

UNIVERSIDADE ESTADUAL PAULISTA “Júlio de Mesquita Filho” - UNESP

INSTITUTO DE BIOCÊNCIAS DE BOTUCATU - IBB

Caio Augusto Gomes Goes

**Análise comparativa da diversidade de satélites nos gêneros da família
Erythrinidae (Characiformes, Teleostei)**

Orientador: Prof. Assoc. Fabio Porto Foresti Co-orientador: Prof. Dr. Ricardo Utsunomia

Botucatu, SP - 2023 –

UNIVERSIDADE ESTADUAL PAULISTA “Júlio de Mesquita Filho” - UNESP

INSTITUTO DE BIOCÊNCIAS DE BOTUCATU - IBB

Caio Augusto Gomes Goes

**Análise comparativa da diversidade de satélites nos gêneros da família
Erythrinidae (Characiformes, Teleostei)**

Tese apresentada ao Instituto de
Biociências de Botucatu, Universidade
Estadual Paulista - UNESP, como parte dos
requisitos para a obtenção do título de
Doutor em Ciências Biológicas (Zoologia).

**Orientador: Prof. Assoc. Fabio Porto Foresti Co-orientador: Prof. Dr.
Ricardo Utsunomia**

Botucatu, SP - 2023 -

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: MARIA CAROLINA A. CRUZ E SANTOS-CRB 8/10188

Goes, Caio Augusto Gomes.

Análise comparativa da diversidade de satélites nos gêneros da família *Erythrinidae* (*Characiformes*, *Teleostei*) / Caio Augusto Gomes Goes. - Botucatu, 2024

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Fábio Porto Foresti

Coorientador: Ricardo Utsunomia

Capes: 20400004

1. Genômica. 2. Peixes de água doce. 3. Citogenética animal.

Palavras-chave: Genômica; Hoplias; Satelitoma.

ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE CAIO AUGUSTO GOMES GOES, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS (ZOOLOGIA), DO INSTITUTO DE BIOCIÊNCIAS - CÂMPUS DE BOTUCATU.

Aos 26 dias do mês de janeiro do ano de 2024, às 09:00 horas, no(a) Departamento de Ciências Biológicas – UNESP Bauru, realizou-se a defesa de TESE DE DOUTORADO de CAIO AUGUSTO GOMES GOES, intitulada **Análise comparativa da diversidade de satélites nos gêneros da família Erythrinidae**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. FABIO PORTO FORESTI (Orientador(a) - Participação Presencial) do(a) Departamento de Ciências Biológicas / Faculdade de Ciências – Universidade Estadual Paulista Júlio de Mesquita Filho – Campus de Bauru, Prof. Dr. MARCELO DE BELLO CIOFFI (Participação Presencial) do(a) Departamento de Genética e Evolução / Universidade Federal de São Carlos (UFSCAR), Prof. Dr. ROBERTO FERREIRA ARTONI (Participação Presencial) do(a) Departamento de Biologia Estrutural, Molecular e Genética / Universidade Estadual de Ponta Grossa , Pós-Doutorando DUILIO MAZZONI ZERBINATO DE ANDRADE SILVA (Participação Virtual) do(a) National Institute of Health, Profa. Dra. FERNANDA DOTTI DO PRADO (Participação Virtual) do(a) Universidade Estadual do Norte do Paraná. Após a exposição pelo doutorando e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, o discente recebeu o conceito final: Aprovado . Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Prof. Dr. FABIO PORTO FORESTI

Documento assinado digitalmente
 **FABIO PORTO FORESTI**
Data: 06/02/2024 08:20:37-0300
Verifique em <https://validar.iti.gov.br>

“Não nos cabe decidir o tempo em que vivemos. Apenas precisamos decidir o que fazer com o tempo que nos é dado.”

J. R. R. Tolkien

Esse trabalho é dedicado à minha filha Clara, que eu ainda nem conheço mas já me faz querer ser uma pessoa melhor.

Agradecimentos

Muitas pessoas me ajudaram a concluir essa etapa, e todas merecem um agradecimento especial.

Agradeço a Deus pelo caminho até aqui, que não foi fácil, mas que eu não mudaria nada.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pelo auxílio concedido.

A Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP, por todo o conhecimento adquirido aqui, seja acadêmico ou pessoal.

Ao Instituto de Biociências, da Universidade Estadual Paulista, campus de Botucatu, pela oportunidade de concluir a pós-graduação.

Ao Laboratório de Genética de Peixes, local que me fascinou e me fez mudar de carreira. Terminamos nossa jornada juntos, e mal posso esperar para o que vem no novo Laboratório de Genômica e Conservação de Peixes.

As instituições parceiras por todo o apoio oferecido a esse trabalho e a mim: Laboratório de Biologia e Genética de Peixes (Botucatu), Laboratório de Citogenética de Peixes (São Carlos), e Laboratório de Biotecnologia de Peixes (Pirassununga).

Ao meu orientador, Prof. Doutor Fábio Porto-Foresti, por ter visto meu potencial, quando nem eu mesmo sabia disso. Com certeza você é uma pessoa chave na minha vida, que alterou totalmente o rumo que ela ia levando. Obrigado por todo o aconselhamento, ensinamentos, amizade, desafios, momentos de descontração e de suporte. Obrigado por sempre estar disponível para uma boa conversa, seja ela por motivos profissionais ou pessoais. Com você eu aprendo todo dia, e em algum momento da minha vida eu gostaria de lidar com pessoas do mesmo jeito que você. Obrigado por me fazer cientista e uma pessoa muito melhor.

Ao meu coorientador, Prof. Doutor Ricardo Utsunomia, por toda a orientação profissional e amizade. Com certeza você é um dos pesquisadores mais influentes da minha carreira, e obrigado por me ensinar tanto.

Ao Prof. Dr. Marcelo Cioffi, e a toda equipe do Laboratório de Citogenética de Peixes da UFSCAR, em especial aos amigos Francisco e Renata. Vocês foram essenciais para o desenvolvimento desse trabalho e são inspirações para mim. Muito obrigado pela parceria, oportunidades e reconhecimento.

A todos os companheiros do Laboratório de Genômica e Conservação de peixes, que foram muito importantes para todo o desenvolvimento do meu doutorado. Zeni, Henrique, Natalia, Amanda, Gabi, Verônica, Camila, Eduardo, Bia, Mel, Vitor, Helio, Kevin, Michel, Auany, Mariana, Letícia, Tadaka, Gabriel, Vinicius, Mingau, José, Gisloti, Lara, Luisa, Giovanna Igepi, Giovana e Manuela. E a todos os incontáveis amigos que participaram desse grupo, mas seguiram seus próprios caminhos. Cada um de vocês participou da minha formação de alguma maneira, e seria impossível passar por todo esse caminho se não fosse na companhia de pessoas tão especiais. Quando a hora chegar, vou guardar com carinho lembranças de cada um de vocês.

A todos os amigos que fiz após minha saída de Pirajuí, em especial Raul, Angela, Ana Julia, Eduardo, Giovanna, Beca, Edgar e Ana Luísa. É graças aos momentos felizes do lado de vocês que suportei as dificuldades que encontrei por aqui.

Aos meus sogros, Daniel e Liene, meus cunhados, Daniel, Natalia, Aline e Saulo, e a toda a minha família de modo geral. Tem um pedacinho de vocês em cada uma dessas páginas e muito obrigado por todo o apoio.

Aos meus maiores e mais antigos amigos, Ewerton, Lucas, Neto, Marcelo, Gleice, Pilpo, Daniel, Adolfo e Ana Clara. Hoje nos vemos muito pouco, e a falta que sinto da companhia de vocês é a maior prova de que construímos uma amizade verdadeira e duradoura. Algumas das minhas memórias mais felizes estão ligadas a vocês, e todos vocês moram em meu coração. É uma alegria para mim saber que nossas aventuras juntos estão muito longe de acabar.

Aos meus pais, Donizeti e Nilva, que são as pedras fundamentais que sustentam a minha vida. Eu mal consigo imaginar o quanto vocês sacrificaram para que eu conseguisse chegar até aqui. Pai, você sempre foi o meu exemplo e a minha bússola para saber se aquele era o caminho certo a se seguir. Muito obrigado por me ensinar toda sua honestidade, sua honra e por ser o espelho de pai que eu gostaria de seguir. Eu já te disse uma vez, mas se conseguir ser a metade da pessoa que o senhor é, ficarei plenamente satisfeito. Mãe, muito obrigado por estar sempre cuidando de mim, seja de perto ou de longe. Eu consigo sentir seu zelo e amor mesmo aqui de Bauru, e sinto sua presença sempre perto de mim. Muito obrigado por vocês me amarem incondicionalmente, e por me darem sempre um lugar para voltar.

A minha amada esposa, Karoline. Da última vez que escrevi um agradecimento, ainda éramos namorados, e uma das minhas maiores alegrias foi ter me casado com você. Esses últimos tempos não foram nada fáceis, e saiba que se você não estivesse ao meu lado, eu não teria chegado até aqui. Obrigado por me causar os sorrisos mais sinceros e enxugar as lágrimas mais verdadeiras. Você é a pessoa que me dá propósito, que me faz levantar e seguir em frente, que me faz enfrentar medos e receios. Todo dia eu quero ser um homem melhor para ser merecedor de estar do seu lado. Tudo o que construímos juntos é especial porque é feito a base do nosso amor, que é o sentimento mais lindo que eu já senti. Tenho muito orgulho da mulher que você é, com posicionamento, contestadora, feminista, gentil e amorosa. Tenho certeza de que você vai conseguir alcançar todos os seus objetivos e metas, seja elas quais forem, por que eu conheço toda a energia que forma a pessoa que você é. Nossa filha não poderia ter escolhido uma pessoa melhor para ser sua mãe. Amo você e todas as suas versões, mais do que ontem e menos do que amanhã.

E finalmente, agradeço a você Clarinha, por ter me escolhido para ser seu pai. Eu nunca fiz isso, mas pode ter certeza de que vou me esforçar para ser o melhor pai que você poderia ter. Eu mal posso esperar para te ter em meus braços, e você é o maior presente que eu poderia receber. Eu e sua mãe escolhemos te trazer para um mundo que, sinceramente, não é um lugar muito amigável. Mas nós temos certeza de que mesmo a menor das pessoas pode fazer do mundo um lugar melhor. Algum dia, quando você estiver lendo isso, saiba que o meu mundo já se tornou melhor só por saber que você está chegando. Seu pai te ama muito, tudo isso é por você.

RESUMO

O genoma dos eucariotos é formado em grande parte de sequências repetitivas, como famílias multigênicas, elementos transponíveis e DNAs satélites (satDNAs). Dentre estas, satDNAs são sequências repetidas *in tandem* que podem formar *arrays* de milhares de repetições de seus monômeros. Tradicionalmente, a evolução de satDNAs segue os preceitos da hipótese *library*, que considera que espécies próximas compartilham uma mesma biblioteca de satDNAs, com divergências na amplificação dessas sequências em cada grupo. Recentemente, o desenvolvimento de ferramentas bioinformáticas permitiu a caracterização de catálogos de satDNAs denominados satelitomas, e essa ferramenta tem trazido importantes informações acerca da evolução desse tipo de sequência em diversos grupos, como plantas, insetos, moluscos, aves e peixes. Neste sentido, a família Erythrinidae abriga os peixes neotropicais dos gêneros *Hoplerythrinus*, *Erythrinus* e *Hoplias*, foco de estudos citogenéticos devido a presença marcante de cromossomos sexuais e diversificação cariotípica, com a presença de diversos cariomorfos para espécies como *Hoplias malabaricus* e *Erythrinus erythrinus*. Sendo assim, o presente trabalho teve como objetivo caracterizar e observar a similaridade entre os satelitomas de *Hoplerythrinus unitaeniatus*, *E. erythrinus* (cariomorfo A), *Hoplias intermedius* e *H. malabaricus* (cariomorfos F e G), a fim de testar a hipótese de haver a presença de satDNAs conservados em todos os grupos. As análises detectaram a presença de seis satDNAs compartilhados entre todos os catálogos descritos neste trabalho. Com exceção de um desses satDNAs, todos os outros demonstraram padrão de clusterização por *Fluorescence in situ hybridization* (FISH) em regiões centroméricas ou teloméricas de todos os indivíduos analisados. Além disso, foi possível observar similaridade entre os catálogos corroborando as relações filogenéticas estabelecidas para o grupo. Nossos resultados demonstram o grande compartilhamento de satDNAs entre a família Erythrinidae, corroborando com a hipótese *library*, além de levantar novas evidências em relação a caracterização de *H. malabaricus* como um complexo de espécies.

ABSTRACT

The genome of eukaryotes is made up largely of repetitive sequences, such as multigenic families, transposable elements, and satellite DNAs (satDNAs). SatDNAs are tandem repeat sequences that can form arrays of thousands of repeats of their monomers. Traditionally, the evolution of satDNAs follows the precepts of the library hypothesis, which considers that closely related species share the same library of satDNAs, with divergences in the amplification of these sequences in each group. Recently, the development of bioinformatics tools has allowed the characterization of catalogs of satDNAs called satellitomes, and this tool has provided important information about the evolution of this type of sequence in several groups, such as plants, insects, mollusks, birds, and fish. In this sense, the family Erythrinidae represents the Neotropical fishes of the genera *Hoplerythrinus*, *Erythrinus*, and *Hoplias*, the focus of cytogenetic studies due to the marked presence of sex chromosomes and karyotypic diversification, with the presence of several karyomorphs for species such as *Hoplias malabaricus* and *Erythrinus erythrinus*. Thus, the present work aimed to characterize and compare the satellitomes of *Hoplerythrinus unitaeniatus*, *E. erythrinus* (karyomorph A), *Hoplias interdemius* and *H. malabaricus* (karyomorphs F and G) to observe the similarity of these catalogs and diagnose the presence of conserved satDNAs in all groups. Our analyses detected the presence of six satDNAs shared among all catalogs described in this work, also represented in the satellitome of *H. malabaricus* (karyomorph D). Except for one of these sequences, all the others showed clustering patterns by FISH in the centromeric or telomeric regions of all the individuals analyzed. Furthermore, it was possible to observe the similarity among the catalogs, corroborating the phylogenetic relationships established for the group, with the catalogs of *H. unitaeniatus* and *E. erythrinus* being closer to each other than to any other catalog of *Hoplias*. Our results demonstrate the great sharing of satDNAs among the Erythrinidae family, corroborating the library hypothesis as well as raising new evidence in relation to the characterization of *H. malabaricus* as a species complex.

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Características gerais de DNAs satélites	12
1.2 Evolução de DNAs satélites	14
1.3 Caracterização de catálogos de DNAs satélites e satelitômica comparativa	16
1.4 Estudo de DNAs satélites em peixes Neotropicais: de marcadores citogenéticos a evolução de satelitomas	18
1.5 Família Erythrinidae e sua importância para citogenômica	21
2 OBJETIVOS	22
3 MATERIAIS E MÉTODOS	23
3.2.1 Extração do DNA genômico total e sequenciamento	24
3.2.2 Caracterização dos satelitomas por TAREAN	24
3.2.3 Filtragem de famílias multigênicas e eliminação de redundâncias	24
3.2.4 Caracterização de abundância e divergência dos satDNAs	25
3.2.5 Comparação entre os satelitomas descritos	25
3.2.6 Seleção dos satDNAs para FISH e desenho de primers	26
3.2.7 Obtenção de cromossomos mitóticos	26
3.2.8 PCRs e marcação das sondas	27
3.2.9 Hibridização <i>in situ</i> fluorescente (FISH)	28
4 RESULTADOS	29
4.1 Análises citogenéticas	29
4.2 Descrição de cinco novos satelitomas para a família Erythrinidae	29
4.3 Comparação entre os catálogos de Erythrinidae refletem as relações filogenéticas do grupo	41
4.4 Compartilhamento de sequências consensos revela seis satDNAs conservados na família Erythrinidae	42
4.5 Mapeamento por FISH de sequências conservadas demonstra o compartilhamento de padrões de clusterização	44
4.6 MSTs revelam diferentes padrões de expansão de haplótipos em satélites conservados de Erythrinidae	49
5 DISCUSSÃO	52
6 CONSIDERAÇÕES FINAIS	56
7 REFERÊNCIAS	57
MATERIAIS SUPLEMENTARES	74

1 INTRODUÇÃO

1.1 Características gerais de DNAs satélites

O genoma dos eucariotos é repleto de sequências repetitivas, tais como elementos transponíveis (TEs), famílias multigênicas, pseudogenes e DNAs satélites (satDNAs) (Kim et al., 2014). Tradicionalmente, as sequências repetitivas são divididas em repetidas *in tandem*, como famílias multigênicas e satDNAs, e dispersas pelo genoma, como TEs (López-Flores e Garrido-Ramos, 2012). SatDNAs são tradicionalmente definidos como *arrays* formados por milhares de repetições em tandem (cabeça-cauda) de unidades altamente similares entre si denominadas monômeros, geralmente localizadas em segmentos cromossômicos heterocromáticos (López-Flores e Garrido-Ramos, 2012; Biscotti et al., 2015). Sua estrutura genômica diverge da estrutura encontrada em famílias multigênicas (DNAs ribossomais e histonas, por exemplo), em que as unidades repetitivas são formadas por vários tipos de sequência, tais como genes e espaçadores, contrastando com a repetição de um único tipo de sequência observada nos satDNAs (Ruiz-Ruano et al., 2016). SatDNAs podem variar em seu tamanho desde poucas bases até milhares de nucleotídeos, sendo classificados de acordo com o seu tamanho em microssatélites, minissatélites ou satélites, apesar de não haver um consenso nos limites de tamanho de cada categoria (Tautz et al., 1993; Kim et al., 2014).

O conceito de satDNAs define essas sequências como parte das regiões heterocromáticas do genoma (Louzada et al., 2020; Thakur et al., 2021). A localização desse tipo de sequência em regiões de heterocromatina é cada vez melhor fundamentada, devido ao mapeamento cromossômico de cada vez mais satDNAs em diversas espécies (Šatović-Vukšić e Plohl, 2023). Entretanto, vale ressaltar que poucas espécies têm a posição cromossômica de satDNAs conhecida, frente a ampla diversidade global. Com o aprofundamento do estudo de satDNAs, clusters desse tipo de sequência também têm sido observados em regiões eucromáticas (Kuhn et al., 2012; Feliciello et al., 2015; Ruiz-Ruano et al., 2016), incluindo casos interessantes como o do besouro *Rhynchophorus ferrugineus*. Nesta espécie, satDNAs correspondem a cerca de 25% do genoma, com as famílias mais abundantes localizadas de forma dominante em regiões eucromáticas, sendo esse um dos catálogos de satDNAs mais diverso em insetos (Montiel et al., 2022).

A noção de que satDNAs podiam ser considerados como “DNA lixo” vem sendo cada vez mais abandonada, devido ao acúmulo de evidências do papel destas sequências na estrutura e evolução cromossômica (Plohl et al., 2014; Escudeiro et al., 2019; Vozdova et al., 2020). O grande acúmulo de satDNAs nos centrômeros de diversas espécies demonstra o papel destas sequências na organização da cromatina centromérica (Plohl et al., 2014; Hartley e O’Neill, 2019; Talbert e Henikoff, 2022) e pareamento meiótico dos cromossomos (Larracuente, 2014). Algumas evidências apontam que a evolução dessas sequências pode estar relacionada a especiação, como a presença de satDNAs centroméricos diferentes para espécies próximas, apontando a substituição dessas sequências durante o processo de especiação (Camacho et al., 2022). Além disso, a função dos satDNAs na formação da heterocromatina tem sido cada mais estabelecida, através de estudos relacionados a transcrição dessas sequências, apontando sua função na manutenção da heterocromatina, definição da identidade do centrômero e preservação da estabilidade genômica (Ahmad et al., 2020; Podgornaya, 2022; Šatović-Vukšić e Plohl, 2023).

Um número crescente de estudos vem deixando cada vez mais destacada a relação intrínseca entre satDNAs e TEs, principalmente relacionado a dispersão das sequências repetidas *in tandem* (Meštrović et al., 2015; Paço et al., 2019; Zattera e Bruschi, 2022). O principal mecanismo de dispersão dos TEs é a *rolling circle replication* (Kapitonov e Jurka, 2007; Thomas e Pritham, 2015), no qual, durante o processo, os TEs podem incorporar sequências *in tandem* em sua estrutura, contribuindo com sua dispersão ao longo do genoma (Yang e Barbash, 2008; Dias et al., 2015; Šatović e Plohl, 2013). Sendo assim, as sequências *in tandem* incorporadas aos TEs podem dar origem a arranjos clássicos de satDNAs, que são mantidos por *crossing over* desigual e outros mecanismos de transferência não-recíproca (Plohl et al., 2012; Garrido-Ramos, 2017;).

A descoberta dos satDNAs deu-se através de um pequeno pico de ultracentrifugação de CsCl (Kit e Hsu, 1961), de onde originou-se o nome “DNA satélite”. Após sua descoberta, diversas metodologias foram utilizadas para a caracterização deste tipo de sequência, como a cinética de renaturação (Hutton e Wetmur, 1973; Botchan, 1974;) e a digestão com enzimas de restrição (Horz e Zachau, 1977; Kopecka et al., 1978), por exemplo. Apesar da precisão em se caracterizar um satDNA através dessas ferramentas, essas metodologias não eram eficientes, devido a dependência da aleatoriedade da enzima selecionada ser compatível com um monômero de algum satDNA. Além disso, satDNAs são sequências muito dinâmicas, apresentando alterações de maneira rápida e frequente, tornando muitas dessas sequências espécie-específicas, dificultando o estudo da evolução dessas sequências em grupos próximos utilizando ferramentas que não sejam eficientes na

caracterização de DNAs satélites. O panorama do estudo de DNAs satélites começou a mudar com o surgimento e popularização do sequenciamento de nova geração (ou NGS, de *next generation sequencing*), permitindo a geração massiva de dados a custos reduzidos (Park e Kim, 2016). A possibilidade de obtenção de dados genômicos de diversas espécies, incluindo as não modelo, e o surgimento de ferramentas e protocolos dedicados ao estudo de sequências repetitivas, tais como o *Repeat explorer* (Novák et al., 2013), TAREAN (Novák et al., 2017) e a *pipeline* SatMiner (RuizRuano et al., 2016), tem possibilitado a caracterização de catálogos de satDNAs de grupos e espécies específicas, contribuindo de forma impactante para o estudo da evolução de satDNAs.

1.2 Evolução de DNAs satélites

SatDNAs representam um grupo de sequências extremamente diverso, acumulando diferenças em conteúdo A+T, tamanho dos monômeros, tamanho de *arrays*, localização na cromatina e distribuição cromossômica (Šatović-Vukšić e Plohl, 2023). Sendo assim, o surgimento de formas de se caracterizar coleções de satDNAs para diversas espécies tem se tornado um fator importante para o desenvolvimento de novos *insights* sobre a origem e evolução dessas sequências.

Recentemente, o estudo do catálogo de satDNAs de gafanhotos demonstrou que os satDNAs de *Locusta migratória* podem ser encontrados de forma clusterizada ou não-clusterizada (dispersa) pelo genoma (Ruiz-Ruano et al., 2016). Esse fator levou os autores a propor um modelo de origem e evolução desse tipo de sequência no genoma, sendo o primeiro passo para o surgimento de um satDNA a duplicação de uma sequência genômica, através de mecanismos como deslizamento na replicação ou *rolling circle replication*, dando origem a pequenos *arrays*. Nesse momento, esses pequenos *arrays* podem ser disseminados pelo genoma, por mecanismos como a ação de TEs e reinserção. Em eucariotos, qualquer um desses novos pequenos *arrays* podem passar por amplificação local por mecanismos como *crossing over* desigual, levando a formação de um cluster que pode ser visualizado por *Fluorescence in situ hybridization* (FISH) (Ruiz-Ruano et al., 2016).

Em termos de evolução de sequências repetidas *in tandem*, duas características são destaque: baixa variabilidade nas sequências das unidades repetitivas (monômeros) e grande variação no número de cópias (Charlesworth et al., 1994; Plohl et al., 2012; Garrido-Ramos, 2017, 2021; Lower et al., 2018; Thakur et al., 2021). Neste sentido, o mecanismo de evolução em concerto exerce um papel importante na evolução de satDNAs, contribuindo para a homogeneização dos monômeros (Elder Jr e Turner, 1995; Smith, 1976; Stephan, 1986). Nesse modelo, monômeros em

arrays de satDNAs evoluem em conjunto, mantendo baixa variabilidade nas sequências entre eles, devido a mutações ocorridas em monômeros específicos serem eliminadas ou distribuídas, através de mecanismos de trocas não-recíprocas (*crossing over* desigual, conversão gênica), num mecanismo denominado *drive* molecular (Dover, 1982, 1986; Strachan et al., 1985). Sendo assim, os monômeros de um determinado satDNA podem ser homogeneizados de maneiras diferentes entre espécies, resultando em sequências homogêneas dentro de cada espécie, mas com rápida divergência entre as espécies, levando à formação de variantes espécie-específicas (Dover, 1987; Bachmann e Sperlich, 1993; Garrido-Ramos et al., 1999; Plohl et al., 2012).

Em contraposição, as mutações recorrentes em sequências repetidas *in tandem* podem se acumular ao longo de um satDNA, diminuindo ou desabilitando os mecanismos de homogeneização de sequências, conforme o proposto no modelo de evolução divergente (Nijman e Lenstra, 2001). Esse modelo pressupõe que um novo monômero divergente pode ser amplificado e formar uma nova subfamília de repetições *in tandem*, relacionando esse modelo evolutivo com o surgimento de novos satDNAs. Em apoio, análises experimentais revelam que os mecanismos de homogeneização são menos eficientes nos monômeros finais da matriz repetitiva, podendo este ser uma fonte do surgimento de novas sequências (Smith, 1976; McAllister e Werren, 1999; Mravinac e Plohl, 2007).

