes showing the Kimura 2-parameter (K2P) divergence of satellite DNA families in three geographic populations of *Coragyps atratus*: (A) Central America (CA), (B) North America (NA), and (C) South America (SA).

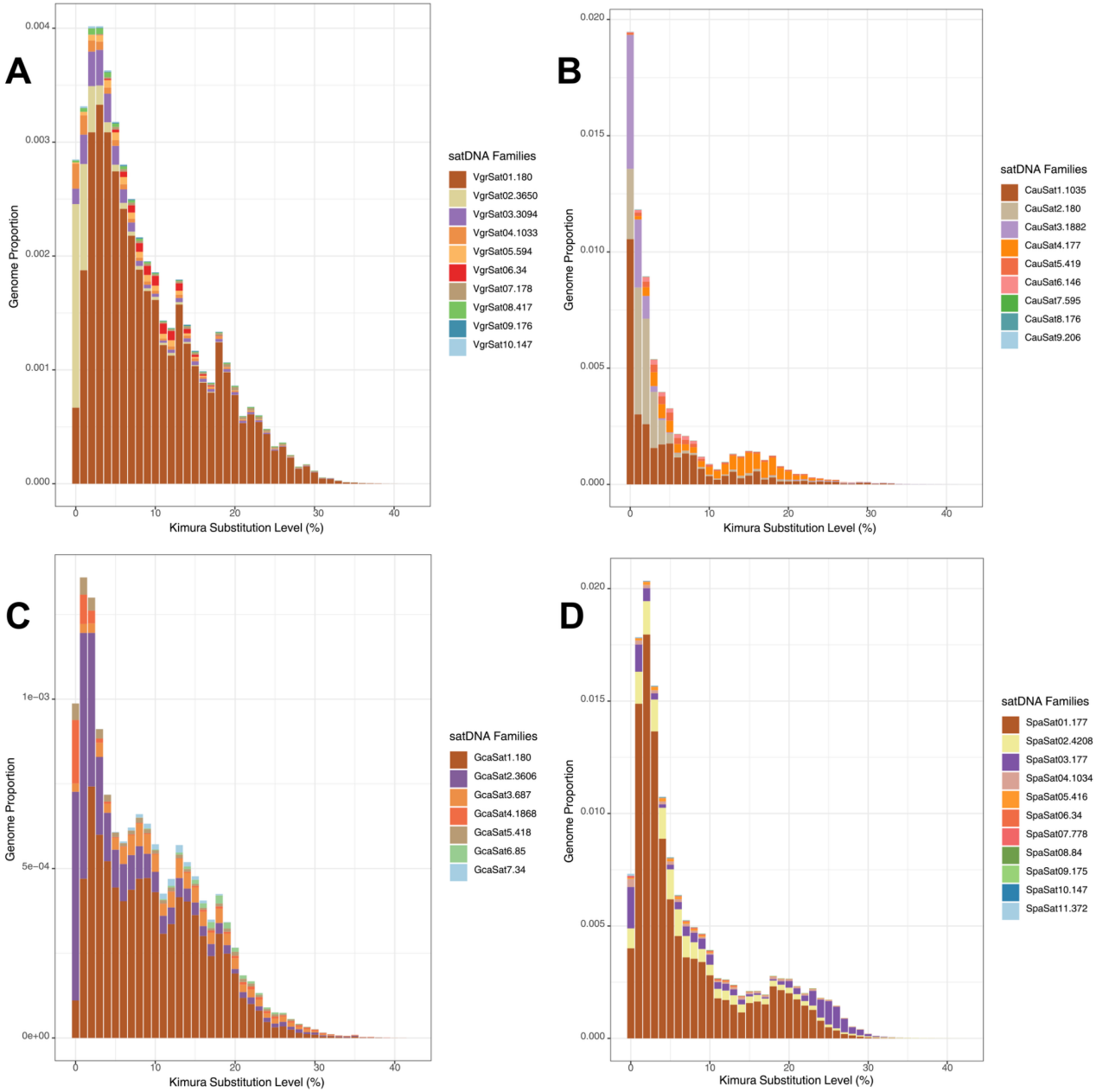

Figure 9: Repeat landscapes showing the Kimura 2-parameter (K2P) divergence of satellite DNA families in four New World Vultures (Cathartidae): A) *Vultur gryphus,* B) *Cathartes aurea*, C) *Gymnogyps californianus,* and D) *Sarcoramphus papa*.

Cross-species comparisons identified nine satDNA groups (Groups 1-9) across the five analyzed species, with sequence similarity ranging from 67.5% to 100% (**Table 2**). Groups 2 and 4 were the most broadly conserved, each comprising multiple variants shared among all five species (**Table 2, Supplementary Table 4**). Within Group 2, four satDNA families (CauSat08-176, CatSASat11-176, SpaSat09-175, and VrgSat09-176) formed a distinct cluster that showed no substantial similarity to other members upon alignment (**Table 2, Supplementary Table 4**). This divergence suggests that, although these sequences exhibit detectable homology with another Group 2 satDNAs, they likely followed a distinct evolutionary trajectory. To emphasize these differences, we designate them as subgroup 2B (**Table 2**). Group 5 was identified in all species except *G. californianus*, with variants exhibiting exceptionally high sequence similarity (95.3 - 100%) and a relatively stable repeat unit length of ˜146 bp (**Table 2, Supplementary Table 4**). Similarly, Group 1 was absent only in the NA population of *C. atratus* and in *G. californianus*. In contrast, Groups 3 and 9 displayed the most restricted distribution among the identified groups: Group 3 was shared exclusively between *C. aura* (CauSat03-1882) and *G. californianus* (GcaSat04-1868), whereas Group 9 was detected in *G. californianus*, *S. papa*, and *V. gryphus*, with sequence similarity ranging from 91.4% to 100% (**Table 2**, **Supplementary Table 4).**

Table 2: Homology groups and sequence similarities among five Cathartidae satellitomes.

