

UNIVERSIDADE FEDERAL DO PAMPA

NAIRO FARIAS DE FARIAS

**MAPEAMENTO CROMOSSOMICO DOS RETROELEMENTOS AvIRTE E CR1 EM
AVES DO SUL DO BRASIL: ENFASE NOS CROMOSSOMOS Z E W**

**São Gabriel
2023**

NAIRO FARIAS DE FARIAS

**MAPEAMENTO CROMOSSOMICO DOS RETROELEMENTOS AvIRTE E CR1 EM
AVES DO SUL DO BRASIL: ENFASE NOS CROMOSSOMOS Z E W**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Orientador: Prof. Dr. Ricardo José Gunski

Coorientador: Prof. Dr. Fabiano Pimentel Torres

**São Gabriel
2023**

Ficha catalográfica elaborada automaticamente com os dados fornecidos
pel(a) autor(a) através do Módulo de Biblioteca do
Sistema GURI (Gestão Unificada de Recursos Institucionais).

D278m De Farias, Nairo Farias

MAPEAMENTO CROMOSSOMICO DOS RETROELEMENTOS AvIRTE E
CR1 EM AVES DO SUL DO BRASIL: ENFASE NOS CROMOSSOMOS Z
E W / Nairo Farias De Farias.

68 p.

Dissertação (Mestrado)-- Universidade Federal do
Pampa, MESTRADO EM CIÊNCIAS BIOLÓGICAS, 2023.

"Orientação: Ricardo Jose Gunski".

1. Retroelementos. 2. AvIRTE. 3. CR1. 4. Aves. 5.
Cromossomos sexuais. I. Título.

NAIRO FARIAS DE FARIAS

**MAPEAMENTO CROMOSSOMICO DOS RETROELEMENTOS AvIRTE E CR1 EM
AVES DO SUL DO BRASIL: ENFASE NOS CROMOSSOMOS Z E W**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Dissertação de Mestrado defendido em 25 de agosto de 2023.

Banca examinadora:

Prof. Dr. Ricardo José Gunski
Orientador
UNIPAMPA

Profa. Dr^a. Vera Lúcia da Silva Valente
UFRGS

Profa. Dr^a. Analía del Valle Garnero
UNIPAMPA

Dedico esta dissertação às pessoas que são as minhas raízes e meu maior exemplo de força, honestidade e coragem. Aos meus pais Florentino e Maria Lúcia.

AGRADECIMENTO

Desde o meu ingresso na graduação, muitas pessoas contribuíram para que minha trajetória acadêmica fosse possível e eu sou eternamente grato. À essas pessoas, todo o meu reconhecimento e admiração!

Ao professor Fabiano pela orientação, amizade, apoio e conselhos ao longo desta jornada. Por transmitir conhecimentos sobre a biologia molecular e ter me apresentado ao universo microcômico que são os elementos transponíveis.

Ao meu orientador e professor Ricardo, pelos preciosos ensinamentos na citogenética, por acreditar no meu trabalho e ser exemplo com a sua dedicação na pesquisa, sobretudo com as aves. Pelas indíadas de campo e também pelos excelentes churrascos que partilhamos.

A professora Analía pela mentora que és, pelo carinho e generosidade no decorrer desses anos que passaram incrivelmente rápido. Também pelo importante incentivo “de livre e espontânea pressão” para que eu conhecesse outro laboratório durante o mestrado.

Ao professor Edivaldo pela acolhida no Pará, por me receber no seu laboratório de Citogenômica do Instituto Evandro Chagas e propiciar os experimentos de FISH. Também agradeço o Fábio pela disponibilidade de me acompanhar e parceria durante esse período.

Ao professor Andrés pelos auxílios e discussões cirúrgicas nas aulas da Pós, no laboratório e corredores da universidade.

À “véia” da autoclave Dr^a Marícia pela parceria, suporte técnico e troca de ideias.

Minha família, que mesmo estando longe - fisicamente - sempre se fez presente, me incentivando e não medindo esforços para que eu pudesse estudar. Meu pai “Florentino/Farias/Netinho” e minha mãe Lucia, minha irmã Sibebe, meu afilhado Erick, minha sobrinha Clara e tia Maria, obrigado por serem meu porto seguro e transbordar meu coração de ternura, amor e bons sentimentos. Amo muito vocês!

A Instituição de ensino superior pública e de qualidade Universidade Federal do Pampa, principalmente do campus São Gabriel e seus colaboradores.

A todos os colegas do grupo de pesquisa em Diversidade Genética Animal, pela companhia diária no laboratório, seja nos experimentos, nos dias de campo, dos desafios que enfrentamos, das idas ao RU, dos momentos descontraídos e de confraternização. Meu agradecimento especial ao Marcelo, pelas vezes que trabalhamos juntos durante o mestrado, por ter me auxiliado com as técnicas de bancada da citogenética clássica e no cultivo de células.

As pessoas que me acolheram em suas casas e me concederam um lar, Dolores “Dona Lola” e Luís Augusto na cidade de Bagé; Vera e Eduardo, Nilda Helena e Dona Cenir em São Gabriel.

Aos colegas de curso que viraram amigos de vida, especialmente a Stéfany e a Melania. E a todos os amigos espalhados pelo “mundo afora”, obrigado por dividir momentos divertidos e gerar boas lembranças e histórias, vocês estarão sempre na minha memória.

Agradeço também o meu fiel e velho companheiro nas horas de estudo, o chimarrão. Nas ocasiões de solidão, a companhia musical das rádios 102.3 FM de Porto Alegre, a rádio Rock 89 de São Paulo e Goiânia, e aos domingos de manhã a Nativa FM da capital farroupilha Piratini pela trilha sonora.

Por fim, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo fomento financeiro da minha bolsa/salário/remuneração. E ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio ao grupo de pesquisa e a ciência brasileira.

“Nasça sempre com as manhãs,
deixe a luz do sol brilhar no céu do seu olhar.”

Luiz Gonzaga

RESUMO

Os elementos transponíveis (TEs) são sequências genéticas repetitivas móveis onipresente em quase todos os seres vivos, possuem a característica de se mover dentro e entre genomas. Os TEs são classificados em duas principais Classes, de acordo com o seu mecanismo de mobilização: I os retrotransposons, que utilizam um intermediário de RNA e, II os transposons via DNA. Nesse trabalho foram estudados dois retrotransposons: AviRTE e CR1, ambos pertencentes a ordem LINE de TEs, a mais abundante no genoma das Aves. Dentre estes, os mais conhecidos e abundantes são os CR1. Os AviRTE foram descobertos mais recentemente no genoma aviário, são menos abundantes e, ainda, menos caracterizados. As Aves possuem cariótipo bimodal com alta variação no número diploide, algumas espécies apresentam cromossomos sexuais incomumente grandes, possivelmente pela presença de sequências repetitivas. Deste modo, o objetivo principal foi mapear os cromossomos sexuais de espécies que possuem tamanho maior do comumente encontrado, a fim de verificar o envolvimento e a contribuição desses elementos nesta diferenciação. No primeiro capítulo, utilizou-se sonda de AviRTE no cariótipo de três espécies: *Progne chalybea* (PCH), *Chloroceryle americana* (CAM) e *Trogon surrucura* (TSU). Adicionalmente, analisamos as relações filogenéticas do elemento no genoma de aves e em outros vertebrados. Nossos resultados mostraram que os padrões de distribuição de AviRTE foram diferentes entre as espécies, não estando restritos a regiões heterocromáticas, com acúmulo no cromossomo Z de CAM e no W de PCH e TSU. A análise filogenética das sequências desse elemento nas aves acompanhou a filogenia proposta para os grupos na maioria dos clados, e permitiu a detecção em maior número de espécies, ampliando a distribuição do elemento. No segundo capítulo, utilizando como sonda o retroelemento CR1, mapeamos a distribuição cromossômica no genoma de *Chloroceryle americana* (CAM) e de *Gallinula melanops* (GME). Nossos dados revelaram diferente distribuição entre as espécies, detectamos a presença de CR1-E nos cromossomos sexuais Z de CAM e no W de GME, o que sugere o provável envolvimento dessas sequências com o aumento do tamanho desses cromossomos.

Palavras-Chave: Retroelementos; AviRTE; CR1; Aves; Cromossomos sexuais.

ABSTRACT

Transposable elements (TEs) are mobile repetitive genetic sequences ubiquitous in almost all living beings, they have the characteristic of moving within and between genomes. TEs are classified into two main classes, according to their mobilization mechanism: I retrotransposons, which use an RNA intermediate, and II transposons via DNA. In this work, two retrotransposons were studied: AviRTE and CR1, both belonging to the LINE order of TEs, the most abundant in the genome of birds. Among these, the most known and abundant are the CR1. AviRTE were discovered more recently in the avian genome, are less abundant and even less characterized. Birds have a bimodal karyotype with high variation in diploid number, some species have unusually large sex chromosomes, possibly due to the presence of repetitive sequences. Thus, the main objective was to map the sex chromosomes of species that are larger than commonly found, in order to verify the involvement and contribution of these elements in this differentiation. In the first chapter, an AviRTE probe was used in the karyotype of three species: *Progne chalybea* (PCH), *Chloroceryle americana* (CAM) and *Trogon surrucura* (TSU). Additionally, we analyzed the phylogenetic relationships of the element in the genome of birds and other vertebrates. Our results showed that AviRTE distribution patterns were different among species, not being restricted to heterochromatic regions, with accumulation on the Z chromosome of CAM and on the W chromosome of PCH and TSU. The phylogenetic analysis of the sequences of this element in birds followed the phylogeny proposed for the groups in most clades, and allowed detection in a greater number of species, expanding the distribution of the element. In the second chapter, using the CR1 retroelement as a probe, we mapped the chromosomal distribution in the genome of *Chloroceryle americana* (CAM) and *Gallinula melanops* (GME). Our data revealed different distribution between species, we detected the presence of CR1-E in the Z sex chromosomes of CAM and in the W of GME, which suggests the probable involvement of these sequences with the increase in the size of these chromosomes.

Keywords: Retroelements; AviRTE; CR1; Birds; Sex chromosomes.

LISTA DE FIGURAS

INTRODUÇÃO

- Figura 1:** Proporção de sequências repetitivas e as subclasses de TEs em vertebrados..... 16
- Figura 2:** As principais superfamílias de elementos transponíveis em aves.. 19
- Figura 3:** Filogenia em escala genômica de aves.....21

CAPITULO 1

- Figure 1:** TE amplification AviRTEfull1, AviRTEfull2, AviRTEfull3 and AviRTEfull4.. 333
- Figure 2:** FISH with AviRTE probe in metaphases..... 344
- Figure 3:** Phylogenetic relationships of avian AviRTE subfamilies.. 355
- Figure 4:** Phylogenetic relationship between AviRTE elements... 377

Supplementary

- Figure S1:** Metaphase of female *T. surrucura* after staining with Giemsa and band C. *P. chalybea* female after staining with Giemsa and band C..... 466

CAPITULO 2

- Figura 4:** Amplificação dos TEs CR1-E e CR1-J3. Fragmentos separados por eletroforese em gel de agarose 1,1% 566
- Figura 5:** Amplificação do PCR de colônias com incerto CR1-E isolado de VSP e CR1-J de PTA utilizando primer universal M13 em gel de agarose 1,1%..... 57
- Figura 6:** Cromossomos metafásicos de *Chloroceryle americana* (CAM). (A) Coloração convencional com Giemsa (B) Bandeamento C e (C) Cariótipo parcial das metáfases A e B..... 58
- Figura 7:** Cromossomos metafásicos de *Gallinula melanops* (GME). (A) Coloração convencional com Giemsa (B) Bandeamento C e (C) Cariótipo parcial das metáfases A e B. 588
- Figura 8:** FISH com sonda do elemento CR1 em (A) *Chloroceryle americana* (CAM) e (B) *Gallinula melanops* (GME).. 599
- Figura 9:** Representação da distribuição de CR1-E nos seis primeiros pares e nos cromossomos sexuais..... 59

LISTA DE TABELAS

Table 1: Detailed information of the sampled individuals.....	299
Tabela 1: Principais informações sobre as coletas.....	53
Tabela 2: Informações dos cariótipos analisados no FISH.....	67

LISTA DE ABREVIATURAS E SIGLAS

µg - Micrograma

µl - Microlitro

CAM - *Chloroceryle americana*

CEUA - Comissão de Ética no Uso de Animais

cDNA - DNA codificador

CR1 - *Chicken repeat 1*

CP - Caçapava do Sul

DAPI - 4',6-diamidino-2-fenilindol

DIRS - Sequência de repetição intermediária de *Dictyostelium*

DNA - Ácido desoxirribonucleico

EDTA - Ácido etilenodiamino tetra-acético

EN - Endonuclease

FISH - Hibridização *in situ* fluorescente

Gb - Giga base = 1.000,000,000 pb

GGA - *Gallus gallus*

GME - *Gallinula melanops*

HT - Transferência horizontal

LINE - Elementos nucleares dispersos longos

LTR - Repetições terminais longas

NaCl - Cloreto de sódio

MCA - *Melanerpes candidus*

MgCl₂ - Cloreto de magnésio

mM - Milimolar

Mya - Milhões de anos atrás

ng – Nanograma

mRNA - RNA mensageiro

ORF - Quadro de leitura aberto

pb - Pares de bases

PBS - Tampão fosfato-salino

PCH - *Progne chalybea*

PCR - Reação em cadeia da polimerase
PCY - *Pygochelidon cyanoleuca*
PLE - Elementos Penelope-like
pMol - Picomolar
PTA - *Progne tapera*
RNA - Ácido ribonucleico
RNase - Ribonuclease
RPM - Rotações por minuto
RT - Transcriptase reversa
RTE - Elementos tipo retrotransposon
SBV - Santana da Boa Vista
SG - São Gabriel
SDS - Dodecil sulfato de sódio
SINE - Elementos nucleares dispersos curtos
SISBIO - Sistema de Autorização e Informação em Biodiversidade
SSC - Solução salina concentrada
TEs - Elementos de transposição
TSU - *Trogon surrucura*
UTR - Regiões não traduzidas
UV - Ultravioleta
VSP - *Veniliornes spilogaster*

SUMÁRIO

1 INTRODUÇÃO	15
1.1 Elementos de transposição.....	15
1.2 Classificação.....	16
1.3 Retrotransposons LINEs.....	17
1.4 Elementos transponíveis e genomas.....	19
1.5 As Aves e suas características genômicas.....	20
1.6 Cromossomos sexuais.....	22
2 OBJETIVOS	244
2.1 Objetivo geral.....	24
2.2 Objetivos específicos.....	24
3 CAPITULO 1	255
3.1 Abstract.....	26
3.2 Introduction.....	27
3.3 Material and Methods.....	29
3.4 Results.....	32
3.5 Discussion.....	38
3.6 Acknowledgments	411
4 CAPITULO 2	511
4.1INTRODUÇÃO.....	51
4.2 MATERIAL E MÉTODOS	522
4.2.1 Coleta de amostras	522
4.2.2 Extrações de DNA.....	522
4.2.3 PCR, purificação e sequenciamento.....	533
4.2.4 Clonagem preparação de sonda.....	544
4.2.5 Preparação cromossômica	544
4.2.6 Citogenética clássica.....	54
4.2.7 Experimentos de FISH.....	55
4.3 RESULTADOS	566
5 CONCLUSÃO	63
6 REFERÊNCIAS	64

1 INTRODUÇÃO

1.1 Elementos de transposição

Os elementos de transposição (TEs) foram descobertos há mais de 80 anos, em meados de 1940 pela geneticista Barbara McClintock, que nos seus estudos com o milho (*Zea mays*) observou quebras no braço curto do cromossomo 9 e um mecanismo dinâmico associado a instabilidade da coloração dos grãos. McClintock havia identificado os primeiros elementos de transposição, no qual nomeou de Ac (ativador) e Ds (dissociador), conhecidos como sistema Ds/Ac. (McClintock, 1950). Até então se acreditava no conceito consolidado proposto por Thomas Hunt Morgan, em suas pesquisas com *Drosophila*, de que os genes estavam dispostos de maneira fixa e inalterável nos cromossomos, um dos motivos que levou a grande descoberta de Barbara ser aceita e reconhecida pela comunidade científica somente nos anos seguintes, culminando com o recebimento do prêmio Nobel de Fisiologia e Medicina em 1983 (Carareto *et al.*, 2015).

Os TEs são sequências genéticas repetitivas, móveis e dispersas capazes de se mobilizarem dentro dos cromossomos e entre genomas. Com os avanços científicos na biologia molecular decorrente do sequenciamento de genomas, foi possível evidenciar a fração e a porcentagem dessas sequências em diferentes grupos de organismos, observou-se uma proporção substancial de TEs no genoma total da maioria dos organismos, embora apenas 3% no genoma da levedura *Saccharomyces cerevisiae*, sua proporção varia em torno de 20% em *Drosophila*, até 45% em humanos e em torno de 80% no genoma do milho (Chénais *et al.*, 2012; Chalopin *et al.*, 2015; Sotero-Caio *et al.*, 2017).

O acúmulo destas sequências móveis aumenta a instabilidade genômica e, para minimizar seus efeitos, mecanismos de recombinação buscam um equilíbrio. Por mais que essa regulação alcance um equilíbrio bem sucedido, a exclusão de TEs dos genomas não parece alcançar o mesmo sucesso, visto que quase todos os genomas eucarióticos sequenciados os contêm, e frequentemente em grandes proporções (Hayward & Gilbert, 2022). Nesse aspecto, as Aves parecem ter sido mais bem sucedidas comparado a outros vertebrados, embora ainda contenham porções significativas de sequências repetitivas, conforme representado na *Figura 1*.