Desde os primeiros estudos voltados a satDNAs, observava-se a presença de um ou mais satDNAs extremamente diferentes em seu número de cópias dentro de um genoma (Šatović-Vukšić e Plohl, 2023). Os mesmos mecanismos de trocas não-recíprocas responsáveis pela evolução em concerto também têm a capacidade de alterar drasticamente os *arrays* de sequências repetidas *in tandem*, tendo como consequência, durante o processo de especiação, a possível diminuição do número de cópias de um satDNA abundante ou a expansão de um satDNA pouco abundante, sendo esse o principal princípio da “hipótese *library*” (Fry e Salser, 1977). De acordo com esse modelo, espécies próximas compartilham parcial ou integralmente um catálogo de satDNAs, mas a abundância de cada um pode variar drasticamente. Cada vez mais evidências reforçam a hipótese *library* em diversos grupos, como peixes (Utsunomia et al., 2017; Goes et al., 2022, 2023) e insetos (Palacios-Gimenez, et al., 2020; Camacho et al., 2022), além de cada vez mais surgirem relatos de satDNAs conservados ao longo de muito tempo em espécies relativamente distantes e, em casos extremos, presentes em toda uma ordem (Plohl et al., 2010; Petraccioli et al., 2015; Halbach et al., 2020; dos Santos et al., 2021;). Entretanto, Belyayev et al., (2020) aponta algumas questões em discordância a hipótese *library*, tais como o surgimento de novos satDNAs, como as bibliotecas

sobrevivem a purificação do repetoma relacionado a especiação, e a subsequente evolução em concerto e o surgimento de novos satDNAs relacionados a TEs.

A descrição de satDNAs conservados ao longo de grupos distantes e por muito tempo ocorre em grupos como peixes (dos Santos et al., 2021), moluscos (Plohl et al., 2010; Petraccioli et al., 2015), e insetos (Halbach et al., 2020). Enquanto a existência dessas sequências pode expandir o conceito da hipótese *library*, expandindo a conservação de satDNAs em espécies distantes, a conexão de algumas dessas sequências com TEs pode afetar as conclusões sobre a ancestralidade comum dessas sequências, abrindo a possibilidade de uma transferência horizontal (Šatović Vukšić e Plohl, 2021). Nesse sentido, o surgimento de tecnologias que permitem a caracterização de catálogos de satDNAs para cada espécie torna-se fundamental para o melhor entendimento da evolução de satDNAs, permitindo a análise completa do conjunto de satDNAs presentes naquele grupo e em espécies relacionadas, para assim corroborar ou desafiar a hipótese *library*.

1.3 Caracterização de catálogos de DNAs satélites e satelitômica comparativa

Com a popularização e disponibilidade cada vez maior de sequenciamento de nova geração e o surgimento de tecnologias voltadas para o estudo de sequências repetitivas, como o *RepeatExplorer* (Novák et al., 2013) e TAREAN (Novák et al., 2017), a descrição de catálogos de satDNAs para cada espécie específicas tornou-se possível. O primeiro catálogo descrito utilizando essas metodologias foi o do gafanhoto *Locusta migratoria* (Ruiz-Ruano et al., 2016), em que os autores propõem o nome satelitoma para a coleção de satDNAs que forem descritos para cada espécie. Neste trabalho pioneiro, houve a descrição de 62 famílias de satDNAs, com padrões clusterizados e dispersos pelos cromossomos (Ruiz-Ruano et al., 2016). Além disso, o padrão de distribuição de cinco satDNAs relacionados em uma superfamília demonstra que satDNAs não clusterizados podem se tornar clusterizados por amplificação local de um dos muitos *arrays* dessa sequência dispersa pelo genoma (Ruiz-Ruano et al., 2016).

Atualmente, cerca de 70 satelitomas foram descritos, principalmente para plantas, peixes, insetos e moluscos, mas com representações em mamíferos, anfíbios e nematoides (revisão em Šatović-Vukšić e Plohl, 2023). A caracterização de satelitomas tem permitido conclusões interessantes relacionadas a relação entre quantidade de satDNAs presentes em uma determinada espécie e a representatividade dos satDNAs no genoma. Em peixes, por exemplo, *Megaleporinus macrocephalus* apresenta 164 satDNAs em seu satelitoma, representando 2,78% de seu genoma

(Utsunomia et al., 2019), enquanto *Astyanax lacustris* tem apenas 33 satDNAs em seu catálogo, com nenhum representando mais de 0,002% de abundância (Goes et al., 2022). Em insetos, os satelitomas variam de apenas um satDNA em *Cydalima perspectalis* (Cabral-de-Mello et al., 2021) até 129 no gafanhoto *Vandiemennella viatica* (Palacios-Gimenez et al., 2020). A abundância neste grupo também é extremamente variável, com *Hippodamia variegata* apresentando 30 satDNAs formando 15% do genoma (Mora et al., 2020) e *Rhodnius prolixus* apresentando 39 satDNAs correspondendo a 8% do genoma (Montiel et al., 2021). O mesmo cenário repete-se em plantas, com satelitomas variando de 11 satDNAs para espécies do gênero *Olea* representando cerca de 24% do genoma (Mascagni et al., 2022) e 23 satDNAs em *Vicia faba* com representatividade de 2,7% do genoma dessa espécie (Robledillo et al., 2018). Outra aplicação do estudo de satelitomas é a prospecção de satDNAs que estão relacionados à evolução e composição de cromossomos sexuais e Bs em diversos grupos, principalmente peixes (Silva et al., 2017; Utsunomia et al., 2019; SerranoFreitas et al., 2020; ; Stornioli et al., 2021; Kretschmer et al., 2022) e insetos (Milani et al., 2018, 2021; Ruiz-Ruano et al., 2018).

Com a caracterização de satelitomas de espécies, surge a possibilidade de se comparar esses catálogos e avaliar a evolução dessas sequências, mesmo em espécies não-modelo. A análise comparativa entre gafanhotos das espécies *L. migratoria* e *Oedaleus decorus*, pertencentes a subfamília Oedipodinae, revelou o compartilhamento de 41 famílias de satDNAs, representando 12 grupos ortólogos, formando uma biblioteca ancestral para esse grupo (Camacho et al., 2022), em conformidade a hipótese *library*. Os autores postulam que as demais famílias de satDNAs devem representar tanto os satDNAs exclusivos de apenas uma espécie, quanto os surgidos durante a especiação desses dois grupos. De forma similar, a comparação entre satelitomas de diferentes linhagens de barbeiros (*Triatoma infestans*) caracterizou 34 satDNAs presentes em todas as linhagens analisadas, de um total de 42 satDNAs presentes no satelitoma (Pita et al., 2017). Mesmo em grupos mais complexos, a comparação entre satelitomas demonstra o compartilhamento de sequências entre espécies próximas, como a comparação realizada em toupeiras da espécie *Talpa aquitania*, em que a maioria dos satDNAs presentes nessa espécie são compartilhados com outras espécies do gênero *Talpa* (Gutiérrez et al., 2022). Peixes neotropicais representam um dos principais grupos alvo de estudos relacionados a evolução de satDNAs e o papel dessas sequências na evolução cromossômica (Silva et al., 2017; Utsunomia et al., 2019; Crepaldi e Parise-Maltempi, 2020; SerranoFreitas et al., 2020; dos Santos et al., 2021; Stornioli et al., 2021; Goes et al., 2022, 2023; Kretschmer et al., 2022). Sendo assim, a comparação de satelitomas envolvendo espécies desse

grupo tem muito a contribuir para o entendimento da evolução dessas sequências, devido à grande diversidade do grupo e presença de diversos sistemas de cromossomos sexuais, cromossomos B e diferentes cariomorfos para a mesma espécie.

1.4 Estudo de DNAs satélites em peixes Neotropicais: de marcadores citogenéticos a evolução de satelitomas

A região Neotropical reúne desde parte do México até o sul da Argentina, e compreende a maior diversidade de peixes do planeta, com aproximadamente 9100 espécies diagnosticadas, sendo cerca de 5160 espécies presentes em águas continentais (Reis et al., 2016). Dentre estas, uma das ordens mais diversas em número espécies são os Characiformes, com 2342 espécies descritas, divididas em 24 famílias (Eschmeyer e Fricke, 2023), sendo um grupo extremamente importante do ponto de vista da citogenética de peixes, devido a diversos casos de diversidade de cariomorfos, presença de diversos sistemas sexuais e cromossomos B.

O estudo de satDNAs voltado para Characiformes teve seu início com a descrição de marcadores citogenéticos clássicos, como o As51 para o gênero *Psalidodon* (Mestriner et al., 2000), SATH1 e SATH2 para *Prochilodus* (Jesus et al., 2003), 5sHindIII para *Hoplias* (Haff et al., 1993) e pPh2004 para *Parodon* (Vicente et al., 2003), levando a conclusões significativas para cada um desses grupos. As51 foi descrito no genoma de *Psalidodon scabripinnis* (Mestriner et al., 2000) e mapeado em populações desta mesma espécie, além de *Psalidodon fasciatus*, *Astyanax janae*, *Psalidodon paranae*, *Astyanax lacustris* e *Astyanax* sp. D (Mantovani et al., 2004; Abel et al., 2006; Vicari et al., 2008; Kante et al., 2009), demonstrando a presença desse satDNA em posição similar dos cromossomos de todos esses grupos, com exceção de *A. lacustris* (Kante et al., 2009; Goes et al., 2022). Além disso, outros trabalhos também permitiram a localização desse satDNA em posição similar ao 18S rDNA, além de sua presença em cromossomos B de *P. scabripinnis* e *P. fasciatus* (Vicari et al., 2008; Silva et al., 2016). No gênero *Prochilodus*, o estudo dos satDNAs SATH1 e SATH2 permitiu inferir a origem intraespecífica dos cromossomos B de *Prochilodus lineatus* (Jesus et al., 2003; Artoni et al., 2006;) e o compartilhamento dessa sequência com o cromossomo supranumerário de duas outras espécies, *Prochilodus argenteus* e *Prochilodus costatus*, sugerindo uma origem ancestral destes elementos genômicos (Melo et al., 2017). O satDNA 5SHindIII, anteriormente chamado de Hop (Haff et al., 1993) foi descrito em *Hoplias malabaricus* e apresenta muitas semelhanças com o gene e os espaçadores da família multigênica 5S rDNA (Martins et al., 2006) sendo esta sua possível origem. Este satélite é exclusivo da espécie *H. malabaricus*, estando ausente até mesmo em espécies

próximas como *Hoplias lacerdae* (Ferreira et al., 2007; Blanco et al., 2010), e é representado por FISH em todos os cariomorfos de *H. malabaricus*, com diferenças entre o número de sítios entre estes (Blanco et al., 2010; Cioffi et al., 2010, 2011, 2013; Cioffi e Bertollo, 2010; Oliveira et al., 2018), tornando este satDNA um bom marcador citogenético para o grupo, demonstrando a evolução cariotípica entre os cariomorfos de *H. malabaricus* (Cioffi et al., 2009).

O surgimento do *Repeat Explorer* (Novák et al., 2013) e mais recentemente do TAREAN (Novák et al., 2017) impulsionou de forma significativa o estudo de satDNAs em peixes da ordem Characiformes. A primeira espécie deste grupo a ter seu satelitoma descrito foi *P. paranae*, contendo um total de 45 famílias de DNAs satélites (Silva et al., 2017). Diversos destes satDNAs descritos foram encontrados por FISH nos cromossomos desta e de duas espécies relacionadas, além de evidenciar as diferenças entre o conteúdo de satDNAs nos cromossomos Bs de *P. paranae*, *P. bockmanni* e *P. fasciatus* (Silva et al., 2017). Em continuidade aos estudos relacionados à cromossomos B, foram caracterizadas 51 famílias de satDNAs satélites para *P. lineatus* (Stornioli et al., 2021). O mapeamento por FISH apresentou um grande acúmulo de DNAs satélites nas três variantes de cromossomos B da espécie, um fato que é explicado devido a sua natureza não recombinante e a predominância de regiões heterocromáticas nestes elementos (Camacho et al., 2000). Outro importante dado levantado por Stornioli et al. (2021) é a possível origem do cromossomo B em *P. lineatus*, a partir do par 4 de autossomos, visto que o satDNA PliSat05-178 foi observado exclusivamente nos cromossomos B e nos dois autossômicos do par.

Outros trabalhos destacam-se na satelitômica de peixes Neotropicais, como por exemplo o catálogo desenvolvido para *M. macrocephalus* formado por 164 satDNAs, sendo este o maior satelitoma caracterizado em peixes neotropicais (Utsunomia et al., 2019). Neste caso, foi possível observar uma maior concentração de DNAs satélites em fêmeas da espécie (1.46% do genoma a mais, ao se comparar ao macho), com três satélites exclusivos da biblioteca feminina, além do satélite mais abundante da espécie, o MmaSat01-566, ser apresentado em uma proporção maior no *landscape* feminino. Corroborando com estes dados, 14 dos 30 DNAs satélites analisados por Utsunomia et al. (2019) apresentaram sinais de FISH somente no cromossomo W da espécie, possibilitando a criação de ferramentas não invasivas para sexagem de indivíduos de *M. macrocephalus*, como a combinação de um satélite pertencente a um par autossômico (MmaSat9837) com um W-específico (MmaSat97-39) na técnica de FISH. Seguindo a abordagem de DNAs satélites em cromossomos sexuais, Kretschmer et al., 2022, realizou a descrição do satelitoma da espécie *Triportheus auritus* e mapeamento por FISH dos satélites com abundância diferencial

entre machos e fêmeas nesta espécie e em mais duas da família Triportheidae. Pela técnica de FISH foram revelados 15 DNAs satélites W-específicos em *T. auritus*, ocorrendo uma acumulação de DNAs satélites neste cromossomo. O FISH comparativo entre as três espécies apresentou a trajetória do cromossomo sexual para este grupo, visto que esta família possui uma origem em comum em seus cromossomos sexuais ZW.

Apesar do crescente interesse em satelitomas de peixes, estudos comparativos ainda são escassos, estando presentes apenas na família Characidae (Goes et al., 2022) e Serrasalminidae (Goes et al., 2023), além da descrição do CharSat01-52, um satélite conservado para toda a ordem Characiformes (dos Santos et al., 2021). A análise comparativa entre três satelitomas de peixes da família Characidae (*P. fasciatus*, *P. bockmanni* e *A. lacustris*), em conjunto ao já publicado satelitoma de *P. paranae* demonstrou uma alta similaridade entre as bibliotecas de DNAs satélites entre esses dois gêneros, com 10 satélites presentes em todas as espécies (Goes et al., 2022). Além do reforço à hipótese *library*, esse trabalho acrescentou informações acerca do DNA satélite As51, que representa o mais abundante DNA satélite em *P. paranae*, *P. fasciatus* e *P. bockmanni*, sendo possível verificar a origem desta variante anteriormente à diferenciação de *Psalidodon* e *Astyanax*, e posterior surgimento de novas variantes em *Psalidodon* (Goes et al., 2022). Recentemente, o estudo comparativo entre os catálogos de *Colossoma macropomum* e *Piaractus mesopotamicus* demonstrou o compartilhamento de 21 satDNAs entre esses dois gêneros, demonstrando que a baixa diferenciação de satDNAs nesse grupo pode se relacionar a estabilidade cariotípica observada nesses peixes (Goes et al., 2023). Além disso, esse trabalho também evidenciou o compartilhamento de sequências consensos de 10 satDNAs descritos em Serrasalminidae com pelo menos outra família de Characiformes, expandindo o conhecimento sobre a manutenção de satDNAs, em especial os curtos, ao longo de grupos divergidos há muito tempo (Goes et al., 2023).

Como demonstrado, a descrição de novos catálogos de satDNAs para peixes neotropicais pode trazer novas informações relacionadas a diversificação de cariótipos e cariomorfos, surgimento de cromossomos B e evolução de sistemas sexuais. Em especial, a comparação de catálogos, além de permitir o melhor entendimento da evolução dessas sequências, revela importantes informações sobre a história evolutiva dos grupos. Nesse sentido, Characiformes ainda guarda famílias que apresentam grande interesse evolutivo. Uma das mais representativas nessa questão é a família Erythrinidae, que apresenta um vasto histórico de estudos citogenéticos e ainda não avaliada em relação ao seu conteúdo ou compartilhamento de satDNAs entre seus gêneros.

1.5 Família Erythrinidae e sua importância para citogenômica

A família Erythrinidae é composta por apenas três gêneros: *Erythrinus* (Scopoli, 1777), *Hoplias* (Gill, 1903) e *Hoplerythrinus* (Gill, 1895), com cerca de 20 espécies descritas (Eschemeyer e Fricke, 2023). Apesar de conter um número de espécies reduzido, os indivíduos desse grupo apresentam ampla distribuição por toda a América do Sul, fato que, em conjunto ao hábito de vida sedentário desses animais, leva a existência de diversas barreiras para o fluxo gênico dessas populações (Reis et al., 2003). Essa condição é particularmente importante para os peixes da família Erythrinidae, que demonstram diferenciação cromossômica específicas de algumas populações (Bertollo et al., 2000; Cioffi et al., 2012). Conhecidos popularmente como “Traíras”, esse grupo é foco de estudos citogenéticos a mais de 40 anos, devido a presença de grande variabilidade cariotípica, variando entre táxons com diversos cariomorfos diferentes e outros com estruturas cariotípica conservadas (Bertollo, 1978, 2000, 2007; Santos et al., 2009; Blanco et al., 2010), além da ampla gama de sistemas sexuais presentes nessa família, sendo estes bem delimitados ou comprovadamente em estágios iniciais de diferenciação (Bertollo et al., 2000, 2007; Cioffi e Bertollo, 2010; Cioffi et al., 2012, 2017; Martinez et al., 2016; Sember et al., 2018).

Especificamente no gênero *Hoplias*, *H. malabaricus* é uma das espécies mais estudadas do ponto de vista citogenético, devido a sua significativa variabilidade cariotípica, com presença de sete cariomorfos distintos (A-G) (Bertollo et al., 2000; Cioffi et al., 2012). A presença de diferentes sistemas sexuais é fundamental para a descrição desses cariomorfos, com $2n = 42$ para o cariomorfo A, sem a presença de um sistema sexual definido (Sember et al., 2018); $2n = 42$ para o cariomorfo B, com sistema sexual XX/XY (Fenocchio et al., 2000; Cioffi et al., 2010); $2n = 40$ para o cariomorfo C, com sistema sexual XY, apesar desses cromossomos não estarem diferenciados morfológicamente (Cioffi e Bertollo, 2010); $2n = 40$ em fêmeas do cariomorfo D, enquanto os machos desse cariomorfo apresentam $2n = 39$, devido a um sistema sexual X_1X_2Y (Bertollo et al., 1997; Cioffi e Bertollo, 2010); $2n = 42$ para o cariomorfo E (Bertollo et al., 2000); $2n = 40$ para o cariomorfo F, com um sistema XY recente (Freitas et al., 2018) e $2n = 40$ para fêmeas do cariomorfo G, enquanto os machos apresentam $2n = 41$, devido a um sistema sexual XY_1Y_2 (Oliveira et al., 2018). Apesar de evidências apontarem uma relação próxima entre os cariomorfos E, F e G (Bertollo et al., 2000), uma gama de análises citogenéticas e moleculares, incluindo DNA *barcode*, levam a hipótese de *H. malabaricus* representar um “complexo de espécies” (Santos et al., 2009; Cavallaro et al., 2000; Cioffi et al., 2012; Marques et al., 2013). Reforçando essa hipótese, apesar da existência de cariomorfos diferentes

vivendo em simpatria, híbridos entre estes não são observados (Dergam et al., 1998, 2002), com a exceção esporádica de um caso de hibridização seguido de aumento da ploidia (Utsunomia et al., 2014).

A mesma variação cromossômica pode ser encontrada em espécies dos gêneros restantes de Erythrinidae, como *Erythrinus erythrinus* e *Hoplerythrinus unitaeniatus*. Nestes casos, existe a variação de número diploide entre $2n = 48$ a $2n = 52$ em *H. unitaeniatus* (Giuliano-Caetano et al., 2001; Diniz e Bertollo, 2003; Rosa et al., 2012) e variação entre $2n = 54$ a $2n = 52$ em *E. erythrinus*, devido a presença de 4 cariomorfos nessa espécie (A, B, C e D), com sistemas sexuais $X_1X_1X_2X_2/X_1X_2Y$ (Bertollo et al., 2004; Cioffi et al., 2010; Martins et al., 2013). Devido ao modo de vida desses animais, que leva a existência de grande variabilidade cariotípica nos três gêneros de Erythrinidae, essa família representa um marco para o estudo da citogenética de peixes neotropicais. Sendo assim, este grupo apresenta grande potencial para o estudo de satelitomas, levando a um maior entendimento da manutenção de satDNAs ao longo de espécies e cariomorfos tão variáveis, resultado da dinâmica evolutiva desses animais

2 OBJETIVOS

Diante do proposto e tendo em vista a relevância crescente de análises comparativas relacionadas a sequências de DNA não codificantes, além da recente caracterização do satelitoma de *H. malabaricus*, cariomorfo D (Toma et al., in prep.), o presente trabalho teve como objetivo aumentar o conhecimento da biologia estrutural e evolutiva das famílias de DNA satélites dentro da família Erythrinidae, através da caracterização de cinco novos satelitomas, representantes dos gêneros *Hoplias*, *Hoplerythrinus* e *Erythrinus*, preenchendo as lacunas existentes para a comparação completa de todos os gêneros deste grupo e melhor entendimento da dinâmica evolutiva dessas sequências e seu papel na evolução cromossômica desses animais. Para isso, pretendemos:

➔ Caracterizar cinco satelitomas adicionais de indivíduos do gênero *Erythrinus* (*E. erythrinus* cariomorfo A), *Hoplerythrinus* (*H. unitaeniatus*) e *Hoplias* (*Hoplias intermedius*, *H. malabaricus* cariomorfos F e G), visando produzir um panorama mais completo da abundância e riqueza de DNAs satélites nos gêneros da família Erythrinidae.

➔ Caracterizar o nível de similaridade através da comparação dos catálogos de DNAs satélites produzidos neste trabalho em conjunto ao catálogo produzido para a espécie *H.*

malabaricus (cariomorfo D), visando diagnosticar o nível de similaridade entre os catálogos de indivíduos de diferentes espécies e gêneros dentro da família Erythrinidae.

→ Identificar, especificamente em *H. malabaricus*, possíveis satélites com abundâncias diferenciais em cada cariomorfo analisado (D, F e G), visando identificar um provável papel dos DNAs satélites na diversificação dos cariomorfos observados nessa espécie.

→ Mapear a localização cromossômica dos satélites que eventualmente estejam conservados para diferentes espécies ou cariomorfos, visando diagnosticar uma possível correlação entre manutenção dessas sequências e localização cromossômica.

3 MATERIAIS E MÉTODOS

3.1 Material

No presente trabalho foram analisados indivíduos representantes dos três gêneros da família Erythrinidae. Foram coletados indivíduos de *E. erythrinus* (cariomorfo A), *H. unitaeniatus*, *H. intermedius* e *H. malabaricus* (cariomorfos F e G). Todas as mostras foram cedidas pela Universidade Federal de São Carlos, através do Prof. Dr. Marcelo de Bello Cioffi.

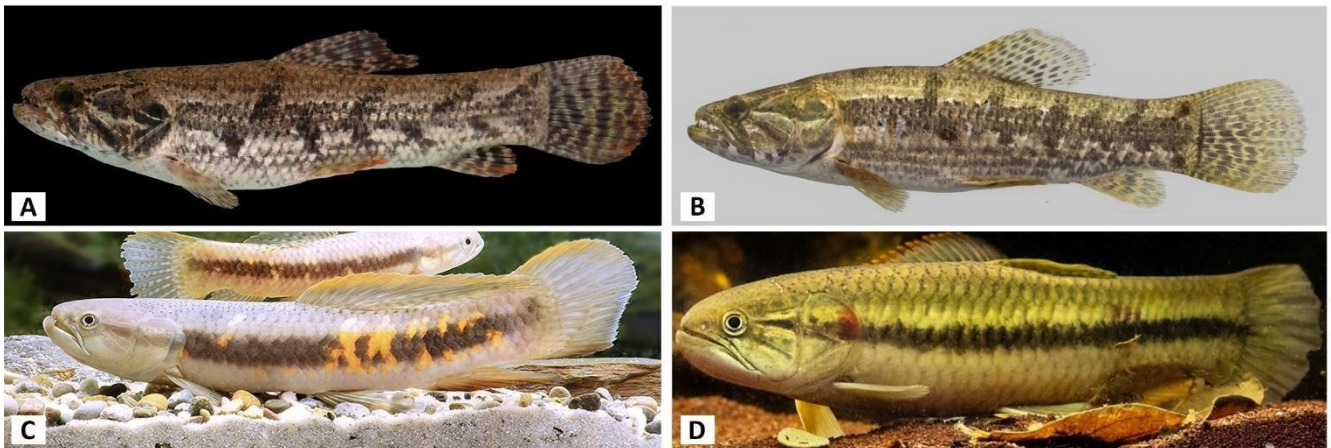


Figura 1: Espécies da família Erythrinidae analisadas, *H. intermedius* (A), *H. malabaricus* (B), *E. erythrinus* (C) e *H. unitaeniatus* (D). Fontes: FishBase (<https://www.fishbase.se/summary/FamilySummary.php?ID=103>) e Encyclopedia of Life (<https://eol.org/pages/6903>).

3.2 Métodos

3.2.1 Extração do DNA genômico total e sequenciamento

A extração de DNA genômico (gDNA) total seguiu o método de fenol-clorofórmio-álcool isoamílico, estabelecido por (Russell e Sambrook, 2001). Para isso, foram utilizados fragmentos de fígado de *H. malabaricus* (F) (macho), *H. malabaricus* (G) (fêmea), *H. intermedius* (macho), *H. unitaeniatus* (macho) e *E. erythrinus* (A) (macho). Os DNAs obtidos passaram por checagem de sua integridade em gel de agarose 1%. Posteriormente, os DNAs obtidos foram enviados para o sequenciamento de *short reads* (2x 150bp paired-end) na plataforma BGISEQ-500, da BGI (BGI Shenzhen Corporation, Shenzhen, China).