Group	***Cathartes aura***	***Coragyps atratus (CA)***	***Coragyps atratus (NA)***	***Coragyps atratus (SA)***	***Gymnogips californianus***	***Sarcoramphus papa***	***Vultur Gryphus***	**Similarity (%)**
Group 1	CauSat01-1035	CatCASat14-1205		CatSASat09-1205		SpaSat04-1034	VgrSat04-1033	82.4 - 93.3
Group 2	CauSat02-180	CatCASat01-177	CatNASat01-177	CatSASat01-177	GcaSat01-180	SpaSat01-177	VgrSat01-180	
	CauSat04-177	CatCASat02-181	CatNASat02-181			SpaSat03-177		74.5 - 100
			CatNASat03-177					
Group 2B*	CauSat08-176			CatSASat11-176		SpaSat09-175	VgrSat09-176	89.8 - 94
Group 3	CauSat03-1882				GcaSat04-1868			90.7
Group 4	CauSat05-419	CatCASat07-422	CatNASat09-346	CatSASat04-414	GcaSat05-418	SpaSat05-416	VgrSat08-417	70.1 - 94
Group 5	CauSat06-146	CatCASat18-146	CatNASat20-146	CatSASat08-146		SpaSat10-147	VgrSat10-147	95.3 - 100
Group 6	CauSat07-595	CatCASat09-506	CatNASat14-569					67.5- 85.2
Group 7	CauSat9-206			CatSASat12-195				92.23%
Group 8		CatCASat12-36	CatNASat16-36	CatSASat06-30				77.8 - 100
Group 9					GcaSat02-3606	SpaSat02-4208	VgrSat03-3094	70.6 - 82.3
Group 10		CatCASat13-34			GcaSat07-34	SpaSat06-34	VgrSat06-34	91.4 - 100
Group 11		CatCASat08-2998		CatSASat03-3477			VgrSat02-3650	83.5 - 90.7

4.3 Chromosome mapping of 18S rRNA gene, microsatellites, and CatSatDNAs

In situ mapping of 18S rRNA genes revealed positive FISH signals on a microchromosome pair. Of the nine microsatellites evaluated, only $(CA)_n$ and $(CAG)_n$ yielded detectable signals, localizing to the second-largest metacentric and a medium-sized acrocentric pair, respectively **(Figure 10).** Most CatSASatDNAs exhibited a predominantly centromeric or pericentromeric distribution, being located in the centromeric regions across all (or multiple) chromosome pairs (e.g., CatSASat2-180, CatSASat1-177, CatSASat3-3477, CatSASat4-414, CatSASat5-357, CatSASat6-30, CatSASat8-146, CatSASat9-1205, and CatSASat11-176) or ii) were confined to one or a few specific chromosome pairs (e.g., CatSASat5-357, CatSASat6-30, CatSASat8-146, CatSASat9-1205, and CatSASat11-176). CatSASat7-541 and CatSASat12-195 deviated from this predominant pattern, being localized to interstitial and terminal regions, respectively, on chromosome pair 2 **(Figure 11).**

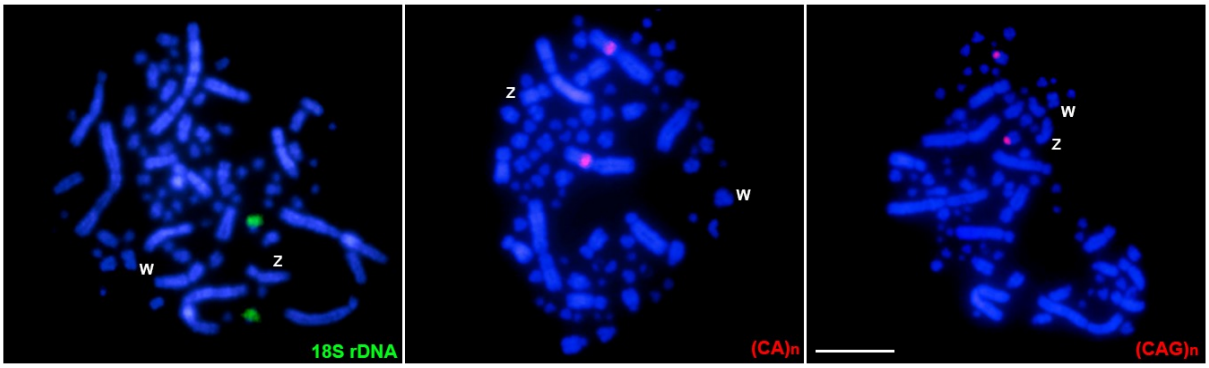

Figure 10: Female metaphases of *Coragyps atratus* showing hybridization signals of distinct repetitive DNA probes. The name of each repeat is detailed in the bottom-right corner in red (ATTO-550-dUTP labeled) or green (ATTO-488-dUTP labeled). The ZW sex chromosomes are indicated. Scale bar = 20 µm.