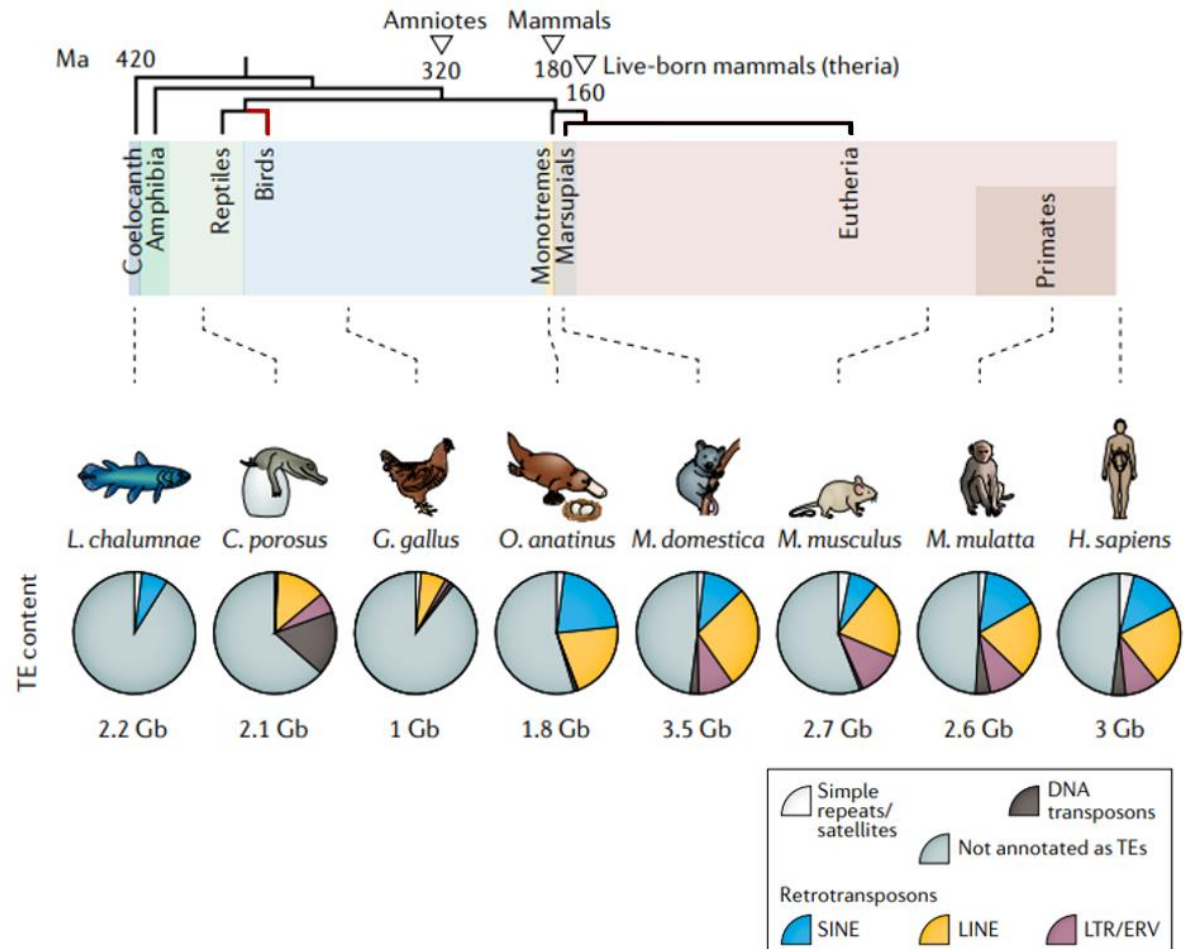


Figura 1: Proporção de sequências repetitivas e as subclasses de TEs encontradas em vertebrados. Adaptado de Senft & Macfarlan, (2021).

1.2 Classificação

Atualmente a fração do genoma composta por TEs é denominada de Mobiloma, a primeira classificação proposta para os TEs foi realizada por Finnegan (1989) que os agrupou em Classe I ou retrotransposons e Classe II ou transposons, de acordo com o mecanismo de transposição, ou seja, a molécula intermediária utilizada para a mobilização (RNA ou DNA).

Devido ao aumento de pesquisas incluindo os TEs, outras classificações foram sugeridas e coexistem. Embora naturalmente apresente algumas divergências sistemáticas, as mais relevantes que se sucederam foram as de (Curcio & Derbyshire, 2003); (Jurka *et al.*, 2005); (Wicker *et al.*, 2007); (Piégu *et al.*, 2015); (Arkhipova, 2017) e recentemente de (Storer *et al.*, 2021). Todas concordam com a primeira divisão da Classe I e II.

A proposta mais difundida e aceita na literatura é a de Wicker e colaboradores (2007), que continua dividida em duas principais Classes conforme o intermediário utilizado para a transposição. As classificações hierárquicas seguintes são: Subclasse, de acordo com a estrutura molecular do elemento e o mecanismo de excisão durante a transposição; Ordem: conforme a organização geral e enzimas dos TEs; Superfamília: pela organização da ORF, presença ou ausência de domínios proteicos; Família: de acordo com sequências de aminoácidos conservadas e Subfamília: elementos são autônomos ou não autônomos.

A Classe I utiliza um mecanismo replicativo de RNA, conhecido como cópia-e-cola, que se subdivide em subclasse I e II. A subclasse I apresenta apenas uma ordem, a LTR (repetição longa terminal), que engloba os retrotransposons propriamente ditos. E a subclasse II engloba quatro ordens sem LTRs, denominados retroposons: os LINEs (elementos nucleares interdispersos longos), SINEs (elementos nucleares interdispersos curtos), DIRS (sequência de repetição intermediária de *Dictyostelium*) e PLE (elementos Penelope-like). (Wicker *et al.*, 2007).

A classe II se transpõe diretamente via DNA, mecanismo conservativo, conhecido como recorta e cola. Divide-se em duas subclasses sendo: a subclasse I, que compreende duas ordens, a TIR (repetição terminal invertida) e a ordem Crypton, em que ocorre a excisão das duas fitas de DNA para a transposição. E a subclasse II contém as ordens Helitron e Maverick, nesses elementos apenas uma das fitas é clivada para transposição, sendo uma transposição do tipo cópia e cola replicativa. (Wicker *et al.*, 2007).

1.3 Retroposons LINEs

A ordem dos LINEs integra retroposons autônomos sem LTRs, suas principais superfamílias são: L1, R2, Jockey, CR1 e RTE. Quando estão completos eles geralmente contêm duas ORFs - ORF1 e ORF2. Enquanto a ORF1 está envolvida com a formação de uma partícula de ribonucleoproteína, a ORF2 codifica para uma transcriptase reversa (RT) e para uma endonuclease (EN) (Wicker *et al.*, 2007; Kapitonov & Jurka, 2008). Os LINEs são altamente truncados no 5' UTR e

conservados no 3' UTR, com uma sequência conservada de microssatélites no terminal 3'.

AviRTE (*Avian retrotransposon like elements*) é uma família pertencente à superfamília RTE, de genomas de metazoários, está distantemente relacionado ao elemento Bov-B (*Bovino-B*), outro elemento referência do clado RTE. O retrotransposon foi descrito nos genomas de nematoides e aves que compartilham cópias autônomas de AviRTE resultante de transferência horizontal (HT), sendo os primeiros eventos HT de um retrotransposon envolvendo espécies de aves (Suh *et al.*, 2016). O elemento AviRTE possui um tamanho total de aproximadamente 4 Kpb, sendo composto por uma única ORF de 3,2 kb, e duas regiões não traduzidas (UTRs), a 5' UTR que têm 800 pb e a porção 3' UTR com 40 pb de tamanho. Cabe ressaltar que o clado RTE tem como característica possuir apenas uma ORF.

Inicialmente foi detectado fragmentos de AviRTE em sequências de DNA associado a sítios de restrição (RAD) no genoma de *Zimmerius chrysops* (Tyrannidae). Uma datação intragenômica de TEs de Suh e colaboradores (2016), em genomas de aves e nematoides parasitas responsáveis por doenças em humanos, como filariose linfática e a loíase, encontrou duas “explosões” de HT, sendo a primeira no período Oligoceno entre 25,0 e 23,6 Mya no ancestral de *Brugia spp* e *Wuchereria bancrofti* com as espécies de aves *Ara macao* e *Melopsittacus undulatus* (Psittaciformes), *Buceros rhinoceros* (Bucerotidae), *Calypte anna* e *Oreotrochilus melanogaster* (Trochilidae) e *Tinamus guttatus* (Tinamidae) e a segunda durante o Mioceno entre 20,2 e 17,7 Mya no ancestral de *Loa loa* e as espécies *Gymnopithys rufigula* e *Manacus vitellinus* (Suboscines), *Apaloderma vittatum* (Trogonidae) e *Mesitornis unicolor* (Mesitornithidae). Outro estudo revelou uma transferência horizontal do retroposon RTE entre nematoides parasitas e o musaranho *Sorex araneus*, no qual nomearam de RTE1_Sar (Dunemann & Wasmuth, 2019).

O retroposon CR1 (*Chicken repeat 1*) é a superfamília LINE dominante de aves, foi inicialmente identificado no genoma de *Gallus gallus*. Difere em estrutura do AviRTE, pois contém duas ORFs, a ORF1 codifica uma proteína tipo *gag* com motivos de dedo de zinco e a ORF2 codifica uma transcriptase reversa e outras enzimas envolvidas na transposição do TE, possui um comprimento total de 4,5 Kpb. O

retroposon CR1 será abordado em maiores detalhes no segundo capítulo deste trabalho.

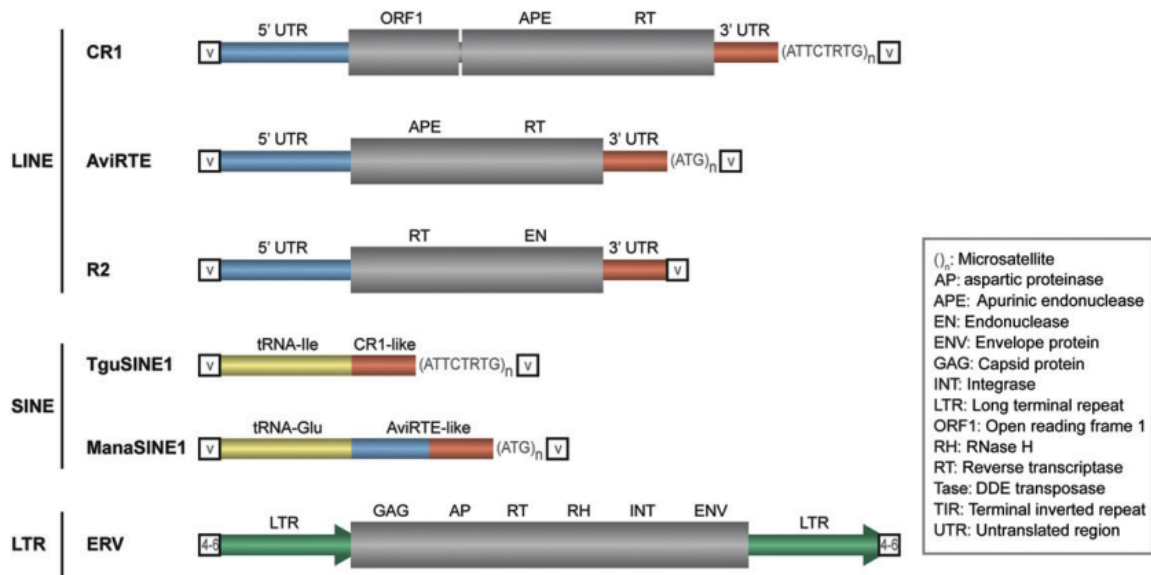


Figura 2: As principais superfamílias de retrotransposons em aves. Elementos estruturais são codificados por cores, ou seja, 5' UTRs (azul), 3' UTRs (vermelho), LTRs (verde), ORFs (cinza) e tRNAs (amarelo). Os quadrados brancos são duplicações de locais de destino, que são variável (v) ou de comprimento específico (como indicado por números e letras). As letras cinzas entre parênteses indicam motivos de microssatélites. Retirado de Kapusta & Suh, (2016).

1.4 Elementos transponíveis e genomas

Os TEs podem impactar os genomas hospedeiros de diversas formas quando se movem, dependendo do local e da região que se encontram, possibilitando que ocorra efeitos benéficos ou deletérios ao genoma hospedeiro. São capazes de ocasionar desde rearranjos cromossômicos a alterações gênicas em nível estrutural e regulatório, ou ainda ser exaptados pelo genoma hospedeiro para nova função.

Se a transposição ocorre em um gene hospedeiro essencial ou em região regulatória, a mutagênese pode ser altamente deletéria ou mesmo fatal para o hospedeiro. Os TEs são também diretamente associado a várias doenças humanas, incluindo distrofia muscular, hemofilia e formas de câncer. (Hayward & Gilbert, 2022). Dados os custos, os genomas desenvolveram estratégias para controlar a atividade

dos elementos, que incluem mecanismos epigenéticos que alteram a conformação da cromatina para reduzir a expressão e degradação direcionada de transcrições.

1.5 As Aves e suas características genômicas

Segundo a União Internacional para a Conservação da Natureza (IUCN) atualmente existem em torno de 11.000 espécies, um estudo populacional recente estima que haja entre 50 e 428 bilhões de aves no planeta (Callaghan *et al.*, 2021). A América do Sul mantém o maior número de espécies, com aproximadamente 3.600 (IUCN, 2020), sendo referida como o “continente das aves”. De acordo com a lista do Comitê Brasileiro de Registros Ornitológicos (CBRO), 1.971 espécies de aves têm ocorrência no Brasil, abrigando uma das mais ricas avifaunas do mundo (Pacheco *et al.*, 2021). O Rio Grande do Sul possui alta biodiversidade avifaunística em seus biomas Pampa e Mata Atlântica, abrigando cerca de 717 espécies, aproximadamente um terço das espécies encontradas no Brasil (Jacobs & Fenalti, 2020).

Compreendem uma Classe de animais muito diversificada dentre os vertebrados atuais, com enorme variação de hábito, tamanho, ecologia e ampla relevância em muitos campos de pesquisa (Zhang *et al.*, 2014). Além do genoma mais compactado, entre 0,89Gb a 2,11Gb, quando comparado a mamíferos, por exemplo, as aves possuem também uma menor proporção de sequências repetitivas no genoma, em torno de 9%, incluindo baixa densidade de elementos transponíveis. (Kapusta & Suh, 2016). De modo majoritário as aves voadoras apresentam genomas menores, mas curiosamente, com mais TEs em comparação com as aves que não voam (Kapusta & Suh, 2016). Dentre os amniotas, é o grupo que apresenta a maior variação no número cromossômico diploide, de $2n=40$ em espécies de Falconiformes a $2n=140$ em Musophagiformes, o que pode ser atribuído ao formato bimodal do cariótipo e a presença de numerosos microcromossomos que se originaram por fissões de macrocromossomos na trajetória evolutiva da classe. (Ellegren, 2010; Kretschmer *et al.*, 2018; Degrandi *et al.*, 2020).

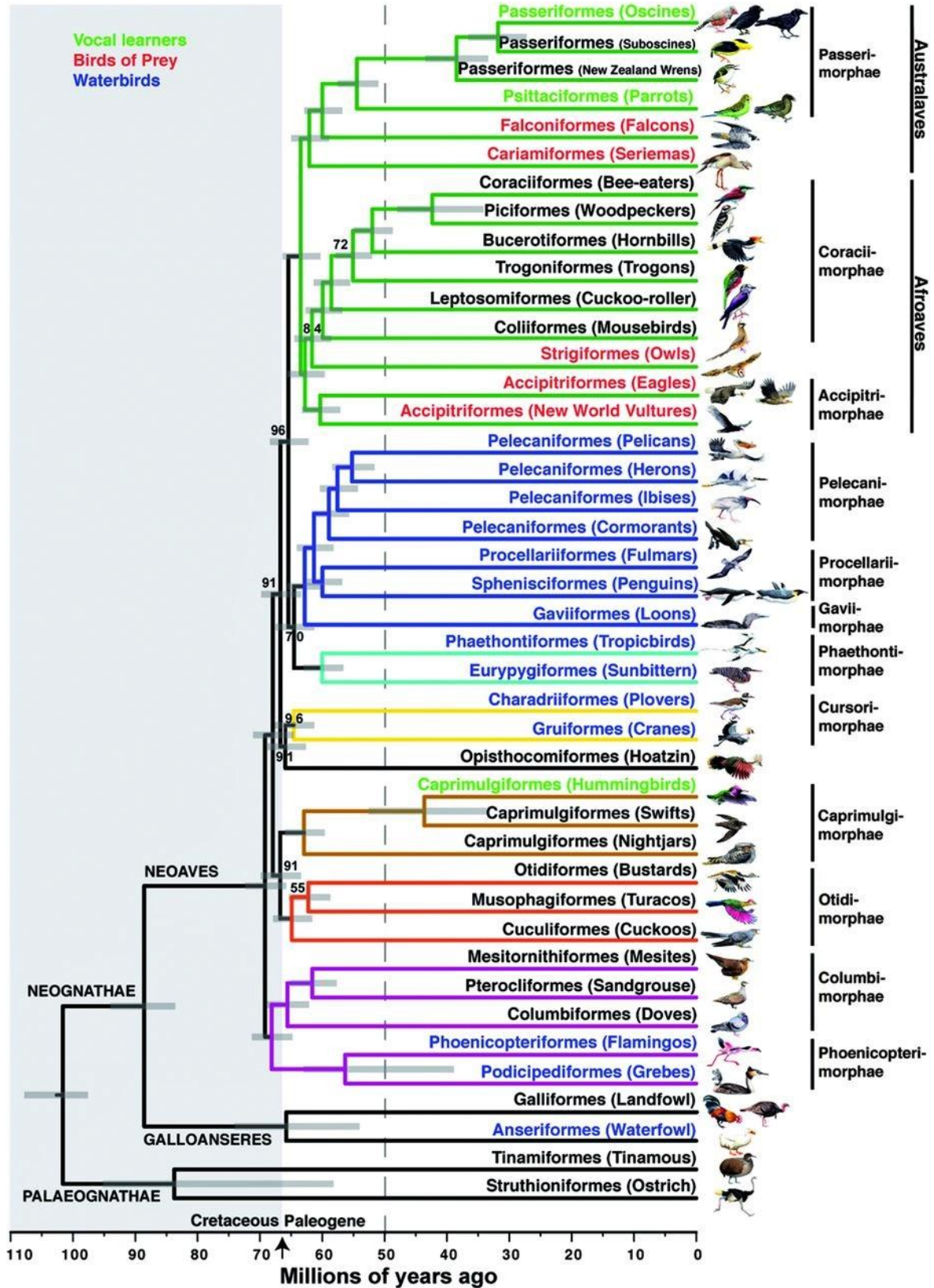


Figura 3: Filogenia em escala genômica de aves. Retirado de Jarvis *et al.*, (2014).

1.6 Cromossomos sexuais

As aves possuem o sistema de determinação sexual ZZ e ZW, em que as fêmeas representam o sexo heterogamético ZW e os machos o sexo homogamético ZZ. A maioria das espécies paleognatas têm cromossomos sexuais homomórficos com cromossomos Z e W citologicamente indistinguíveis, em contraste com os cromossomos sexuais altamente heteromórficos de neognatas e mamíferos. (Wang *et al.*, 2021). Homologamente ao cromossomo Y em mamíferos, o W das Neoaves é por via de regra rico em heterocromatina, preenchido com repetições de DNA satélite e com um número limitado de genes. Mostram diferenças significativas no tamanho entre os cromossomos Z e W, com o W geralmente sendo menor do que Z. Entretanto, existem exceções dessas características, há casos em que o cromossomo W apresenta tamanho semelhante ou maior do que o cromossomo Z.