3.2.2 Caracterização dos satelitomas por TAREAN

A caracterização dos satelitomas de *H. malabaricus* (F), *H. malabaricus* (G), *H. intermedius*, *H. unitaeniatus* e *E. erythrinus* foram realizadas utilizando o software TAREAN (Novák et al., 2017). Todas as bibliotecas utilizadas passaram por uma etapa de filtragem de qualidade e correção nos tamanhos dos *reads*, utilizando o software Trimmomatic (Bolger et al., 2014). Após essa etapa, foi realizada uma seleção aleatória de 2x 500.000 *reads*, para a clusterização no TAREAN. Posteriormente, as sequências apontadas como sendo em *tandem* pelo *output* do TAREAN foram filtradas da biblioteca original, utilizando o software DeconSeq (Schmieder e Edwards, 2011). Após a retirada das sequências já identificadas pelo TAREAN, uma nova amostragem de 2x 500.000 *reads* foi realizada, com posterior clusterização pelo TAREAN. Cada ciclo que resulta em um *output* do TAREAN com satélites putativos é denominado iteração, e esse processo é repetido até que nenhuma sequência em *tandem* seja identificada pelo TAREAN.

3.2.3 Filtragem de famílias multigênicas e eliminação de redundâncias

Após a finalização das iterações no TAREAN, foi realizada uma filtragem buscando famílias multigênicas que porventura possam ter sido caracterizadas pela metodologia. Para isso, foram utilizadas sequências das famílias multigênicas 5S rDNA, 18S rDNA, 25s rDNA, Histona H1 e genes U1, U2, U4, U5 e U6, utilizando o software Geneious v. 8.1 e qualquer sequência similar as famílias multigênicas citadas foram descartadas. A seguir, foi realizada uma análise comparativa dentro de

cada catálogo obtido pelo TAREAN visando verificar a similaridade entre suas próprias sequências, buscando possíveis redundâncias. Para isso, foi utilizado o software RepeatMasker (https://github.com/fjruizruano/satminer/blob/master/rm_homology.py) seguindo os padrões estabelecidos por (Ruiz-Ruano et al., 2016), com sequências com acima de 95% de similaridade sendo consideradas as mesmas variantes de um satélite, similaridade entre 80% e 95% consideradas variantes de um mesmo satélite e similaridade entre 50% e 80% consideradas superfamílias.

3.2.4 Caracterização de abundância e divergência dos satDNAs.

Após a caracterização pelo TAREAN e a devida filtragem para famílias multigênicas e redundâncias, foi utilizado o *software RepeatMasker* (Smit et al., 2017) utilizando a ferramenta “*cross_match*”, para o estudo da abundância e divergência de cada um dos satDNAs encontrados nos novos satelitomas. Para isso, foi utilizado uma amostra de 10 milhões de *reads* de cada biblioteca e o resultado da abundância foi calculado através de uma proporção entre a quantidade de *reads* mapeados de cada satélite dividido pelo total de nucleotídeos utilizado na análise. Com esse resultado, os satDNAs foram classificados em ordem decrescente de abundância, e renomeados com uma abreviação para cada espécie (Hin para *H. intermedius*, Hun para *H. unitaeniatus* e Eer para *E. erythrinus* A) em adição ao termo “Sat” e o número do catálogo, em ordem decrescente de abundância, além de seu tamanho em pares de base (HmfSat01-139, por exemplo) como proposto por Ruiz-Ruano et al., 2016. Para os catálogos descritos para os cariomorfos de *Hoplias*, houve uma adaptação da regra de nomenclatura visando diferenciar os catálogos. Os nomes formados pela primeira letra do gênero, em maiúsculo, seguido da primeira letra epíteto específico em conjunto a letra do cariomorfo específico (Hmf para *H. malabaricus* F, Hmg para *H. malabaricus* G,). Também utilizando o *RepeatMasker*, a divergência média das sequências foi calculada utilizando a distância entre essas sequências. Para isso, foi utilizado o parâmetro Kimura-2, tendo como base o *script calcDivergenceFromAlign.py*, pertencente ao *RepeatMasker* (Smit et al., 2017).

3.2.5 Comparação entre os satelitomas descritos

De forma semelhante a busca por redundâncias nos catálogos descritos, realizamos uma comparação global de todos os satélites descritos em *H. malabaricus* (F), *H. malabaricus* (G), *H. intermedius*, *H. unitaeniatus* e *E. erythrinus* (A), em conjunto com o satelitoma de *H. malabaricus* (D)

(Toma et al., *in. prep.*). Para isso, também utilizamos o software RM_homology (https://github.com/fjruiaruano/satminer/blob/master/rm_homology.py), com uma etapa adicional de alinhamento utilizando o software Geneius 8.1, utilizando a opção *de novo assemble*. Os critérios para a confirmação do compartilhamento de um satDNA entre dois satelitomas foi o mesmo descrito por (Ruiz-Ruano et al., 2016), sendo considerados satDNAs compartilhados aqueles que estabelecem no mínimo uma relação de superfamília (acima de 50% de similaridade).

3.2.6 Seleção dos satDNAs para FISH e desenho de primers

A seleção dos satDNAs para o mapeamento cromossômico seguiu alguns critérios, sendo o primeiro deles a exclusão de satDNAs exclusivos de apenas um grupo. A partir daí, selecionamos satDNAs para a FISH de acordo com cinco categorias: 1- satDNAs compartilhados entre os seis grupos; 2- satDNAs compartilhados apenas entre os três cariomorfos de *Hoplias*; 3- satDNAs compartilhados entre todos os catálogos do gênero *Hoplias*; 4- satDNAs compartilhados entre *H. unitaeniatus* e *E. erythrinus*.

O desenho de *primers* foi realizado utilizando o *software* Geneious 8.1, seguindo critérios relacionados ao tamanho do satDNA. Para satélites com menos de 600bp foram desenhados *primers* divergentes, enquanto para satDNAs maiores de 600bp foram desenhados primers convergentes (Figura 2). A qualidade dos primers foi checada utilizando o *Multiple Primer Analyzer* (ThermoFisher Scientific), visando evitar o auto pareamento e formação de grampos.

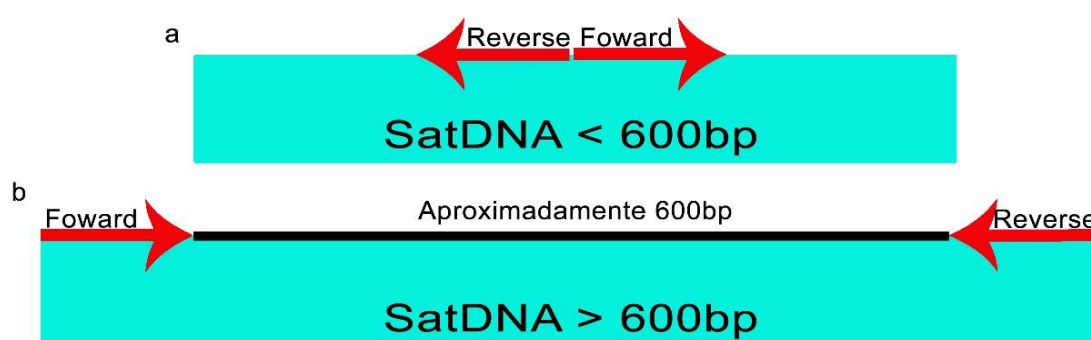


Figura 2: Confeção de primers divergentes para satélites menores de 600bp (a) e convergentes para satélites maiores de 600bp (b).

3.2.7 Obtenção de cromossomos mitóticos

A obtenção de cromossomos mitóticos seguiu os procedimentos estabelecidos por Bertollo et al., 2015 e Foresti et al., 1981. Previamente, foi realizado o estímulo a mitose, com aplicação de fermento biológico na musculatura dorsolateral dos animais, seguindo o protocolo estabelecido por Oliveira et al., 1988. Após 48 horas, os animais foram submetidos ao tratamento com Colchicina (1 ml de solução a 0,005%/ 100 gramas de peso), por 45 minutos. Após anestesiados, os animais foram sacrificados e fragmentos do rim anterior foram retirados, visando obter uma suspensão celular e hipotonização das células, em solução de cloreto de Potássio (KCL) 0,075M durante 21 minutos a 36°C. A fixação do material foi feita utilizando solução de metanol (3 partes) e ácido acético (1 parte).

3.2.8 PCRs e marcação das sondas

As reações de amplificação em cadeia da polimerase (PCR) foram realizadas para cada um dos satDNAs, a partir do DNA genômico de *H. malabaricus* (F) e *E. erythrinus* (A), devido a maioria dos primers desenhados servirem para a todas as espécies que compartilhavam os satDNAs. Diferentes diluições de DNAs foram utilizadas (10%, 1% e 0,1%), e as reações seguiram os seguintes parâmetros: 1x PCR Buffer, 1,5 mM de MgCl₂, 200 µM de dNTP, 0,1 µM de Primer A + Primer B, 0,1U de Taq polimerase e 100ng de DNA genômico, em uma reação total de 12,5 µl. A marcação das sondas foi realizada com kit de Nick-Translation mix (Roche, Manheim, Alemanha), utilizando Atto550-dUTP (vermelho), Atto488-dUTP (verde), Atto425-dUTP (ciano) e Atto680-dUTP (amarelo).

Devido ao alto número de satélites, as reações em termociclador variaram de acordo com os parâmetros exigidos para cada DNA satélite. Todas as amplificações foram checadas em gel de agarose 2%. Um resumo das reações aplicadas encontra-se no esquema abaixo:

95°C → 5min

*95°C → 20s

*54°C – 58°C → 15s

*72°C → 15 – 36s

72°C → 14min.

* Repetidos em ciclos por 30 vezes

3.2.9 Hibridização *in situ* fluorescente (FISH)

O procedimento de hibridação seguiu o protocolo de alta adstringência (Yano et al., 2017; Sassi et al., 2022). O tratamento inicial das lâminas envelhecidas a temperatura ambiente *overnight* ou a 60 °C por 1h, foi realizado utilizando 100 µl de solução de RNase (10 µg/ml) em câmara úmida a 37°C por 1h. Após esta etapa, uma lavagem em PBS 1x por agitação foi realizada, com o subsequente tratamento em 100 µl de solução de Pepsina (495 µl de H₂O miliq autoclavada + 5 µl HCl 1M + 1,5 µl Pepsina (20mg/µl)). Após essa etapa, uma nova lavagem com PBS 1x foi realizada, seguida de uma lavagem em série alcoólica, utilizando álcool a 70%, 85% e 100%, por dois minutos em cada. As lâminas foram desnaturadas utilizando o tratamento com Formamida 70%, por 3min e 15s a 72°C, com posterior lavagem em série alcoólica como anteriormente, com o álcool 70% gelado. O mix de hibridização conteve 2 µl de cada sonda desejada, além de uma quantidade de *hybridization buffer* (2g sulfatodextrano em 10 ml de 50% formamida em 2xSSC/50 mM), completando 20 µl de solução, desnaturada em termociclador a 86°C por 10 min e resfriado a 4 °C por pelo menos 2 min antes da aplicação. Após a aplicação do mix de hibridação em cada lâmina, estas foram incubadas a 37°C em câmara úmida, por um período *overnight*. Após isso, as lâminas foram lavadas em 1xSSC a 65°C e solução de 4xSSC/Twen (4xSSC, 0,05% Tween 20) a temperatura ambiente, ambos por 5 minutos em agitação. Finalmente, após uma lavagem em 1xPBS por um minuto e desidratação em série alcóolica, as lâminas foram coradas com 20 µl de DAPI + Antifading.

3.2.10 Produção de *Minimum spanning trees* (MST).

Visando analisar os padrões de abundância e divergência de haplótipos, satélites conservados para os seis grupos analisados foram selecionados para construção de *minimum spaning trees* (MSTs). Devido a limitações técnicas, essa análise incluiu apenas satélites com sequências consensus menores do que 150 pb. Para isso, foram realizadas amostragens aleatórias de 2 x 500.000 reads das bibliotecas genômicas de cada grupo, para a filtragem de monômeros de cada satDNA selecionado, com a exceção de satDNA pouco abundantes (HmfSat45-56 e Hmf52-49), em que 2 x 1.000.000 de reads foram utilizados na amostragem aleatória. Após a obtenção de reads correspondentes a monômeros dos satDNAs selecionados, foi realizado o alinhamento destes com dímeros dos respectivos satDNAs, e a região central do monômero correspondente foi extraída para as análises seguintes. Após isso, os *singletons* foram retirados, utilizando o software CD-HIT (Fu et al., 2012), e

os monômeros restantes foram utilizados para a construção das MSTs, utilizando o software PHYLOVIZ (Nascimento et al., 2016).

4 RESULTADOS

4.1 Análises citogenéticas

Indivíduos de *H. malabaricus* (F), *H. malabaricus* (G), *H. intermedius*, *H. unitaeniatus* e *E. erythrinus* (A) foram amostrados e submetidos aos métodos de obtenção de cromossomos metafásicos *in vivo*. As amostras utilizadas foram selecionadas independente do sexo. Os números diploides obtidos foram: $2n = 40$ para *H. malabaricus* (F), $2n = 40$ para *H. malabaricus* (G), $2n = 50$ para *H. intermedius*, $2n = 52$ para *H. unitaeniatus* e $2n = 54$ para *E. erythrinus* (D) (Figura 3).

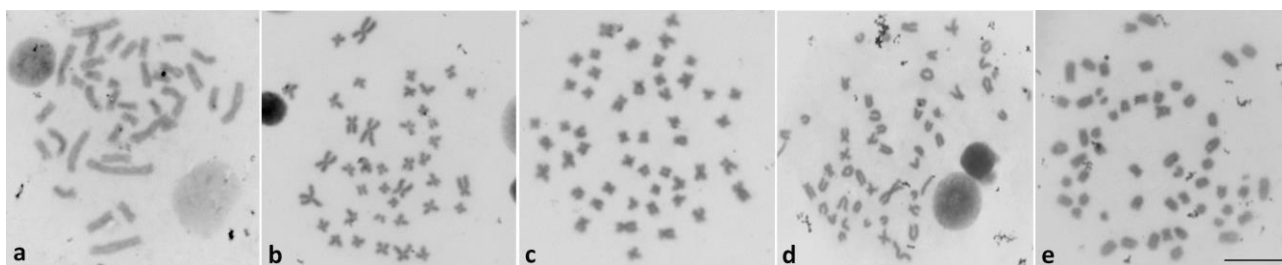


Figura 3: Metáfases dos cariomorfos de *H. malabaricus* (F) (a) e *H. malabaricus* (G) (b), *H. intermedius* (c), *H. unitaeniatus* (d), *E. erythrinus* (e). Barra = 10 μ m

4.2 Descrição de cinco novos satelitomas para a família Erythrinidae

A caracterização dos satelitomas propostos seguiu a pipeline desenvolvida para o software TAREAN, utilizando as opções padrão. As iterações foram realizadas até que nenhum satDNA fosse recuperado, sendo realizadas três iterações para *H. malabaricus* (F), *H. malabaricus* (G), *H. intermedius* e *H. unitaeniatus*, e quatro iterações para *E. erythrinus* (A). Sendo assim, foram caracterizadas 56 famílias de satDNAs para *H. malabaricus* (F), 45 famílias de satDNAs para *H. malabaricus* (G), 73 famílias de satDNAs para *H. intermedius*, 40 famílias de satDNAs para *H.*

unitaeniatus e 49 famílias de satDNAs para *E. erythrinus*. As características gerais dos satelitomas caracterizados encontram-se nas tabelas 2 a 6.

Tabela 2: Características gerais do satelitoma de *H. malabaricus* (F).

SatDNA	RUL	A+T (%)	Abundância	Divergência (%)
HmfSat01-139	139	65	0.009206394	10.56
HmfSat02-1894	1894	65	0.008210824	13.51
HmfSat03-46	46	44.4	0.008194474	8.86
HmfSat04-513	513	60.6	0.006027545	3.78
HmfSat05-2936	2936	55	0.00556974	8.13
HmfSat06-453	453	63.4	0.002245702	3.02
HmfSat07-149	149	50	0.001863612	9.13
HmfSat08-2944	2944	55	0.001419167	8.32
HmfSat09-31	31	45.2	0.001339327	10.01
HmfSat10-28	28	53.6	0.001292539	9.4
HmfSat11-1922	1922	60	0.001099455	6.6
HmfSat12-58	58	51.7	0.001007309	2.89
HmfSat13-212	212	58	0.000827472	6.09
HmfSat14-49	49	55.6	0.000759121	5.64
HmfSat15-1192	1192	60	0.000754676	6.07
HmfSat16-702	702	55	0.000701037	1.57
HmfSat17-292	292	62.7	0.000697888	2.61
HmfSat18-84	84	60	0.000564623	14.43
HmfSat19-719	719	60	0.000530547	1.45
HmfSat20-42	42	52.4	0.000526884	5.64
HmfSat21-42	42	50	0.000501426	18.34
HmfSat22-34	34	47.1	0.000488391	13.02
HmfSat23-707	707	48.4	0.000438815	0.81
HmfSat24-217	217	60	0.000424977	7.49
HmfSat25-941	941	55	0.000422212	0.91
HmfSat26-378	378	55	0.000396309	12.49
HmfSat27-53	53	64.2	0.000394314	5.28
HmfSat28-142	142	66.9	0.000391119	7.17
HmfSat29-141	141	60	0.000386276	6.59
HmfSat30-684	684	50.1	0.000366284	1.63
HmfSat31-93	93	55	0.000311843	5.63
HmfSat32-403	403	50	0.00028446	4.25

HmfSat33-177	177	55	0.000257428	13.84
HmfSat34-636	636	60	0.000240682	4.28
HmfSat35-1313	1313	55	0.000239686	3.19
HmfSat36-168	168	60	0.00022959	7.96
HmfSat37-841	841	55	0.000222987	10.46
HmfSat38-1394	1394	55	0.00022106	1.29
HmfSat39-57	57	50.9	0.000208419	1.71
HmfSat40-457	457	49.7	0.00020737	6.92
HmfSat41-177	177	60	0.00020577	11
HmfSat42-192	192	60	0.000178224	5
HmfSat43-129	129	50	0.000169116	5.61
HmfSat44-186	186	55	0.00016579	2.09
HmfSat45-56	56	53.6	0.000163946	4.67
HmfSat46-50	50	58	0.000154555	1.45
HmfSat47-326	326	60	0.000149251	2.86
HmfSat48-42	42	54.8	0.000137447	4.73
HmfSat49-322	322	60	0.000136745	5.57
HmfSat50-199	199	58.8	0.000125397	4.82
HmfSat51-83	83	50	0.000115893	1.86
HmfSat52-49	49	44.4	0.00011306	6.2
HmfSat53-41	41	61.1	0.000111406	4.78
HmfSat54-35	35	48.6	9.17E-05	4.57
HmfSat55-47	47	51.1	8.07E-05	7.8
HmfSat56-126	126	52.4	7.03E-05	9.66

Tabela 3: Características gerais do satelitoma de *H. malabaricus* (G).

SatDNA	RUL	A+T (%)	Abundância	Divergência (%)
HmgSat01-58	58	50	0.027450575	2.9
HmgSat02-513	513	60.8	0.01049602	3.91
HmgSat03-46	46	47.8	0.006170622	9.51
HmgSat04-139	139	63.3	0.004969837	12.44
HmgSat05-63	63	38.9	0.002287971	6.81
HmgSat06-260	260	46.5	0.002229842	11.36

HmgSat07-212	212	58	0.001908067	6.47
HmgSat08-152	152	63.2	0.001585596	15.85
HmgSat09-1240	1240	59	0.000874338	10.46
HmgSat10-705	705	61.7	0.000862115	16.09
HmgSat11-49	49	46.9	0.00076433	5.22
HmgSat12-292	292	62.7	0.000702072	2.32
HmgSat13-58	58	51.7	0.000687842	2.91
HmgSat14-31	31	41.9	0.000645255	9.18
HmgSat15-1140	1140	48.4	0.000576824	2.25
HmgSat16-696	696	48.6	0.000564546	1.07
HmgSat17-719	719	51.5	0.00054403	1.32
HmgSat18-84	84	53.6	0.000504293	15.29
HmgSat19-141	141	65.2	0.000461919	5.91
HmgSat20-42	42	54.8	0.00045691	18.28
HmgSat21-206	206	60	0.000445914	7.04
HmgSat22-941	941	63.9	0.00035215	0.73
HmgSat23-168	168	70.8	0.000329272	7.34
HmgSat24-42	42	52.4	0.000304269	5.77
HmgSat25-322	322	74.8	0.00029605	4.37
HmgSat26-620	620	55.6	0.00027497	2.9
HmgSat27-1380	1380	45.3	0.000271407	0.74
HmgSat28-1312	1312	52.7	0.000270847	1.42
HmgSat29-684	684	50.1	0.000265859	1.06
HmgSat30-27	27	44.4	0.000248715	10.86
HmgSat31-28	28	53.6	0.000235337	8.3
HmgSat32-827	827	55	0.00020717	2.07
HmgSat33-177	177	71.2	0.000203107	11.09
HmgSat34-403	403	57.8	0.000191261	5.93
HmgSat35-192	192	64.1	0.000186896	6.01

HmgSat36-177	177	65.5	0.000173289	16.05
HmgSat37-467	467	44.4	0.000173259	9.41
HmgSat38-142	142	67.6	0.0001542	8.01
HmgSat39-34	34	47.1	0.000153018	7.88
HmgSat40-49	49	49	0.000144479	5.82
HmgSat41-21	21	71.4	0.000140705	6.76
HmgSat42-23	23	43.5	0.000133235	14.83
HmgSat43-142	142	60	0.000129129	5.96
HmgSat44-56	56	53.6	0.000115051	6.17
HmgSat45-588	588	56.1	0.000114585	12.67

Tabela 4: Características gerais do satelitoma de *H. intermedius*.

SatDNA	RUL	A+T (%)	Abundância	Divergência (%)
HinSat01-294	294	62.6	0.042113345	6.27
HinSat02-193	193	60.6	0.028559905	5.8
HinSat03-138	138	28.3	0.009379041	6.25
HinSat04-152	152	62.3	0.00815127	9.72
HinSat05-174	174	70.1	0.005576086	9.31
HinSat06-66	66	40.9	0.004948455	8.54
HinSat07-159	159	27.7	0.004870451	5.32
HinSat08-152	152	65.8	0.003852308	10.25
HinSat09-38	38	47.4	0.002464525	11.44
HinSat10-305	305	64.9	0.002305419	7.44
HinSat11-3739	3739	43.1	0.002278421	6.07
HinSat12-267	267	60.3	0.002039765	5.8
HinSat13-69	69	55.1	0.001677834	8.97
HinSat14-33	33	30.3	0.00128553	15.93
HinSat15-1034	1034	55	0.001243605	2.22
HinSat16-73	73	60.3	0.00121805	6.13
HinSat17-153	153	62.7	0.001115466	9.43
HinSat18-505	505	44.2	0.001055075	9.76
HinSat19-58	58	53.4	0.000984144	1.97
HinSat20-1284	1284	55	0.000808446	5.76
HinSat21-1856	1856	60.8	0.000791266	10.22

HinSat22-294	294	63.9	0.000766709	5.3
HinSat23-923	923	60	0.000703258	1.45
HinSat24-1122	1122	55.9	0.000546758	2.76
HinSat25-49	49	46.9	0.000539588	6.29
HinSat26-152	152	67.1	0.000539559	7.6
HinSat27-1773	1773	58.4	0.000520975	8.14
HinSat28-34	34	55.9	0.000509009	16.39
HinSat29-1013	1013	49.8	0.000452142	1.02
HinSat30-1591	1591	61.2	0.00043092	0.56
HinSat31-906	906	55	0.000424033	1.44
HinSat32-84	84	53.6	0.000408506	17.24
HinSat33-296	296	64.9	0.00038673	7.67
HinSat34-765	765	55	0.00035709	1.04
HinSat35-27	27	55.6	0.000353442	5.76
HinSat36-21	21	52.4	0.000351965	6.56
HinSat37-30	30	60	0.000347533	8.31
HinSat38-1972	1972	60	0.000345755	2.4
HinSat39-405	405	56.5	0.000312836	4.71
HinSat40-42	42	50	0.000302744	14.03
HinSat41-50	50	64	0.000285113	2.95
HinSat42-168	168	70.2	0.000284341	8.38
HinSat43-1062	1062	54.7	0.00028281	0.71
HinSat44-686	686	49.3	0.000281238	0.81
HinSat45-598	598	56	0.000267935	4.39
HinSat46-885	885	51.4	0.000262126	5.73
HinSat47-1393	1393	45.8	0.000262046	0.63
HinSat48-54	54	48.1	0.000253195	8.09
HinSat49-273	273	57.9	0.000245606	7.27
HinSat50-175	175	69.1	0.000224826	11.24
HinSat51-30	30	46.7	0.000212737	2.37
HinSat52-42	42	50	0.00019782	3.94
HinSat53-106	106	58.5	0.000196661	1.98
HinSat54-685	685	52.6	0.000193752	7.57
HinSat55-61	61	65.6	0.000180517	16.86
HinSat56-42	42	50	0.000178757	8.67
HinSat57-815	815	67.6	0.000175168	4.14
HinSat58-378	378	55.3	0.000174956	4.81
HinSat59-219	219	53.4	0.000170101	6.71
HinSat60-1510	1510	60	0.000163511	1.16

HinSat61-20	20	55	0.000163015	17.48
HinSat62-51	51	66.7	0.000157135	10.23
HinSat63-512	512	60.7	0.000128542	4.82
HinSat64-143	143	71.3	0.000123427	6.23
HinSat65-47	47	59.6	0.000122971	3.51
HinSat66-56	56	53.6	0.000121791	5.12
HinSat67-60	60	66.7	0.000113381	6.33
HinSat68-49	49	44.9	0.000106068	3.41
HinSat69-41	41	61	0.000101313	3.01
HinSat70-50	50	58	9.63E-05	2.14
HinSat71-322	322	73.3	9.50E-05	10.78
HinSat72-21	21	47.6	9.01E-05	0.48
HinSat73-232	232	56.5	8.31E-05	7.99

Tabela 5: Características gerais do satelitoma de *H. unitaeniatus*.