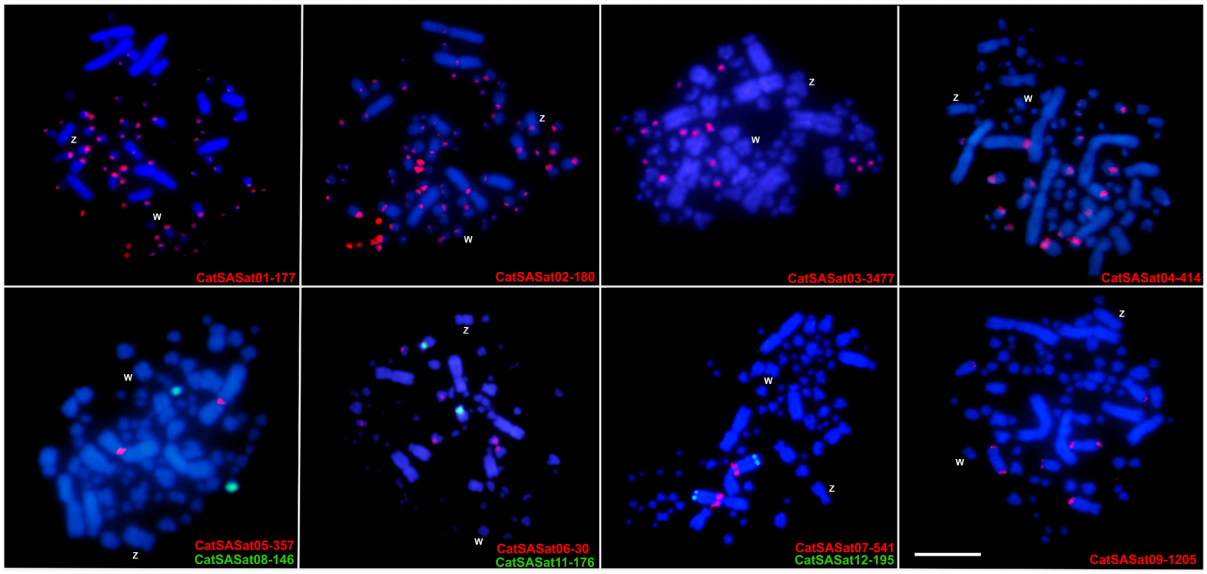

Figure 11: Female metaphases of *Coragyps atratus* showing hybridization signals of CatSatDNAs as probes. The names of each satellite DNA family are detailed in the bottom-right corner in red (Atto-550-dUTP labeled) or green (ATTO-488-dUTP labeled). The ZW sex chromosomes are indicated. Scale bar = 20 µm.

4.4 *In silico* profiling of conserved satDNAs across genome assemblies

Genome-wide mapping of six conserved satDNA families (two representatives of Group 2 and one each of Groups 4, 5, 9, and 10) (**Table 2**) displayed significant distribution patterns in the genome of *S. papa* (**Figure 12**). These satDNA families were predominantly clustered in pericentromeric regions across all chromosomes, with occasional subtelomeric signals, consistent with the FISH patterns observed in *C. atratus* (**Figure 11).** In contrast, *in silico* mapping in the two other available chromosome-level assemblies of New World vultures (*G. californianus* and *V. gryphus*) yielded only very faint, dispersed signals without evident clustering, likely reflecting the incomplete assembly of repetitive DNA regions **(Figure 13**).