Em estudos citogenéticos feitos em Caprimulgiformes por Nieto *et al.*, (2012) e De Souza *et al.*, (2020), evidenciou-se uma variação incomum na morfologia do cromossomo sexual W do urutau *Nyctibius griseus*, que apresentou ampliação de tamanho, muito semelhante ao cromossomo Z. A espécie de surucuá-variado *Trogon surrucura* também possui um W de tamanho aproximado ao Z, de acordo com Degrandi *et al.*, (2017). Outra observação equivalente que foge do modelo tradicional do W foi descrita por Barcellos *et al.*, (2019), em pesquisas com três espécies da família Hirundinidae (Passeriformes), na qual relatou-se um alargamento no W das andorinhas. E ainda, estudos realizados por Gunski *et al.*, (2019), na galinha d'água *Gallinula melanops*, ralídeo da ordem Gruiformes, foi descrito um cromossomo W raro e incomum, maior do que o Z. A esses fatos os autores atribuem a relação da heterocromatina e o envolvimento de DNAs repetitivos microssatélites.

Outros exemplos de diferenciação cariotípica e de cromossomos sexuais é observado em espécies que possuem o Z muito grande, maior que o primeiro par. Conforme detalhado no trabalho de Oliveira *et al.*, (2017), com três espécies de pica-paus da família Picidae, que exibem esta característica, na qual foi atribuído a presença repleta de DNAs satélites e sequências de microssatélites. E na espécie martim-pescador-pequeno *Chloroceryle americana*, da família Alcedinidae com o cromossomo sexual Z submetacêntrico sendo o maior do cariótipo, sugerindo a provável ocorrência de fusão cromossômica. (Degrandi *et al.*, 2018).

Nesta perspectiva, Bertocchi *et al.*, (2018) constaram que em quatro espécies de pica-paus (*Colaptes campestris*, *Colaptes melanocloros*, *Melanerpes candidus* e *Veniliornis spilogaster*) ocorre uma acumulação do retrotransposon CR1-E no cromossomo sexual Z, indicando que, o tamanho cromossômico e o acúmulo de sequências repetitivas estejam relacionadas entre si.

Nesse contexto, este trabalho se propôs a estudar o possível envolvimento e a contribuição de elementos transponíveis retrotransposons, na diferenciação de tamanho dos cromossomos sexuais de espécies que apresentam estas características.

2 OBJETIVOS

2.1 Objetivo geral

Contribuir para o conhecimento sobre a distribuição, impacto e evolução de elementos transponíveis nos genomas das aves.

2.2 Objetivos específicos

Investigar o envolvimento e a contribuição de AviRTE na morfologia incomum do cromossomo sexual W das espécies *Trogon surrucura* e *Progne chalybea* e no Z de *Chloroceryle americana*.

Analisar as relações filogenéticas da família AviRTE nos genomas sequenciados de aves e outros vertebrados.

Verificar o envolvimento e a contribuição de CR1 no cromossomo W de *Gallinula melanops* e Z de *Chloroceryle americana*.

3 CAPITULO 1

Manuscrito submetido à Revista Genome

Chromosome mapping of retrotransposon *AviRTE* in three Neotropical bird species

Short title: AviRTE retrotransposon in birds

Nairo Farias de Farias^{1,2} (0000-0002-3033-4963), Ricardo José Gunski^{1,2*}, (0000-0002-7315-0590), Analía Del Valle Garneró¹² (0000-0003-4252-8228), Andrés Delgado Cañedo² (0000-0002-8377-6204), Edivaldo Herculano Correa de Oliveira^{3,4} (0000-0001-6315-3352), Fábio Augusto Oliveira Silva⁵, and Fabiano Pimentel Torres¹ (0009-0009-5514-0522).

¹ Laboratório de Diversidade Genética Animal, Universidade Federal do Pampa, Campus São Gabriel, RS, Brazil.

² Programa de Pós-graduação em Ciências Biológicas, Universidade Federal do Pampa, Campus São Gabriel, RS, Brazil.

³ Laboratório de Citogenômica e Mutagênese ambiental, SEAMB, Instituto Evandro Chagas, Ananindeua, PA, Brazil.

⁴ Faculdade de Ciências Naturais, ICEN, Universidade Federal do Pará, Belém, PA, Brazil.

⁵ Programa de Pós-Graduação em Neurociências e Biologia Celular, ICB, Universidade Federal do Pará, Belém, PA, Brazil.

***Corresponding author:** Ricardo José Gunski, Federal University of Pampa, Rua Aluizio Barros Macedo, São Gabriel, RS, Brazil, Email:ricardogunski@unipampa.edu.br

3.1 ABSTRACT

Avian genomes are characterized as being more compact than other amniotes, with less diversity and density of transposable elements (TEs). In addition, birds usually show bimodal karyotypes, exhibiting a great variation in diploid numbers. Some species present unusually large sex chromosomes, possibly due to the accumulation of repetitive sequences. AviRTE is a long interspersed nuclear element (LINE) recently discovered in the genomes of birds and nematodes, and it is still poorly characterized in terms of chromosomal mapping and phylogenetic relationships. In this study, we mapped AviRTE isolated from the genome of *Trogon surrucura* in the karyotype of three species that present atypical sex chromosomes: *Progne chalybea* (PCH), *Chloroceryle americana* (CAM) and *Trogon surrucura* (TSU). Additionally, we analyzed the phylogenetic relationships of this LINE in birds and other vertebrates. Our results showed that the patterns of distribution of AviRTE were different among the species, and not restricted to heterochromatic regions, with accumulation on the Z chromosome of CAM and on the W of PCH and TSU. The phylogenetic analysis of AviRTE sequences in birds agreed with the proposed phylogeny of species in most clades, and allowed the detection of this sequence in other species, expanding the distribution of the element.

Keywords: AviRTE; Birds; Sex chromosomes; FISH.

3.2 INTRODUCTION

Long and short interspersed nuclear elements (LINEs and SINEs, respectively) are retrotransposons widely distributed in Vertebrate genomes, comprising a substantial proportion of total genomic content (Hayward and Gilbert, 2022). LINEs are also known as autonomous retrotransposons without LTRs, presenting one or two ORFs, a normally truncated 5' UTR region, and a conserved 3' UTR, with a conserved microsatellite sequence at the 3' terminal portion (Suh et al. 2014). In the case of transposable elements, their mobility can directly or indirectly influence the organization, functioning and evolution of the host genome, promoting everything from genomic rearrangements and dimensioning, to alterations in the gene expression network, and sometimes imposing various negative impacts to their hosts.

Transposable elements (TEs) have been considered to be an important genome evolution factor in eukaryotes (Richard et al. 2008). However, the accumulation of these mobile repetitive sequences increases genomic instability, an effect that is minimized by some mechanism, such as reciprocal recombination (Hayward & Gilbert, 2022). Nevertheless, as much as this regulation achieves a successful balance, the deletion of TEs from genomes does not seem to succeed in the same level, since almost all sequenced eukaryotic genomes contain large proportions of TEs. In this sense, birds seem to have been more successful in this issue.

Birds present the smallest proportions of repetitive sequences among amniotes (4,1 a 9,8%). Consequently, birds usually show highly compact genome, which correspond to the smallest DNA contents among Vertebrates - 0,89 Gb to 2,11 Gb (Kapusta and Suh, 2016). However, there are some exceptions, such as species of the Picidae family, which have a proportion of around 22% of TEs (Kapusta and Suh, 2016). This increase in the amount of satellite DNA is related to the accumulation of these sequences in sex chromosomes, increasing their size.

Hence, although the W chromosome in Neoaves is generally smaller than the Z, there are exceptions, in which the W chromosome has a similar size or is larger than the Z chromosome, such as in *Nyctibius griseus* (Caprimulgiformes), *Trogon surrucura* (Trogoniformes) some species of swallows (Passeriformes) and *Gallinula melanops* (Gruiformes) (Nieto et al. 2012; Degrandi et al. 2017; Barcellos et al. 2019; Gunski et

al. 2019; de Souza et al. 2020). In some of these studies, the use of probes corresponding to microsatellite sequences showed these uncommon large sex chromosomes were a consequence of the accumulation of repetitive sequences (Barcellos et al. 2019; de Souza et al. 2020).

Likewise, species that have a Z chromosome larger than the first autosomal pair, have also been observed, such as in woodpeckers (Piciformes) due to the accumulation of microsatellites and the CR1 retrotransposon (Oliveira et al. 2017; Bertocchi et al. 2018). CR1 is the most abundant and well-known transposable element family in birds, a LINE comprising 39-88% of all TEs.

A more recently discovered repetitive sequence, AviRTE (Avian Retrotransposon like Elements) was originally described in the genomes of birds and nematodes that share high similarity of autonomous copies, resulting from repeated horizontal transfer of TEs (HTT), being one of the first HTT events of a retrotransposon involving bird species (Suh et al. 2016). This sequence is a LINE belonging to the RTE superfamily of metazoans, distantly related to LINE Bov-B (another RTE reference of the clade). The AviRTE element it has a size of approximately 4 kbp, composed of a single ORF of approximately 3.2 kb, a 5' UTR of 800 bp and the 3' UTR portion of 40 bp in size. However, other aspects of the AviRTE element, such as its phylogenetic relationships and chromosomal mapping in birds, are still poorly known.

In this study, in order to verify if the AviRTE element has some participation in the enlargement of sex chromosomes, we characterized this sequence from *Trogon surrucura*, and mapped it physically in the chromosomes of three species with atypical large sex chromosomes: *Chloroceryle americana* (Coraciiformes) which present the large Z chromosome, and *Progne chalybea* (Passeriformes) and *T. surrucura* (Trogoniformes) both presenting a large W chromosome. Additionally, we analyzed the distribution of AviRTE and its phylogenetic relationships in birds and other vertebrates.

3.3 MATERIAL AND METHODS

3.3.1 Samples

Samplings were carried out in nature, in the municipalities of Caçapava do Sul (geographical coordinates 30°23' 00.12" S 53°21' 45.40" W), Santana da Boa Vista (30°55' 36.9" S 53°01' 55.6" W) and São Gabriel (30°20' 01.23" S 54°21' 45.40" W), located in the State of Rio Grande do Sul, Brazil, under the license number 61047-3 (SISBIO). The experiments followed the protocols approved by the Animal Ethics Committee (CEUA) of the Federal University of Pampa, number 019/2020. The species collected for this study and other information are detailed in Table 1.

Table 1: Detailed information of the sampled individuals.

Family	Species/Vulgar name	Number and sex	Location	Analyses		
				PCR	Seq	FISH
Alcedinidae	<i>Chloroceryle americana</i> (TSU) Green Kingfisher	2 ♀	SG / CP	X	-	X
Hirundinidae	<i>Progne chalybea</i> (PCH) Grey-breasted Martin	1 ♂ / 1 ♀	SG	-	-	X
	<i>Pygochelidon cyanoleuca</i> (PCY) Blue-and-white Swallow	1 ♀	SG	X	-	-
Trogonidae	<i>Trogon surrucura</i> (TSU) Surucua Trogon	2 ♀	CP	X	X	X
Picidae	<i>Melanerpes candidus</i> (MCA) White Woodpecker	1 ♀	SBV	X	-	-
Thamnophilidae	<i>Thamnophilus ruficapilus</i> (TRU) Rufous-capped Antshrike	1 ♀	SG	X	X	-
Rallidae	<i>Gallinula melanops</i> (GME) Spot-flanked Gallinule	2 ♂	SG / SBV	X	X	-

Caption: CP = Caçapava do Sul, SBV = Santana da Boa Vista and SG = São Gabriel. Seq = sequencing. Marked with X were analyzed.

3.3.2 Screening, isolation, and sequencing of AviRTE sequences

Genomic DNA was extracted from blood and/or tissue samples previously stored in 15% EDTA, by the phenol-chloroform method adapted from Sassi et al. (2005), and quantified in a NanoVue spectrophotometer (GE Healthcare Life Science). AviRTE elements were amplified by PCR in the sampled species, with four sets of primers, which isolate four overlapping amplicons: AviRTEfull-1F/R (5'-TCGTGGGAAAGAGCTTG-3' / 5'-AATACAATCGGAATGACCTGTC-3'), AviRTEfull-2F/R (5'-AGGCATCTCTCAGGAGTTGG-3' / 5'-CATAGAATCCCTTGTGGTCACC-3'), AviRTEfull-3F/R (5'-CAAGTGGTGGATCAACCTAGC-3' / 5'-TGATTTAGGGTCTTGGTGTGG-3'), and AviRTEfull-4F/R (5'-CCTATTCAATCTAAGGCGACTG-3' / 5'-ATCATCATGGCTTTGGCTTC-30), according to Suh et al. (2016). PCR parameters were an initial denaturation for 120 s at 94°C, 35 cycles of 94°C for 30 s, 52°C for 30 s and 72°C for 80 s, followed by final elongation for 300 s at 72°C. Amplicons of the species *T. surrucura*, of approximately 1000 bp, were purified from the 1% agarose gel band with the PureLink Quick Gel Extraction & PCR Kit (Invitrogen) following the manufacturer's recommendations, and sequenced at ACTGene (RS/Brazil), allowing the generation of a consensus sequence.

Using the resulting sequence product, we conducted BLAST searches using the Rebase database (Jurka et al. 2005) and GenBank (<https://www.ncbi.nlm.nih.gov/>) to confirm the identity of the element. Also, ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) tools were used to identify the TE. Therefore, the fragment from *T. surrucura* was used for phylogenetic inference, and labeled to be used as a probe in the fluorescence in situ hybridization (FISH) analyses. The AviRTE sequence from *T. surrucura* was deposited in GenBank (AviRTE_Tsu3).

3.3.3 Chromosome preparations and FISH experiments

Metaphase chromosomes were obtained from bone marrow short duration cultures (Garnero and Gunski, 2000) or fibroblast cultures (Sasaki et al. 1968), followed by colchicine (0.05%) incubation for 1 hour, and hypotonization (KCl 0.075M) for 25 minutes at 37°C. Chromosomes were washed and fixed with methanol and acetic acid 3:1. An amount of approximately 1 to 1.5 µg of the amplicon AviRTE from *T. surrucura* (full3) was labeled by nick translation (DIG-nick translation mix kit,

ROCHE), according to manufacturer's instructions, and used as probes in the FISH experiments, following Kubat et al. (2008), with modifications. After mounting the slides, metaphases were analyzed and photographed using a Zeiss AxioImager Z2 fluorescent microscope.

3.3.4 Analyses of Similarity and distance

AviRTE genomic presence and similarity searches (with AviRTE_Tsu3) were performed by BLASTn in the NCBI database WGS (whole genome shotgun contigs) (<https://www.ncbi.nlm.nih.gov/>), taxa: Birds; Actinopterygii, Sarcopterygii, Chondrichthyes, Amphibia, Reptilia Lepidosauria, Reptilia Testudines and Reptilia Crocodylidae; Mammalia. Similarities within the 80 – 80 – 80 rule (Wicker *et al.*, 2007) and e-value = 0 were considered for analysis. Adjusted according to the first sequence (AviRTE_Tsu3) (Kato and Standley, 2013). The resulting alignments were inspected manually by Aliview (Larsson, 2014), and ambiguously aligned sites were excluded from the alignment. Pairwise nucleotide distances between AviRTE consensus sequences (Table 2) were calculated in MEGA11 (Tamura, Stecher and Kumar, 2021) under the Kimura 2-parameter model⁵¹ with uniform rates among sites and pairwise deletion of gaps/missing data.

3.3.5 Phylogenetic analysis

We reconstructed the phylogenetic relationships of bird AviRTE subfamilies under maximum likelihood using the Phylogenetic reconstruction tool with RAxML (GTRGAMMA model, 1,000 bootstrap inferences) on the Galaxy platform (usegalaxy.org) from the part of the ORF2 RT from all bird AviRTE consensus generated sequences. We used the RTE-4_Ami sequence of the *Alligator mississippiensis* (crocodilians) as our outgroup (Green et al. 2014). According to the phylogeny of Suh et al. (2016), the AviRTE element sequence from the group of non-avian reptiles formed a clade outside the avian group.

Galaxy parameters: Phylogenetic reconstruction with RAxML - Maximum Likelihood based inference of large phylogenetic trees (Galaxy Version 8.2.12+galaxy0). Model type nucleotide; Substitution model GTRGAMMA; Random

seed used for the parsimony inferences 1234567890; RAxML options to use required. For AviRTE sequences in other metazoans that showed high similarity with the entire AviRTE_Tsu3 sequence, we reconstructed the phylogenetic relationships by the maximum likelihood method with FASTTREE (Galaxy Version 2.1.10+galaxy1), GTR+CATmodel. To visualize the trees, the Newick tool from Galaxy was used, and to edit them, FigTree v1.4.4.

3.4 RESULTS

3.4.1 Screening, isolation, and sequencing of AviRTE sequences

Four amplicons, each one with the expected size of approximately 1000 bp, constituting the complete AviRTE element, were amplified in all the sampled species. We have observed variation in the size of some fragments among the species, especially the AviRTEfull1 (Fig. 1). Purified amplicons Avifull2, Avifull3 and Avifull4 sequencing allowed the assembly of the entire ORF and 3' UTR terminal portion of the AviRTE element. However, due to the low quality of Avifull 1, the 5' UTR portion was not included in the final assembly. Element identification was confirmed by similarity analysis in the Repbase database, showing a similarity above 96% with AviRTE from *Apaloderma vittatum*, and African Trogonidae.

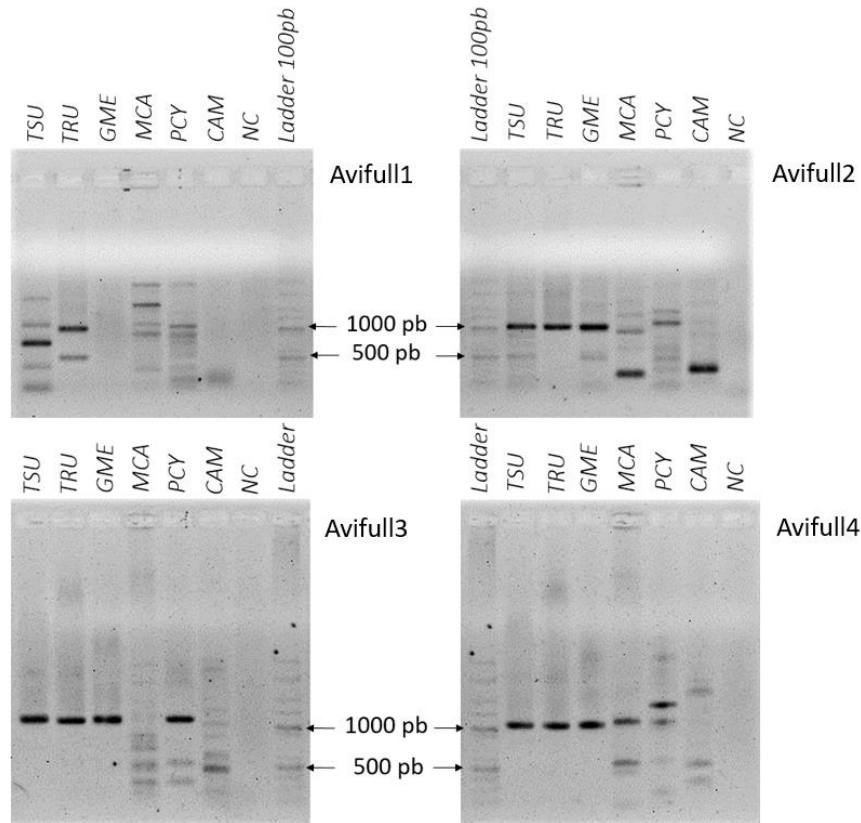


Figure 1: TE amplification AviRTEfull1, AviRTEfull2, AviRTEfull3 and AviRTEfull4. Fragments separated by 1.1% agarose gel electrophoresis. Molecular Marker 100pb DNA Ladder; *T. sarrucura* (TSU); *T. ruficapillus* (TRU); *G. melanops* (GME); *M. candidus* (MCA); *P. cyanoleuca* (PCY); *C. americana* (CAM); negative control (Water).