SatDNA	RUL	A+T (%)	Abundância	Divergência (%)
HunSat01-130	130	68.5	0.007945853	4.2
HunSat02-141	141	61.7	0.005153603	8.04
HunSat03-85	85	60	0.005087716	10.12
HunSat04-45	45	38.9	0.004078272	2.71
HunSat05-44	44	47.7	0.00237464	10.96
HunSat06-86	86	48.8	0.001962165	6.95
HunSat07-33	33	63.6	0.00146289	4.15
HunSat08-51	51	64.7	0.000767729	3.86
HunSat09-43	43	55.8	0.000687789	4.14
HunSat10-177	177	64.4	0.000586401	11.46
HunSat11-49	49	49	0.000533146	6.57
HunSat12-41	41	48.8	0.00051128	11.25
HunSat13-56	56	55.4	0.00047154	4.46
HunSat14-31	31	41.9	0.000456787	8
HunSat15-48	48	52.1	0.000433383	6.06
HunSat16-820	820	53.5	0.000403238	15.63

HunSat17-923	923	53.4	0.000382169	4.25
HunSat18-1053	1053	55.2	0.000365779	5.11
HunSat19-958	958	64.4	0.000360095	2.38
HunSat20-42	42	52.4	0.000311904	4.63
HunSat21-700	700	51.9	0.000321884	9.11
HunSat22-62	62	54.8	0.000273588	2.15
HunSat23-152	152	61.8	0.00026705	18.15
HunSat24-92	92	66.3	0.000249978	7.74
HunSat25-653	653	55	0.000246734	8.77
HunSat26-43	43	58.1	0.000214082	3.29
HunSat27-1752	1752	49.7	0.000207505	4.33
HunSat28-23	23	52.2	0.000207017	5.88
HunSat29-764	764	47	0.000185375	1.5
HunSat30-1096	1096	60	0.000168512	2.17
HunSat31-31	31	48.4	0.000150241	5.61
HunSat32-427	427	56	0.000122164	6.05
HunSat33-49	49	51	0.00012124	6.11
HunSat34-98	98	50	0.000115088	2.1
HunSat35-214	214	52.3	0.000114939	7.71
HunSat36-142	142	71.8	0.000113669	10.68
HunSat37-32	32	50	0.000113562	7.74
HunSat38-63	63	50.8	0.00011342	1.38
HunSat39-369	369	61	8.96E-05	7.12
HunSat40-58	58	55.2	8.89E-05	10.22

Tabela 6: Características gerais do satelitoma de *E. erythrinus* (A).

SatDNA	RUL	A+T (%)	Abundância	Divergência (%)
EerSat01-242	242	56.2	0.009269251	6.78
EerSat02-130	130	68.5	0.00878756	4.28
EerSat03-141	141	61.7	0.006156787	7.59
EerSat04-45	45	44.4	0.0003664956	2.79
EerSat05-473	473	62.2	0.001704209	10.14
EerSat06-111	111	60	0.000128471	4.8
EerSat07-31	31	41.9	0.000937735	7.26
EerSat08-51	51	64.7	0.000935156	3.74
EerSat09-41	41	51.2	0.000610733	11.17
EerSat10-43	43	55.8	0.000596528	4.28
EerSat11-871	871	53.6	0.000586566	4.8
EerSat12-49	49	49	0.000533559	6.67
EerSat13-691	691	52.8	0.00053318	9.28
EerSat14-177	177	44.4	0.00053005	11.03
EerSat15-835	835	53.8	0.000484409	12.85
EerSat16-84	84	53.6	0.000446449	16.52
EerSat17-56	56	60	0.000438513	4.74
EerSat18-1053	1053	55.1	0.000403777	4.9
EerSat19-958	958	63.5	0.000396183	2.4
EerSat20-48	48	52.1	0.000395122	5.41
EerSat21-94	94	60	0.000342821	7.96
EerSat22-152	152	59.9	0.000324135	16.92
EerSat23-62	62	44.4	0.000307026	1.97

EerSat24-42	42	52.4	0.000279482	4.56
EerSat25-1767	1767	49.6	0.000262658	3.7
EerSat26-31	31	64.5	0.000238926	10.64
EerSat27-43	43	58.1	0.00023639	3.28
EerSat28-1019	1019	45.9	0.0002301	1.73
EerSat29-772	772	47	0.000197752	1.46
EerSat30-21	21	57.1	0.000182298	7.28
EerSat31-14	14	64.3	0.000172444	10.36
EerSat32-940	940	59.5	0.000164376	6.99
EerSat33-214	214	53.3	0.000155149	6.51
EerSat34-921	921	50	0.000155028	2.62
EerSat35-76	76	55	0.000147384	5.66
EerSat36-1270	1270	55	0.000141402	2.47
EerSat37-905	905	55	0.000141098	2.36
EerSat38-31	31	48.4	0.000134339	6.07
EerSat39-49	49	49	0.00013065	5.41
EerSat40-59	59	44.4	0.000126351	1.05
EerSat41-583	583	55	0.00012056	4.46
EerSat42-39	39	48.7	0.000119666	7.84
EerSat43-427	427	55.7	0.000115252	3.11
EerSat44-80	80	65	0.000114391	2.12
EerSat45-28	28	50	0.000109812	6.04
EerSat46-50	50	65	0.000109477	9.61
EerSat47-32	32	50	0.000104942	7.13
EerSat48-52	52	65	9.57E-05	2.79
EerSat49-389	389	61.7	7.09E-05	7.99

A maior variação de comprimento das unidades repetidas (RUL – *repeat unit length*) (Tabela 7) foi observada em *H. intermedius*, com o tamanho das unidades variando entre 3739bp e 20bp (mediana = 168), sendo o HinSat11-3739, presente em *H. intermedius* o maior satélite observado dentre os catálogos descritos. Para as demais espécies, a variação do tamanho das unidades observadas foram: 2944bp a 28bp para *H. malabaricus* (F) (mediana = 177), 1380bp a 21bp para *H. malabaricus* (G) (mediana = 177), 1752bp a 23bp para *H. unitaeniatus* (mediana = 85) e 1767bp a 14bp para *E. erythrinus* (mediana = 84), sendo o EerSat31-14 o menor satélite caracterizado para os catálogos descritos. Houve a predominância de satélites longos (> 100bp) para todos os catálogos descritos do gênero *Hoplias*, com esta fração das famílias de satDNAs representando 64,28% do satelitoma de *H. malabaricus* (F), 64,44% do satelitoma de *H. malabaricus* (G) e 61,64% do satelitoma de *H. intermedius*. Entretanto, observou-se um equilíbrio entre satDNAs longos e curtos em *H. unitaeniatus* e *E. erythrinus* (A), com satDNAs longos representando 57,51% e 53,06% destes catálogos, respectivamente.

A análise do conteúdo A+T demonstrou uma leve tendência a satDNAs ricos em A+T para os cinco catálogos descritos, com médias de 55.5 para *H. malabaricus* (F), 55.65 para *H. malabaricus* (G), 56.08 para *H. intermedius*, 55.08 para *H. unitaeniatus* e 54.96 para *E. erythrinus* (A). A maior variação foi observada em *H. intermedius*, com conteúdo A+T variando entre 27,7% entre 71,3%. Os dados encontram-se resumidos na tabela 7 e Figura 4.

Tabela 7: Médias e variações de conteúdo A+T dos satelitomas descritos.

Espécie	A+T (%) médio	Variação A+T (%)	Variação RUL
<i>H. malabaricus</i> (F)	55.5	44.4 – 66.9	28bp – 2944bp
<i>H. malabaricus</i> (G)	55.65	38.9 – 74.8	21bp – 1380bp
<i>H. intermedius</i>	56.08	27.7 – 71.3	20bp – 3739bp
<i>H. unitaeniatus</i>	55.08	38.9 – 71.8	23bp – 1752bp
<i>E. erythrinus</i> (A)	56.96	41.9 – 68.5	14bp – 1767bp

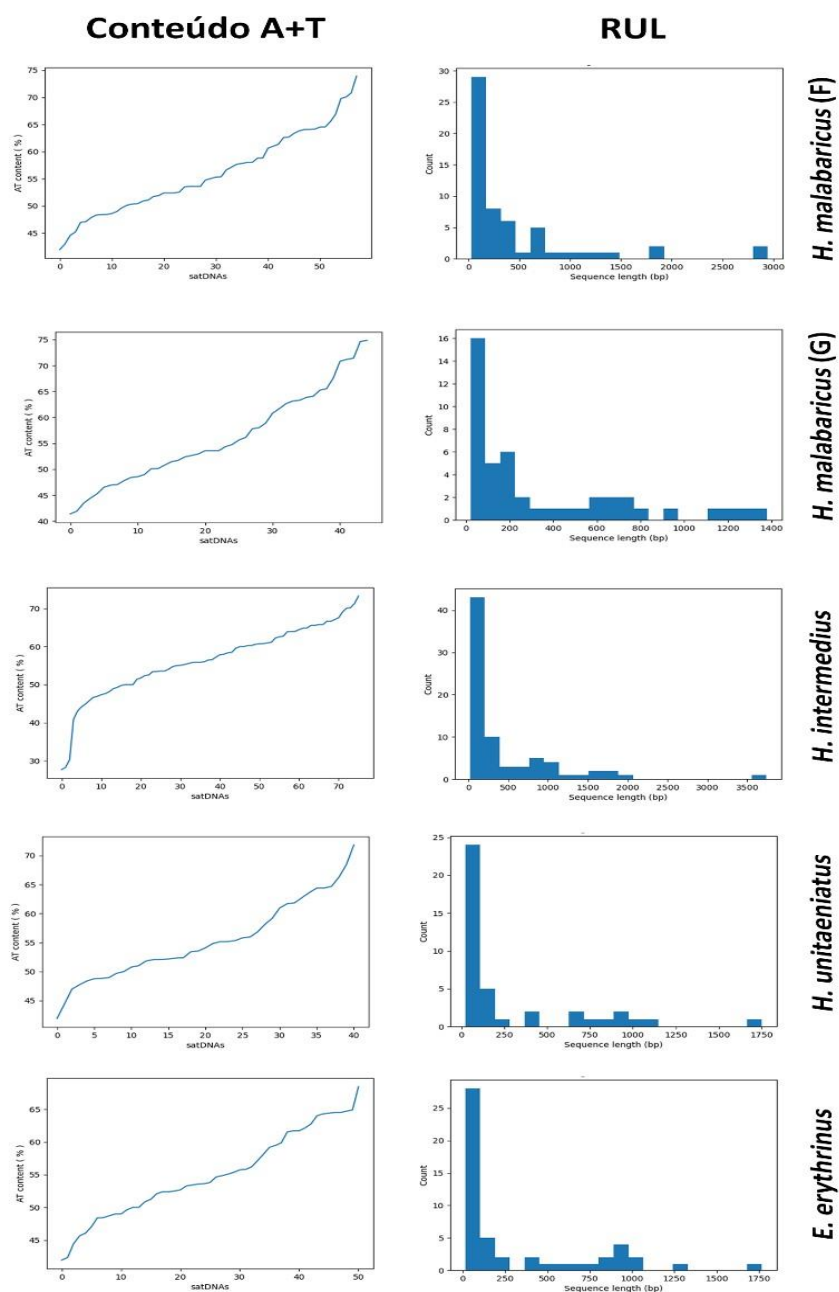


Figura 4: Representação gráfica da variação do conteúdo A+T e RUL dos satelitomas descritos neste trabalho. A esquerda, satDNAs, em ordem crescente de conteúdo A+T (eixo X) e conteúdo A+T em porcentagem (eixo Y). A direita, tamanho das sequências (eixo X) e contagem de satDNAs daquele tamanho (eixo Y).

As análises comparativas entre os satDNAs de cada catálogo revelaram a presença de superfamílias (SF) para todos os grupos. Foram descritas sete superfamílias para *H. intermedius*, seis superfamílias para *H. malabaricus* (G), cinco superfamílias para *H. unitaeniatus* e quatro superfamílias para *H. malabaricus* (F) e *E. erythrurus*. Em todos os catálogos, destaca-se o

agrupamento de satélites longos em uma mesma superfamília, além da similaridade entre satélites curtos e regiões de satélites longos. Todos os alinhamentos observados para as superfamílias encontram-se no material suplementar. Além disso, os satélites HmgSat05-63 e HmgSat14-31, agrupados na SF3, foram caracterizados como uma *high order repeat* (HOR) (Figura 5).

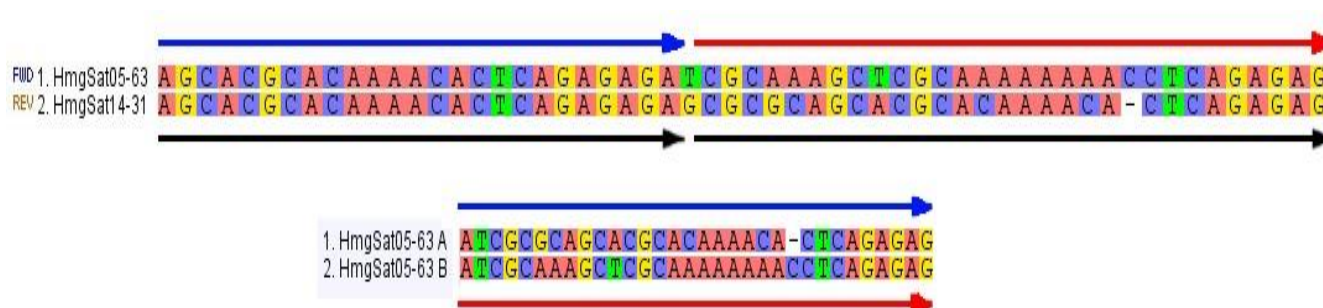


Figura 5: HOR encontrada entre dois satDNAs em *H. malabaricus* (G). As setas pretas representam a repetição dos monômeros do HmgSat14-31, enquanto as setas vermelhas e azuis representam as sub-repartições do HmgSat05-63.

4.3 Comparação entre os catálogos de Erythrinidae refletem as relações filogenéticas do grupo

Utilizando a ferramenta *rm_homology*, em conjunto a alinhamentos do software Geneious, utilizando as opções *de novo assemble*, foi possível realizar análises comparativas entre as sequências consensos dos cinco novos catálogos descritos, em conjunto ao satelitoma de *H. malabaricus* (D), já descrito em Erythrinidae (Toma et al., *in prep.*). Foram considerados satélites conservados aqueles que apresentaram similaridade entre 50% e 100%, sendo considerados partes de uma mesma superfamília (50% - 80%), variantes de um mesmo satélite (80% - 95%) ou mesmas variantes (95% -100%).

As análises comparativas demonstraram a presença de famílias de satDNAs isoladas por TAREAN apenas em cariomorfos específicos de *H. malabaricus*, apesar desses grupos estarem reunidos em uma mesma espécie. Foram observados 15 satDNAs isolados apenas no catálogo de *H. malabaricus* (D), 10 satDNAs apenas em *H. malabaricus* (F) e 3 satDNAs apenas em *H. malabaricus* (G). Destaca-se a grande similaridade entre os catálogos descritos para os cariomorfos F e G de *H. malabaricus*, sendo que 82,2% das sequências consensos de satDNAs descritas para *H. malabaricus* (G) encontram similaridade com alguma sequência consensos no catálogo de *H. malabaricus* (F), agrupadas em, no mínimo, superfamílias (> 50% de similaridade) (Tabela 8). Apesar de ainda

apresentar grande similaridade com os catálogos do cariomorfos F e G, o catálogo descrito para *H. malabaricus* (D) apresenta menos compartilhamento de consensos com os outros cariomorfos, com 55,35% de suas sequências compartilhadas com o cariomorfo F e 50% de suas sequências compartilhadas com o cariomorfo G (Tabela 8).

Tabela 8: Matriz de similaridade entre o conteúdo de um satelitoma (horizontal) contra todos os outros satelitomas analisados (vertical).

	HmaSat	HmfSat	HmgSat	HinSat	HunSat	EerSat
HmaSat	100	55,35	62,2	32,87	22,5	18,36
HmfSat	55,35	100	82,2	39,72	22,5	20,40
HmgSat	50	66,07	100	36,98	20	16,32
HinSat	42,85	51,78	60	100	30	24,48
HunSat	16,07	16,07	17,77	16,43	100	63,26
EerSat	16,07	17,85	17,77	16,43	77,5	100

O catálogo de *H. intermedius* compartilha mais sequencias consensus com os catálogos descritos para *H. malabaricus*, com compartilhamento de 29 satDNAs entre esta espécie e *H. malabaricus* (D) e *H. malabaricus* (F), e 27 satDNAs compartilhados com *H. malabaricus* (G) (Figura 6). Apesar de um número representativo de sequências compartilhadas nesses grupos, o alto número de satélites descritos por TAREAN apenas em *H. intermedius* (34 satDNAs), faz com que a porção do satelitoma de *H. intermedius* representada em outros catálogos do gênero *Hoplias* varie entre 32% (cariomorfo D) e 39% (cariomorfo F).

Os satelitomas de *H. unitaeniatus* e *E. erythrinus* (A) mostraram o compartilhamento de 31 sequências consensos de satDNAs, sendo a segunda maior taxa de compartilhamento de sequências, com 77.5% dos satélites de *H. unitaeniatus* sendo compartilhados com *E. erythrinus* (A) ao nível de superfamília, no mínimo. *E. erythrinus* (A) e *H. unitaeniatus* representaram os catálogos mais divergentes em relação aos cariomorfos de *H. malabaricus* e *H. intermedius*, com, no máximo, 22.5% de sequências compartilhadas com um desses catálogos (12 satDNAs entre *H. unitaeniatus* e *H. intermedius*).

4.4 Compartilhamento de sequências consensos revela seis satDNAs conservados na família Erythrinidae

As análises comparativas para os catálogos descritos de Erythrinidae revelaram a ocorrência de seis satDNAs conservados para esta família, sendo quatro dessas sequências consideradas satélites curtos (<100bp) e as duas restantes, satélites longos (>100bp) (Tabela 9). Destaca-se o fato de que, na grande maioria dos casos, os satDNAs considerados conservados para todas as espécies apresentaram uma relação de variantes do mesmo satélite (80% a 95% de similaridade) ou mesma variante (>95%), com exceção dos satélites HunSat35-214 e EerSat33-214, que apresentam similaridade de, em média, 70% com os satélites HmaSat51-219, HmfSat24-217 HmgSat21-206 e HinSat59-219, agrupando-se como uma superfamília. A maioria das divergências observadas entre os satDNAs considerados conservados para todas as espécies dá-se através de substituições de nucleotídeos, destacando-se os grupos HmaSat20-49, HmfSat14-49, HmgSat11-49, HinSat25-49, HunSat11-49 e EerSat12-49; e HmaSat53-56, HmfSat45-56, HmgSat44-56, HinSat66-56, HunSat1356 e EerSat17-56, onde as sequências divergem apenas por substituição de uma base em *H. unitaeniatus* e *E. erythrinus* (A). Entretanto, no grupo HmaSat23-42, HmfSat20-42, HmgSat24-42, HinSat52-42, HunSat15-48 e EerSat20-48, a principal divergência entre as sequências é devida a uma inserção de seis bases, ocorrida em *H. unitaeniatus* e *E. erythrinus* (A). Todos os alinhamentos encontram-se no material suplementar.

Além disso, foi possível observar a conservação de satDNAs em grupos específicos, sendo eles: satDNAs conservados em *H. malabaricus*, satDNAs conservados no gênero *Hoplias*, satDNAs conservados entre os gêneros *Hoplerythrinus* e *Erythrinus*, e satDNAs conservados em diversas espécies, mas ausentes em pelo menos um grupo (Figura 6), baseados na análise das sequências consensus.

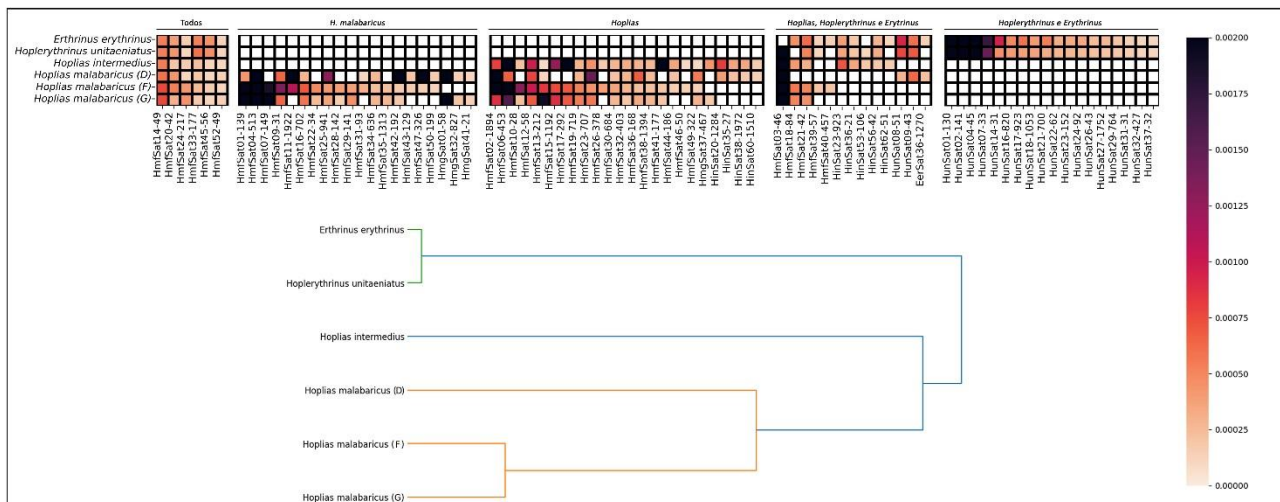


Figura 6: Mapa de calor dos satélites conservados, divididos em 5 grupos: todas as espécies analisadas, presentes apenas nos cariomorfos de *H. malabaricus*, presentes em satelitomas do gênero *Hoplias*, incluindo *H. intermedius*, satDNAs conservados em diversos grupos, mas não presentes em todos, e apenas compartilhados entre *H. unitaeniatus* e *E. erythrinus*. Em conjunto, uma filogenia baseada na quantidade de satDNAs compartilhados entre os grupos, analisando-se apenas as sequências consensus.

4.5 Mapeamento por FISH de sequências conservadas demonstra o compartilhamento de padrões de clusterização.

O mapeamento por FISH foi realizado em indivíduos de *H. malabaricus* dos cariomorfos F e G, *H. intermedius*, *H. unitaeniatus* e *E. erythrinus*. Para o mapeamento das sequências, selecionamos todos os satDNAs considerados conservados para todas as espécies, sendo representados pelos satDNAs HmfSat14-49, HmfSat24-217, HmfSat45-56, HmfSat52-49, HunSat15-48 e HmfSat33-177 (Figura 7). O mapeamento dessas sequências demonstra o mesmo padrão de clusterização para todos os grupos mapeados, com acúmulo dessas sequências principalmente em regiões teloméricas, com HmfSat14-49 demonstrando clusters em todos os telômeros de todos os grupos analisados, além da região pericentromérica de um par de cromossomos em *H. malabaricus* F e *H. malabaricus* G. HmfSat45-56 demonstra um padrão de clusterização disperso pelos telômeros de todos os grupos, com maior evidência em 2 pares nos cariomorfos de *H. malabaricus*. HmfSat52-49 e HunSat15-48 apresentam variação no número de cromossomos marcados, com 14 a 6 pares marcados (Figura 7). HmfSat24-217 demonstra um padrão de distribuição cromossômica mais restritos, presente nos telômeros de 2 pares em *H. malabaricus* F e 1 par nas demais espécies. O único exemplo de satDNA não clusterizado foi o HmfSat33-177, exibindo o mesmo padrão para todas as espécies.

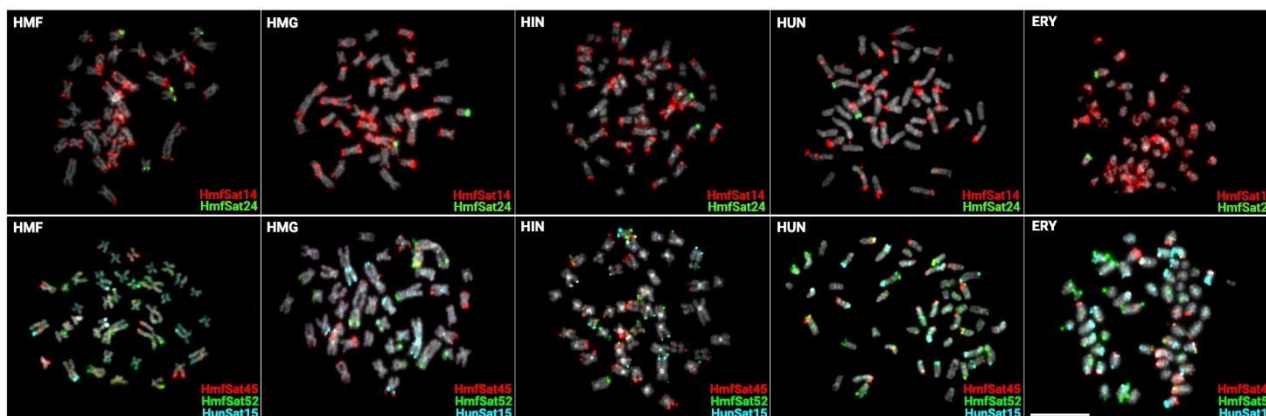


Figura 7: Mapeamento das famílias de satDNAs conservados para todos os grupos analisados. Em *H. malabaricus* F (HMF), *H. malabaricus* G (HMG), *H. intermedius* (HIN), *H. unitaeniatatus* (HUN) e *E. erythrinus* (ERY). Barra = 10 μ m.

Além disso, foram realizados mapeamentos direcionados a satDNAs conservados em catálogos específicos, sendo esses HmfSat01-139, HmfSat04-513 e HmfSat42-192, específicos para os três cariomorfos de *H. malabaricus*; HmfSat02-1894, HmfSat06-453, HmfSat12-58, HmfSat13-212 e HmgSat08-152, específicos para o gênero *Hoplias*; HunSat21-700, EerSat02-130 e EerSat03-141, específicos para *H. unitaeniatatus* e *E. erythrinus*; (Figuras 8, 9 e 10, respectivamente).