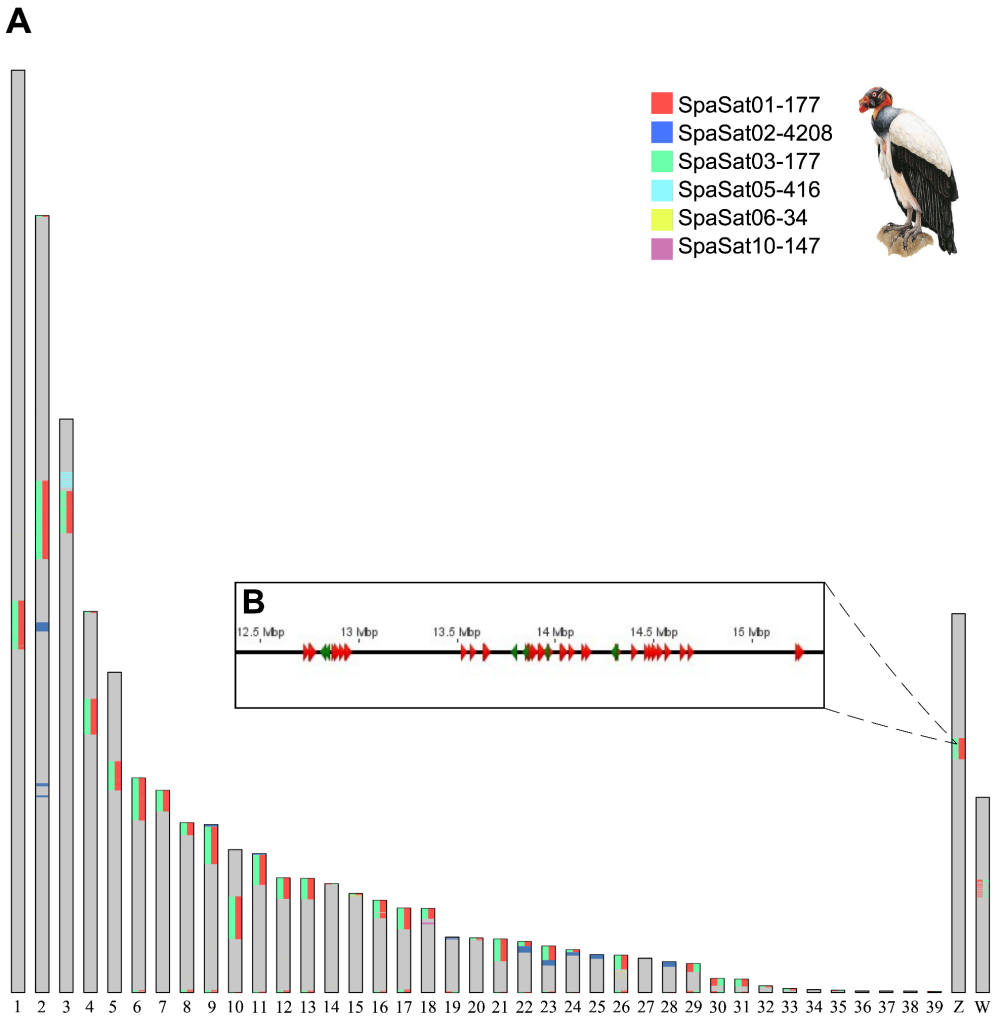

Figure 12: A) *In silico* mapping of highly conserved satDNAs in the *Sarcoramphus papa* genome. Colors denote individual SatDNA families, as indicated in the upper legend. B) Close-up view of the spatial organization of SpaSat01-177 and SpaSat03-177 within their shared genomic locus.

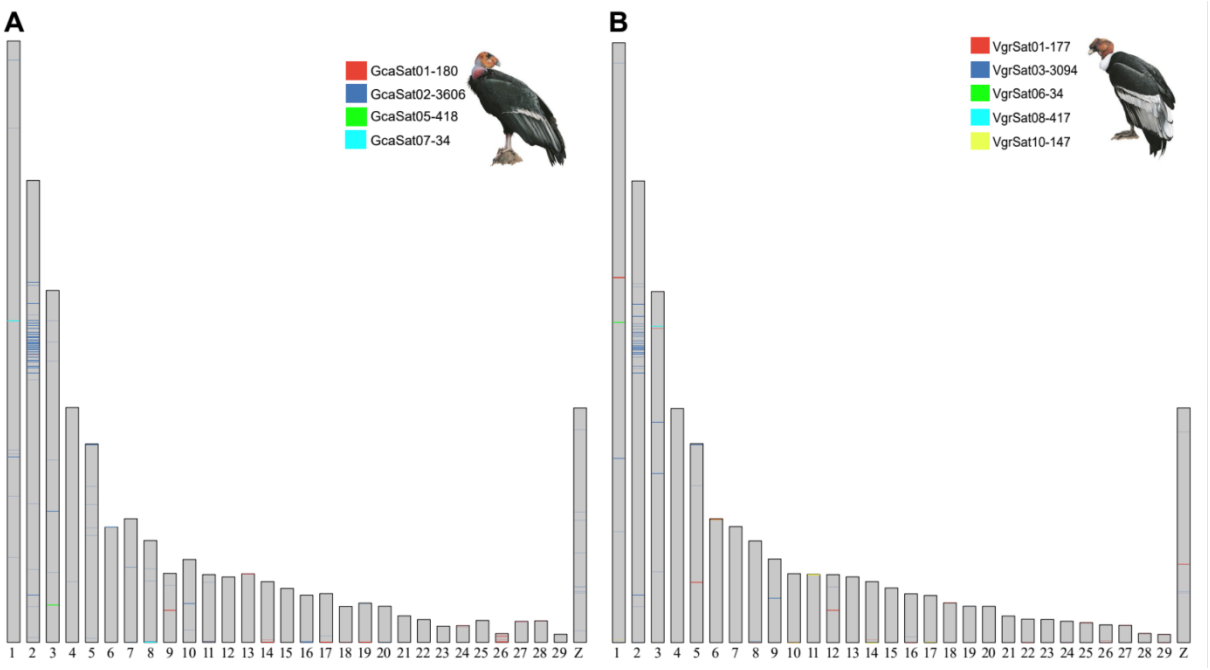

Figure 13: *In silico* mapping of highly conserved satDNAs in the *Gymnogyps californianus* (A) and *Vultur gryphus* (B) genomes.