3.4.2 *AviRTE* Chromosome Mapping

The physical mapping of the *AviRTE* element indicated different patterns of distribution in the karyotypes of the sampled species - *P. chalybea* (PCH, Passeriformes), *C. americana* (CAM, Coraciiformes) and *T. sarrucura* (TSU, Trogoniformes) (Fig. 2). In PCH, hybridization signals were observed in almost all the elements of the karyotype. For instance, signals were observed the first metacentric pair in the terminal region of one arm and interstitial in the other; the second pair had signals in the entire small arm and interstitial region of the long arm; the third pair, in the interstitial region of the arm small and terminal of the long one, and in the sex chromosomes, in the pericentromeric region of the Z, while the W chromosome had signals in the terminal region of the short arm and interstitial of the long one. CAM showed a hybridization signal in the terminal region of the short arm of the Z chromosome and in some microchromosomes. In TSU, this sequence hybridized in

the pericentromeric region of the short arm and terminal part of the long arm of the first pair, in a few microchromosomes, in the short arm of the Z, and produced conspicuous signals in the centromeric region of the W chromosome.

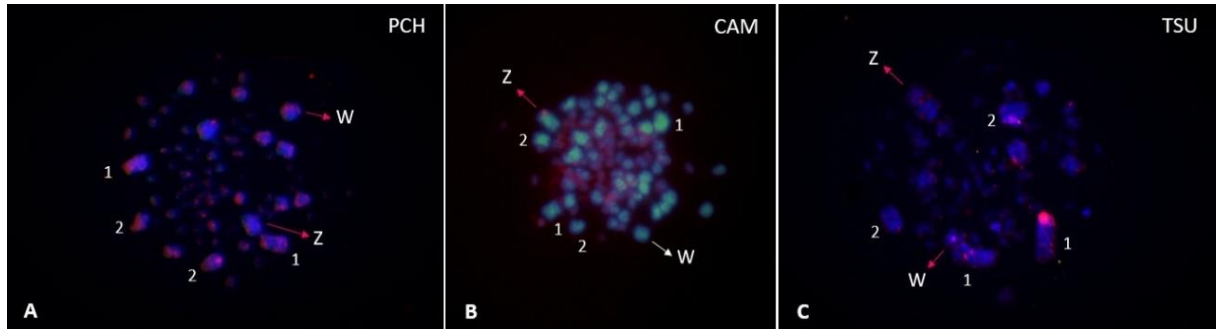


Figure 2: FISH with AviRTE probe in metaphases of (A) *Progne chalybea* (PCH), (B) *Chloroceryle americana* (CAM), (C) *Trogon surrucura* (TSU). The chromosomes were counterstained with DAPI (blue), and the probes with anti-digoxigenin-rhodamine (red). The numbers 1 and 2 indicate the first and second pair, respectively, and the arrows indicate the Z and W sex chromosomes.

3.4.3 *In silico* analyses

Rebase and NCBI analyses revealed similarity with the sequences deposited in these databases, and the ORF finder found a correspondence with the RT protein. In birds, similarity was found with more than 5000 genomic sequences, in a total of 144 species. We compiled a sample of 207 consensus sequences from bird genomes, including the full AviRTEs sequences described by Suh et al. (2016). In addition to the clades already confirmed by these authors, the presence of AviRTE was revealed in two more orders, Coraciiformes and Accipitriformes.

We reconstructed the phylogenetic relationships of bird AviRTE subfamilies under maximum likelihood using the Phylogenetic reconstruction tool with RAxML (GTRGAMMA model, 1,000 bootstrap inferences) on the Galaxy platform (usegalaxy.org) from the part of the ORF2 RT from all bird AviRTE consensus sequences generated. The resulting AviRTE tree (Fig. 3) was rooted to the reptile outgroup, *A. mississippiensis*, which showed a topology with birds according to their hosts.

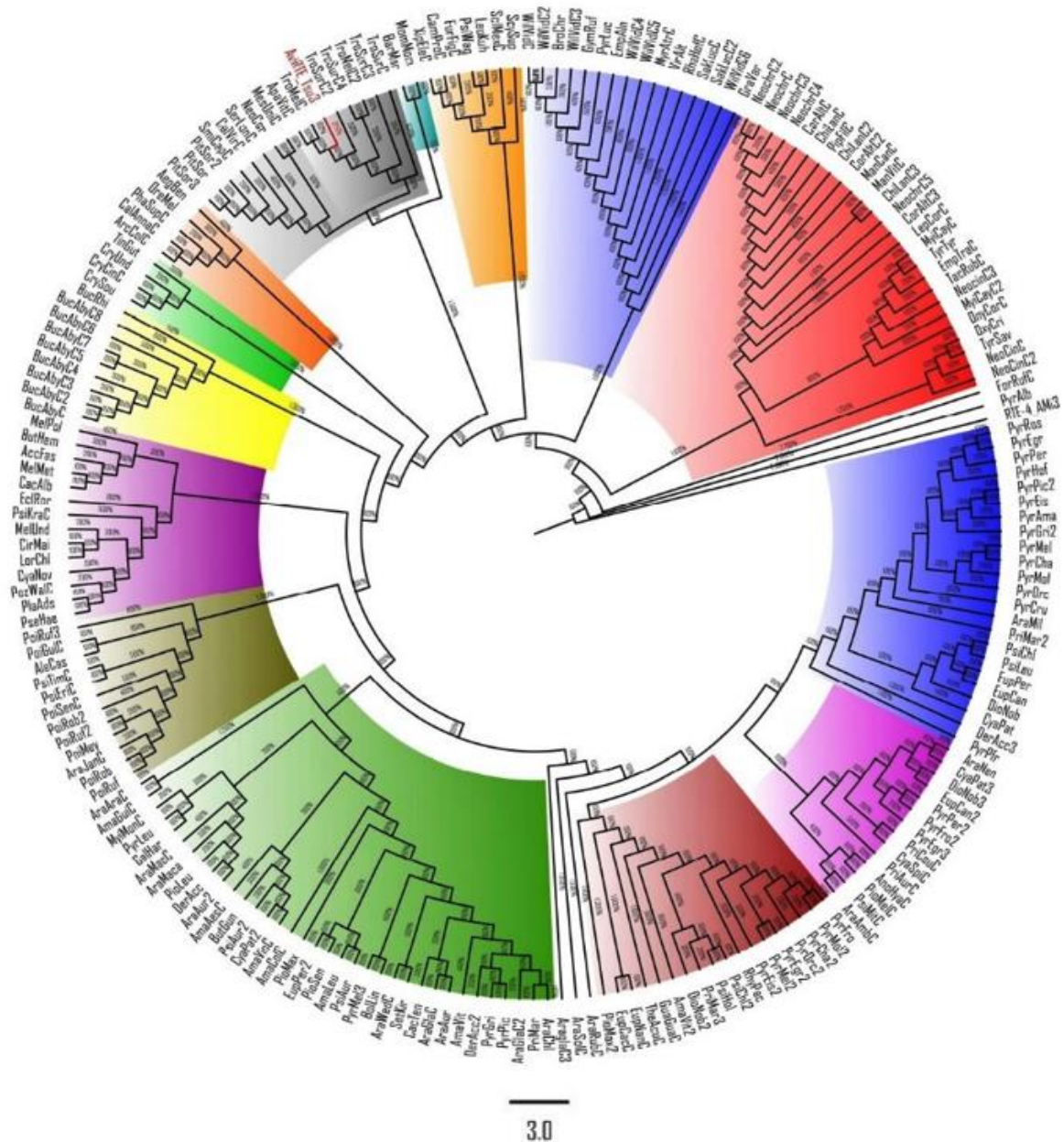


Figure 3: Phylogenetic relationships of avian AviRTE subfamilies. Evolutionary history was inferred using the Maximum Likelihood method. The percentage of replicate trees in which taxa clustered in the bootstrap test (1000 replicates) are shown next to the branches. Evolutionary distances were calculated using the p-distance method. This analysis involved 5068 sequences, which were reduced by consensus to 207 nucleotide sequences. All ambiguous positions were removed for each pair of sequences (paired deletion option). There were a total of 1.119 positions in the final dataset. The evolutionary analyzes were performed on the MEGA11. Species: *Accipiter fasciatus*= **AccFas**; *Aegotheles bennettii*= **AegBen**; *Alethe castanea*= **AleCas**; *Amazona aestiva*= **AmaAes**; *Amazona collaria*= **AmaCol**; *Amazona guildingii*= **AmaGuil**; *Amazona leucocephala*= **AmaLeu**; *Amazona vinacea*= **AmaVin**; *Amazona vittata*= **AmaVit**; *Anodorhynchus hyacinthinus*= **AnoHya**; *Apaloderma vittatum*= **ApaVit**; *Ara ambiguus*= **AraAmb**; *Ara ararauna*= **AraAra**; *Ara chloropterus*= **AraChi**; *Ara glaucogularis*= **AraGla**; *Ara macao*= **AraMaca**; *Ara militaris*= **AraMil**; *Ara rubrogenys*= **AraRub**; *Aratinga auricapillus*= **AraAur**; *Aratinga jandaya*= **AraJan**; *Aratinga maculata*=

AraMac; *Aratinga nenday*= **AraNen**; *Aratinga solstitialis*= **AraSol**; *Aratinga weddellii*= **AraWed**; *Archilochus colubris*= **ArcCol**; *Baryphthengus martii*= **BarMar**; *Bolborhynchus lineola*= **BolLin**; *Buceros bicornis*= **BucBic**; *Bucorvus abyssinicus*= **BucAby**; *Brotogeris chrysoptera*= **BroChr**; *Buteo hemilasius*= **ButHem**; *Buteogallus gundlachii*= **ButGun**; *Cacatua alba*= **CacAlb**; *Cacatua tenuirostris*= **CacTen**; *Calypte anna*= **CalAnna**; *Calyptomena viridis*= **CalVir**; *Campylorhamphus procurvodes*= **CamPro**; *Chiroxiphia lanceolata*= **ChiLan**; *Circus maillardi*= **CirMai**; *Corapipo altera*= **CorAlt**; *Cryptoleucopteryx plumbea*= **CryPlu**; *Crypturellus cinnamomeus*= **CryCin**; *Crypturellus soui*= **CrySou**; *Crypturellus undulatus*= **CryUnd**; *Cyanoliseus patagonus*= **CyaPat**; *Cyanopsitta spixii*= **CyaSpi**; *Deropterus accipitrinus*= **DerAcc**; *Diopsittaca nobilis*= **DioNob**; *Eclectus roratus*= **EclRor**; *Empidonax alnorum*= **EmpAln**; *Empidonax traillii*= **EmpTra**; *Eupsittula cactorum*= **EupCac**; *Eupsittula canicularis*= **EupCan**; *Eupsittula nana*= **EupNan**; *Eupsittula pertinax*= **EupPer**; *Formicarius rufipectus*= **ForRuf**; *Furnarius figulus*= **FurFig**; *Gallinago hardwickii*= **GalHar**; *Grallaria varia*= **GraVar**; *Guaruba guarouba*= **GuaGua**; *Harpagus bidentatus*= **HarBid**; *Lepidothrix coronata*= **LepCor**; *Leucopternis kuhli*= **LeuKuh**; *Lorius chlorocercus*= **LorChl**; *Manacus candei*= **ManCan**; *Manacus vitellinus*= **ManVit**; *Melierax metabates*= **MelMet**; *Melierax poliopterus*= **MelPol**; *Mesitornis unicolor*= **MesUni**; *Momotus momota*= **MomMom**; *Myiopsitta monachus*= **MyiMon**; *Myiozetetes cayanensis*= **MyiCay**; *Myrmophylax atrothorax*= **MyrAtr**; *Neodrepanis coruscans*= **NeoCor**; *Neopelma chrysocephalum*= **NeoChr**; *Neopipo cinnamomea*= **NeoCin**; *Onychorhynchus coronatus*= **OnyCor**; *Oxyruncus cristatus*= **OxyCri**; *Pezoporus wallicus*= **PezWal**; *Phaethornis superciliosus*= **PhaSup**; *Phoeniculus purpureus*= **PhoPur**; *Pionites leucogaster*= **PioLeu**; *Pionites melanocephalus*= **PioMel**; *Pionus maximiliani*= **PioMax**; *Pionus senilis*= **PioSen**; *Pipra filicauda*= **PipFil**; *Pitta sordida*= **PitSor**; *Platycercus adscitus*= **PlaAds**; *Poicephalus gulielmi*= **PoiGul**; *Poicephalus meyeri*= **PoiMey**; *Poicephalus robustus*= **PoiRob**; *Poicephalus rufiventris*= **PoiRuf**; *Poicephalus senegalus*= **PoiSen**; *Primolius auricollis*= **PriAur**; *Primolius couloni*= **PriCou**; *Primolius maracana*= **PriMar**; *Psephotus haematonotus*= **PseHae**; *Psilopsiagon aurifrons*= **PsiAur**; *Psittacara chloropterus*= **PsiChl**; *Psittacara holochlorus*= **PsiHol**; *Psittacara leucophthalmus*= **PsiLeu**; *Psittacara mitratus*= **PsiMit**; *Psittacara wagleri*= **PsiWag**; *Psittacula krameri*= **PsiKra**; *Psittacus erithacus*= **PsiEri**; *Psittacus timneh*= **PsiTim**; *Pyrrhura albipectus*= **PyrAlb**; *Pyrrhura amazonum*= **PyrAma**; *Pyrrhura chapmani*= **PyrCha**; *Pyrrhura cruentata*= **PyrCru**; *Pyrrhura egregia*= **PyrEgr**; *Pyrrhura eisenmanni*= **PyrEis**; *Pyrrhura frontalis*= **PyrFro**; *Pyrrhura griseipectus*= **PyrGri**; *Pyrrhura hoffmanni*= **PyrHof**; *Pyrrhura leucotis*= **PyrLeu**; *Pyrrhura lucianii*= **PyrLuc**; *Pyrrhura melanura*= **PyrMel**; *Pyrrhura molinae*= **PyrMol**; *Pyrrhura orcesi*= **PyrOrc**; *Pyrrhura perlata*= **PyrPer**; *Pyrrhura pfrimeri*= **PyrPfr**; *Pyrrhura picta*= **PyrPic**; *Pyrrhura roseifrons*= **PyrRos**; *Rhegmatorhina hoffmannsi*= **RheHof**; *Rhynchopsitta pachyrhyncha*= **RhyPac**; *Sakesphorus luctuosus*= **SakLuc**; *Sclerurus mexicanus*= **ScIMex**; *Scytalopus superciliaris*= **ScySup**; *Serilophus lunatus*= **SerLun**; *Setophaga kirtlandii*= **SetKir**; *Smithornis capensis*= **SmiCap**; *Tachuris rubrigastra*= **TacRub**; *Thectocercus acuticaudatus*= **TheAcu**; *Tinamus guttatus*= **TinGut**; *Trogon melanurus*= **TroMel**; *Trogon surrucura*= **TroSur**; *Tyrannus savana*= **TyrSav**; *Tyrannus tyrannus*= **TyrTyr**; *Vireo altiloquus*= **VirAlt**; *Willisornis vidua*= **WilVid**; *Xiphorhynchus elegans*= **XipEle**.

In addition to birds, analyses in the NCBI wgs database of other metazoan clades revealed significant similarities with two species of Osteichthyan Actinopterygii fishes (*Hypoplectrus providencianus*, 84% and *Genidens machadoi*, 84%) and four

species of Squamata reptiles (*Sceloporus occidentalis*, 81%, *Sceloporus tristichus*, 81%; *Sceloporus undulatus*, 80% and *Phrynosoma blainvilli*, 80%). Significant similarities, although below 80% (72 - 75%), with coverage above 90%, were found in Chondrichthyes fish species (*Mobula birostris*, 75% and *Hypanus sabinus*, 73%); Testudine (*Apalone spinifera*, 74%) and a Cnidaria - Anthozoa (*Ricordea florida*, 72%). No similarity was found in mammals.

A distance p was generated for the sequences that showed similarity within the rule (fish and reptiles), together with AviRTE of the nematode *Loa loa* and the representative species of birds *Bucorvus abyssinicus*, *T. sarrucura*, *Willisornis vidua*, *Tinamus guttatus* (Fig. 4- A). In addition, we also generated a phylogeny to understand the relationships between these subfamilies (Fig. 4-B). The phylogenetic relationships of the AviRTE subfamilies in other metazoans that showed high similarity with the entire AviRTE_Tsu3 sequence, resulted in a tree rooted with the crocodilian *A. mississippiensis*, as an outgroup presenting a topology according to its hosts.

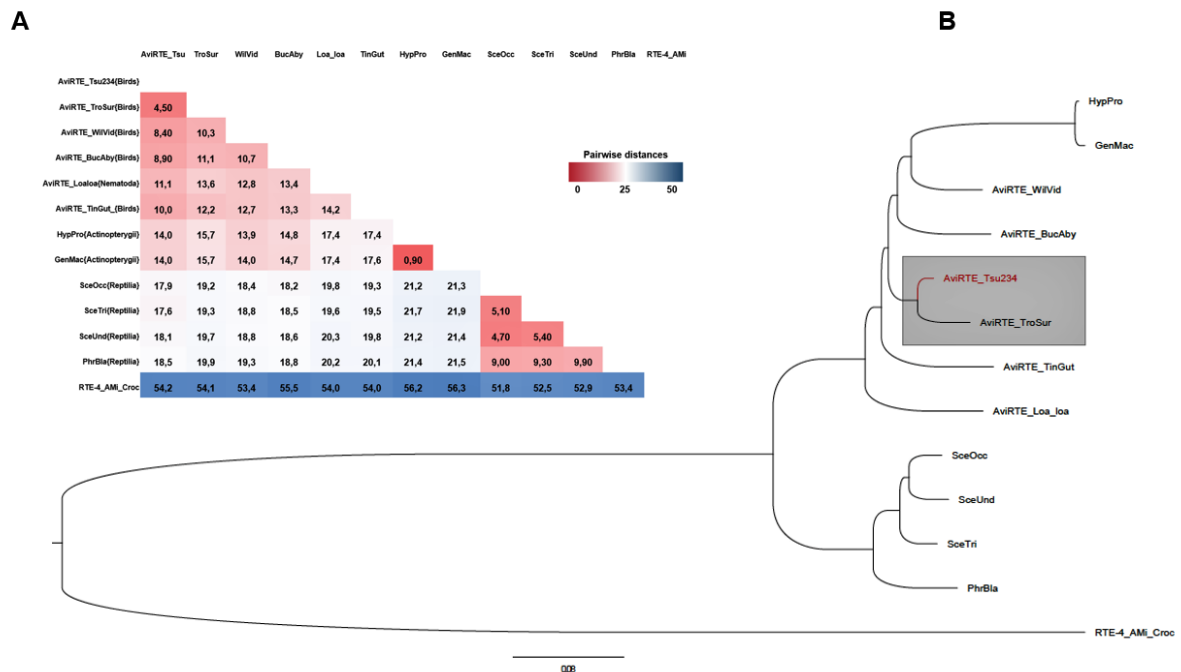


Figure 4: Phylogenetic relationship between AviRTE elements. (A) Evolutionary distances were calculated using the p-distance method and are in units of the number of base differences per site. This analysis involved 13 nucleotide sequences. All ambiguous positions were removed for each pair of sequences (paired deletion option). There were a total of 3411 positions in the final dataset. (B) The evolutionary cladogram was inferred using the Maximum Likelihood method, the taxa grouped in the bootstrap test with 1000 replicates. The evolutionary analyzes were performed on the MEGA11.