Dentre os satDNAs específicos para os cariomorfos de *H. malabaricus*, HmfSat01-139, HmfSat04-513 e HmfSat42-192 demonstraram clusterização somente em *H. malabaricus* F e G, ausente nas demais espécies (Figura 8). HmfSat01-139 e HmfSat42-513 demonstraram clusters em regiões centroméricas dos dois cariomorfos, com variação no número de cromossomos marcados para HmfSat01-139 (14 cromossomos em *H. malabaricus* (F) e 12 cromossomos em *H. malabaricus* (G)), e conservação do número de cromossomos marcados para HmfSat42-513 (1 par de cromossomos em ambos os cariomorfos) (Figura 08). HmfSat04-513 demonstrou clusters em regiões teloméricas em 22 cromossomos de ambos os cariomorfos (Figura 08).

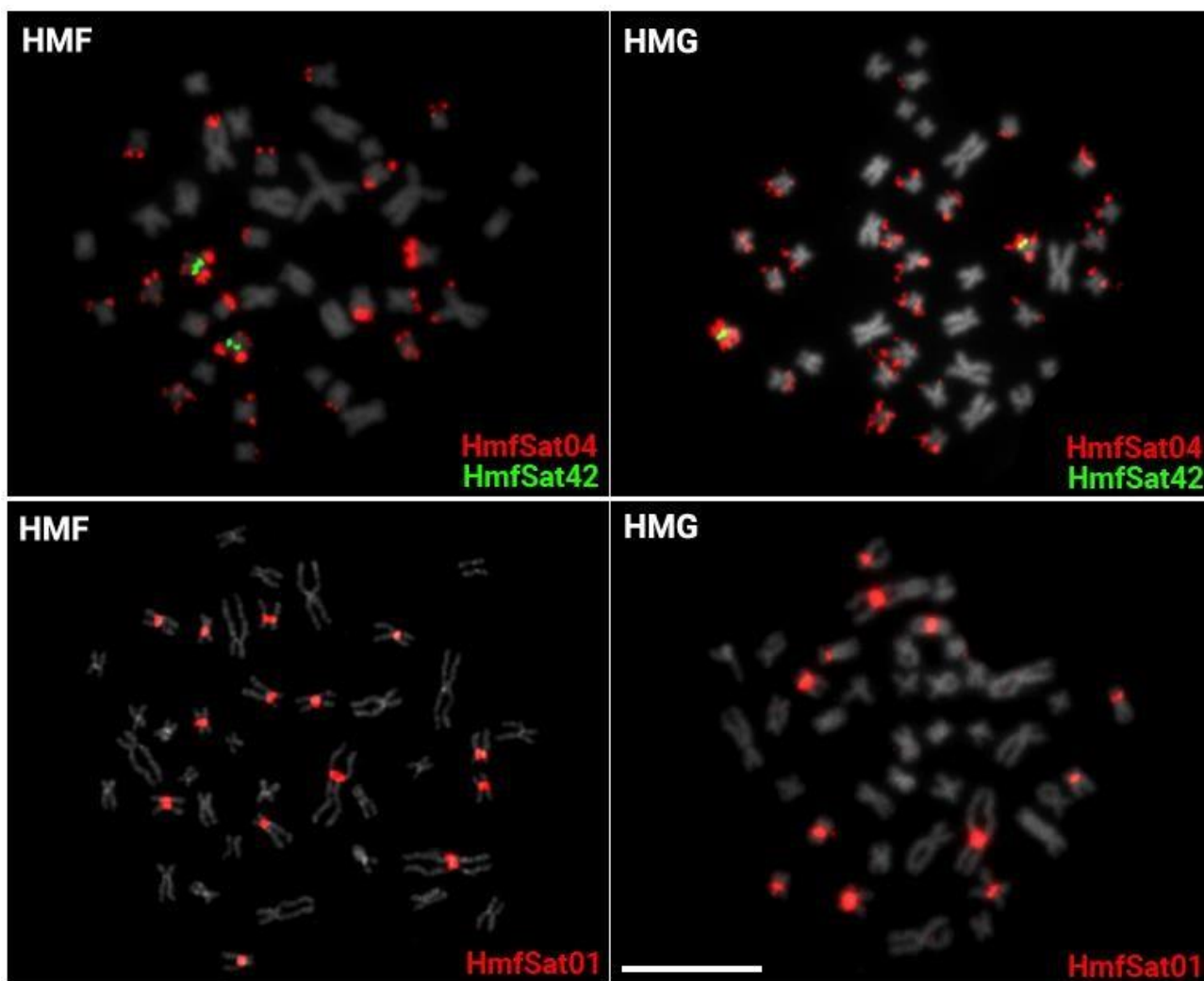


Figura 08: Mapeamento de satDNAs específicos dos cariomorfos de *H. malabaricus* F (HMF) e G (HMG)). Os satDNAs mapeados e suas respectivas cores estão descritos em cada imagem. Barra = 10 μ m.

Dentre os satDNAs específicos do gênero *Hoplias*, HmfSat02-1894, HmgSat08-152 e HmfSat06-453 demonstraram clusters nas porções centroméricas dos cromossomos. Nestes casos, HmgSat08-152 e HmfSat06-453 foram conservados no número de cromossomos marcados para todas as espécies (1 par e 2 pares, respectivamente), enquanto HmfSat02 demonstrou clusterização em 5 pares em *H. malabaricus* (F), 3 pares em *H. malabaricus* (G), com uma sintenia em um dos pares com HmgSat08-152 e apenas 1 par em *H. intermedius*. O caso mais discrepante foi o HmfSat13212, centromérico nos dois cariomorfos de *H. malabaricus* e clusterizado nos telômeros de um par em *H. intermedius* (Figura 09). HmfSat12-58 demonstrou clusterização em regiões teloméricas de todos os cromossomos, mantendo esse padrão em todas as espécies. Nenhum sinal desses satélites foi observado em *H. unitaeniatus* e *E. erythrinus*.

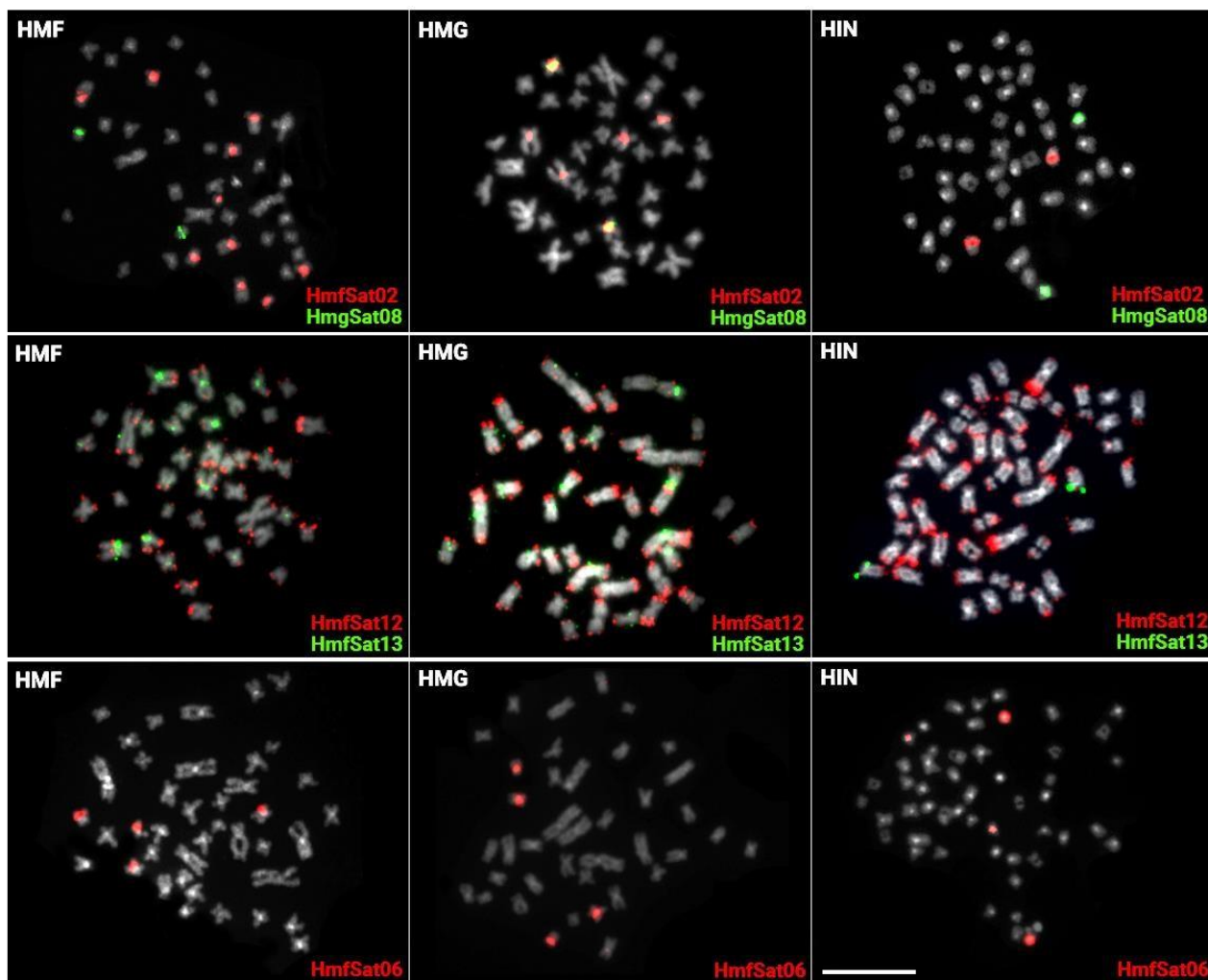


Figura 09: Mapeamento de satDNAs específicos do gênero *Hoplias* (*H. malabaricus* (F) (HMF), G (HMG) e *H. intermedius* (HIN)). Os satDNAs mapeados e suas respectivas cores estão descritos em cada imagem. Barra = 10 µm.

Todos os satDNAs selecionados para o mapeamento entre *H. unitaeniatus* e *E. erythrinus* foram localizados em regiões pericentromérica para ambas as espécies (Figura 10). Neste caso, chama atenção de um satDNA presente em praticamente todos os centrômeros das duas espécies (EerSat02-130), um satDNA presente em todos os centrômeros de *E. erythrinus* e presente em 17 pares de *H. unitaeniatus* (EerSat03-141) e a conservação de 1 par marcado em ambas as espécies para o HunSat21-700 (Figura 10).

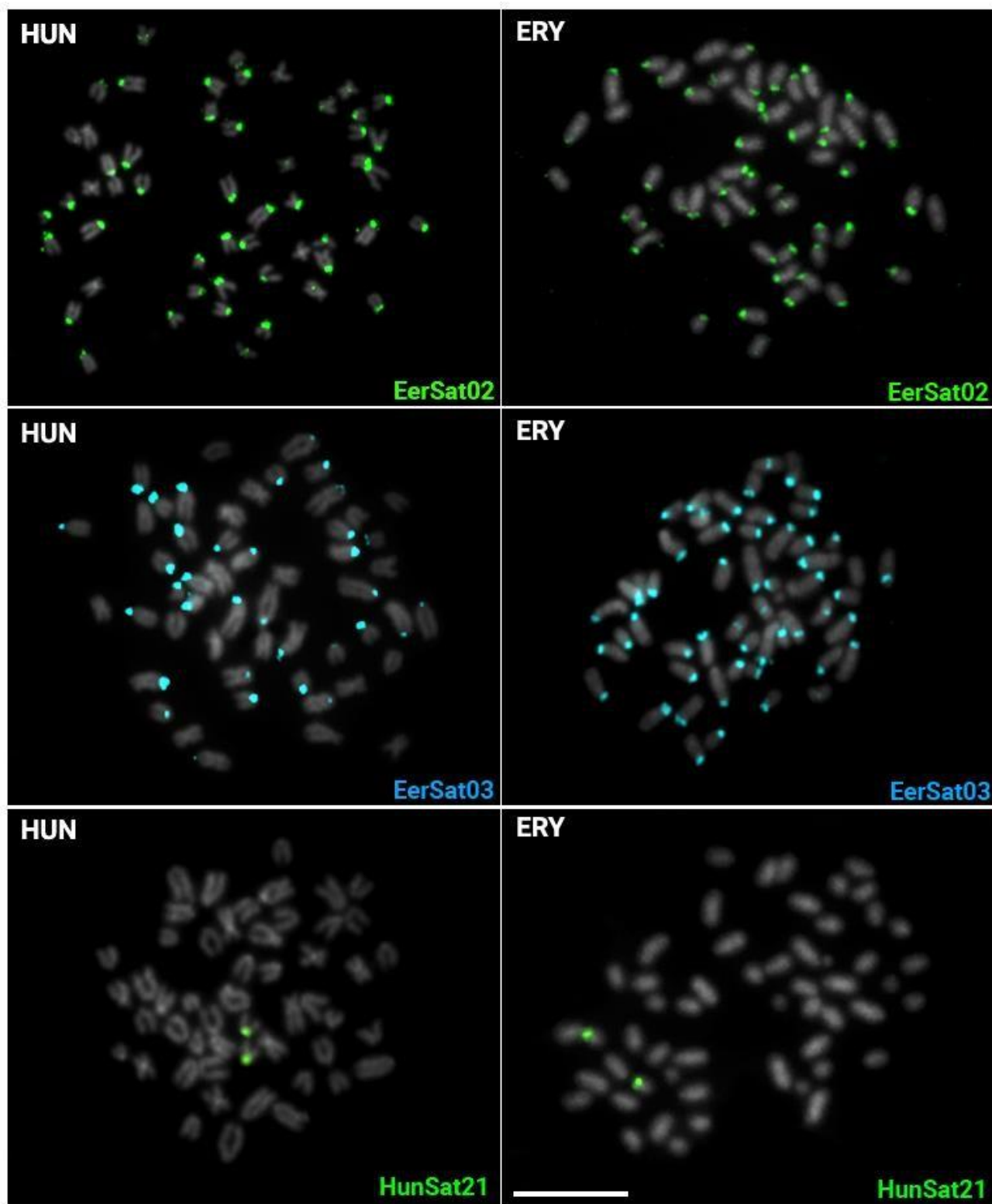


Figura 10: Mapeamento de satDNAs específicos de *H. unitaeniatus* e *E. erythrinus*. Os satDNAs mapeados e suas respectivas cores estão descritos em cada imagem. Barra = 10 μ m.

4.6 MSTs revelam diferentes padrões de expansão de haplótipos em satélites conservados de *Erythrinidae*

Para investigar a homogeneização e compartilhamento de haplótipos entre os grupos, selecionamos os satDNAs HmfSat14-49, HmfSat20-42, HmfSat45-56 e HmfSat52-49, todos presentes nos seis satelitomas analisados. Todos esses satDNAs demonstraram clusterização em regiões teloméricas de todas as espécies analisadas, com variação entre presença em todos os cromossomos (HmfSat14-49 e HmfSat45-56) ou apenas alguns cromossomos (HmfSat20-42 e HmfSat52-49). Foi possível observar dois padrões principais: satDNAs que compartilham haplótipos entre todas as espécies (HmfSat14-49 e HmfSat45-56) e satDNAs com haplótipos apenas compartilhados entre as espécies do gênero *Hoplias* ou entre *Erythrinus* e *Hoplerythrinus* (HmfSat20-42 e HmfSat52-49). HmfSat14-49 demonstrou dois haplótipos principais, sendo um deles compartilhado entre todos os gêneros (Figura 11). Neste caso, é possível observar a amplificação do segundo haplótipo principal apenas em *Hoplias*, com pouca abundância em *Erythrinus* e *Hoplerythrinus*. De modo similar, HmfSat45-56 demonstrou dois haplótipos principais, compartilhados entre todos os grupos analisados. Entretanto, nesse caso foi possível observar a predominância de sequências providas de *H. unitaeniatus* e *E. erythrinus*, devido a maior abundância desse satDNA nesses dois grupos (Figura 12).

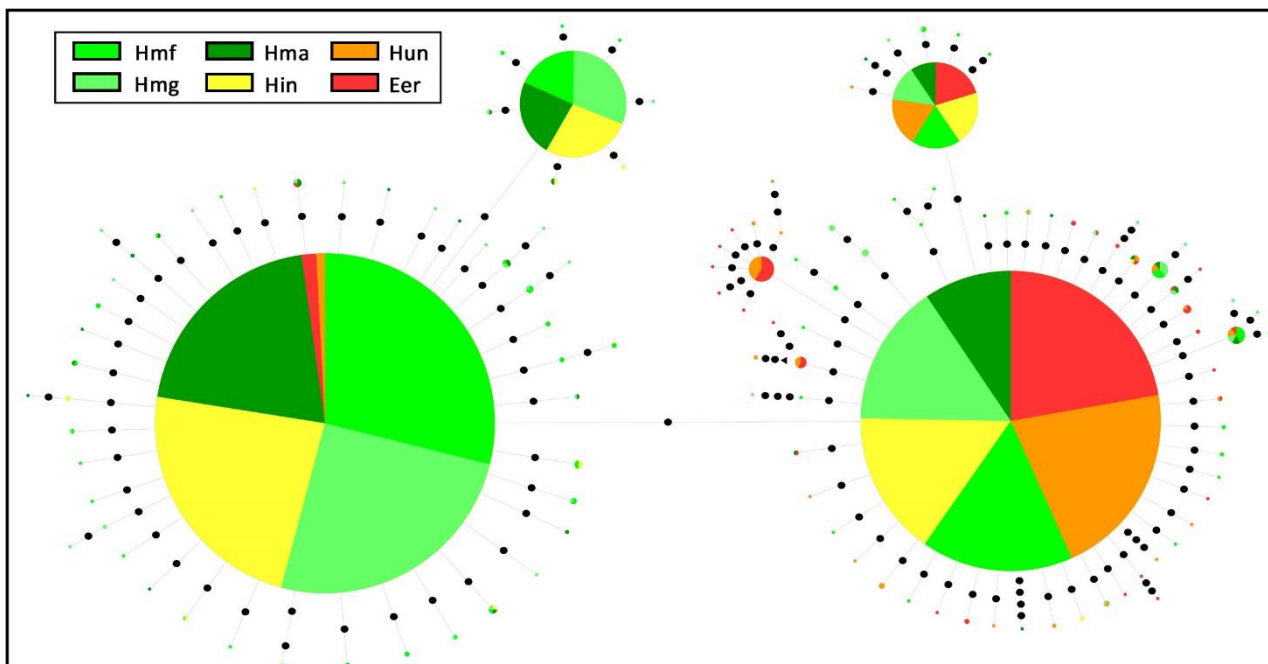


Figura 11: MST linear obtida através de *reads* correspondentes ao HmfSat14-49 dos seis grupos presentes neste trabalho. O diâmetro do círculo são proporcionais a abundância do haplótipo. Círculos pretos representam 1 base de divergência entre os haplótipos, enquanto triângulos pretos representam 5 bases de divergência.

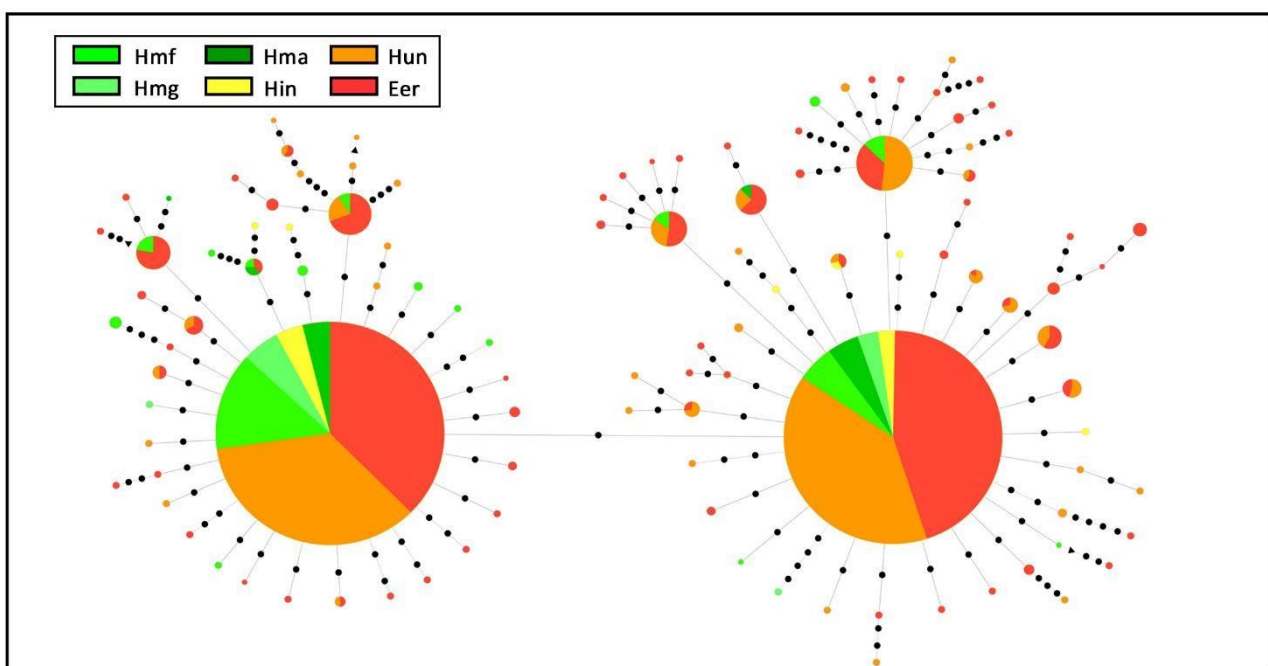


Figura 12: MST linear obtida através de *reads* correspondentes ao HmfSat45-56 dos seis grupos presentes neste trabalho. O diâmetro do círculo são proporcionais a abundância do haplótipo. Círculos pretos representam 1 base de divergência entre os haplótipos, enquanto triângulos pretos representam 5 bases de divergência.

As MSTs obtidas através dos monômeros dos satDNAs HmfSat20-42 e HmfSat52-49 não demonstraram nenhum haplótipo compartilhado entre *Hoplias* e os outros dois gêneros da família Erythrinidae (Figuras 13 e 14). Para HmfSat20-42, o alinhamento das sequências consensus dos seis grupos analisados demonstram uma inserção de 6 pares de bases exclusivas de *H. unitaeniatus* e *E. erythrinus* (ver material suplementar), o que corrobora a ausência de compartilhamento de haplótipos entre esses grupos e *Hoplias* (Figura 13). HmfSat52-49 demonstrou-se um satélite mais divergente, com a ausência de haplótipos dominantes, provavelmente devido sua baixa abundância em todos os grupos analisados.

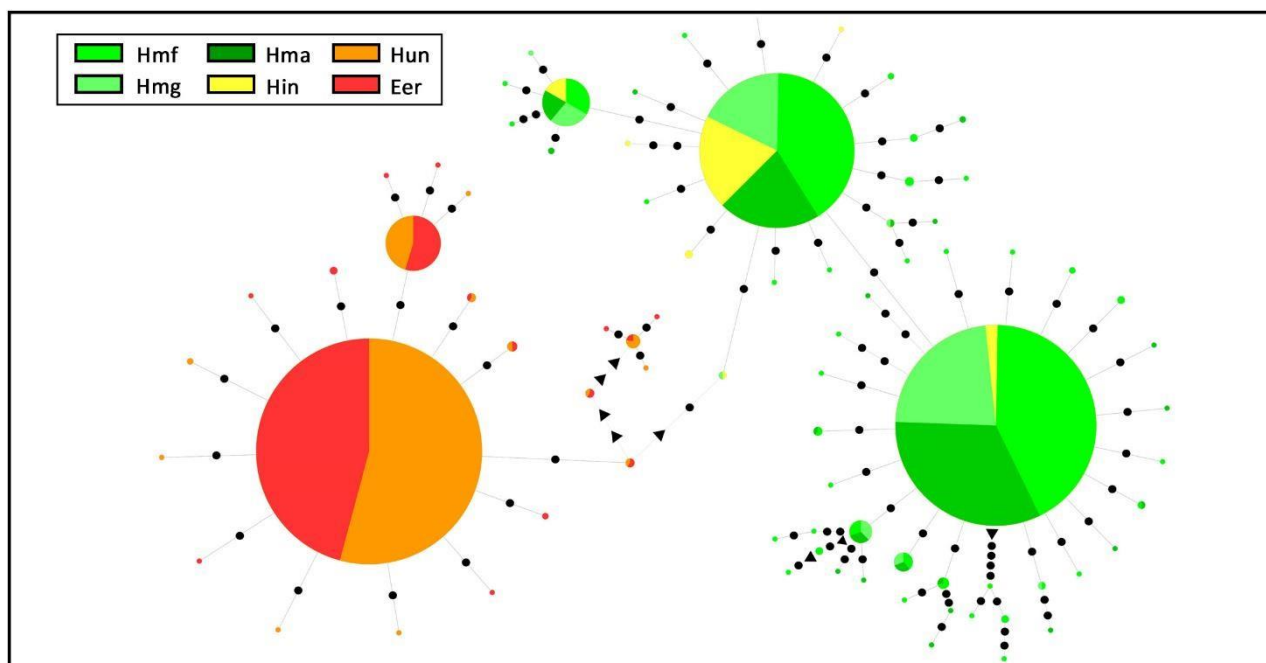


Figura 13: MST linear obtida através de *reads* correspondentes ao HmfSat20-42 dos seis grupos presentes neste trabalho. O diâmetro do círculo são proporcionais a abundância do haplótipo. Círculos pretos representam 1 base de divergência entre os haplótipos, enquanto triângulos pretos representam 5 bases de divergência.