5. DISCUSSION

Repetitive DNAs are a major driver of genome evolution, yet their contribution to genome architecture remains poorly understood in lineages with conserved karyotypes. Here, by integrating repeatome profiling with cytogenetics, we show that New World vultures combine remarkable chromosomal stability with dynamic and lineage-specific repeat landscapes. This apparent contradiction demonstrates that significant genomic diversity may happen without major karyotypic changes, mostly owing to satellite DNA turnover and variable transposable element activity. New World vultures diverged around 15 million years ago, and the difference in repeat content within Cathartidae reflects unique evolutionary pathways of repeated elements subsequent to lineage splitting. The variations among *C. atratus* populations indicate that repeatome evolution might also occur on a minor scale, considering the constraints on gene flow and the specialized nature of demographic history. The primarily sedentary habit of black vultures (Jackson, 1988) likely leads to restricted gene exchange among continental populations, thus providing a foundation for the observed difference in repetitive DNA profiles. Such population-level differentiation supports the view that repetitive DNA evolves rapidly and can reflect recent demographic and evolutionary processes (Belyayev, 2014). Considering that repetitive elements can follow distinct evolutionary paths after lineage separation, variation in repeat composition is predicted even among closely related species (Biscotti et al., 2015).

The composition, structure, and evolution of repetitive DNA across species diversification remain a central topic in genome biology. Despite their structural and functional roles, repetitive sequences are often referred to as the “dark matter of the genome.” Avian genomes have historically been considered to contain low proportions of repetitive DNA, typically around 10%, with only a few species deviating from this pattern (e.g., Piciformes (Zhang et al., 2014) and Accipitriformes (Canesin et al., 2024; Mota-Souza et al., submitted). This percentage is comparably low to other vertebrates, whose genomes often contain more than 40% of repetitive DNA (Reichwald et al., 2009; Kapusta et al., 2017; Kosch et al., 2024; Moreira et al., 2025). However, as non-model species have been increasingly analyzed, the complexity of the structure and distribution of repetitive DNA in avian genomes has become more evident. In particular, centromeric regions and sex chromosomes are enriched with diverse repetitive sequences, especially satellite DNAs, which show high rates of diversification even among closely related species (Crepaldi et al., 2020; Courret et al., 2024; Lisachova et al., 2025). Within this dynamic and context-dependent landscape, certain satDNA families nevertheless exhibit remarkable conservation, raising questions about the forces that constrain their evolution.

FISH experiments using LAL probes **(Figure 4)** revealed a high degree of chromosomal stability, as the hybridization signals in Coragyps atratus corresponded to chromosomal regions previously described in Gallus gallus. Because the chromosomal organization of these loci is well established in chicken, the conserved distribution pattern observed here suggests the maintenance of the putative ancestral caryotupe and a low rate of large-scale chromosomal rearrangements during the evolutionary history of this species. The following FISH experiments (**Figure 11**) revealed a clear enrichment of CatSASat02 in pericentromeric regions across all chromosomes, indicating a recurrent spatial association with centromere-proximal domains. Among the conserved satellites identified, this satellite from group 2 is particularly noteworthy due to its consistent presence across the analyzed taxa. Complementary *in silico mapping* in *S. papa* **(Figure 12)** further supports this distribution pattern, and although faint, the genome mapping signals in *G. californianus* and *V. gryphus* also follow this trend (**Figure 13**). This contrast in resolution highlights how repetitive DNA is often underrepresented in genome assemblies. Both genomes present a significant amount of unplaced scaffolds, while *S. papa* has very few, which could indicate how a more precise genome-assembly might avoid this underrepresentation.

In addition, the remaining satellites within group 2 **(Table 2)** correspond to the most abundant repeats in the other species examined, reinforcing a pattern of high genomic representation. This observation aligns with previous studies in birds showing that highly abundant satellite families are often located at or near centromeric regions (Talbert & Henikoff, 2020; Peona et al., 2021; Kretschmer et al., 2024; Souza et al., 2024; Pozzobon et al., 2025). Taken together, the combination of high abundance, conservation across species, and recurrent localization in pericentromeric regions suggests that group 2 satellites may be associated with centromere-related genomic compartments. However, given the dynamic nature of avian centromeres and the absence of direct functional assays, additional evidence is required to more accurately resolve the nature of this association. By identifying satellite families conserved across multiple species, our results challenge the expectation of rapid and unconstrained turnover of centromeric repeats. This suggests that, despite the well-established epigenetic determination of centromere identity (Henikoff et al., 2001; Talbert & Henikoff, 2022), sequence-level constraints may operate within specific lineages. Although the convergence of cytogenetic and *in silico* evidence supports a consistent association of these sequences with centromere-proximal regions, an important question remains as to whether they are integral components of the functional centromere or instead predominantly located within adjacent

pericentromeric heterochromatin. Three non-mutually exclusive mechanisms may explain this conservation: (i) structural roles in pericentromeric heterochromatin organization, potentially playing roles in chromatin compaction, cohesion, or the establishment of a genomic environment conducive to proper centromere function (Garrido-Ramos 2012; Saksouk et al., 2015; Hartley & O'Neill 2019; Thakur et al., 2021); (ii) linkage to functionally constrained centromeric regions under reduced recombination, where reduced recombination and linked selection can limit sequence turnover (Zafar et al., 2017; Giunta et al., 2021); and (iii) involvement in maintaining centromere identity through epigenetic interactions, such as CENP-A deposition or by delimiting centromeric boundaries (Plohl et al., 2012; Talbert & Henikoff, 2022; Andrade Ruiz et al., 2024).