3.5 DISCUSSION

Our PCR amplifications of AviRTE (Fig. 1) in Neoaves species showed positive results in representatives of expected clades such as *T. sarrucura* (Trogoniformes) and *Thamnophilus ruficapillus* (Thamnophilidae - Suboscine) and others of special interest from this study, such as *Gallinula melanops* (Gruiformes), *Pygochelidon cyanoleuca* (Hirundinidae - Oscine), *Melanerpes candidus* (Piciformes) and *Chloroceryle americana* (Coraciiformes), expanding the distribution of this family of elements in birds, in addition to those already described by Suh et al. (2016) and the ones with available sequenced genomes, which were included herein. Our chromosomal mapping using the AviRTE transposable element probe integrates the first data including repetitive sequences for these species, since sequences from microsatellites or other TEs have not yet been chromosomally mapped by FISH in CAM and TSU species.

The FISH mapping results of PCH (Fig. 2A) revealed a wide distribution of the element in the karyotype, with conspicuous signals both in macro and microchromosomes, and not restricted to heterochromatic regions. The signals on the W chromosome confirmed AviRTE are co-located with microsatellite sequences describe by Barcellos et al. (2019) in telomeres and in the pericentromeric region, corroborating the tendency of accumulation of repetitive sequences in the W, which are probably involved in the enlargement of this chromosome in this species.

In CAM, the AviRTE hybridization signals observed in the Z chromosome indicate that its increased size is due to the accumulation of repetitive sequences (Fig. 2B), corroborating what was described in woodpeckers by Bertocchi et al. (2018), who found an accumulation of CR1 retroposon sequences in the Z chromosome of these species. Interestingly, both CAM and woodpeckers are included in the clade Coraciimorphae, which also comprises Coraciiformes, Piciformes, Bucerotiformes, Trogoniformes, Leptosomiformes and Coliiformes (Jarvis et al. 2014; Prum et al. 2015). However, in woodpeckers, in addition to the Z chromosome, a similar accumulation of TEs was also evidenced in pair 1 (Bertocchi et al. 2018), which was not observed in CAM.

In TSU, we noticed a different pattern of distribution (Fig. 2C), compared to the previous species, especially the conspicuous signal in the centromere of the W chromosome, which shows the association of the AviRTE element with heterochromatin in this chromosome, while in other macrochromosome it seems to be associated to euchromatin.

Concerning the results of the phylogenetic tree built with AviRTE sequences of birds available in the NCBI wgs database, our proposal was similar to that published by Suh et al. (2016) in the description of this element, but with a higher number of species showing the presence of the element, probably due to an increase in the number of sequenced genomes currently available. Our isolated sequence AviRTE_Tsu3, highlighted in red in the gray group (Fig. 3), grouped with other species of Trogoniformes present in this study, which points to the very high homology between the sequences of the element within the order.

Some arrangements observed in our avian cladogram are shown in Figure 3. A parallel group to the Trogons was formed by a mesito species *Mesitornis unicolor* and a passerine endemic to Madagascar, *Neodrepanis coruscans*, together with four Suboscines: *Smithornis capensis*, *Calyptomena viridis*, *Serilophus lunatus* and *Pitta sordida* (gray), possibly indicating that the insertion of the AviRTE element in the genome of these species occurred in a very close period, about 18 Mya ago during the second wave of HT, which includes Madagascar and occurred predominantly in the Neotropics (Suh et al. 2016). Eight consensuses of the Bucerotidae *Bucorvus abyssinicus* formed a single group highlighted in yellow; four species of paleognathous Tinamidae representatives were grouped together: *Tinamus guttatus*, *Crypturellus undulatus*, *Crypturellus soui* and *Crypturellus cinnamomeus* (light green); and in as sister groups, a clade formed by two species of Coraciiformes *Momotus momota* and *Baryphthengus martii* (blue) and the other composed of four species of Apodiformes, except *Aegotheles bennettii*, which demonstrated less variability of the element in these orders of birds. The Passeriformes suboscines formed three groups, the largest is composed of species of the family Tyrannidae and Pipridae, highlighted in red, the second houses the Thamnophilidae, which is in light blue and the third corresponded to a small one, constituted by the families Furnariidae and Dendrocolaptidae, highlighted in light orange. The phylogeny of the element in the Tyranni suborder

reflects on subdivisions of the Passeriformes and follows the proposed phylogeny for the species (Ohlson et al. 2013).

Other groups that comprise the lower part of the circle (Fig. 3), are composed mainly by the orders Psittaciformes and Accipitriformes that form several groups distributed in different branches of the tree. It is worthy emphasize that Psittaciformes was previously grouped close to Falconiformes (Jarvis et al. 2014) and recently close to Accipitriformes (Feng et al. 2020). We highlight the distribution of three AviRTE sequences from Psittacidae *Deropterus accipitrinus* (DerAcc, DerAcc2 and DerAcc3), as well as the genus *Pyrrhura*, which showed greater sequence variability, being arranged in different clades of the phylogeny. This can be attributed to mutations of the element in the host genome due to new insertions; in addition, in the second case, the high number of species that the genus *Pyrrhura* harbors can also be one of the factors. The analyses of these sequences allowed to expand the distribution of the AviRTE element to the orders Coraciiformes and Accipitriformes, which corroborates its irregular distribution in the avian tree, being found in orders that diverged at different times during the evolution of the group (Suh et al. 2016; Wallau et al. 2018).

Considering other groups of Vertebrates (Fig. 4-B), including species from two classes of vertebrates in which we have found significant similarity with our sequence AviRTE_Tsu234, we observed that the fish Actinopterygii *Hypoplectrus providenus* and *Genidens machadoi* had a high degree of similarity to one another, and were oddly closer to birds than the four analyzed reptilian species *Sceloporus occidentalis*, *S. tristichus*, *S. undulatus*, and *Phrynosoma blainvillei* (Fig. 4-A and B).

Due to its discontinuous distribution in other vertebrate groups, and its absence in mammals and the high similarity between distantly related species, we suggest that the AviRTE element may possibly be involved in other horizontal transfer events, as those from which it was originally discovered, between birds and nematodes. However, the veracity of this claim necessitates further substantiation through comprehensive study in order to prove the presence of HTT in these particular species.

3.6 ACKNOWLEDGMENTS

The authors are grateful to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES); National Council for Scientific and Technological Development – CNPq (process 407285/2021-0) and Natasha Ávila Bertocchi for her contributions to the phylogenetic analyses.

3.7 CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

3.8 REFERENCES

- Barcellos S, Kretschmer R, De Souza MS, *et al.* (2019) Karyotype Evolution and Distinct Evolutionary History of the W Chromosomes in Swallows (Aves, Passeriformes). *Cytogenetic and Genome Research*. 158(2):98-105. <https://doi.org/10.1159/000500621>
- Bertocchi NA, De Oliveira TD, Garnero ADV, *et al.* (2018) Distribution of CR1-like transposable element in woodpeckers (Aves Piciformes): Z sex chromosomes can act as a refuge for transposable elements. *Chromosome Res* 26, 333–343. <https://doi.org/10.1007/s10577-018-9592-1>
- Degrandi TM, Barcellos SA, Costa AL, Garnero ADV, Hass I, Gunski RJ (2020) Introducing the Bird Chromosome Database: An Overview of Cytogenetic Studies in Birds. *Cytogenetic and Genome Research*. 160(4):199-205. <https://doi.org/10.1159/000507768>
- Degrandi TM, de Oliveira JCP, de Araújo Soares A, Ledesma MA, Hass I, Garnero ADV, Gunski RJ (2018) Karyotype description and comparative analysis in Ringed Kingfisher and Green Kingfisher (Coraciiformes, Alcedinidae). *Comparative Cytogenetics* 12(2): 163-170. <https://doi.org/10.3897/compcytogen.v12i2.23883>
- Degrandi TM, Garnero ADV, O'Brien PCM, *et al.* (2017) Chromosome Painting in *Trogon s. surrucura* (Aves, Trogoniformes) Reveals a Karyotype Derived by Chromosomal Fissions, Fusions, and Inversions. *Cytogenet Genome Res* 151 (4): 208–215. <https://doi.org/10.1159/000471782>
- de Oliveira TD, Kretschmer R, Bertocchi NA, Degrandi TM, de Oliveira EHC, Cioffi MB, Garnero AV, Gunski RJ (2017) Genomic organization of repetitive DNA in woodpeckers (Aves, Piciformes): implications for karyotype and ZW sex chromosome differentiation. *PLoS One*. <https://doi.org/10.1371/journal.pone.0169987>
- de Souza MS, Kretschmer R, Barcellos AS, Costa AL, Cioffi MB, De Oliveira EHC, Garnero ADV, Gunski RJ (2020) Repeat Sequence Mapping Shows Different W Chromosome Evolutionary Pathways in Two Caprimulgiformes Families. *Birds*. <https://doi.org/10.3390/birds1010004>
- Feng S, Stiller J, Deng Y. *et al.* (2020) Dense sampling of bird diversity increases power of comparative genomics. *Nature* 587, 252–257. <https://doi.org/10.1038/s41586-020-2873-9>
- Garnero ADV, Gunski RJ (2000) Comparative analysis of the karyotypes of *Nothura maculosa* and *Rynchotus rufescens* (Aves, Tinamidae). A case of chromosomal polymorphism. *Nucl* 43:64–70.

- Green RE, Braun EL, Armstrong J, Earl D, Nguyen N, Hickey G, *et al.* (2014) Three crocodylian genomes reveal ancestral patterns of evolution among archosaurs. *Science* 346(80): Issue 6215 <https://www.science.org/doi/10.1126/science.1254449>
- Gunski RJ, Kretschmer R, Souza MS, Furo IO, *et al.* (2019) Evolution of Bird Sex Chromosomes Narrated by Repetitive Sequences: Um usual W Chromosome Enlargement in *Gallinula melanops* (Aves: Gruiformes: Rallidae). *Cytogenetic and Genome Research*. <https://doi.org/10.1159/000501381>
- Hayward A, Gilbert C (2022) Transposable elements. *Current Biology Magazine* 32, Sep 12; R904-R909. <https://doi.org/10.1016/j.cub.2022.07.044>
- Jarvis ED, Mirarab S, Aberer AJ, *et al.* (2014) Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*. 346(6215):1320-1331. <https://www.science.org/doi/10.1126/science.1253451>
- Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J (2005) Repbase update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110:462– 467. <https://doi.org/10.1159/000084979>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Kapusta A, Suh A (2017) Evolution of bird genomes-a transposon's-eye view. *Ann N Y Acad Sci* 1389:164–185. <https://doi.org/10.1111/nyas.13295>
- Kretschmer R, Ferguson-Smith MA, De Oliveira EHC (2018) Karyotype Evolution in Birds: From Conventional Staining to Chromosome Painting. *Genes*, 9(4): 181. <https://doi.org/10.3390/genes9040181>
- Kubat Z, Hobza R, Vyskot B, and Kejnovsky E (2008) Microsatellite accumulation on the Y chromosome in *Silene latifolia*. *Genome*. 51(5): 350-356. <https://doi.org/10.1139/G08-024>
- Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30(22), 3276-3278. <https://doi.org/10.1093/bioinformatics/btu531>
- Kretschmer R, Nieto L, Ledesma M, Garnero A, Gunski R (2012) Karyotype morphology suggests that the *Nyctibius griseus* (Gmelin, 1789) carries an ancestral ZW-chromosome pair to the order Caprimulgiformes (Aves). *Comparative Cytogenetics* 6(4): 379-387. <https://doi.org/10.3897/compcytogen.v6i4.3422>
- National Research Council (2011) *Guide for the Care and Use of Laboratory Animals: Eighth Edition*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/12910>

Nieto LM, Kretschmer R, Ledesma MA, Garnero ADV, & Gunski RJ (2012) Karyotype morphology suggests that the *Nyctibius griseus* (Gmelin, 1789) carries an ancestral ZW-chromosome pair to the order Caprimulgiformes (Aves). *Comparative cytogenetics*, 6(4), 379–387. <https://doi.org/10.3897/CompCytogen.v6i4.3422>

Ohlson JI, Irestedt M, Ericson PG, & Fjeldså J (2013) Phylogeny and classification of the New World suboscines (Aves, Passeriformes). *Zootaxa*, 3613, 1–35. <https://doi.org/10.11646/zootaxa.3613.1.1>

Prum RO, Berv JS, Dornburg A, *et al.* (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature*. 526 (7574):569-573. <https://doi.org/10.1038/nature15697>

Rambaut A (2020) FigTree v1.4.4. Available at: <http://tree.bio.ed.ac.uk/software/figtree>

Richard GF, Kerrest A, Dujon B (2008) Comparative Genomics and Molecular Dynamics of DNA Repeats in Eukaryotes. *Microbiology and Molecular Biology Reviews* 72: 686 - 727. doi: 10.1128/MMBR.00011-08

Sassi AK, Herédia F, Loreto EL, *et al.* (2005) Transposable elements P and gypsy in natural populations of *Drosophila willistoni*. *Genetics and Molecular Biology*. v. 28 (4). <https://doi.org/10.1590/S1415-47572005000500013>

Sasaki M, Ikeuchi T, Makino S (1968) A feather pulp culture technique for avian chromosomes, with notes on the chromosomes of the peafowl and the ostrich. *Experientia* 24: 1292–1293. <https://doi.org/10.1007/BF02146680>

Suh A, Churakov G, Ramakodi MP, Platt RN II, Jurka J, Kojima KK, *et al.* (2014) Multiple lineages of ancient CR1 retroposons shaped the early genome evolution of amniotes. *Genome Biol Evol* 7:205–217. <https://doi.org/10.1093/gbe/evu256>

Suh A, Witt C, Menger J, *et al.* (2016) Ancient horizontal transfers of retrotransposons between birds and ancestors of human pathogenic nematodes. *Nature Communications* 7, 11396. <https://doi.org/10.1038/ncomms11396>

Tamura K, Stecher G, and Kumar S (2021) MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* 38 (3022–3027). <https://doi.org/10.1093/molbev/msab120>

Wallau GL, Vieira C, & Loreto ÉLS (2018) Genetic exchange in eukaryotes through horizontal transfer: connected by the mobilome. *Mobile DNA* 9, 6. <https://doi.org/10.1186/s13100-018-0112-9>

Wang ZJ, Chen GJ, Zhang GJ, Zhou Q (2021) Dynamic evolution of transposable elements, demographic history, and gene content of paleognathous birds. *Zool Res*. 42(1), 51–61. <https://doi.org/10.24272/j.issn.2095-8137.2020.175>

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, *et al.* (2007) A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, v. 8, n. 12, p. 973-982. <https://doi.org/10.1038/nrg2165>

Appendicle A:

The chromosome number of TSU and PCH was confirmed as described in previous publications. C-banding showed constitutive heterochromatin in centromeric regions of macrochromosomes and in some microchromosomes analyzed from TSU and PCH (Fig. S1), as expected. In addition, it contributed to confirm the sex of the species, since the W chromosome is mostly heterochromatic and of special interest in this work.

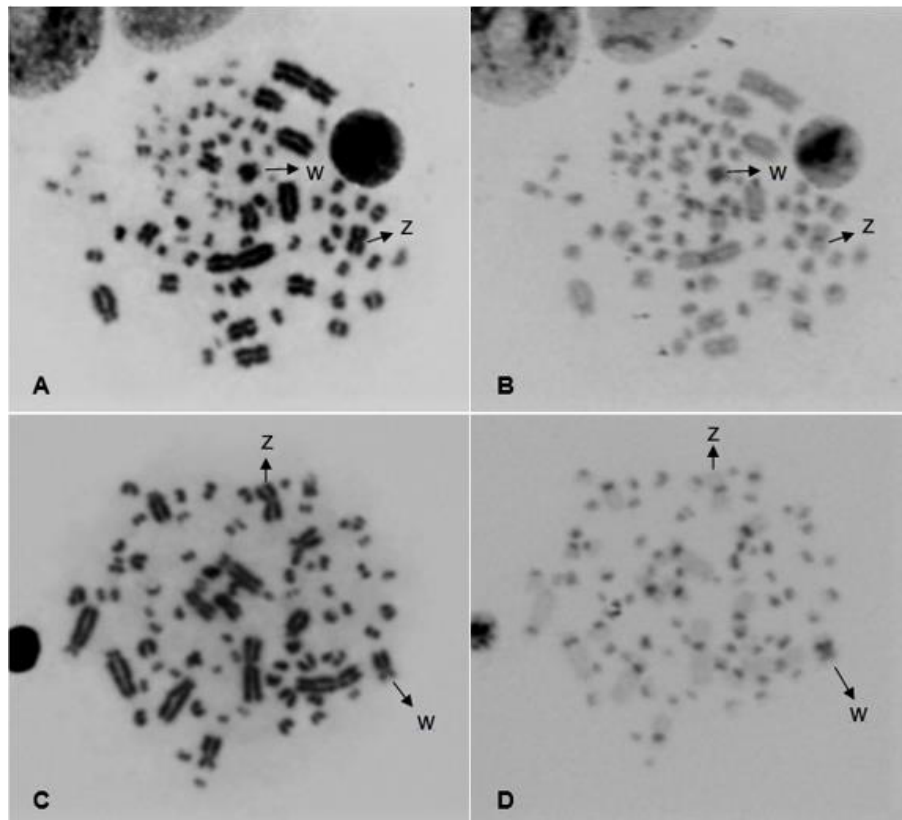


Figure S1: Metaphase of female *Trogon surrucura* after staining with Giemsa (A) and band C (B). Female *Progne chalybea* after staining with Giemsa (C) and band C (D). Arrows indicate the Z and W sex chromosomes.