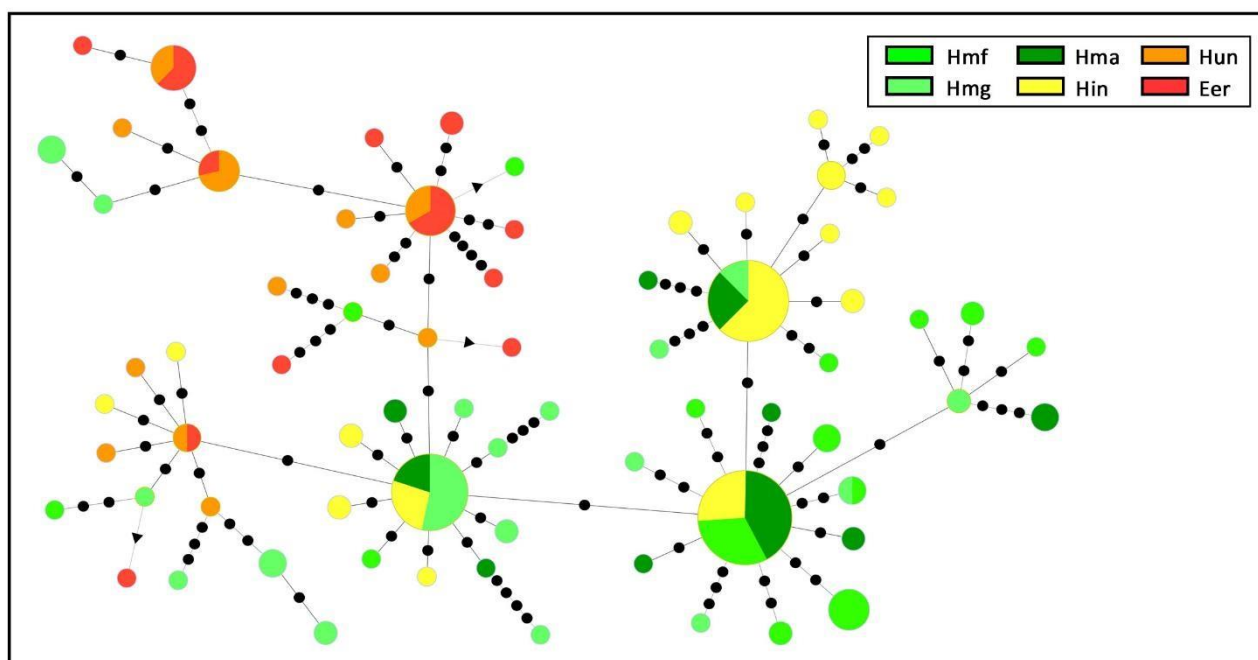


Figura 14: MST linear obtida através de *reads* correspondentes ao HmfSat52-49 dos seis grupos presentes neste trabalho. O diâmetro do círculo são proporcionais a abundância do haplótipo. Círculos pretos representam 1 base de divergência entre os haplótipos, enquanto triângulos pretos representam 5 bases de divergência.

5 DISCUSSÃO

A descrição de satelitomas em peixes Neotropicais, especialmente nas espécies pertencentes a ordem dos Characiformes vem trazendo novas informações em relação a diferenciação de cromossomos sexuais e cromossomos B, principalmente nos trabalhos de Silva et al. (2017), Utsunomia et al. (2019), Serrano-Freitas et al. (2020), Stornioli et al. (2021) e Kretschmer et al. (2022). Entretanto, poucos são os trabalhos dedicados ao estudo da evolução de satDNAs nesse grupo através da comparação de catálogos dessas sequencias, principalmente frente a grande diversidade da ictiofauna Neotropical. Estudos relacionados a comparação de satelitomas em peixes Neotropicais atualmente estão restritos a ordem dos Characiformes, com comparação entre espécies de dois gêneros da família Serrasalminidae (Goes et al., 2023) e dois gêneros da família Characidae (Goes et al., 2022). Em outros grupos, a satelitômica comparativa tem sido estudada principalmente em insetos; tais como gafanhotos (Palacios-Gimenez, et al., 2020; Camacho et al., 2022) e barbeiros (Pita et al., 2017), nematóides (Despot-Slade et al., 2022), mamíferos (Gutiérrez et al., 2022), plantas (Amosova et al., 2022) e aves (Peona et al., 2022).

Todas essas análises comparativas seguem as predições da hipótese *library*, que dita que espécies próximas compartilham um mesmo catálogo de satDNAs provindos de um ancestral em comum, com diferenças na amplificação desses satDNAs em cada grupo (Fry e Salser, 1977). Nossas análises também demonstraram seguir essa predição, com seis satDNAs compartilhados entre os seis grupos analisados, além de diversos outros compartilhados entre os cariomorfos de *H. malabaricus*, entre o gênero *Hoplias*, ou entre os gêneros *Hoplerythrinus* e *Erythrinus*, mais próximos filogeneticamente (Conde-Saldanha 2022). Diversos autores sugerem que *H. malabaricus* seja considerado um complexo de espécies (Cavallaro et al., 2000; Santos et al., 2009; Cioffi et al., 2012; Marques et al., 2013) e a análise dos cariomorfos D, F e G revelam a presença de muitos satDNAs isolados em apenas um cariomorfo por TAREAN, demonstrando a existência de um catálogo ancestral e variação na abundância dessas sequências. Entretanto, nestes casos a hipótese *library* também se mostrou significativa, com grande parte dos catálogos sendo compartilhados entre os cariomorfos, além de grande variação de abundância entre alguns satDNAs considerados compartilhados, como por exemplo HmaSat24-142 ser correspondente ao HmfSat01-139 e HmgSat04-139. A mesma situação pode ser observada entre os satDNAs compartilhados entre todos os integrantes analisados do gênero *Hoplias*, como o HinSat73-232 ser o menos abundante de *H. intermedius*, mas ser correspondente ao HmfSat13-212, HmgSat07-212 e HmaSat52-211, e entre *H. unitaeniatus* e *E. erythrinus*, com o HunSat21-700 corresponder ao EerSat13-691.

Através das MSTs produzidas, foi possível acompanhar, em alguns casos, os padrões de expansão que ocorreram ao longo da história evolutiva desse grupo, conforme predito pela evolução em concerto. Alguns exemplos disponíveis são o surgimento de um haplótipo muito abundante restrito às espécies do gênero *Hoplias* em HmfSat14-49 e a maior abundância do HmfSat45-56 em *H. unitaeniatus* e *E. erythrinus*, em comparação a qualquer indivíduo de *Hoplias*. Entretanto, não foi possível correlacionar os padrões de diversificação de haplótipos com a posição cromossômica desses satDNAs, como observado em outras espécies de peixes, em específico nos cromossomos sexuais de *T. auritus* (Kretschmer et al., 2022). Neste caso, observou-se satélites menos diversificados posicionados nos telômeros dos sexuais, enquanto mais diversos em regiões pericentroméricas. Os dados obtidos nesse trabalho apontam a presença nos telômeros de satélites com menos diversidade haplotípica, compartilhada entre diferentes grupos, quanto de satélites mais diversos, com ausência de compartilhamento de haplotipos entre os grupos.

A ocorrência de perfis espécie específicos de satelitomas derivados de amplificações ou retrações de uma biblioteca ancestral está previsto dentro da hipótese *library* (Fry e Salser, 1977).

SatDNAs não compartilhados podem ser remanescentes de uma biblioteca ancestral, amplificados somente em apenas um grupo, ou então ter surgido *de novo*, como proposto por Camacho et al., 2022. Nossas análises demonstram satDNAs que podem se enquadrar nessas duas possibilidades, com satDNAs presentes em apenas espécie grupos relacionados (ver tabela 9) podendo ter surgido *de novo* durante a diversificação de seus respectivos grupos, ao contrário de casos como HmfSat40457, também encontrado em *E. erythrinus* (EerSat37-905), que possivelmente estava presente na biblioteca ancestral de Erythrinidae, e foi amplificado somente nestes dois grupos, ou então suprimido dos restantes. Camacho et al., 2022 sugerem que os satDNAs evoluem em ciclos de amplificação-degeneração, evidenciando a possibilidade de eliminação de satDNAs ao longo do tempo. Em resumo, os autores definem que após a amplificação de um satDNA em um genoma, essa sequência pode se manter por muito tempo, integrando longos *arrays* ou *high order repeats*. Entretanto, mutações de ponto agem ao longo do tempo, levando a degeneração desses *repeats*, podendo ser completamente degenerados ou eliminados por *crossing over* desigual. Entretanto, nesses casos os autores sugerem que os satDNAs mais curtos estão mais suscetíveis a eliminação, contrariando nossas observações, já que as sequencias consensos compartilhadas entre os grupos analisados neste trabalho são, em sua maioria, curtas.

O número de satDNAs conservados ao longo do tempo vem crescendo conforme a descrição de satelitomas avança (Mravinac et al., 2002; Robles et al., 2004; Petraccioli et al., 2015; Lorite et al., 2017; Halbach et al., 2020). Nossas análises demonstraram a presença de seis satDNAs presentes em todos os catálogos analisados, com grande nível de homogeneização entre os gêneros *Hoplias*, *Hoplerythrinus* e *Erythrinus* (Tabela 9), demonstrando a persistência desses satDNAs por pelo menos 50.8 milhões de anos, segundo tempo de divergência desses gêneros (Conde-Saldanha 2022). Nenhum destes corresponde ao CharSat01-52, corroborando os dados apresentados em dos Santos et al. (2021), caracterizando a existência dessa sequência em diversas famílias de Characiformes, mas não em Erythrinidae. Todos os satDNAs observados sendo conservados para todos os grupos demonstraram padrões de clusterização por FISH em regiões pericentroméricas ou teloméricas, com exceção do HmfSat33-177. Apesar de trabalhos de satelitômica comparativa detectarem o compartilhamento de sequencias sem evidências de que estas tenham alguma função (Camacho et al., 2022), a seleção pode operar positivamente em repetições através de recrutamento funcional para importantes funções, como organização centromérica e teloméricas, como o caso de satDNAs recrutados por genes envolvidos na organização centromérica em primatas (Muro et al., 1992; Haaf

et al., 1995; Masumoto et al., 1989, 2004). A posição cromossômica de satDNAs conservados em regiões teloméricas ou centroméricas também é observada em outros grupos de peixes

Neotropicais, entre os gêneros *Astyanax* e *Psalidodon* (Goes et al., 2022), *Piaractus* e *Colossoma* (Goes et al., 2023), demonstrando a possibilidade de satDNAs participando da organização dessas regiões.

E. erythrinus e *H. unitaeniatus*, compartilham um satDNA com clusterização em regiões pericentroméricas em todos os cromossomos, apesar destes terem se diversificado a 30 milhões de anos (Conde Saldanha 2022). De acordo com Melters et al., 2013, o satDNA com maior abundância deve realizar alguma participação na organização do centrômero. Entretanto, casos de divergência entre repetições centroméricas em espécies próximas ocorrem em plantas do gênero *Solanum*, (Zhang et al., 2014) e gafanhotos (Camacho et al., 2022), sugerindo ciclos de rápida expansão anteriormente a um satDNA se tornar predominante nos centrômeros de uma espécie. SatDNAs centroméricos podem estar envolvidos no *drive* meiótico e segregação cromossômica (Plohl et al., 2014; Hartley e O'Neill, 2019; Talbert e Henikoff, 2022), sugerindo que sua evolução pode levar a isolamento reprodutivo. Nesse contexto, apesar de mais experimentos serem necessários, chama a atenção não haver um satDNA centromérico presente em todos os cromossomos e compartilhado entre os cariomorfos de *H. malabaricus*, apesar de nossos critérios de seleção de mapeamento de sequências por FISH não terem incluído todos os satDNAs mais abundantes desse grupo, devido as hipóteses desse grupo formar um complexo de espécies e a ausência de indivíduos híbridos, mesmo entre cariomorfos que vivem em simpatria (Bertollo et al., 2000).

Diversas sequências repetitivas foram mapeadas em espécies da família Erythrinidae, como famílias multigênicas 5S rDNA e 18S rDNA, microssatélites, e satDNAs, especificamente o 5S*HindIII* nos cariomorfos de *H. malabaricus* (Blanco et al., 2010; Cioffi et al., 2010, 2011, 2013; Cioffi e Bertollo, 2010; Oliveira et al., 2018). Todos os mapeamentos de satDNAs realizados neste trabalho seguiram a previsão de clusterização em cada grupo, baseado na presença e ausência da sequência no satelitoma do grupo e padrões de clusterização conservados para as espécies, com algumas variações no número de cromossomos marcados. Todos os satélites mapeados encontram-se em regiões centroméricas ou teloméricas, corroborando dados de CGH gerados por (Sember et al., 2018), demonstrando o acúmulo de sequências repetitivas nessas regiões em *H. malabaricus*. O presente trabalho não verificou nenhum acúmulo diferencial de satDNAs conservados em cromossomos sexuais, amplamente estudados neste grupo (Oliveira et al., 2018).

6 CONSIDERAÇÕES FINAIS

O presente trabalho descreve, pela primeira vez, o satelitoma de espécies dos gêneros *Erythrinus* e *Hoplerythrinus*, além de acrescentar três novos satelitomas para o gênero *Hoplias*. Nossas análises permitiram observar que a evolução desses catálogos segue a hipótese *library*, com grande compartilhamento de satDNAs entre esses catálogos, de forma condizente a filogenia estabelecida para o grupo. Além disso, seis satDNAs foram encontrados com grande conservação de suas sequências consensos para todas as espécies analisadas, com clusterização em regiões centroméricas ou teloméricas. Apesar deste trabalho não observar indício de função para essas sequências, é possível que elas estejam envolvidas na organização dessas regiões cromossômicas. Nossos resultados também reforçam a caracterização de *H. malabaricus* como um complexo de espécies, devido a menor similaridade do catálogo descrito para *H. malabaricus* D em relação aos outros, e a ausência de um marcador centromérico comum para os cariomorfos.

O presente trabalho demonstra o potencial do estudo de satelitomas em peixes neotropicais, principalmente em grupos com problemáticas citogenéticas e taxonômicas, como é a família *Erythrinidae*. Nossos resultados corroboraram teorias acerca da evolução de satDNAs e agregam informações acerca da dinâmica evolutiva dessas sequências, além de gerar diversos novos marcadores citogenéticos em potencial para o estudo da evolução cariotípica de *Erythrinidae*, em especial dos cariomorfos de *H. malabaricus*, contribuindo de forma significativa para o melhor entendimento dos processos relacionados a diversificação cariotípica deste grupo tão emblemático para a citogenética brasileira.

7 REFERÊNCIAS

- Abel, L. D. S., Mantovani, M., & Moreira-Filho, O. (2006). Chromosomal distribution of the As51 satellite DNA in two species complexes of the genus *Astyanax* (Pisces, Characidae). *Genetics and Molecular Biology*, 29(3), 448–452.
- Ahmad, S. F., Singchat, W., Jehangir, M., Suntronpong, A., Panthum, T., Malaivijitnond, S., & Srikulnath, K. (2020). Dark matter of primate genomes: Satellite DNA repeats and their evolutionary dynamics. *Cells*, 9(12), 2714.
- Amosova, A. V., Yurkevich, O. Y., Bolsheva, N. L., Samatadze, T. E., Zoshchuk, S. A., & Muravenko, O. V. (2022). Repeatome Analyses and Satellite DNA Chromosome Patterns in *Deschampsia sukatschewii*, *D. cespitosa*, and *D. antarctica* (Poaceae). *Genes*, 13(5), 762.
- Artoni, R. F., Vicari, M. R., Endler, A. L., Cavallaro, Z. I., de Jesus, C. M., de Almeida, M. C., Moreira-Filho, O., & Bertollo, L. A. C. (2006). Banding pattern of A and B chromosomes of *Prochilodus lineatus* (Characiformes, Prochilodontidae), with comments on B chromosomes evolution. *Genetica*, 127(1), 277–284.
- Bachmann, L., & Sperlich, D. (1993). Gradual evolution of a specific satellite DNA family in *Drosophila ambigua*, *D. tristis*, and *D. obscura*. *Molecular Biology and Evolution*, 10(3), 647–659.
- Belyayev, A., Jandová, M., Josefiová, J., Kalendar, R., Mahelka, V., Mandák, B., & Krak, K. (2020). The major satellite DNA families of the diploid *Chenopodium album* aggregate species: Arguments for and against the “library hypothesis.” *PLoS One*, 15(10), e0241206.
- Bertollo, L. A. C. (1978). Cytotaxonomic considerations on *Hoplias lacerdae* (Pisces, Erythrinidae). *Braz. J. Genet.*, 1, 103–120.
- Bertollo, L. A. C., Cioffi, M., & Moreira-Filho, O. (2015). Direct chromosome preparation from freshwater teleost fishes. *Fish Cytogenetic Techniques (Chondrichthyans and Teleosts)*; OzoufCostaz, C., Pisano, E., Foresti, F., Almeida Toledo, L. F., Eds, 21–26.
- Bertollo, L. A. C., Oliveira, C., Molina, W. F., Margarido, V. P., Fontes, M. S., Pastori, M. C., Falcao, J. N., & Fenocchio, A. S. (2004). Chromosome evolution in the erythrinid fish, *Erythrinus erythrinus* (Teleostei: Characiformes). *Heredity*, 93, 228–233. <https://doi.org/10.1038/sj.hdy.6800511>

- Bertollo, Luiz A C, Born, G. G., Dergam, J. A., Fenocchio, A. S., & Moreira-Filho, O. (2000). A biodiversity approach in the neotropical Erythrinidae fish, *Hoplias malabaricus*. Karyotypic survey, geographic distribution of cytotypes and cytotaxonomic considerations. *Chromosome Research*, 8, 603–613. <https://doi.org/10.1023/A:1009233907558>
- Bertollo, Luiz A C, Fontes, M. S., Fenocchio, A. S., & Cano, J. (1997). The X 1 X 2 Y sex chromosome system in the fish *Hoplias malabaricus*. I. G-, C-and chromosome replication banding. *Chromosome Research*, 5, 493–499.
- Bertollo, Luiz Antonio Carlos. (2007). Chromosome evolution in the Neotropical Erythrinidae fish family: an overview. *Fish Cytogenetics*, 195, 211.
- Biscotti, M. A., Olmo, E., & Heslop-Harrison, J. S. P. (2015). *Repetitive DNA in eukaryotic genomes*. Springer.
- Blanco, D R, Lui, R. L., Vicari, M. R., Bertollo, L. A. C., & Moreira-Filho, O. (2010). Comparative cytogenetics of giant trahiras *Hoplias aimara* and *H. intermedius* (Characiformes, Erythrinidae): chromosomal characteristics of minor and major ribosomal DNA and cross-species repetitive centromeric sequences mapping differ among morphologically. *Cytogenetic and Genome Research*, 132(1–2), 71–78.
- Blanco, Daniel Rodrigues, Lui, R. L., Bertollo, L. A. C., Diniz, D., & Filho, O. M. (2010). Characterization of invasive fish species in a river transposition region: evolutionary chromosome studies in the genus *Hoplias* (Characiformes, Erythrinidae). *Reviews in Fish Biology and Fisheries*, 20, 1–8.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.
<https://doi.org/https://doi.org/10.1093/bioinformatics/btu170>
- Botchan, M. R. (1974). Bovine satellite I DNA consists of repetitive units 1,400 base pairs in length. *Nature*, 251(5473), 288–292.
- Cabral-de-Mello, D. C., Zrzava, M., Kubičková, S., Rendon, P., & Marec, F. (2021). The role of satellite DNAs in genome architecture and sex chromosome evolution in Crambidae moths. *Frontiers in Genetics*, 12, 661417.

- Camacho, J. P. M., Cabrero, J., López-León, M. D., Martín-Peciña, M., Perfectti, F., Garrido-Ramos, M. A., & Ruiz-Ruano, F. J. (2022). Satellitome comparison of two oedipodine grasshoppers highlights the contingent nature of satellite DNA evolution. *BMC Biology*, 20(1), 1–24.
- Camacho, J. P. M., Sharbel, T. F., & Beukeboom, L. W. (2000). B-chromosome evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 355(1394), 163–178.
- Cavallaro, Z. I., Bertollo, L. A. C., Perfectti, F., & Camacho, J. P. M. (2000). Frequency increase and mitotic stabilization of a B chromosome in the fish *Prochilodus lineatus*. *Chromosome Research*, 8(7), 627–634.
- Charlesworth, B., Sniegowski, P., & Stephan, L. W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*, 371(6494), 215–220.
- Cioffi, M. B., & Bertollo, L. A. C. (2010). Initial steps in XY chromosome differentiation in *Hoplias malabaricus* and the origin of an X1X2Y sex chromosome system in this fish group. *Heredity*, 105, 554–561. <https://doi.org/https://doi.org/10.1038/hdy.2010.18>
- Cioffi, M. B., Liehr, T., Trifonov, V., Molina, W. F., & Bertollo, L. A. C. (2013). Independent sex chromosome evolution in lower vertebrates: a molecular cytogenetic overview in the Erythrinidae fish family. *Cytogenetic and Genome Research*, 141(2–3), 186–194.
- Cioffi, M. B., Martins, C., & Bertollo, L. A. C. (2010). Chromosome spreading of associated transposable elements and ribosomal DNA in the fish *Erythrinus erythrinus*. Implications for genome change and karyoevolution in fish. *BMC Evolutionary Biology*, 10(1), 1–9. <https://doi.org/https://doi.org/10.1186/1471-2148-10-271>
- Cioffi, M. B., Martins, C., Centofante, L., Jacobina, U., & Bertollo, L. A. C. (2009). Chromosomal variability among allopatric populations of Erythrinidae fish *Hoplias malabaricus*: Mapping of three classes of repetitive DNAs. *Cytogenetic and Genome Research*, 125, 132–141. <https://doi.org/https://doi.org/10.1159/000227838>
- Cioffi, M. B., Martins, C., Vicari, M. R., Rebordinos, L., & Bertollo, L. A. C. (2010). Differentiation of the XY sex chromosomes in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae): unusual accumulation of repetitive sequences on the X chromosome. *Sexual Development*, 4, 176–185. <https://doi.org/https://doi.org/10.1159/000309726>

- Cioffi, M. B., Molina, W. F., Moreira-Filho, O., & Bertollo, L. A. C. (2011). Chromosomal distribution of repetitive DNA sequences highlights the independent differentiation of multiple sex chromosomes in two closely related fish species. *Cytogenetic and Genome Research*, 134(4), 295–302.
- Cioffi, M. de B., Yano, C. F., Sember, A., & Bertollo, L. A. C. (2017). Chromosomal evolution in lower vertebrates: Sex chromosomes in neotropical fishes. *Genes*, 8(10), 258.
- Conde Saldanha, 2022. Biogeografia histórica e filogenômica de Erythrinidae e taxonomia integrativa para o estudo do gênero neotropical *Hoplerethrinus* (Characiformes: erythrinidae)
- Crepaldi, C., & Parise-Maltempi, P. P. (2020). Heteromorphic sex chromosomes and their DNA content in fish: an insight through satellite DNA accumulation in *Megaleporinus elongatus*. *Cytogenetic and Genome Research*, 160(1), 38–46.
<https://doi.org/https://doi.org/10.1159/000506265>
- Da Rosa, R., Rubert, M., Martins-Santos, I. C., & Giuliano-Caetano, L. (2012). Evolutionary trends in *Hoplerethrinus unitaeniatus* (Agassiz 1829)(Characiformes, Erythrinidae). *Reviews in Fish Biology and Fisheries*, 22, 467–475.
- de Bello Cioffi, M., Franco, W., Ferreira, R., & Bertollo, L. (2012). Chromosomes as tools for discovering biodiversity—The case of Erythrinidae fish family. *Recent Trends Cytogenet Stud—Methodol Appl.*
- de Freitas, N. L., Al-Rikabi, A. B. H., Bertollo, L. A. C., Ezaz, T., Yano, C. F., de Oliveira, E. A., Hatanaka, T., & Cioffi, M. B. (2018). Early stages of XY sex chromosomes differentiation in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae) revealed by DNA repeats accumulation. *Current Genomics*, 19(3), 216–226.
<https://doi.org/https://doi.org/10.2174/1389202918666170711160528>
- De Jesus, C. M., Galetti, P. M., Valentini, S. R., & Moreira-Filho, O. (2003). Molecular characterization and chromosomal localization of two families of satellite DNA in *Prochilodus lineatus* (Pisces, Prochilodontidae), a species with B chromosomes. *Genetica*, 118(1), 25–32.
- de la Herrán, R., Fontana, F., Lanfredi, M., Congiu, L., Leis, M., Rossi, R., Ruiz-Rejón, C., Ruiz-Rejón, M., & Garrido-Ramos, M. A. (2001). Slow rates of evolution and sequence homogenization in an ancient satellite DNA family of sturgeons. *Molecular Biology and Evolution*, 18(3), 432–436.

- Sassi, FMC., Perez, M. F., Oliveira, V. C. S., Deon, G. A., de Souza, F. H. S., Ferreira, P. H. N., de Oliveira, E. A., Hatanaka, T., Liehr, T., & Bertollo, L. A. C. (2021). High genetic diversity despite conserved karyotype organization in the giant trahiras from genus *Hoplias* (Characiformes, erythrinidae). *Genes*, 12(2), 252.
- de Oliveira, E. A., Sember, A., Bertollo, L. A. C., Yano, C. F., Ezaz, T., Moreira-Filho, O., Hatanaka, T., Trifonov, V., Liehr, T., Al-Rikabi, A. B. H., Ráb, P., Pains, H., & Cioffi, M. B. (2018). Tracking the evolutionary pathway of sex chromosomes among fishes: characterizing the unique XX/XY1Y2 system in *Hoplias malabaricus* (Teleostei, Characiformes). *Chromosoma*, 127, 115–128. <https://doi.org/https://doi.org/10.1007/s00412-017-0648-3>
- Dergam, J. A., Paiva, S. R., Schaeffer, C. E., Godinho, A. L., & Vieira, F. (2002). Phylogeography and RAPD-PCR variation in *Hoplias malabaricus* (Bloch, 1794)(Pisces, Teleostei) in southeastern Brazil. *Genetics and Molecular Biology*, 25, 379–387.
- Dergam, J. A., Suzuki, H. I., Shibatta, O. A., Duboc, L. F., Júlio Jr, H. F., Giuliano-Caetano, L., & Black IV, W. C. (1998). Molecular biogeography of the Neotropical fish *Hoplias malabaricus* (Erythrinidae: Characiformes) in the Iguaçu, Tibagi, and Paraná rivers. *Genetics and Molecular Biology*, 21, 493–496.
- Despot-Slade, E., Širca, S., Mravinac, B., Castagnone-Sereno, P., Plohl, M., & Meštrović, N. (2022). Satellitome analyses in nematodes illuminate complex species history and show conserved features in satellite DNAs. *BMC Biology*, 20(1), 1–19.
- Dias, G. B., Heringer, P., Svartman, M., & Kuhn, G. C. S. (2015). Helitrons shaping the genomic architecture of *Drosophila*: enrichment of DINE-TR1 in α - and β -heterochromatin, satellite DNA emergence, and piRNA expression. *Chromosome Research*, 23, 597–613.
- Diniz, D., & Bertollo, L. A. C. (2003). Karyotypic studies on *Hoplerethrinus unitaeniatus* (Pisces, Erythrinidae) populations. A biodiversity analysis. *Caryologia*, 56(3), 303–313.
- Dos Santos, R. Z., Calegari, R. M., Silva, D. M. Z. de A., Ruiz-Ruano, F. J., Melo, S., Oliveira, C., Foresti, F., Uliano-Silva, M., Porto-Foresti, F., & Utsunomia, R. (2021). A long-term conserved satellite DNA that remains unexpanded in several genomes of Characiformes fish is actively transcribed. *Genome Biology and Evolution*, 13(2), evab002.
- Dover, G. A. (1982). Molecular drive: a cohesive mode of species evolution. *Nature*, 299(5879), 111–117.