However, against this background of broadly conserved repeat landscapes, notable deviations emerge within Cathartidae. The absence of highly conserved satellite families in *G. californianus* is particularly striking given their persistence across other Cathartidae lineages. One plausible explanation lies in the species' well-documented demographic history: a severe population bottleneck reduced its effective population size to fewer than 20 individuals, from which the current population descends (U.S. Fish and Wildlife Service, 2026). Under such extreme conditions, genetic drift is expected to override selection, accelerating both the stochastic loss and fixation of repetitive elements. This process can profoundly reshape satellitome composition over relatively short evolutionary timescales, as previously demonstrated for centromeric repeats in other bottlenecked or inbred systems (Schneider et al., 2016). In this context, the apparent erosion of shared satellite families in *G. californianus* may not reflect lineage-specific innovation but rather the outcome of drift-driven homogenization and repeat turnover. More broadly, this observation highlights the sensitivity of satellite DNA landscapes to demographic perturbations. Unlike coding regions, which are constrained by functional requirements, satellite DNAs may respond rapidly to changes in effective population size, recombination regime, and genomic architecture (Belyayev, 2014; Plohl et al., 2012). In Cathartidae, the contrast between *G. californianus* and the remaining species suggests that repeat evolution is not solely shaped by phylogenetic divergence but can be profoundly modulated by species-specific demographic trajectories. This finding reinforces the view that repeatomes are dynamic genomic compartments in which drift, selection, and molecular drive interact to produce lineage-specific outcomes (Plohl et al., 2012; Schneider et al., 2016).

6. FINAL REMARKS

Overall, our integrative approach reveals that repetitive DNA evolution can be unrelated to karyotypic stability, generating substantial genomic diversity without altering chromosome number or structure. Our findings highlight satellite DNA as a key component of genome evolution in birds and emphasize the necessity of combining sequence-based and cytogenetic approaches to resolve the functional organization of complex genomic regions. The predominant pericentromeric localization of satellite DNA families in *C. atratus* could be presumed across other species like *S. papa*, based on homology and *in silico* mapping, demonstrating that cytogenetic analysis provides consistent support for genomic predictions. Repetitive DNA regions present challenges for genome assembly, evidenced by weak dispersed signals in *Gymnogyps californianus* and *Vultur gryphus* versus clear pericentromeric clusters in *S. papa*, which is a discrepancy that reflects incomplete assembly rather than biological differences. This highlights the importance of cytogenetic validation for genomic predictions, as computational methods alone remain limited in resolving the physical distribution of repetitive sequences within chromosomes. The population-level differences in satellitome composition within *Coragyps atratus* and the apparent loss of conserved satellite families in the bottlenecked *Gymnogyps californianus* show that population structuring and genetic drift could heavily influence repeat landscapes, even if lacking phylogenetic divergence. For future studies, we highlight genome evolution research should integrate cytogenetic and genomic methodologies. Cytogenetics has been historically essential for understanding the chromosomal distribution of non-coding DNA when it was still referred to as “junk DNA”, and it continues to gain relevance through its integration with modern sequencing technologies. When we combine modern bioinformatic tools with classical cytogenetic approaches, we achieve a more accurate understanding of genome organization, reducing errors that arise from genome-assembly limitations, and take full advantages of the strengths of novel technologies.

7. REFERENCES

ANDRADE RUIZ, L.; KOPS, G. J.; SACRISTAN, C. Vertebrate centromere architecture: from chromatin threads to functional structures. **Chromosoma**, v. 133, n. 3, p. 169–181, 2024.

AVISE, J. C.; NELSON, W. S.; SIBLEY, C. G. DNA sequence support for a close phylogenetic relationship between some storks and New World vultures. **Proceedings of the National Academy of** Sciences, v. 91, n. 11, p. 5173–5177, 1994.

BACHTROG, D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. **Nature Reviews Genetics**, v. 14, n. 2, p. 113–124, 18 jan. 2013.

BELYAYEV, A. Bursts of transposable elements as an evolutionary driving force. **Journal of Evolutionary Biology**, v. 27, n. 12, p. 2573–2584, 2014.

BERNAT GEL; SERRA, E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. **Bioinformatics**, v. 33, n. 19, p. 3088–3090, 2017.

BISCOTTI, M. A.; OLMO, E.; HESLOP-HARRISON, J. S. Repetitive DNA in eukaryotes. **Chromosome Research**, v. 23, n. 3, p. 415–420, 2015.

BLACK, B. E.; CLEVELAND, D. W. Epigenetic centromere propagation and the nature of CENP-a nucleosomes. **Cell**, v. 144, n. 4, p. 471–479, 2011.

CAMACHO, J. P. M. et al. Satellitome comparison of two oedipodine grasshoppers highlights the contingent nature of satellite DNA evolution. **BMC Biology**, v. 20, p. 36, 2022.

CHARLESWORTH, B.; SNIEGOWSKI, P.; STEPHAN, W. The evolutionary dynamics of repetitive DNA in eukaryotes. **Nature**, v. 371, n. 6494, p. 215–220, set. 1994.

COURRET, C. et al. Turnover of retroelements and satellite DNA drives centromere reorganization over short evolutionary timescales in Drosophila. **PLOS Biology**, v. 22, n. 11, e3002911, 2024.