B- AviRTE_Tsu234 contig sequence query in RepBase, local alignments.

Name	From	To	Name	From	To	Dir	Sim	Pos/Mm:Ts	Score
av234	1	2843	AviRTE_ApVit	1079	3920	c	0.9680	3.6667	24506

```

1 CCGATGCGGGATAGGCAAGTCCGGTTGCAAAAGGCACAGCGAAAGTCTCCTTAGGCGGTATTGGCGAGG 70
|:|
3920 CCGATGCGGGATAGGCAAGTCCGGTTGCAAAAGGCACAGTGAAAGTCTCCTTAGGCGGTATTGGCGAGG 3851

71 CACGGTCTTTCTGCGTTGTCTTTTCTCCTCGAGACTAATTCTACGTGCGTCTCAAAGGAGGCAGCAGC 140
|:|
3850 CACGGTCTTTCTGCGTTGTCTTTTCTCCTCGAGACTAATTCTACGTGCGTCTCAAAGGAGGCAGCAGC 3781

141 GTTATGGATGGTGTGTCTCCACGTCTCCCGATTGGAGGCCAGAGTGGACCACTGATGGTAGTCAATATGG 210
|:|
3780 GTTATGGATGGTGTGTCTCCACGTCTCCCGATTGGAGGCCAGAGTGGACCACTGATGGTAGTCAATATGG 3711

211 CCAAAACTGAGGTATTGTTTCAGAGAGTCTTGTATCTCCTCTTTGGGGCTCCTCTGTTGCGGCAGCCGG 280
|:|
3710 CCAAAACTGAGGTATTGTTTCAGAGAGTCTTGTATCTCCTCTTTGGGGCTCCTCTGTTGCGGCAGCCGG 3641

281 TGGCGAGTTCACCACAGAGCACAACTTAGGGAGGCAGTATCCTCCATCTCGAGACGTGCCCTGCCCA 350
|:|
3640 TGGCGAGTTCACCACAGAGCACAACTTAGGGAGGCAGTATCCTCCATCTCGAGACGTGCCCTGCCCA 3571

351 GCGCAGCTGCGTCTCAGCAGCATGGCCTCAATACTGGTGACCCCTGCCTGTTCAAGGACAGTAACATTA 420
|:|
3570 GCGCAGCTGCGTCTCAGCAGCATGGCCTCAATACTGGTGACCCCTGCCTGTTCAAGGACAGTAACATTA 3501

421 GTCACGTAGTCAGTCCAGTGGATGTTTAGGATTGTACGGAGGCAGCGCTGATGGAAGCGTTCGAGGAGTC 490
|:|
3500 GTCACGTAGTCAGTCCAGTGGATGTTTAGGATTGTACGGAGGCAGCGCTGATGGAAGCGTTCGAGGAGTC 3431

491 GCAGGTGGTGGCGGTAGATGACCCATGATTAGACCCATATAAAAGAGCAGACAGTACAATGGCTCTGTA 560
|:|
3430 GCAGGTGGTGGCGGTAGATGACCCATGATTAGACCCATATAAAAGAGTACAGTACAATGGCTCTGTA 3361

561 GACACTAATCTTTGACTTTTCTCAGGTGTTATTGCACCAGACTCTTTTATGGAGTTTCCAAAAGCT 630
|:|
3360 GACACTAATCTTTGACTTTTCTCAGGTGTTATTGCACCAGACTCTTTTATGGAGTTTCCAAAAGCT 3291

631 CTGTATGCCTTTGCTAACCTGTTGTCTATCTCTTTGTCAATCTTACCATCCGAGGAAATGATACTTCTTA 700
|:|
3290 CTGTATGCCTTTGCTAACCTGTTGTCTATCTCTTTGTCAATCTTACCATCCGAGGAAATGATACTTCTTA 3221

701 GATAGTTGAACTGCTGGACTGACTTAAGCTCTGATTTGCCTATGGTGATGTGGGGATGATGGGAGACTTC 770
|+|
3220 GRTAGTTGAACTGCTGGACTGACTTAAGCTCTGATTTGCCTATGGTGATGTGGGGATGATGGGAGACTTC 3151

771 CTGAGGTGCAGGTTGGTAAAGAACTTCTGTCTTCTCAAGCTGACTTCCAGCCAAAAGCTCAGCAGCC 840
|:|
3150 CTGAGGTGCAGGTTGGTAAAGAACTTCTGTCTTCTCAAGCTGACTTCCAGCCAAAAGCTCAGCAGCC 3081

841 TCTGCAAAGCAGGATGTTAAGCGCTGCAGAGCTGCTTCTGTGTGAGCAACGAGGGCGGCATCGTCAGCAA 910
|:|
3080 TCTGCAAAGCAGGATGTTAAGCGCTGCAGAGCTGCTTCTGTGTGAGCAACGAGGGCGGCATCGTCAGCAA 3011

911 AAAGCAGCTCGTGGACAAGGTGATTTAGGGTCTTAGTGTGGGCTTCAAGTGCCTTAGGTTGAATAGGTT 980
|:|
3010 AAAGCAGCTCGTGGACAAGGTGATTTAGGGTCTTAGTGTGGGCTTCAAGTGCCTTAGGTTGAATAGGTT 2941

981 ACCATCGGTACGATANCGGATGTAATGCCGTTATCTTCATCGAGGTCTGCCGTGGCCCTTTGGAGCATC 1050
|:|
2940 ACCATCGGTACGATANCGGATGTAATGCCGTTATCTTCATCGAGGTCTGCCGTGGCCCTTTGGAGCATC 2871

1051 ATGCTGAAGAAGATTGTGAATAGAGTTGGTGCCGAGAACGCAACCTGTTTTCACACCATTGGTTATTGAGA 1120
|:|
2870 ATGCTGAAGAAGATTGTGAATAGAGTTGGTGCCGAGAACGCAACCTGTTTTCACACCATTGGTTATTGAGA 2801

```



```

|||||
1610 TGAAGATTGTTAACTTGGAGCTCTCTCTTGGAAATGCCGCCCTCTTAGGTTTGGGCTTAAAGTGGAGGT 1541

|||||
2380 TAGGTTTGCATCACACAAGACAGTGGTCTGTTTGCATTCTGCACTAGGCATCNCTCGGGTATGACGGAC 2449
|||:|||||:|||||:|||||
1540 TAAGTTTGCATCGACAAGGCGGTGTTCTGTTTGCATTCTGCACTAGGCATCACTCGGGTATGACGGAC 1471

|||||
2450 ATCGCGAATATTGCTCTGNTNGCACTAAGACGTAGTCAATGAGGTGCCAGTGTCTGGATCGAGGATGCAT 2519
|||||
1470 ATCGCGAATATTGCTCTGG-CGCACTAAGACGTAGTCAATGAGGTGCCAGTGTCTGGATCGAGGATGCAT 1402

|||||
2520 CCAGGTTGCCTTCAAGCTGTCTTTCTGCCNNAAGANAGTGTGGTGTGGTGGAGNTGCTGCTCTGCACA 2589
|||||:|||||:|||||:|||||
1401 CCAGGTTGTCTTCAAGCTGTCTTTCTGC-TGAAAGATAGTGTGGTGTGGTGGAGCTGCTGCTCTGCACA 1333

|||||
2590 AAANTCTAGCAAGANNCCANTGTCATTGCAGTTTCCANCNCANGCTTNCNTAGTACTCCTTTCCAG 2659
|||
1332 AAANTCTAGCAAGAGGCGACCATTGTCATTGCAGTTTCCAACACCATGCTTGCCTAGTACTCCTTTCCAG 1263

|||||
2660 NCTTCAGAGTTCTTNCCTCTCTGGNGTGAAGTCNCCAAGGATTNTGATCTTATCATCTGCAGGANCT 2729
|||||
1262 GCTTCAGAGTTCTTCTACTCTGGCGTTGAAGTCACCAAGGATTATGATCTTATCATCTGCAGGAACCT 1193

|||||
2730 TTTGNGTGAGGNGCGCAGGTCTNNTGTAGAAGTGTCTTTTCCNGAGGGTCAGCTTGGAGAGTTGGGGC 2799
|||
1192 TTTGNGTGAGGCGCGCAGGTCTGGTGTAGAAGTGTCTTTTCCGAGGGTCAGCTTGGAGAGTTGGGGC 1123

2800 ATANATGCTGAAGAGAACAACATGTTGCTTGTGTGTAGTGGGA 2843
|||
1122 ATATATGCTGAAGAGAACAACATGTTGCTTGTGTGTAGTGGGA 1079

```

C- AvIRTE_Tsu3 contig sequence query in RepBase, local alignments.

Name	From	To	Name	From	To	Dir	Sim	Pos/Mm:Ts	Score
TSUAVI3	1	920	AvIRTE_ApVit	1998	2917	d	0.9837	1.5000	8073

```

1 GGCTCTGAAAGCAGTGTACGGACCCACACACCACGTTCAAAGCCAGCTGCTAAGCGCAGACGGCCAAATG 70
|||||
1998 GGCTCTGAAAGCAGTGTACGGACCCACACACCACGTTCAAAGCCAGCTGCTAAGCGCAGACGGCCAAATG 2067

71 CTTCTCACAGATAAAGCCTCCATCCTGAATCGATGGTCTGAACATTTCCAGACTCTCTTCAGTGTCAACC 140
|||||
2068 CTTCTCACAGATAAAGCCTCCATCCTGAATCGATGGTCTGAACATTTCCAGACTCTCTTCAGTGTCAACC 2137

141 GTGTAGTTCAAGGCTCAGCAATTCGGCACATTACACAACAACCGGTGAAAACGAATGGATATAGCCCC 210
|||||
2138 GTGTAGTTCAAGGCTCAGCAATTCGGCACATTACACAACAACCGGTGAAAACGAATGGATATAGCCCC 2207

211 TACTATGGGAGAGACTCTCAAGGCCATACAGCAGGTGAAAACGGCAAGGAGCTGGGATTGATGGAATC 280
|||||
2208 TACTATGGGAGAGACTCTCAAGGCCATACAGCAGGTGAAAACGGCAAGGAGCTGGGATTGATGGAATC 2277

281 CCACCTGAAGTCTGGAAGCATGGAGGCCAAGCACTCCATGCTAAATTTCCAGGACTCATTGTGCGTTGCT 350
|||||
2278 CCACCTGAAGTCTGGAAGCATGGAGGCCAAGCACTCCATGCTAAATTTCCAGGACTCATTGTGCGTTGCT 2347

351 GGGAAAAAGGCCAACTACCACCAGATCTCCGTGATGCACTCATCATCACCTGTACAAGAAAAAGGAAA 420
|||||:|||||:|||||:|||||:|||||:
2348 GGGAAAAAGGCCAACTACCACCAGATCTCCGTGATGCACTCATCATCACCTGTACAAGAAAAAGGAGA 2417

421 AAAATCCAACGTGTTCAAATTACCAGGATTAATCTTTGCTCTTTATTGCTGGTAAAATCCTTGCAAGAATA 490
|||||:|||||:|||||:|||||:|||||:
2418 AAAATCAGACTGTTCAAATTACCAGGATTAATCTTTGCTCTTTATTGCTGGTAAAATCCTTGCAAGAATA 2487

491 CTTTTGAACAGACTAGTACCAGCTATCTCAGAAGAAATTTACCTGAAAGCCAATGTGGTTTCAGAGCCA 560
|||||:|||||:|||||:|||||:|||||:
2488 CTTTTGAACAGACTAGTACCAGCTATCTCAGAAGAAATTTACCTGAAAGCCAATGTGGTTTCAGAGCCA 2557

```

```

561 ACAGGAGTACCACAGACATGGTATTTGTTTCAGACAACCTGCAAGAGAAGTGTAGGGAACAGAACAAAGG 630
   |||:|||||+|||||+|||||
2558 ACAGGAGTACCACAGACATGGTATTTGTTCTCAGACARCTGCAAGAGAAGTGYAGGGAACAGAACAAAGG 2627

631 ACTCTATGTAACTTTCGTTGACCTCACCAAAGCTTTTGACACTGTGAGCAGAAAAGGCCGTGGCAGATC 700
   |||:|||||+|||||+|||||
2628 ACTCTATGTAACTTTCGTTGACCTCACCAAAGCTTTTGACACTGTGAGCAGAAAAGGCCGTGGCAGATC 2697

701 TTGGAACGATTAGGATGTCCTCCCAAGTTCCTCAAATGATTATCCTGTTACACGAGGATCAGCGCGGCC 770
   |||:|||||+|||||+|||||
2698 TTGGAACGATTAGGATGTCCTCCCAAGTTCCTCAGAATGATTATCCTGTTACACGAGGATCAGCGCGGCC 2767

771 AAGTCAGATATGGCGATGCACTCTTTGAGCCCTTCTCAATAACCAATGGTGTGAAACAAGGTTGCGTTCT 840
   |||:|||||+|||||+|||||
2768 AAGTCAGATATGGCGATGCACTCTTGAGCCCTTCTCAATAACCAATGGTGTGAAACAAGGTTGCGTTCT 2837

841 CGCACCAACTCTATTCACAATCTTCTTCAGCATGATGCTCAAAGGGCCACGGCAGACCTCGATGAAGAT 910
   |||:|||||+|||||+|||||
2838 CGCACCAACTCTATTCACAATCTTCTTCAGCATGATGCTCAAAGGGCCACGGCAGACCTCGATGAAGAT 2907

911 AACGGCATT 920
   |||
2908 AACGGCATT 2917

```

4 CAPITULO 2

Mapeamento cromossômico do retrotransposon *CR1* em *Chloroceryle americana* (Coraciiformes) e *Gallinula melanops* (Gruiformes)

4.1 INTRODUÇÃO

Os retroelementos *Chicken repeat 1* (CR1) são a superfamília de TE mais dominante na linhagem dos amniotas. Isso se aplica aos genomas de aves, crocodilianos, tartarugas e serpentes. Enquanto genomas de mamíferos são exceções e exibem dominância de outros LINEs como L1 e L2 (Suh *et al.*, 2014). A maior parte dos TEs no genoma de *Gallus gallus* (galinha doméstica), primeira ave sequenciada e utilizada como espécie modelo em estudos aviários, são cópias degradadas de CR1, sugerindo inicialmente um genoma estável no grupo (Wicker *et al.*, 2005).

Posteriormente, o sequenciamento de 48 genomas de aves pelo *Avian Phylogenomics Project* confirmou os CR1s como o TE dominante em todas as aves não passeriformes, e uma expansão de retrovírus endógenos (ERVs) em passeriformes oscines e suboscines (Zhang *et al.*, 2014).

Atualmente, existem mais de 700 genomas de aves sequenciados disponíveis, principalmente pelo projeto Bird 10.000 Genomas (B10K), o que contribui para o aumento de dados sobre TEs em aves. Outro ponto importante a ser destacado, é a variação cariotípica que vem sendo observada, mostrando muitos rearranjos intra e intercromossômicos, que pode ser impulsionada pela presença dos TEs (Zhang *et al.*, 2014; Gunski *et al.*, 2019; Feng *et al.*, 2020; Furo *et al.*, 2021).

Embora as aves tipicamente possuam baixa densidade de TEs em seus genomas 4-10% (Kapusta & Suh, 2016), existem exceções, como os pica-paus, conhecidos por possuírem maior porcentagem de TEs em relação a outras aves (Jarvis *et al.*, 2014; Zhang *et al.*, 2014; Feng *et al.*, 2020). No entanto, um recente estudo mostrou aumento em outro grupo, em torno de 23% do genoma total de *Rhinopomastus cyanomelas* (Bucerotiformes) é constituído por CR1s (Feng *et al.*, 2020). Visto que o acúmulo dessas sequências é atribuído a elevação do tamanho de alguns macrocromossomos e cromossomos sexuais Z devido ao acúmulo de CR1 (De Oliveira *et al.*, 2017; Bertocchi *et al.*, (2018).

Com o objetivo de conhecer o padrão de distribuição cromossômica e o possível envolvimento de TEs com o tamanho cromossômico, neste estudo, analisamos cromossomicamente duas espécies previamente apresentadas no primeiro capítulo, seguindo a linha dos cromossomos sexuais incomuns em tamanho, o Z no martim-pescador-pequeno - *Chloroceryle americana* (Alcedinidae) e o W na galinha-d'água-carijó - *Gallinula melanops* (Rallidae), utilizando desta vez, o LINE CR1-E isolado do genoma de *Veniliornis spilogaster* (Picidae) como sonda nos experimentos de FISH.

4.2 MATERIAL E MÉTODOS

4.2.1 Coleta de amostras

As amostragens foram realizadas na natureza, nos municípios de Caçapava do Sul, coordenada geográfica 30°23' 00.12" S 53°21' 45.40" W, Santana da Boa Vista, coordenada 30°55' 36.9" S 53°01' 55.6" W e São Gabriel, coordenada 30°20' 01.23" S 54°21' 45.40" W, localizados no Estado do Rio Grande do Sul, Brasil, conforme licença do Sistema de Autorização e Informação em Biodiversidade (SISBIO), registro 61047-3. Os experimentos seguiram os protocolos aprovados pelo Comitê de Ética no Uso de Animais (CEUA) da Universidade Federal do Pampa, número 019/2020. As espécies coletadas para este trabalho e demais informações estão detalhadas na Tabela 1.

4.2.2 Extrações de DNA

As extrações de DNA genômico foram realizadas a partir de amostras de sangue e/ou tecido previamente armazenados em microtubos de 1,5 ml estéreis com o uso do protocolo fenol-clorofórmio adaptado de Sassi *et al.*, (2005), que utiliza um tampão de lise contendo (Tris-HCl 1M pH 8,0, SDS 10%, EDTA 0,5M, NaCl 1M e H₂O milliQ); maceração mecânica; banho-maria a 65°C; centrífuga a 13.000 rpm; fenol equilibrado (pH 8,0) e clorofórmio; álcool etílico absoluto e álcool 70%. E quantificado em espectrofotômetro NanoVue (GE Healthcare Life Sciences).

Tabela 1: Principais informações sobre as coletas.

Família	Espécie e nome popular	Número e sexo	Localização
Alcedinidae	<i>Chloroceryle americana</i> (CAM) Martim-pescador-pequeno	2 ♀	SG e CP
Hirundinidae	<i>Progne tapera</i> (PTA) Andorinha-do-campo	2 ♂	SG e CP
	<i>Pygochelidon cyanoleuca</i> (PCY) Andorinha-pequena-de-casa	1 ♀	SG
Trogonidae	<i>Trogon surrucura</i> (TSU) Surucuá-variado	1 ♂ e 2 ♀	CP
Picidae	<i>Veniliornes spilogaster</i> (VSP) Pica-pau-verde-carijó	1 ♂ e 1 ♀	SG
Phasianidae	<i>Gallus gallus</i> (GGA) Galinha-doméstica	1 ♀	SG
Ardeidae	<i>Ardea alba</i> (AAL) Garça-branca-grande	2 ♂	SBV
	<i>Ardea cocoi</i> (ACO) Garça-moura	1 ♂	SBV
Rallidae	<i>Galinulla melanops</i> (GME) Galinha-d'água-carijó	2 ♂	SG e SBV

Legenda: CP = Caçapava do Sul, SBV = Santana da Boa Vista e SG = São Gabriel.