- Dover, G. A. (1986). Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends in Genetics*, 2, 159–165.
- Dover, G. A. (1987). DNA turnover and the molecular clock. *Journal of Molecular Evolution*, 26(1), 47–58. <https://doi.org/https://doi.org/10.1007/BF02111281>
- Du, K., Stöck, M., Kneitz, S., Klopp, C., Woltering, J. M., Adolphi, M. C., Feron, R., Prokopov, D., Makunin, A., & Kichigin, I. (2020). The sterlet sturgeon genome sequence and the mechanisms of segmental rediploidization. *Nature Ecology & Evolution*, 4(6), 841–852.
- Elder Jr, J. F., & Turner, B. J. (1995). Concerted evolution of repetitive DNA sequences in eukaryotes. *The Quarterly Review of Biology*, 70(3), 297–320. <https://doi.org/DOI: 10.1086/419073>
- Eschmeyer WNR, Fricke R. ESCHMEYER'S CATALOG OF FISHES: GENERA, SPECIES, van der Laan (eds) 2023. <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>. Electro nic version accessed dd mmm 2023.
- Escudeiro, A., Ferreira, D., Mendes-da-Silva, A., Heslop-Harrison, J. S., Adegas, F., & Chaves, R. (2019). Bovine satellite DNAs—a history of the evolution of complexity and its impact in the Bovidae family. *The European Zoological Journal*, 86(1), 20–37. <https://doi.org/https://doi.org/10.1080/24750263.2018.1558294>
- Feliciello, I., Akrap, I., Brajković, J., Zlatar, I., & Ugarković, Đ. (2015). Satellite DNA as a driver of population divergence in the red flour beetle *Tribolium castaneum*. *Genome Biology and Evolution*, 7(1), 228–239.
- Fenocchio, A. S., Bertollo, L. A. C., Takahashi, C. S., & Camacho, J. P. M. (2000). B chromosomes in two fish species genus *Rhamdia* (Siluriformes, Pimelodidae). *FOLIA BIOLOGICA-KRAKOW*, 48(3/4), 105–110.
- Ferreira, I. A., Bertollo, L. A. C., & Martins, C. (2007). Comparative chromosome mapping of 5S rDNA and 5SHindIII repetitive sequences in Erythrinidae fishes (Characiformes) with emphasis on the *Hoplias malabaricus* ‘species complex’’. *Cytogenetic and Genome Research*, 118, 78–83.
- Foresti, F., Almeida-Toledo, L. F., & Toledo-Filho, S. A. (1981). Polymorphic nature of nucleolus organizer regions in fishes. *Cytogenetic and Genome Research*, 31(3), 137–144. <https://doi.org/https://doi.org/10.1159/000131639>

- Fry, K., & Salser, W. (1977). Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell*, 12(4), 1069–1084.
- Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23), 3150-3152.
<https://doi.org/10.1093/bioinformatics/bts565>
- Garrido-Ramos, M. A. (2017). Satellite DNA: an evolving topic. *Genes*, 8(9), 230. <http://www.mdpi.com/2073-4425/8/9/230htm>
- Garrido-Ramos, M. A. (2021). The genomics of plant satellite DNA. *Satellite DNAs in Physiology and Evolution*, 103–143.
- Garrido-Ramos, M. A., de la Herrán, R., Jamilena, M., Lozano, R., Ruiz-Rejón, C., & Ruiz-Rejón, M. (1999). Evolution of centromeric satellite DNA and its use in phylogenetic studies of the Sparidae family (Pisces, Perciformes). *Molecular Phylogenetics and Evolution*, 12(2), 200–204.
- Giuliano-Caetano, L., LC, J., Moreira-Filho, O., & LAC, B. (2001). Comparative cytogenetic studies on *Hoplerhynchus unitaeniatus* populations (Pisces, Erythrindae). *Cytologia*, 66(1), 39–43.
- Goes, C. A. G., dos Santos, N., Rodrigues, P. H. de M., Stornioli, J. H. F., Silva, A. B. da, dos Santos, R. Z., Vidal, J. A. D., Silva, D. M. Z. de A., Artoni, R. F., & Foresti, F. (2023). The Satellite DNA Catalogues of Two Serrasalminae (Teleostei, Characiformes): Conservation of General satDNA Features over 30 Million Years. *Genes*, 14(1), 91.
- Goes, C. A. G., dos Santos, R. Z., Aguiar, W. R. C., Alves, D. C. V., Silva, D. M. Z. de A., Foresti, F., Oliveira, C., Utsunomia, R., & Porto-Foresti, F. (2022). Revealing the satellite DNA history in *Psalidodon* and *Astyanax* characid fish by comparative satellitomics. *Frontiers in Genetics*, 13, 1567. <https://doi.org/10.3389/fgene.2022.884072>
- Gutiérrez, J., Aleix-Mata, G., Montiel, E. E., Cabral-de-Mello, D. C., Marchal, J. A., & Sánchez, A. (2022). Satellitome Analysis on *Talpa aquitania* Genome and Inferences about the satDNAs Evolution on Some Talpidae. *Genes*, 14(1), 117.
- Haaf, T., Mater, A. G., Wienberg, J., & Ward, D. C. (1995). Presence and abundance of CENP-B box sequences in great ape subsets of primate-specific α -satellite DNA. *Journal of Molecular Evolution*, 41, 487–491.

- Haff, T., Schmid, M., Steinlein, C., Galetti, P. M., & Willard, H. F. (1993). Organization and molecular cytogenetics of a satellite DNA family from *Hoplias malabaricus* (Pisces, Erythrinidae). *Chromosome Research*, 1, 77–86.
- Halbach, R., Miesen, P., Joosten, J., Taşköprü, E., Rondeel, I., Pennings, B., Vogels, C., Merklings, S., Koenraadt, C., & Lambrechts, L. (2020). An ancient satellite repeat controls gene expression and embryonic development in *Aedes aegypti* through a highly conserved piRNA. *Nature*, 580(7802), 274–277. <https://doi.org/10.1038/s41586-020-2159-2>
- Hartley, G., & O'Neill, R. J. (2019). Centromere repeats: hidden gems of the genome. *Genes*, 10(3), 223.
- HÖRZ, W., & ZACHAU, H. G. (1977). Characterization of distinct segments in mouse satellite DNA by restriction nucleases. *European Journal of Biochemistry*, 73(2), 383–392.
- Hutton, J. R., & Wetmur, J. G. (1973). Effect of chemical modification on the rate of renaturation of deoxyribonucleic acid. Deaminated and glyoxalated deoxyribonucleic acid. *Biochemistry*, 12(3), 558–563.
- Kantek, D. L. Z., Vicari, M. R., Peres, W. A. M., Cestari, M. M., Artoni, R. F., Bertollo, L. A. C., & Moreira-Filho, O. (2009). Chromosomal location and distribution of As51 satellite DNA in five species of the genus *Astyanax* (Teleostei, Characidae, Incertae sedis). *Journal of Fish Biology*, 75(2), 408–421.
- Kapitonov, V. V., & Jurka, J. (2007). Helitrons on a roll: eukaryotic rolling-circle transposons. *TRENDS in Genetics*, 23(10), 521–529.
- Kim, Y. B., Oh, J. H., McIver, L. J., Rashkovetsky, E., Michalak, K., Garner, H. R., Kang, L., Nevo, E., Korol, A. B., & Michalak, P. (2014). Divergence of *Drosophila melanogaster* repeatomes in response to a sharp microclimate contrast in Evolution Canyon, Israel. *Proceedings of the National Academy of Sciences*, 111(29), 10630–10635. <http://www.pnas.org/content/111/29/10630.short>
- Kit, S., & Hsu, T. C. (1961). Relative stability to thermal denaturation of deoxyribonucleic acid (DNA) preparations containing bromodeoxyuridine. *Biochemical and Biophysical Research Communications*, 5(2), 120–124.

- Kopecka, H., Macaya, G., Cortadas, J., Thiéry, J., Bernardi, G. (1978). Restriction enzyme analysis of satellite DNA components from the bovine genome. *European Journal of Biochemistry*, 84(1), 189–195.
- Kretschmer, R., Goes, C. A. G., Bertollo, L. A. C., Ezaz, T., Porto-Foresti, F., Toma, G. A., Utsunomia, R., & Cioffi, M. B. (2022). Satellitome analysis illuminates the evolution of ZW sex chromosomes of Triportheidae fishes (Teleostei: Characiformes). *Chromosoma*, 131, 29–45. <https://doi.org/https://doi.org/10.1007/s00412-022-00768-1>
- Kuhn, Gustavo C S, Sene, F. M., Moreira-Filho, O., Schwarzacher, T., & Heslop-Harrison, J. S. (2008). Sequence analysis, chromosomal distribution and long-range organization show that rapid turnover of new and old pBuM satellite DNA repeats leads to different patterns of variation in seven species of the *Drosophila buzzatii* cluster. *Chromosome Research*, 16, 307–324.
- Kuhn, Gustavo Campos Silva, Küttler, H., Moreira-Filho, O., & Heslop-Harrison, J. S. (2012). The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Molecular Biology and Evolution*, 29(1), 7–11.
- Larracuente, A. M. (2014). The organization and evolution of the Responder satellite in species of the *Drosophila melanogaster* group: dynamic evolution of a target of meiotic drive. *BMC Evolutionary Biology*, 14(1), 1–12.
- López-Flores, I., & Garrido-Ramos, M. A. (2012). The Repetitive DNA Content of Eukaryotic Genomes. In *Repetitive DNA* (Vol. 7, pp. 1–28). Karger Publishers.
- Lorite, P., Muñoz-López, M., Carrillo, J. A., Sanllorente, O., Vela, J., Mora, P., Tinaut, A., Torres, M. I., & Palomeque, T. (2017). Concerted evolution, a slow process for ant satellite DNA: study of the satellite DNA in the *Aphaenogaster* genus (Hymenoptera, Formicidae). *Organisms Diversity & Evolution*, 17(3), 595–606.
- Louzada, S., Lopes, M., Ferreira, D., Adegas, F., Escudeiro, A., Gama-Carvalho, M., & Chaves, R. (2020). Decoding the role of satellite DNA in genome architecture and plasticity—An evolutionary and clinical affair. *Genes*, 11(1), 72. <https://doi.org/https://doi.org/10.3390/genes11010072>
- Lower, S. S., McGurk, M. P., Clark, A. G., & Barbash, D. A. (2018). Satellite DNA evolution: old ideas, new approaches. *Current Opinion in Genetics & Development*, 49, 70–78.

- Mantovani, M., Abel, L. D. dos S., Mestriner, C. A., & Moreira-Filho, O. (2004). Evidence of the differentiated structural arrangement of constitutive heterochromatin between two populations of *Astyanax scabripinnis* (Pisces, Characidae). *Genetics and Molecular Biology*, 27, 536–542.
- Marques, D. F., Santos, F. A. dos, Silva, S. S. da, Sampaio, I., & Rodrigues, L. R. R. (2013). Cytogenetic and DNA barcoding reveals high divergence within the trahira, *Hoplias malabaricus* (Characiformes: Erythrinidae) from the lower Amazon River. *Neotropical Ichthyology*, 11, 459–466.
- Martinez, J. F., Lui, R. L., Traldi, J. B., Blanco, D. R., & Moreira-Filho, O. (2016). Comparative cytogenetics of *Hoplerethrinus unitaeniatus* (Agassiz, 1829)(Characiformes, Erythrinidae) species complex from different brazilian hydrographic basins. *Cytogenetic and Genome Research*, 149(3), 191–200.
- Martins, C., Ferreira, I. A., Oliveira, C., Foresti, F., & Galetti Jr., P. M. (2006). A tandemly repetitive centromeric DNA sequence of the fish *Hoplias malabaricus* (Characiformes: Erythrinidae) is derived from 5S rDNA. *Genetica*, 127(1–3), 133–141.
- Martins, N. F., Bertollo, L. A. C., Troy, W. P., Feldberg, E., de Souza Valentin, F. C., & de Bello Cioffi, M. (2013). Differentiation and evolutionary relationships in *Erythrinus erythrinus* (Characiformes, Erythrinidae): comparative chromosome mapping of repetitive sequences. *Reviews in Fish Biology and Fisheries*, 23, 261–269.
- Mascagni, F., Barghini, E., Ceccarelli, M., Baldoni, L., Trapero, C., Díez, C. M., Natali, L., Cavallini, A., & Giordani, T. (2022). The Singular Evolution of Olea Genome Structure. *Frontiers in Plant Science*, 13.
- Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., & Okazaki, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *The Journal of Cell Biology*, 109(5), 1963–1973.
- Masumoto, H., Nakano, M., & Ohzeki, J. (2004). The role of CENP-B and α -satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. *Chromosome Research*, 12, 543–556.
- McAllister, B. F., & Werren, J. H. (1999). Evolution of tandemly repeated sequences: What happens at the end of an array? *Journal of Molecular Evolution*, 48, 469–481.

- Melo, S., Utsunomia, R., Penitente, M., Sobrinho-Scudeler, P. E., Porto-Foresti, F., Oliveira, C., Foresti, F., & Dergam, J. A. (2017). B chromosome dynamics in *Prochilodus costatus* (Teleostei, Characiformes) and comparisons with supernumerary chromosome system in other *Prochilodus* species. *Comparative Cytogenetics*, 11(2), 393.
- Melters, D. P., Bradnam, K. R., Young, H. A., Telis, N., May, M. R., Ruby, J. G., Sebra, R., Peluso, P., Eid, J., & Rank, D. (2013). Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biology*, 14(1), 1–20. <https://doi.org/https://doi.org/10.1186/gb-2013-14-1-r10>
- Mestriner, C. A., Galetti Jr., P. M., Valentini, S. R., Ruiz, I. R. G., Abel, L. D. S., Moreira-Filho, O., & Camacho, J. P. M. (2000). Structural and functional evidence that a B chromosome in the characid fish *Astyanax scabripinnis* is an isochromosome. *Heredity*, 85(1), 1–9.
- Meštrović, N., Castagnone-Sereno, P., & Plohl, M. (2006). Interplay of selective pressure and stochastic events directs evolution of the MEL172 satellite DNA library in root-knot nematodes. *Molecular Biology and Evolution*, 23(12), 2316–2325. <https://doi.org/https://doi.org/10.1093/molbev/msl119>
- Meštrović, N., Mravinac, B., Pavlek, M., Vojvoda-Zeljko, T., Šatović, E., & Plohl, M. (2015). Structural and functional liaisons between transposable elements and satellite DNAs. *Chromosome Research*, 23, 583–596.
- Milani, D., Bardella, V. B., Ferretti, A. B. S. M., Palacios-Gimenez, O. M., Melo, A. de S., Moura, R. C., Loreto, V., Song, H., & Cabral-de-Mello, D. C. (2018). Satellite DNAs unveil clues about the ancestry and composition of B chromosomes in three grasshopper species. *Genes*, 9(11), 523.
- Milani, D., Ruiz-Ruano, F. J., Camacho, J. P. M., & Cabral-de-Mello, D. C. (2021). Out of patterns, the euchromatic B chromosome of the grasshopper *Abracris flavolineata* is not enriched in highcopy repeats. *Heredity*, 127(5), 475–483.
- Montiel, E. E., Mora, P., Rico-Porras, J. M., Palomeque, T., & Lorite, P. (2022). Satellitome of the Red Palm Weevil, *Rhynchophorus ferrugineus* (Coleoptera: Curculionidae), the Most Diverse Among Insects. *Front. Ecol. Evol*, 10, 826808.
- Montiel, E. E., Panzera, F., Palomeque, T., Lorite, P., & Pita, S. (2021). Satellitome analysis of *rhodnius prolixus*, one of the main chagas disease vector species. *International Journal of Molecular Sciences*, 22(11), 6052.

- Mora, P., Vela, J., Ruiz-Ruano, F. J., Ruiz-Mena, A., Montiel, E. E., Palomeque, T., & Lorite, P. (2020). Satellitome analysis in the ladybird beetle *Hippodamia variegata* (Coleoptera, Coccinellidae). *Genes*, 11(7), 783.
- Mravinac, B., & Plohl, M. (2007). Satellite DNA junctions identify the potential origin of new repetitive elements in the beetle *Tribolium madens*. *Gene*, 394(1–2), 45–52.
- Mravinac, B., Plohl, M., Mestrović, N., & Ugarković, Đ. (2002). Sequence of PRAT Satellite DNA “Frozen” in Some Coleopteran Species. *Journal of Molecular Evolution*, 54(6), 774–783.
- Mravinac, B., Plohl, M., & Ugarković, Đ. (2005). Preservation and high sequence conservation of satellite DNAs suggest functional constraints. *Journal of Molecular Evolution*, 61(4), 542–550. <https://doi.org/https://doi.org/10.1007/s00239-004-0342-y>
- Muro, Y., Masumoto, H., Yoda, K., Nozaki, N., Ohashi, M., & Okazaki, T. (1992). Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. *The Journal of Cell Biology*, 116(3), 585–596.
- Nascimento, M., Sousa, A., Ramirez, M., Francisco, A. P., Carriço, J. A., Vaz, C. (2016). PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics*, 33, 128–129. <https://doi.org/10.1093/bioinformatics/btw582>
- Nijman, I. J., & Lenstra, J. A. (2001). Mutation and recombination in cattle satellite DNA: a feedback model for the evolution of satellite DNA repeats. *Journal of Molecular Evolution*, 52, 361–371.
- Novák, P., Ávila Robledillo, L., Koblížková, A., Vrbová, I., Neumann, P., & Macas, J. (2017). TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Research*, 45(12), e111–e111. <https://doi.org/https://doi.org/10.1093/nar/gkx257>
- Novák, P., Neumann, P., Pech, J., Steinhaisl, J., & Macas, J. (2013). RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from nextgeneration sequence reads. *Bioinformatics*, 29(6), 792–793.
- Ohta, T., & Dover, G. A. (1984). The cohesive population genetics of molecular drive. *Genetics*, 108(2), 501–521. <https://doi.org/https://doi.org/10.1093/genetics/108.2.501>
- Oliveira, C., Almeida-Toledo, L. F., Foresti, F., Britski, H. A., & Toledo-Filho, S. A. (1988). Chromosome formulae of Neotropical freshwater fishes. *Rev. Brasil. Genet*, 11(3), 577–624.

- Paço, A., Freitas, R., & Vieira-da-Silva, A. (2019). Conversion of DNA sequences: From a transposable element to a tandem repeat or to a gene. *Genes*, *10*(12), 1014.
- Palacios-Gimenez, O. M., Koelman, J., Palmada-Flores, M., Bradford, T. M., Jones, K. K., Cooper, S. J. B., Kawakami, T., & Suh, A. (2020). Comparative analysis of morabine grasshopper genomes reveals highly abundant transposable elements and rapidly proliferating satellite DNA repeats. *BMC Biology*, *18*(1), 1–21.
- Palacios-Gimenez, O. M., Milani, D., Song, H., Marti, D. A., López-León, M. D., Ruiz-Ruano, F. J., Camacho, J. P. M., & Cabral-de-Mello, D. C. (2020). Eight million years of satellite DNA evolution in grasshoppers of the genus *Schistocerca* illuminate the ins and outs of the library hypothesis. *Genome Biology and Evolution*, *12*(3), 88–102.
- Park, S. T., & Kim, J. (2016). Trends in next-generation sequencing and a new era for whole genome sequencing. *International Neuourology Journal*, *20*(Suppl 2), S76.
- Peona, V., Kutschera, V. E., Blom M. P. K., Irestedt M., Suh A. (2022). Satellite DNA evolution in Corvoidea inferred from short and long reads. *Molecular Ecology*, *32*, 1288–1305.
- Petraccioli, A., Odierna, G., Capriglione, T., Barucca, M., Forconi, M., Olmo, E., & Biscotti, M. A. (2015). A novel satellite DNA isolated in *Pecten jacobaeus* shows high sequence similarity among molluscs. *Molecular Genetics and Genomics*, *290*(5), 1717–1725.
- Pita Mimbacas, S., Panzera Arballo, F., Mora, P., Vela, J., Cuadrado, A., Sánchez, A., Palomeque, T., & Lorite, P. (2017). Comparative repeatome analysis on *Triatoma infestans* Andean and NonAndean lineages, main vector of Chagas disease. *PloS One*, *2017 12* (7), p. E0181635.
- Plohl, M, Petrović, V., Luchetti, A., Ricci, A., Šatović, E., Passamonti, M., & Mantovani, B. (2010). Long-term conservation vs high sequence divergence: the case of an extraordinarily old satellite DNA in bivalve mollusks. *Heredity*, *104*(6), 543–551.
- Plohl, Miroslav, Luchetti, A., Meštrović, N., & Mantovani, B. (2008). Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero) chromatin. *Gene*, *409*(1), 72–82. <https://doi.org/https://doi.org/10.1016/j.gene.2007.11.013>
- Plohl, Miroslav, Meštrović, N., & Mravinac, B. (2012). Satellite DNA evolution. In *Repetitive DNA* (Vol. 7, pp. 126–152). Karger Publishers.

- Plohl, Miroslav, Meštrović, N., & Mravinac, B. (2014). Centromere identity from the DNA point of view. *Chromosoma*, 123(4), 313–325.
- Podgornaya, O. I. (2022). Nuclear organization by satellite DNA, SAF-A/hnRNPU and matrix attachment regions. *Seminars in Cell & Developmental Biology*.
- Reis, R E, Albert, J. S., Di Dario, F., Mincarone, M. M., Petry, P., & Rocha, L. A. (2016). Fish biodiversity and conservation in South America. *Journal of Fish Biology*, 89(1), 12–47.
- Reis, Roberto E, Kullander, S. O., & Ferraris, C. J. (2003). *Check list of the freshwater fishes of South and Central America*. Edipucrs.
- Robledillo, L. A., Koblížková, A., Novák, P., Böttinger, K., Vrbová, I., Neumann, P., Schubert, I., & Macas, J. (2018). Satellite DNA in *Vicia faba* is characterized by remarkable diversity in its sequence composition, association with centromeres, and replication timing. *Scientific Reports*, 8(1), 1–11.
- Robles, F., de la Herrán, R., Ludwig, A., Ruiz-Rejón, C., Ruiz-Rejón, M., & Garrido-Ramos, M. A. (2004). Evolution of ancient satellite DNAs in sturgeon genomes. *Gene*, 338(1), 133–142.
- Ruiz-Ruano, F. J., Castillo-Martínez, J., Cabrero, J., Gómez, R., Camacho, J. P. M., & López-León, M. D. (2018). High-throughput analysis of satellite DNA in the grasshopper *Pyrgomorpha conica* reveals abundance of homologous and heterologous higher-order repeats. *Chromosoma*, 127(3), 323–340.
- Ruiz-Ruano, F. J., López-León, M. D., Cabrero, J., & Camacho, J. P. M. (2016). High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Scientific Reports*, 6.
- Russell, D. W., & Sambrook, J. (2001). *Molecular cloning: a laboratory manual* (Vol. 1). Cold Spring Harbor Laboratory Cold Spring Harbor, NY.
- Santos, U., Völcker, C. M., Belei, F. A., Cioffi, M. B., Bertollo, L. A. C., Paiva, S. R., & Dergam, J. A. (2009). Molecular and karyotypic phylogeography in the Neotropical *Hoplias malabaricus* (Erythrinidae) fish in eastern Brazil. *Journal of Fish Biology*, 75(9), 2326–2343.
- Sassi, F. de M. C., Toma, G. A., & de Bello Cioffi, M. (2022). FISH—in Fish Chromosomes. *Cytogenetics and Molecular Cytogenetics*, 281–296. <https://doi.org/10.1201/9781003223658-24>
- Šatović-Vukšić, E., & Plohl, M. (2023). Satellite DNAs—From Localized to Highly Dispersed Genome Components. *Genes*, 14(3), 742.