CREPALDI, C.; PARISE-MALTEMPI, P. P. Heteromorphic Sex Chromosomes and Their DNA Content in Fish: An Insight through Satellite DNA Accumulation in *Megaleporinus elongatus.* **Cytogenetic and Genome Research**, v. 160, n. 1, p. 38–46, 2020.

DEGRANDI, T. M. et al. Introducing the Bird Chromosome Database: An Overview of Cytogenetic Studies in Birds. **Cytogenetic and Genome Research**, v. 160, n. 4, p. 199–205, 2020.

DO NASCIMENTO MOREIRA, C. et al. Genomic landscape of repetitive DNAs in Neotropical electric fishes. **Molecular Genetics and Genomics**, v. 300, p. 40, 2025.

DUDKA, D. et al. Satellite DNA shapes dictate pericentromere packaging in female meiosis. **Nature**, 8 jan. 2025.

EDUARDO, L. et al A reference genome for the Harpy Eagle reveals steady demographic decline and chromosomal rearrangements in the origin of Accipitriformes. **Scientific Reports**, v. 14, 2024.

ERICSON, P. G. P. et al. A 14,000-year-old genome sheds light on the evolution and extinction of a Pleistocene vulture. **Communications Biology**, v. 5, p. 857, 2022.

FIRE, A. et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. **Nature**, v. 391, n. 6669, p. 806–811, fev. 1998.

FURO, I. et al. Revising the Chromosome-Specific Probes of White Hawk (*Leucopternis albicollis).* **The 1st International Electronic Conference on Genes: Theoretical and Applied Genomics**, p. 1, 2 nov. 2020.

GARRIDO-RAMOS, M. A.; PLOHL, M.; ŠATOVIĆ-VUKŠIĆ, E. Satellite DNA Genomics: The Ongoing Story. **International Journal of Molecular Sciences**, v. 26, n. 23, p. 11291, 22 nov. 2025.

GRIFFIN, D. K. et al. Avian cytogenomics: Small Chromosomes, Long Evolutionary History. **Genes**, v. 16, p. 1001, 2025.

HAO, Z. et al. RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. **PeerJ Computer Science**, v. 6, e251, 2020.

HARTLEY, G.; O'NEILL, R. J. Centromere Repeats: Hidden Gems of the Genome. **Genes**, v. 10, n. 3, p. 223, 2019.

HELLER, R. et al. Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage. **Proceedings of the National Academy of Sciences**, v. 93, n. 14, p. 7125–7130, 9, 1996.

HENIKOFF, S.; AHMAD, K.; MALIK, H. S. The centromere paradox: stable inheritance with rapidly evolving DNA. **Science**, v. 293, n. 5532, p. 1098–1102, 2001.

JACKSON, J. A. American black vulture. In: PALMER, R. (ed.). ***Handbook of North American birds***. New Haven: Yale University Press, 1988.

JARVIS, E. D. et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. **Science**, v. 346, n. 6215, p. 1320–1331, 11 dez. 2014

JOHSON, J. A. et al. Multi-locus phylogenetic inference among New World Vultures (Aves: Cathartidae). **Molecular Phylogenetics and Evolution**, v. 105, p. 193–199, 2016.

KAPUSTA, A.; SUH, A.; FESCHOTTE, C. Dynamics of genome size evolution in birds and mammals. **Proceedings of the National Academy of Sciences**, v. 114, n. 8, p. E1460–E1469, 2017.

KIT, S. Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. **Journal of Molecular Biology,** v. 3, n. 6, p. 711-IN2, dez. 1961.

KOSCH, T. A. et al. Comparative analysis of amphibian genomes: An emerging resource for basic and applied research. **Molecular Ecology Resources**, v. 25, 2025.

KRETSCHMER, R. et al. Satellitome Analysis in the Southern Lapwing (*Vanellus chilensis)* Genome: Implications for SatDNA Evolution in Charadriiform Birds. **Genes**, v. 15, n. 2, p. 258, 2024.

KRETSCHMER, R.; FERGUSON-SMITH, M.; DE OLIVEIRA, E. Karyotype Evolution in Birds: From Conventional Staining to Chromosome Painting. **Genes**, v. 9, n. 4, p. 181, 27 mar. 2018.

LIFTON, R. P. et al. The organization of the histone genes in *Drosophila melanogaster:* functional and evolutionary implications. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 42 Pt 2, p. 1047–1051, 1978.

LADA LISACHOVA et al. Concerted Evolution of Genus-Specific Centromeric Satellite DNA in Eremias (Lacertidae, Reptilia). Cytogenetic and Genome Research, p. 1–13, 17 mar. 2025.

LISNEY, T. J. et al. Comparison of eye morphology and retinal topography in two species of New World vultures (Aves: Cathartidae). **Anatomical Record**, v. 296, n. 12, p. 1954–1970, 2013.

LÓPEZ-FLORES, I.; GARRIDO-RAMOS, M. A. The Repetitive DNA Content of Eukaryotic Genomes. **Genome Dynamics**, p. 1–28, 2012.

LUI, R. et al. Propidium iodide for making heterochromatin more evident in the C-banding technique. **Biotechnic & Histochemistry**, v. 87, n. 7, p. 433–438, 2 jul. 2012.