4.2.3 PCR, purificação e sequenciamento

As reações em cadeia da polimerase (PCR) foram realizadas em volume final 15 ul para as de checagem, contendo aproximadamente 25ng de concentração de DNA, tampão 10X, MgCl₂, 10 pmol de cada primer, 10mM de cada desoxirribonucleotídeo (dNTP) e 1U de Taq DNA polimerase. O produto amplificado foi submetido à eletroforese em gel de agarose 1,1%, corado com GelRed, visualizado e fotografado em fotodocumentador com luz UV L-Pix (Loccus Biotecnologia).

As condições de PCR para CR1-E e CR1-J foram: primeira desnaturação a 95°C por 2 min, 35 ciclos, desnaturação a 95°C por 40s, anelamento a 55°C por 30s e extensão de 72°C por 1 min, seguido de uma extensão a 72°C por 5 min, de acordo com Bertocchi *et al.*, 2018.

Primers: CR1_E500S: 5'- ATGGACAGTGGCATTGAGTG -3' e CR1_E500A: 5'- GCCCTCTGATCATCTTTGTG -3'. E CR1_J3S: 5'- TTGACAGGCTGGACAGATG -3' e CR1_J3A: 5'- GTTGCCTGGGAGAAGAGAC -3'.

Para purificação todos os reagentes foram escalonados para volume final de 100 μ l. Posteriormente, o fragmento de interesse selecionado foi recortado do gel de agarose 1% com SYBR no transiluminador de luz azul Safe Imager 2.0 (Invitrogen) e purificado com o Kit PureLink Quick Gel Extraction & PCR (Invitrogen). Os elementos CR1-E e CR1-J3 foram clonados utilizando o kit TOPO TA Cloning for Sequencing (Invitrogen) seguindo recomendações do fabricante, pelo método de choque térmico em células quimiocompetentes de *Escherichia coli* da linhagem *Top10*, de acordo com Sambrook & Russell (2001), com adaptações. Os clones selecionados foram sequenciados seguindo as recomendações do serviço da ACTGene (Alvorada/RS).

4.2.4 Preparação de sonda

Para produzir a sonda foi utilizado de 1 a 1,5 μ g do isolado de DNA nos experimentos de FISH através do kit DIG-nick translation mix (ROCHE) de acordo com o fabricante.

4.2.5 Preparação cromossômica

A obtenção cromossômica das metáfases ocorreram através da técnica de cultura de medula óssea de curta duração segundo Garnero & Gunski (2000) e cultura de longa duração para o cultivo de fibroblastos de acordo com Sasaki *et al.*, (1968).

Para a coleta de ambas foi adicionado colchicina 0,05% por 1 hora, posteriormente as amostras hipotonizadas em solução de cloreto de potássio KCl, 0,075M por 25 minutos a 37°C e fixadas com metanol e ácido acético 3:1, lavadas por no mínimo três vezes.

4.2.6 Citogenética Clássica

Com o auxílio de micropipeta e ponteira P200, foi pingado cerca de 20 μ l de material fixado em cada lâmina. Posteriormente, as lâminas foram coradas com Giemsa 5% em tampão fosfato no pH 6,8 na proporção 9:1 (Tampão e Giemsa), por 4 minutos, analisadas e fotografadas.

A distribuição da heterocromatina constitutiva foi verificada seguindo Ledesma *et al.*, (2002), com modificações. Para o bandeamento C, primeiramente incubou-se a lâmina em Ácido Clorídrico (HCl) 0,2 N à 37°C por 10 minutos, após a lâmina foi mergulhada em Hidróxido de Bário (Ba(OH)₂) também a 37°C por 2min45s. Em seguida a lâmina foi colocada em HCl 0,01 N à temperatura ambiente por 30 segundos, lavada em água destilada e incubada em solução salina concentrada 2xSSC à 60°C por 1 hora. Coradas com Giemsa 5% em tampão fosfato no pH 6,8 por 40 minutos e analisadas no microscópio. As metáfases foram fotografadas com auxílio do microscópio ótico Olympus BX 53.

4.2.7 Experimentos de FISH

1° etapa: As lâminas com os cromossomos fixados foram envelhecidas em estufa na temperatura de 60°C por 30 minutos, em seguida foram incubadas a 37°C em câmara úmida com 200ul de RNase A 10mg/ul + 2xSSC por 30 minutos, após, lavadas por 3 vezes de 5 minutos em 2xSSC e incubadas com ácido clorídrico 10 mM e pepsina 0,005% a 37°C por 10 minutos. Posteriormente, repetiu-se a etapa de lavagem e incubou-se em formaldeído 1% com PBS 1x por 10 minutos, depois lavadas novamente e então desidratadas em série de etanol 50%, 70% e 100% por 2 minutos cada. As amostras nas lâminas foram desnaturadas em formamida 70% diluída com 2xSSC a 72 °C por 1 minuto e 20 segundos, e imediatamente desidratadas em série de etanol gelado 70%, 90% e 100% por 2 minutos cada. O hybmix contendo a sonda e tampão de hibridização foi desnaturado concomitantemente a 86°C por 10 minutos e transferido imediatamente para o gelo, após aplicado e vedado a hibridização foi realizada em câmara úmida a 37°C overnight.

2° etapa: As lâminas foram incubadas em 2xSSC por duas vezes de 5 minutos e uma vez em 1xSSC, lavadas em PBS 1x com mergulho rápido de 3 vezes e desidratadas em série de etanol gelado 70%, 90% e 100% por 2 minutos cada. Incubou-se a lâmina com a solução de detecção 1ul + 200ul de 4xSSTween por lâmina com parafilm em câmara úmida por 30 minutos, e então lavadas duas vezes em 4xSSTween por 5 minutos e uma em 2xSSC. Por fim foi adicionado o DAPI+ Antifade. As metáfases foram observadas e fotografadas com auxílio de microscópio fluorescente Zeiss Axio Imager Z2.

4.3 RESULTADOS

Reação em cadeia da Polimerase

Os elementos CR1-E e CR1-J (Fig. 3) foram detectados por PCR, com tamanho em torno de 540pb, em todas as espécies analisadas, como esperado devido a sua ampla distribuição nos genomas das aves.

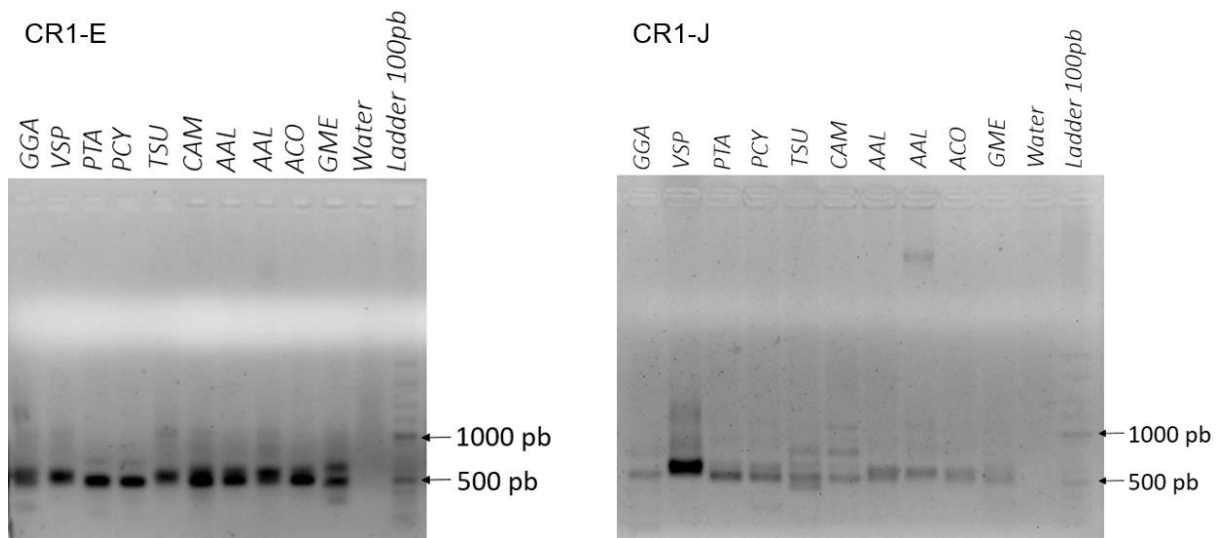


Figura 4: Amplificação dos TEs CR1-E e CR1-J3. Fragmentos separados por eletroforese em gel de agarose 1,1%. *G. gallus* (GGA); *V. spilogaster* (VSP); *P. tapera* (PTA) *P. cyanoleuca* (PCY); *T. surrucura* (TSU); *C. americana* (CAM); *A. alba* (AAL); *A. cocoi* (ACO); *G. melanops* (GME); controle negativo (Water); marcador molecular 100pb DNA Ladder.

A PCR de colônias realizada para confirmar a clonagem dos amplicons purificados foi realizada com primers universais M13, obteve-se amplificação em 5 clones confirmados, sendo um de CR1-E e quatro de CR1-J (Fig 4). Foi sequenciado um amplicon de cada elemento, possibilitando gerar duas sequências contig, confirmada no Replibase.

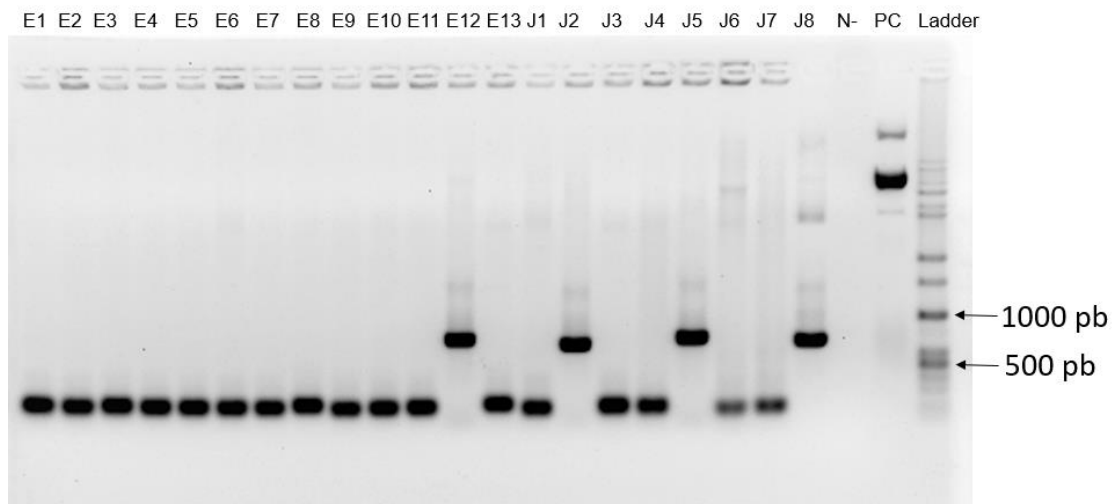


Figura 5: Amplificação do PCR de colônias com incerto CR1-E isolado de VSP e CR1-J de PTA utilizando primer universal M13 em gel de agarose 1,1%. E1 a E13 = CR1-E; J1 a J8 = CR1-J; N- = controle negativo; PC = plasmídeo clonado de CR1-J; DNA Ladder 100pb.

Citogenética clássica

O cariótipo do *Chloroceryle* (CAM) é composto de $2n=94$ cromossomos e *Gallinula* (GME) com $2n=80$, corroborando os estudos anteriores (Degrandi *et al.*, 2018; Gunski *et al.*, 2019). O bandeamento C mostrou heterocromatina constitutiva em regiões centroméricas de macrocromossomos e em alguns microcromossomos analisados de CAM e GME (Fig. 5 e 6), respectivamente. Além disso, contribuiu para confirmar o sexo das espécies, evidenciar o W grande de GME e a presença da constrição secundária no braço longo.

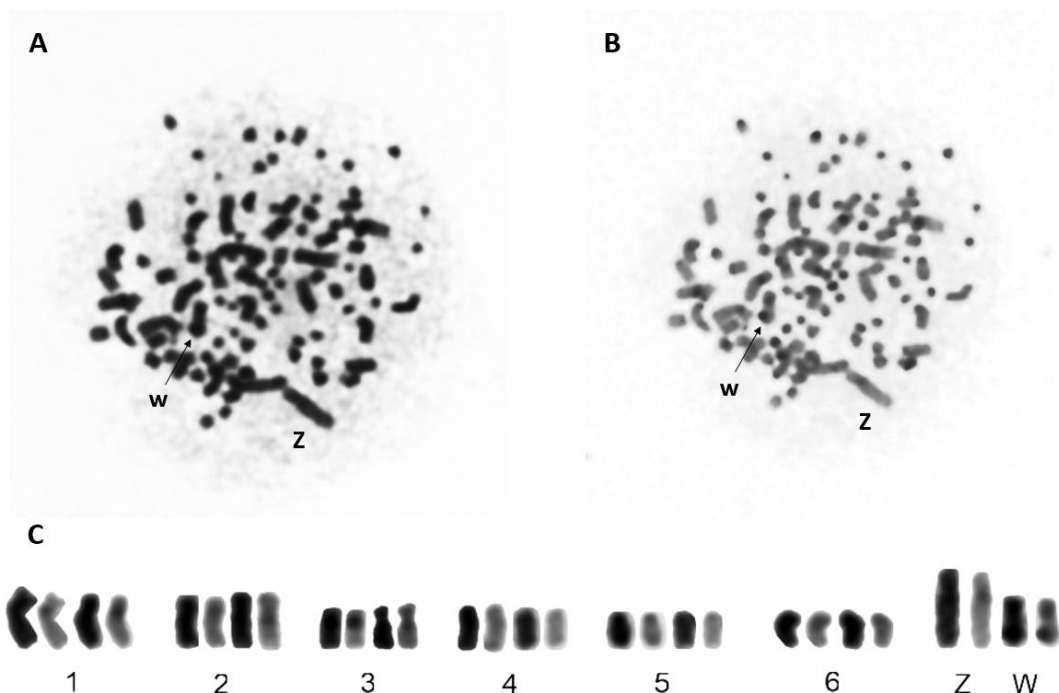


Figura 3: Cromossomos metafásicos de *Chloroceryle americana* (CAM). (A) Coloração convencional com Giemsa (B) Bandeamento C e (C) Cariótipo parcial das metáfases A e B.

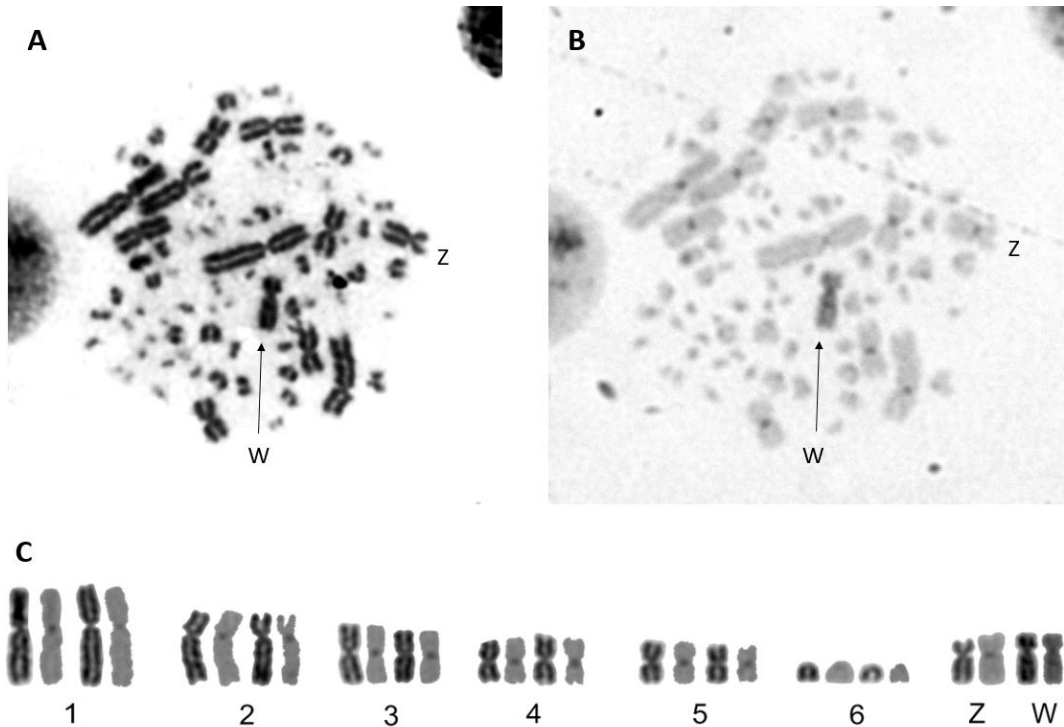


Figura 7: Cromossomos metafásicos de *Gallinula melanops* (GME). (A) Coloração convencional com Giemsa (B) Bandeamento C e (C) Cariótipo parcial das metáfases A e B.

Hibridização in situ fluorescente

Na análise de FISH, as duas espécies mostraram sinais de hibridização no mapeamento utilizando sonda do elemento CR1_E. Em CAM observamos uma ampla distribuição nos cromossomos, com marcações nas regiões terminais e intersticiais da maioria dos autossomos e no sexual Z, sem marcação no W.

Em GME na região terminal e intersticial do primeiro par, alguns macros e microcromossomos e no braço longo do cromossomo sexual W, sem marcação no Z (Fig 7). Características sobre número diploide e a morfologia dos seis primeiros pares cromossômicos e os cromossomos sexuais Z e W do cariótipo das espécies analisadas no *FISH* neste trabalho, estão descritos na Tabela 3 em Apêndice.