MALIK, H. S.; HENIKOFF, S. Major evolutionary transitions in centromere complexity. **Cell**, v. 138, n. 6, p. 1067–1082, 2009.

MITCHELL, A. et al. The organisation of repetitive DNA sequences on human chromosomes with respect to the kinetochore analysed using a combination of oligonucleotide primers and CREST anticentromere serum. **Chromosoma**, v. 101, n. 5-6, p. 333–341, 1 mar. 1992.

MULLER, H. J. The Gene Material as the Initiator and the Organizing Basis of Life. **The American Naturalist,** v. 100, n. 915, p. 493–517, 1966.

MULLIGAN, R.; BERG, P. Expression of a bacterial gene in mammalian cells. **Science**, v. 209, n. 4463, p. 1422–1427, 19 1980.

MULLIS, K. et al. Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 51, n. 0, p. 263–273, 1 jan. 1986.

NOVÁK, P.; NEUMANN, P.; MACAS, J. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. **Nature Protocols**, v. 15, n. 11, p. 3745–3776, 23 out. 2020.

OHNO, S. So much “junk” DNA in our genome. **Brookhaven Symposia in Biology**, v. 23, p. 366–370, 1972.

PEONA, V.; WEISSENSTEINER, M. H.; SUH, A. How complete are “complete” genome assemblies?—An avian perspective. **Molecular Ecology Resources**, v. 18, n. 6, p. 1188–1195, 2018.

PEONA, V. et al. Identifying the causes and consequences of assembly gaps using a multiplatform genome assembly of a bird-of-paradise. **Molecular Ecology Resources**, v. 21, n. 1, p. 263–286, 2021.

PETY, A. M. et al. Whole chromosome painting in two populations of the weakly electric fish Gymnotus mamiraua (Teleostei, Gymnotiformes). **Scientific Reports**, v. 15, n. 1, 15 jul. 2025.

PLOHL, M.; MEŠTROVIĆ, N.; MRAVINAC, B. Satellite DNA evolution. **Genome Dynamics**, v. 7, p. 126–152, 2012.

REICHWALD, K. et al. High tandem repeat content in the genome of the short-lived annual fish Nothobranchius furzeri: a new vertebrate model for aging research. **Genome Biology**, v. 10, p. R16, 2009.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences**, v. 74, n. 12, p. 5463–5467, 1 dez. 1977.

SAKSOUK, N.; SIMBOECK, E.; DÉJARDIN, J. Constitutive heterochromatin formation and transcription in mammals. **Epigenetics & Chromatin**, v. 8, p. 3, 2015.

SCHNEIDER, K. L. et al. Inbreeding drives maize centromere evolution. **Proceedings of the National Academy of Sciences**, v. 113, n. 8, p. E987–E996, 2016.

SCHRADER, L.; SCHMITZ, J. The impact of transposable elements in adaptive evolution. **Molecular Ecology**, v. 28, n. 6, p. 1537–1549, 2019.

SHANG, W. H. et al. The Centromere: Chromatin Foundation for the Kinetochore Machinery. **Developmental Cell**, v. 24, n. 6, p. 635–648, 2013.

SOUZA, G. M. et al. Cytogenomic analysis in Seriemas (Cariamidae): Insights into an atypical avian karyotype. **Journal of Heredity**, v. 116, n. 4, p. 441–452, 11 mar. 2025

SOUZA, G. M. et al. Satellitome analysis on the pale-breasted thrush Turdus leucomelas (Passeriformes; Turdidae) uncovers the putative co-evolution of sex chromosomes and satellite DNAs. **Scientific Reports**, v. 14, 2024.

TALBERT, P. B.; HENIKOFF, S. The genetics and epigenetics of satellite centromeres. **Genome Research**, v. 32, n. 4, p. 608–615, 2022.

THAKUR, J.; PACKIARAJ, J.; HENIKOFF, S. Sequence, Chromatin and Evolution of Satellite DNA. **International Journal of Molecular Sciences**, v. 22, n. 9, p. 4309, 2021.

V, S. et al. CENP-A chromatin prevents replication stress at centromeres to avoid structural aneuploidy. **Proceedings of the National Academy of Sciences**, v. 118, n. 10, 2021.

VESELJAK, D. et al. Tribolium madens satellitome reveals a network of highly abundant satellite DNAs in megabase-sized regions hallmarked by macro-dyad symmetries. **Genome Biology**, 2026.

WARREN, W. C. et al. A New Chicken Genome Assembly Provides Insight into Avian Genome Structure. **G3: Genes, Genomes, Genetics**, v. 7, n. 1, p. 109–117, 2017.

WINK, M. Phylogeny of Old and New World Vultures (Aves: Accipitridae and Cathartidae) Inferred from Nucleotide Sequences of the Mitochondrial Cytochrome b Gene. **Zeitschrift für Naturforschung** C, v. 50, n. 11-12, p. 868–882, 1995.

WINKER, K. A Guide to the Birds of Mexico and Northern Central America S. N. G. Howell S. Webb. The Condor, v. 97, n. 4, p. 1088–1089, 1995.

ZHANG, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. **Science**, v. 346, n. 6215, p. 1311–1320, 2014.

ZWICK, M. et al. A rapid procedure for the isolation of C0t-1 DNA from plant. **Genome**, v. 40, p. 138–142, 1997.