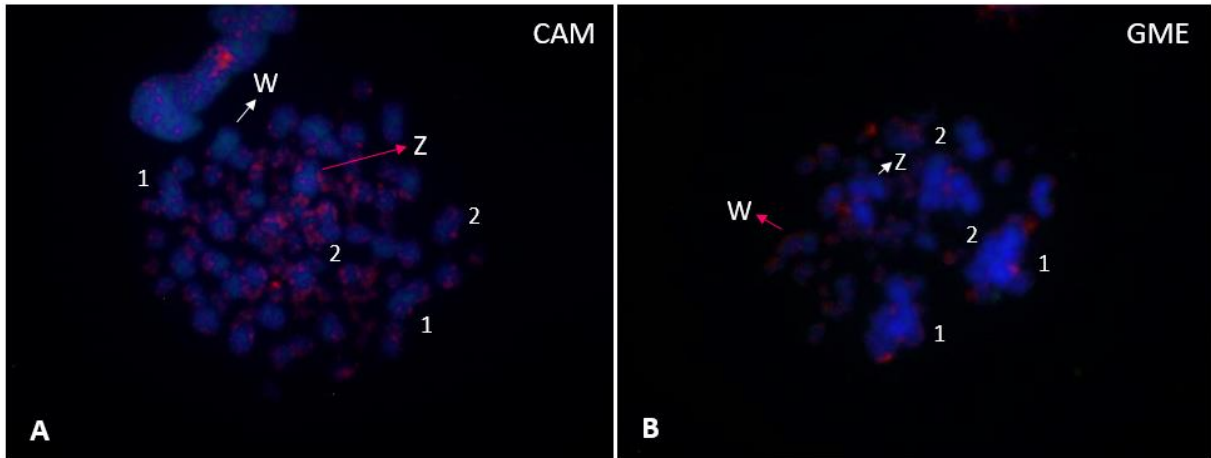


Figura 8: FISH com sonda do elemento CR1 em (A) *Chloroceryle americana* (CAM) e (B) *Gallinula melanops* (GME). Os cromossomos foram contracorados com DAPI (azul) e as sondas com anti-digoxigenina-rodamina (vermelho). Os números 1 e 2 indicam o primeiro e segundo par, respectivamente, e as setas indicam os cromossomos sexuais Z e W.

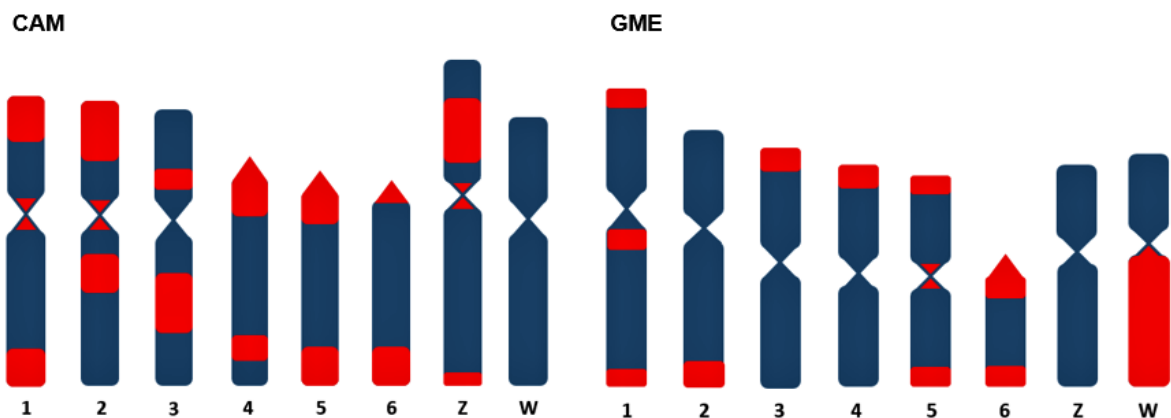


Figura 9: Representação da distribuição de CR1-E nos seis primeiros pares e nos cromossomos sexuais em *Chloroceryle americana* (CAM) e *Gallinula melanops* (GME). As regiões em vermelho correspondem ao local de hibridização.

4.4 DISCUSSÃO

As amplificações de CR1 por PCR (Fig. 3) em espécies de diferentes clados da árvore aviária confirma a ampla distribuição das duas subfamílias CR1-E e CR1-J, presente em Galloanserae e Neoaves, desde *Gallus gallus* (Galliformes), *Gallinula melanops* (Gruiformes), *Ardea alba* e *A. cocoi* (Pelecaniformes), *Trogon surrucura* (Trogoniformes), *Veniliornes spilogaster* (Piciformes), *Chloroceryle americana*

(Coraciiformes) até *Progne tapera* e *Pygochelidon cyanoleuca* (Passeriformes - Oscines).

No mapeamento cromossômico com CR1-E_Vsp, verificamos sinais de hibridização em CAM e GME. Os resultados do FISH (Fig. 7) revelaram ampla distribuição do elemento no cariótipo de CAM, muito maior comparada à encontrada com AviRTE_Tsu3, o que concorda com a densidade de cada elemento no genoma (Kapusta & Suh, 2016; Feng *et al.*, 2020). Além disso, foi observado também um padrão de marcação e distribuição próximo ao encontrado por Bertocchi *et al.*, (2018) em pica-paus, nos macrocromossomos e no marcante cromossomo sexual Z dessas espécies, porém com alguns microcromossomos também apresentando sinais de hibridização com este elemento.

Na espécie GME, as marcações no braço longo do cromossomo W mostraram estar localizada com a heterocromatina, próximo às sequências de microssatélites e de rDNA 18S encontradas por Gunski *et al.*, (2019) e Furo *et al.*, (2021). Comparado com CAM mostrou menos sinais de hibridização e menos densidade, o que pode indicar menor quantidade de cópias do TE neste genoma.

Nossos dados mostram que o CR1 está presente nos cromossomos sexuais e que provavelmente estas sequências possam contribuir para o tamanho cromossômico, independente da cromatina.

4.5 REFERÊNCIAS

Bertocchi NA, De Oliveira TD, Garnero ADV, et al. (2018) Distribution of CR1-like transposable element in woodpeckers (Aves Piciformes): Z sex chromosomes can act as a refuge for transposable elements. *Chromosome Res* 26, 333–343.

Degrandi TM, de Oliveira JCP, de Araújo Soares A, Ledesma MA, Hass I, Garnero ADV, Gunski RJ (2018) Karyotype description and comparative analysis in Ringed Kingfisher and Green Kingfisher (Coraciiformes, Alcedinidae). *Comparative Cytogenetics* 12(2): 163-170.

de Oliveira TD, Kretschmer R, Bertocchi NA, Degrandi TM, de Oliveira EHC, Cioffi MB, Garnero AV, Gunski RJ (2017) Genomic organization of repetitive DNA in woodpeckers (Aves, Piciformes): implications for karyotype and ZW sex chromosome differentiation. *PLoS One* 12:e0169987.

Feng S, Stiller J, Deng Y. et al. (2020) Dense sampling of bird diversity increases power of comparative genomics. *Nature* 587, 252–257.

Furo, I. O., Kretschmer, R., O'Brien, P. C. M., Pereira, J. C. D. C., Gunski, R. J., Garnero, A. D. V., O'Connor, R. E., Griffin, D. K., Ferguson-Smith, M. A., & Oliveira, E. H. C. (2021). Cytotaxonomy of *Gallinula melanops* (Gruiformes, Rallidae): Karyotype evolution and phylogenetic inference. *Genetics and molecular biology*, 44(2).

Garnero ADV, Gunski RJ (2000) Comparative analysis of the karyotypes of *Nothura maculosa* and *Rynchotus rufescens* (Aves, Tinamidae). A case of chromosomal polymorphism. *Nucl* 43:64–70.

Gunski, RJ, Kretschmer, R, Souza M.S., Furo I.O., et al. (2019) Evolution of Bird Sex Chromosomes Narrated by Repetitive Sequences: Um usual W Chromosome Enlargement in *Gallinula melanops* (Aves: Gruiformes: Rallidae). *Cytogenetic and Genome Research*.

Jarvis ED, Mirarab S, Aberer AJ, et al. (2014) Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*. 346(6215):1320-1331.

Kapusta, A. & Suh, A. (2016) Evolution of bird genomes – A transposon'seyevuew. *Annals of the New York Academy of Sciences*.

Ledesma, M, A.; Martinez P, A.; Calderón, O, S.; Boeris, J, M and Meriles, J, M. (2002) Descrição do cariótipo e padrões de bandas C e NOR em *Pheucticus aureoventris* (Emberizidae:Cardinalinae). *Revista Brasileira de Ornitologia* 14 (1) 5962.

Sambrook, J. & Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*. 3rd Edition, Vol. 1, Cold Spring Harbor Laboratory Press, New York.

Sasaki M, Ikeuchi T, Makino S (1968) A feather pulp culture technique for avian chromosomes, with notes on the chromosomes of the peafowl and the ostrich. *Experientia* 24: 1292–1293.

Sassi AK, Herédia F, Loreto EL, et al. (2005) Transposable elements P and gypsy in natural populations of *Drosophila willistoni*. *Genetics and Molecular Biology*. v. 28 (4).

Suh A, Churakov G, Ramakodi MP, Platt RN 2nd, Jurka J, Kojima KK, Caballero J, Smit AF, Vliet KA, Hoffmann FG, Brosius J, Green RE, Braun EL, Ray DA, Schmitz J. (2014) Multiple lineages of ancient CR1 retroposons shaped the early genome evolution of amniotes. *Genome Biol Evol*. 7 (1):205-17.

Wicker, T. et al. The repetitive landscape of the chicken genome. *Genome Research*, v. 15, n. 1, p. 126–136, 2005.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D.M., Lee, C., et al. (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, v. 346, n. 6215, p. 1311-1320.

5 CONCLUSÃO

No primeiro capítulo, com o mapeamento cromossômico foi possível inferir que as sequências de AviRTE são constituintes dos cromossomos sexuais de espécies que apresentam polimorfismos, reforçando a correlação entre sequências repetitivas e o aumento do tamanho cromossômico, ajudando a explicar o Z grande em CAM e o W de PCH e TSU.

Na abordagem *in silico* da análise filogenética, a diversidade de cópias e a distribuição de AviRTE no clado das aves, acompanhou a filogenia proposta para o grupo em grande parte, além de ampliar a distribuição para outras espécies. A distribuição descontínua em outros grupos de vertebrados, a ausência em mamíferos e a alta similaridade entre espécies distantemente relacionadas, sugere que o elemento AviRTE possivelmente pode estar envolvido em outros eventos de transferência horizontal.

No segundo capítulo, as amplificações de CR1 corroboram a distribuição contínua do elemento pela árvore aviária. Através das hibridizações por FISH, foi possível determinar a localização de elementos transponíveis e o acúmulo nos cromossomos sexuais das espécies estudadas, bem como, corroborar sua participação com o aumento do tamanho cromossômico.

Deste modo, com este trabalho apresentamos novos dados sobre elementos transponíveis em Aves, acreditando ter contribuído para o conhecimento científico nesta linha de pesquisa.

REFERÊNCIAS

- ARKHIPOVA, I.R. Using bioinformatic and phylogenetic approaches to classify transposable elements and understand their complex evolutionary histories. **Mobile DNA** 8, 19, 2017.
- BARCELLOS, S.; KRETSCHMER, R.; DE SOUZA, M.S.; et al. Karyotype Evolution and Distinct Evolutionary History of the W Chromosomes in Swallows (Aves, Passeriformes). **Cytogenetic and Genome Research**. 2019.
- BERTOCCHI, N.A., DE OLIVEIRA, T.D., GARNERO, A. D. V. et al. Distribution of CR1-like transposable element in woodpeckers (Aves Piciformes): Z sex chromosomes can act as a refuge for transposable elements. **Chromosome Res** 26, 333–343, 2018.
- CALLAGHAN, CT., NAKAGAWA, S., CORNWELL, WK. Global abundance estimates for 9,700 bird species. **Genome Biology and Evolution**. Proceedings of the National Academy of Sciences, 2021.
- CARARETO CMA, MONTEIRO-VITORELLO CB & SLUYS M-A Van. Elementos de transposição: diversidade, evolução, aplicações e impacto nos genomas dos seres vivos, 1st ed. **FIOCRUZ**, Rio de Janeiro. 2015.
- CHALOPIN D, NAVILLE M, PLARD F, GALIANA D, VOLFF JN. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. **Genome Biol Evol**. 2015.
- CHÉNAIS, B.; CARUSO, A.; HIARD, S.; CASSE, N. The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. **Gene**, 509, 7–15, 2012.
- CURCIO, M., DERBYSHIRE, K. The outs and ins of transposition: from Mu to Kangaroo. **Nat Rev Mol Cell Biol** 4, 865–877, 2003.
- DEGRANDI, T. M.; BARCELLOS, S. A.; COSTA, A. L.; GARNERO, A. D. V.; HASS, I.; GUNSKI, R. J. Introducing the Bird Chromosome Database: An Overview of Cytogenetic Studies in Birds. **Cytogenetic and Genome Research**. 160(4):199-205, 2020.
- DEGRANDI TM, DE OLIVEIRA JCP, SOARES ADA, LEDESMA MA, HASS I, GARNERO ADV, GUNSKI RJ. Karyotype description and comparative analysis in Ringed Kingfisher and Green Kingfisher (Coraciiformes, Alcedinidae). *Comparative Cytogenetics* 12(2): 163-170, 2018.
- DEGRANDI TM, GARNERO ADV, O'BRIEN PCM, et al. Chromosome Painting in Trogon s. surrucura (Aves, Trogoniformes) Reveals a Karyotype Derived by Chromosomal Fissions, Fusions, and Inversions. *Cytogenet Genome Res* 151 (4): 208–215, 2017.

DE OLIVEIRA TD, KRETSCHMER R, BERTOCCHI NA, DEGRANDI TM, DE OLIVEIRA EHC, et al. Genomic Organization of Repetitive DNA in Woodpeckers (Aves, Piciformes): Implications for Karyotype and ZW Sex Chromosome Differentiation. **PLOS ONE** 12(1): e0169987, 2017.

DE SOUZA, M.S.; KRETSCHMER, R.; BARCELLOS, A.S.; COSTA, A.L.; CIOFFI, M.B.; DE OLIVEIRA, E.H.C.; GARNERO, A.D.V.; GUNSKI, R.J. Repeat Sequence Mapping Shows Different W Chromosome Evolutionary Pathways in Two Caprimulgiformes Families. **Birds**, 2020.

DUNEMANN, S. M., & WASMUTH, J. D. Horizontal transfer of a retrotransposon between parasitic nematodes and the common shrew. **Mobile DNA**, 10, 24, 2019.

ELLEGREN H. Evolutionary stasis: the stable chromosomes of birds. **Trends Ecol Evol**. 2010.

FINNEGAN D.J. Eukaryotic transposable elements and genome evolution. Trends in genetics: **TIG**, 5(4), 103–107, 1989.

GUNSKI, RJ, KRETSCHMER, R, SOUZA M.S., FURO I.O., et al. Evolution of Bird Sex Chromosomes Narrated by Repetitive Sequences: Um usual W Chromosome Enlargement in Gallinula melanops (Aves: Gruiformes: Rallidae). **Cytogenetic and Genome Research**. 2019.

HAYWARD A, GILBERT C. Transposable elements. **Current Biology Magazine** 32, Sep 12; R904-R909, 2022.

IUCN [International Union for Conservation of Nature] (2020) The IUCN red list of threatened species. Version 2020–3. Acessado 26 de maio de 2023.

JACOBS, F, FENALTI, P. Aves do Rio Grande do Sul: Guia de Identificação. Editora Aratinga. 2020.

JARVIS ED *et al.* Whole-genome analyses resolve early branches in the tree of life of modern birds. **Science** 346, 1320-1331, 2014.

JURKA J., KAPITONOV V.V., PAVLICEK A., KLONOWSKI P., KOHANY O., WALICHIEWICZ J. Repbase Update, a database of eukaryotic repetitive elements. **Cytogenetic and Genome Research**. 110 (1-4): 462–467, 2005.

KAPITONOV V.V, JURKA J. A universal classification of eukaryotic transposable elements implemented in Repbase. **Nat Rev Genet** 9(5):411–412, 2008.

KAPUSTA, A. & SUH, A. Evolution of bird genomes – A transposon's eyeview. **Annals of the New York Academy of Sciences**. 2016.

KRETSCHMER, R.; FERGUSON-SMITH, M. A.; DE OLIVEIRA, E. H. C. Karyotype Evolution in Birds: From Conventional Staining to Chromosome Painting. **Genes**, 9(4): 181, 2018.

MCCLINTOCK, B. The Origin and Behavior of Mutable Loci in Maize. **Genetics**, v. 36, n. 6, p. 344–355, 1950.

NIETO LM, KRETSCHMER R, LEDESMA MA, GARNERO ADV, GUNSKI RJ: Karyotype morphology suggests that the *Nyctibius griseus* (Gmelin, 1789) carries an ancestral ZW-chromosome pair to the order Caprimulgiformes (Aves). **Comp Cytogenet** 6:379–387, 2012.

PACHECO, J.F., SILVEIRA, L.F., ALEIXO, A. *et al.* Annotated checklist of the birds of Brazil by the Brazilian Ornithological Records Committee—second edition. **Ornithol. Res.** 29, 94–105, 2021.

PIÉGU, B, BIRE, S, ARENSBURGER, P, BIGOT, Y, A survey of transposable element classification systems – A call for a fundamental update to meet the challenge of their diversity and complexity, **Molecular Phylogenetics and Evolution**, Vol. 86, 2015.

SENF, A.D., MACFARLAN, T.S. Transposable elements shape the evolution of mammalian development. **Nat Rev Genet** 22, 691–711, 2021.

SOTERO-CAIO CG, PLATT RN, SUH A, RAY D.A. Evolution and Diversity of Transposable Elements in Vertebrate Genomes, **Genome Biology and Evolution**, Volume 9, Issue 1, 2017.

SUH, A., WITT, C., MENGER, J. *et al.* Ancient horizontal transfers of retrotransposons between birds and ancestors of human pathogenic nematodes. **Nature Communications** 7, 11396, 2016.

STORER, J., HUBLEY, R., ROSEN, J. *et al.* The Dfam community resource of transposable element families, sequence models, and genome annotations. **Mobile DNA** 12, 2, 2021.

WANG ZJ, CHEN GJ, ZHANG GJ, ZHOU Q. Dynamic evolution of transposable elements, demographic history, and gene content of paleognathous birds. **Zool Res.** 2021.

WICKER, T., SABOT, F., HUA-VAN, A. *et al.* A unified classification system for eukaryotic transposable elements. **Nat Rev Genet** 8, 973–982, 2007.

ZHANG, G., LI, C., LI, Q., LI, B., LARKIN, D.M., LEE, C., *et al.* Comparative genomics reveals insights into avian genome evolution and adaptation. **Science**, v. 346, n. 6215, p. 1311-1320, 2014.

APÊNDICES

A- Tabela 2: Informações dos cariótipos analisados no FISH.

Espécie	2n	1	2	3	4	5	6	Z	W	Referência
<i>Chloroceryle americana</i>	94	SM	SM	SM	T	T	T	SM	SM	Degrandi <i>et al.</i> , 2018.
<i>Trogon surrucura</i>	82	SM	A	A	A	M	SM	M	SM	Degrandi <i>et al.</i> , 2017.
<i>Gallinula melanops</i>	80	SM	SM	M	M	M	T	SM	SM	Gunski <i>et al.</i> , 2019.
<i>Progne chalybea</i>	76	M	A	A	M	SM	T	M	SM	Barcellos <i>et al.</i> , 2019.

Legenda: 2n= número diploide; Morfologia cromossômica: M= metacêntrico, S= submetacêntrico, A= Acrocêntrico e T= Telocêntrico.