- Šatović, E., & Plohl, M. (2013). Tandem repeat-containing MITEs in the clam *Donax trunculus*. *Genome Biology and Evolution*, 5(12), 2549–2559.
- Šatović Vukšić, E., & Plohl, M. (2021). Exploring satellite DNAs: specificities of bivalve mollusks genomes. *Satellite DNAs in Physiology and Evolution*, 57–83.
- Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27(6), 863–864.
<https://doi.org/https://doi.org/10.1093/bioinformatics/btr026>
- Sember, A., Bertollo, L. A. C., Ráb, P., Yano, C. F., Hatanaka, T., De Oliveira, E. A., & Cioffi, M. de B. (2018). Sex chromosome evolution and genomic divergence in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae). *Frontiers in Genetics*, 9, 71.
- Serrano-Freitas, É. A., Silva, D. M. Z. A., Ruiz-Ruano, F. J., Utsunomia, R., Araya-Jaime, C., Oliveira, C., Camacho, J. P. M., & Foresti, F. (2020). Satellite DNA content of B chromosomes in the characid fish *Characidium gomesi* supports their origin from sex chromosomes. *Molecular Genetics and Genomics*, 295(1), 195–207.
- Silva, D. M. Z. A., Daniel, S. N., Camacho, J. P. M., Utsunomia, R., Ruiz-Ruano, F. J., Penitente, M., Pansonato-Alves, J. C., Hashimoto, D. T., Oliveira, C., & Porto-Foresti, F. (2016). Origin of B chromosomes in the genus *Astyanax* (Characiformes, Characidae) and the limits of chromosome painting. *Molecular Genetics and Genomics*, 291(3), 1407–1418.
- Silva, D. M. Z. A., Utsunomia, R., Ruiz-Ruano, F. J., Daniel, S. N., Porto-Foresti, F., Hashimoto, D. T., Oliveira, C., Camacho, J. P. M., & Foresti, F. (2017). High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*. *Scientific Reports*, 7, 12726. <https://www.nature.com/articles/s41598-017-12939-7>
- Smit, A. F. A., Hubley, R., & Green, P. (2017). 1996–2010. *RepeatMasker Open-3.0*.
- Smith, G. P. (1976). Evolution of repeated DNA sequences by unequal crossover. *Science*, 191(4227), 528–535.
- Stephan, W. (1986). Recombination and the evolution of satellite DNA. *Genetics Research*, 47(3), 167–174.

- Stornioli, J. H. F., Goes, C. A. G., Calegari, R. M., dos Santos, R. Z., Giglio, L. M., Foresti, F., Oliveira, C., Penitente, M., Porto-Foresti, F., & Utsunomia, R. (2021). The B Chromosomes of *Prochilodus lineatus* (Teleostei, Characiformes) Are Highly Enriched in Satellite DNAs. *Cells*, 10(6), 1527.
- Strachan, T., Webb, D., & Dover, G. A. (1985). Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *The EMBO Journal*, 4(7), 1701–1708.
- Talbert, P. B., & Henikoff, S. (2022). The genetics and epigenetics of satellite centromeres. *Genome Research*, 32(4), 608–615.
- Tautz, D., Pena, S. D. J., Chakraborty, R., Epplen, J. T., & Jeffreys, A. J. (1993). DNA fingerprinting: state of the science. *Birkhäuser, Basel*.
- Thakur, J., Packiaraj, J., & Henikoff, S. (2021). Sequence, chromatin and evolution of satellite DNA. *International Journal of Molecular Sciences*, 22(9), 4309.
- Thomas, J., & Pritham, E. J. (2015). Helitrons, the eukaryotic rolling-circle transposable elements. *Mobile DNA iii*, 891–924.
- Utsunomia, R., Pansonato Alves, J. C., Paiva, L. R. S., Costa Silva, G. J., Oliveira, C., Bertollo, L. A. C., & Foresti, F. (2014). Genetic differentiation among distinct karyomorphs of the wolf fish *Hoplias malabaricus* species complex (Characiformes, Erythrinidae) and report of unusual hybridization with natural triploidy. *Journal of Fish Biology*, 85, 1682–1692. <https://doi.org/10.1111/jfb.12526>
- Utsunomia, Ricardo, Ruiz-Ruano, F. J., Silva, D. M. Z. A., Serrano, É. A., Rosa, I. F., Scudeler, P. E. S., Hashimoto, D. T., Oliveira, C., Camacho, J. P. M., & Foresti, F. (2017). A glimpse into the satellite DNA library in characidae fish (Teleostei, Characiformes). *Frontiers in Genetics*, 8(AUG). <https://doi.org/10.3389/fgene.2017.00103>
- Utsunomia, Ricardo, Silva, D. M. Z. A., Ruiz-Ruano, F. J., Goes, C. A. G., Melo, S., Ramos, L. P., Oliveira, C., Porto-Foresti, F., Foresti, F., & Hashimoto, D. T. (2019). Satellitome landscape analysis of *Megaleporinus macrocephalus* (Teleostei, Anostomidae) reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome. *Scientific Reports*, 9(1), 1–10.
- Vicari, M. R., Artoni, R. F., Moreira-Filho, O., & Bertollo, L. A. C. (2008). Colocalization of repetitive DNAs and silencing of major rRNA genes. A case report of the fish *Astyanax janae*iroensis. *Cytogenetic and Genome Research*, 122, 67–72.

- Vicente, V. E., Bertollo, L. A. C., Valentini, S. R., & Moreira-Filho, O. (2003). Origin and differentiation of a sex chromosome system in *Parodon hilarii* (Pisces, Parodontidae). Satellite DNA, G- and Cbanding. *Genetica*, 119(2), 115–120.
- Vozdova, M., Kubickova, S., Cernohorska, H., Fröhlich, J., Martínková, N., & Rubes, J. (2020). Sequence analysis and FISH mapping of four satellite DNA families among Cervidae. *Genes*, 11(5), 584.
- Yang, H.-P., & Barbash, D. A. (2008). Abundant and species-specific DINE-1 transposable elements in 12 *Drosophila* genomes. *Genome Biology*, 9(2), 1–14.
- Yano, C. F., Bertollo, L. A. C., Ezaz, T., Trifonov, V., Sember, A., Liehr, T., & Cioffi, M. (2017). Highly conserved Z and molecularly diverged W chromosomes in the fish genus *Triportheus* (Characiformes, Triportheidae). *Heredity*, 118(3), 276–283.
- Zattera, M. L., & Bruschi, D. P. (2022). Transposable Elements as a Source of Novel Repetitive DNA in the Eukaryote Genome. *Cells*, 11(21), 3373.
- Zhang, H., Koblížková, A., Wang, K., Gong, Z., Oliveira, L., Torres, G. A., Wu, Y., Zhang, W., Novák, P., & Buell, C. R. (2014). Boom-bust turnovers of megabase-sized centromeric DNA in *Solanum* species: rapid evolution of DNA sequences associated with centromeres. *The Plant Cell*, 26(4), 1436–1447. [https://doi.org/https://doi.org/10.1105/tpc.114.123877](https://doi.org/10.1105/tpc.114.123877)

MATERIAIS SUPLEMENTARES

Alinhamentos das famílias de satDNAs conservados entre os cariomorfos de *Hoplias malabaricus* (D, F e G), *Hoplias intermedius*, *Hoplerthrinus unitaeniatus* e *Erythrinus erythrinus* (cariomorfo A).

Satélites compartilhados entre todos os grupos analisados

HmaSat20-49, HmfSat14-49, HmgSat11-49, HinSat25-49, HunSat11-49 e EerSat12-49

Consensus	1 10 20 30 40 49
Identity	
FWD 1. EerSat12-49	CACTCCAGAGAACACAGTTCCACTGTTCCACACACCAGAGGAACCTCCAT
FWD 2. HunSat11-49	CACTCCAGAGAACACAGTTCCACTGTTCCACACACCAGAGGAACCTCCAT
FWD 3. HmaSat20-49	CACTCCAGAGAACACAGTTCCACTGTTCCACACACCAGAGGAACCTCCAT
FWD 4. HinSat25-49	CACTCCAGAGAACACAGTTCCACTGTTCCACACACCAGAGGAACCTCCAT
FWD 5. HmfSat14-49	CACTCCAGAGAACACAGTTCCACTGTTCCACACACCAGAGGAACCTCCAT
REV 6. HmgSat11-49	CACTCCAGAGAACACAGTTCCACTGTTCCACACACCAGAGGAACCTCCAT

HmaSat23-42, HmfSat20-42, HmgSat24-42, HinSat52-42, HunSat15-48 e EerSat20-48

Consensus	1 10 20 30 40 48
Identity	
REV 1. HmaSat23-42	GGTTAGTGTCTCTCCTCAGAGCGTGTGAGTCTGATGAA-----CAGA
REV 2. HinSat52-42	GGTTAGTGTCTCTCCTCAGAGCGTGTGAGTCTGATGAA-----CAGA
REV 3. EerSat20-48	GGTTAGTGTCTCTCCTCAGAGCGTGTGAGTCTGATGAA--CAGAA--CAGA
REV 4. HunSat15-48	GGTTAGTGTCTCTCCTCAGAGCGTGTGAGTCTGATGAA--CAGAA--CAGA
FWD 5. HmfSat20-42	GGTTAGTGTCTCTCCTCAGAGCGTGTGAGTCTGATGAA-----CAGA
FWD 6. HmgSat24-42	GGTTAGTGTCTCTCCTCAGAGCGTGTGAGTCTGATGAA-----CAGA

HmaSat41-176, HmfSat33-177 HmgSat36-177, HinSat50-175, HunSat10-177 e EerSat14-177

Consensus	1 20 40 60 80 100 106 120 140 160 178
Identity	
FWD 1. HmaSat41-176	-----
REV 2. HmfSat33-177	-----
REV 3. HmgSat36-177	-----
REV 4. HunSat10-177	-----
FWD 5. HinSat50-175	-----
FWD 6. EerSat14-177	-----

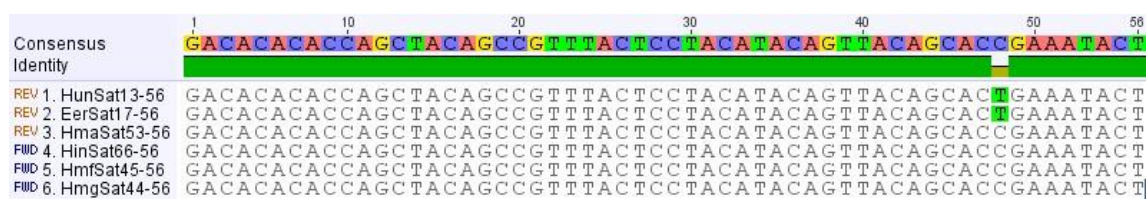
HmaSat48-49, HmfSat52-49, HmgSat40-49, HinSat68-49, HunSat33-49 e EerSat39-49

Consensus	1 10 20 30 40 49
Identity	
REV 1. HinSat68-49	TCCGCTCTGATAGAACCGGGTCAGAACACACACCGCTGAAAGAGTTAAC
FWD 2. EerSat39-49	ACCGCTCTGATAGAACCGGGTCAGAACACACACCGCTGAAAGAGTTAAC
REV 3. HunSat33-49	ACCGCTCTGATAGAACCGGGTCAGAACACACACCGCTGAAAGAGTTAAC
FWD 4. HmaSat48-49	TCCGCTCTGATAGAACCGGGTCAGAACACACACCGCTGAAAGAGTTAAC
FWD 5. HmfSat52-49	TCCGCTCTGATAGAACCGGGTCAGAACACACACCGCTGAAAGAGTTAAC
FWD 6. HmgSat40-49	TCCGCTCTGATAGAACCGGGTCAGAACACACACCGCTGAAAGAGTTAAC

HmaSat51-219, HmfSat24-217, HmgSat21-206, HinSat59-219, HunSat35-214 e EerSat33-214



HmaSat53-56, HmfSat45-56, HmgSat44-56, HinSat66-56, HunSat13-56 e EerSat17-56

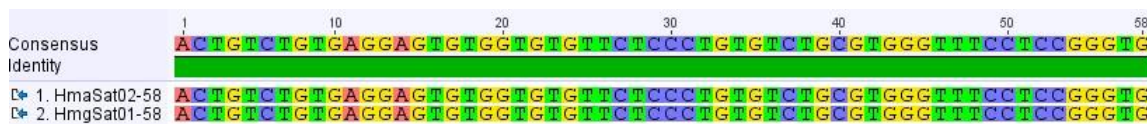


Satélites compartilhados entre apenas entre os cariomorfos de *H. malabaricus*.

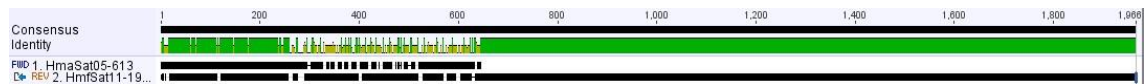
HmaSat01-513, HmfSat04-513 e HmgSat02-513



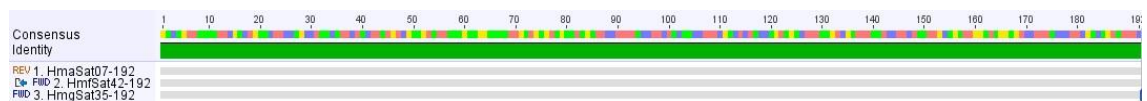
HmaSat02-58 e HmgSat01-58



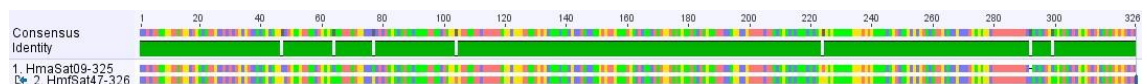
HmaSat05-613 e HmfSat11-1922



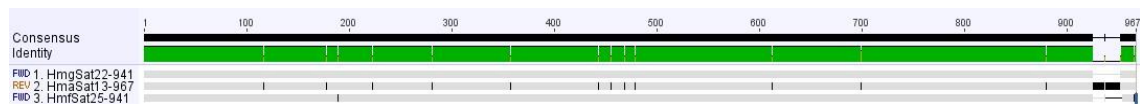
HmaSat07-192, HmfSat42-192 e HmgSat35-192



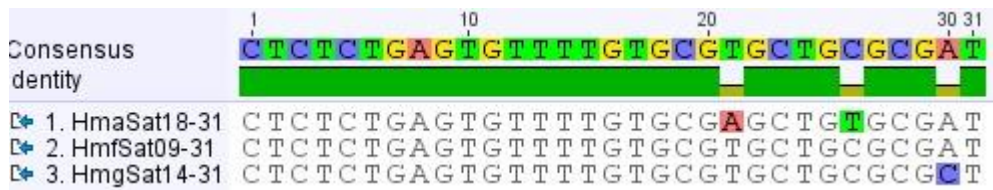
HmaSat09-325 e HmfSat47-326



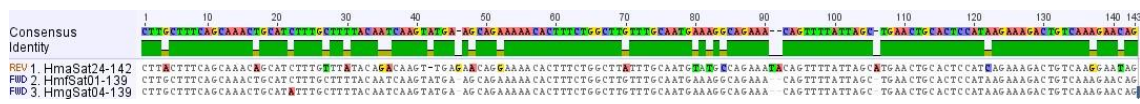
HmaSat13-967, HmfSat25-941 e HmgSat22-941



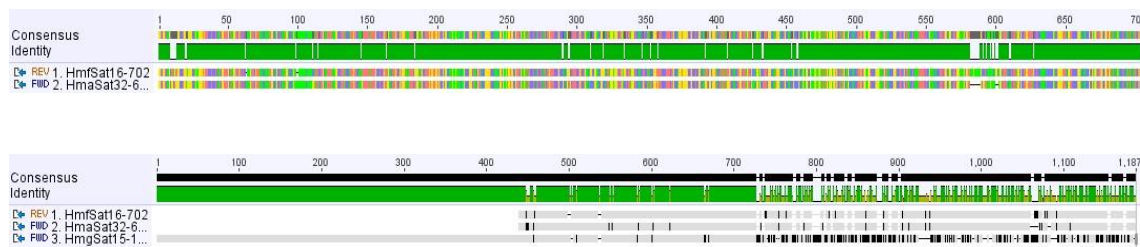
HmaSat18-31, HmfSat09-31 e HmgSat14-31



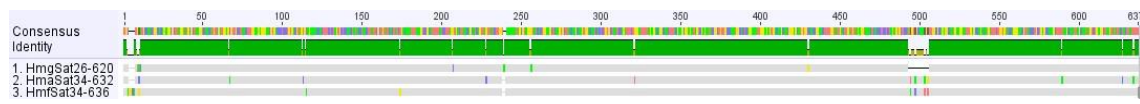
HmaSat24-142, HmfSat01-139 e HmgSat04-139



HmaSat32-695, HmfSat16-702 e HmgSat15-1140



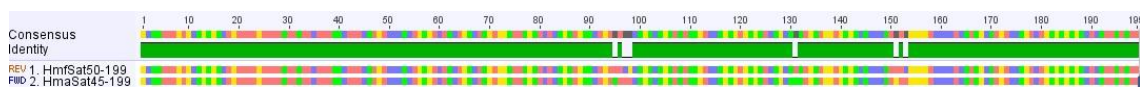
HmaSat34-632, HmfSat34-636 e HmgSat26-620



HmaSat40-129 e HmfSat43-129



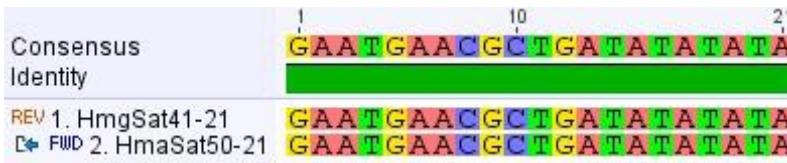
HmaSat45-199 e HmfSat50-199



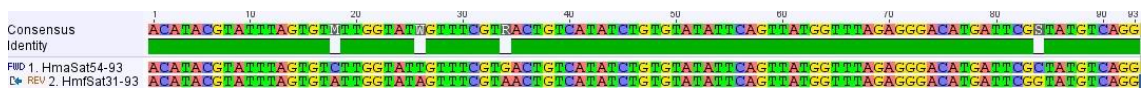
HmaSat46-760 e HmgSat32-827



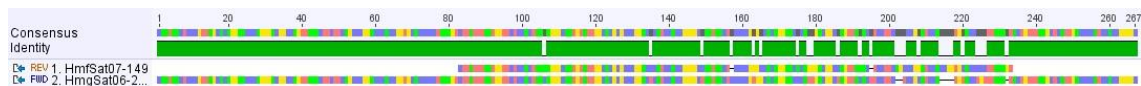
HmaSat50-21 e HmgSat41-21



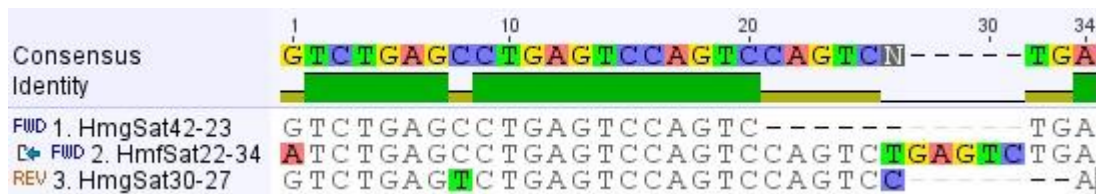
HmaSat54-93 e HmfSat31-93



HmfSat07-149 e HmgSat06-260



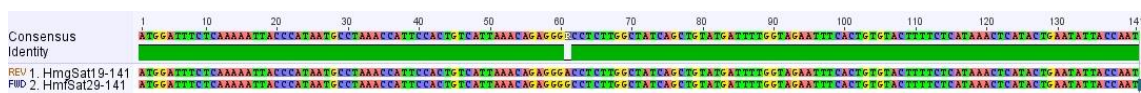
HmfSat22-34, HmgSat30-27 e HmgSat42-23



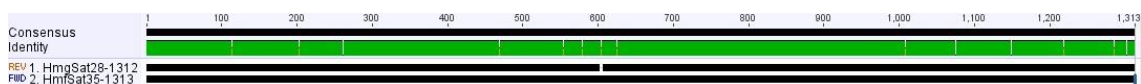
HmfSat28-142 e HmgSat38-142



HmfSat29-141 e HmgSat19-141

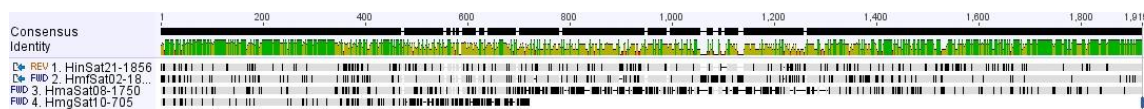


HmfSat35-1313 e HmgSat28-1312

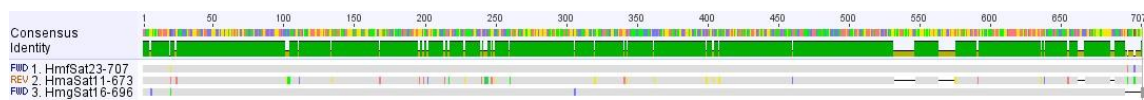


. Satélites compartilhados entre apenas entre o gênero *Hoplias*.

HmaSat08-1750, HmfSat02-1894, HmgSat10-705 e HinSat21-1856



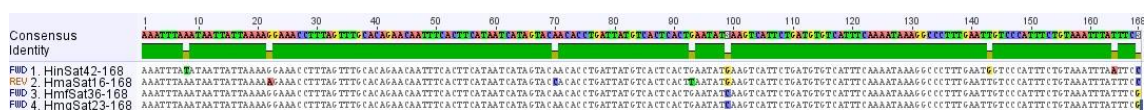
HmaSat11-673, HmfSat23-707, HmgSat16-696, HinSat29-1013



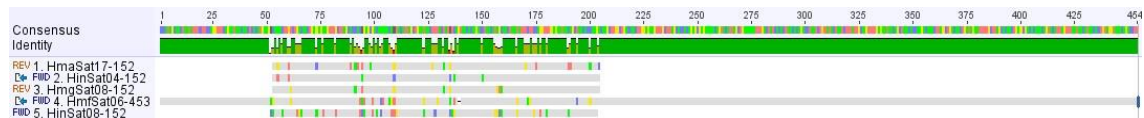
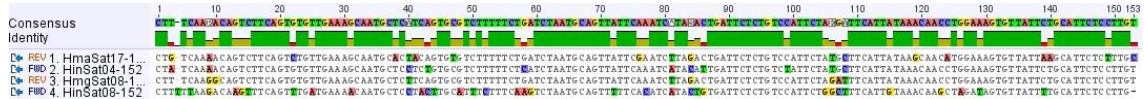
HmaSat15-58, HmfSat12-58, HmgSat13-58, HinSat19-58



HmaSat16-168, HmfSat36-168, HmgSat23-168, HinSat42-168



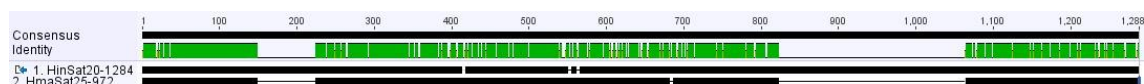
HmaSat17-152, HmfSat06-453, HmgSat08-152 e HinSat04-152/HinSat08-152



HmaSat21-322, HmfSat49-322, HmgSat25-322 e HinSat71-322



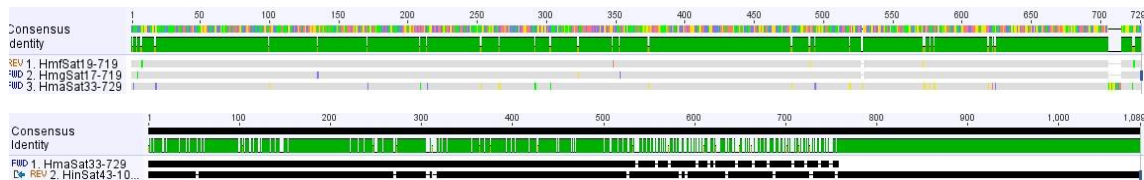
HmaSat25-972 e HinSat20-1284



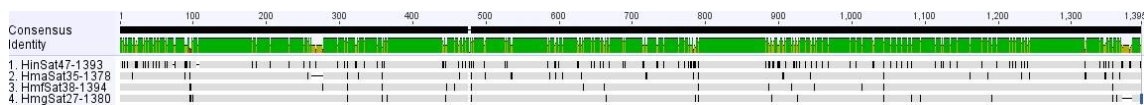
HmaSat30-403, HmfSat32-403, HmgSat34-403 e HinSat39-405



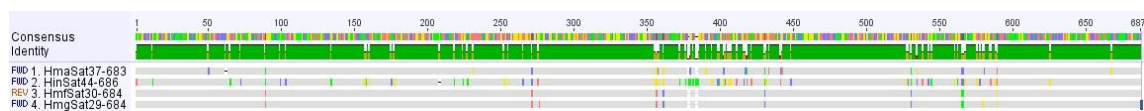
HmaSat33-729, HmfSat19-719, HmgSat17-719 e HinSat43-1062



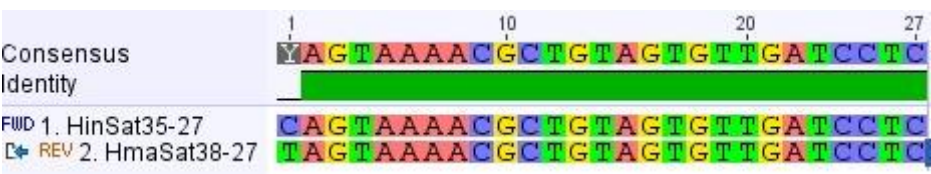
HmaSat35-1378, HmfSat38-1394, HmgSat27-1380 e HinSat47-1393



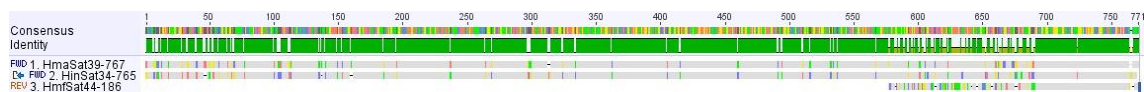
HmaSat37-683, HmfSat30-684, HmgSat29-684 e HinSat44-686



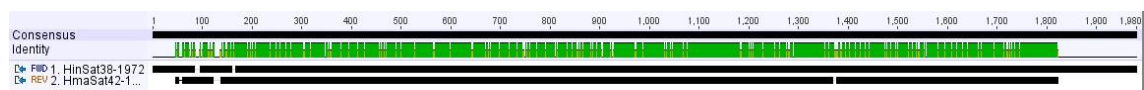
HmaSat38-27 e HinSat35-27



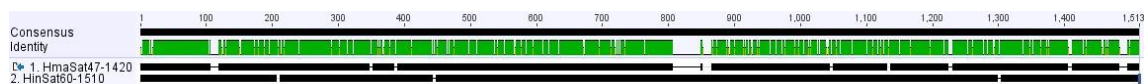
HmaSat39-767, HmfSat44-765 e HinSat34-186



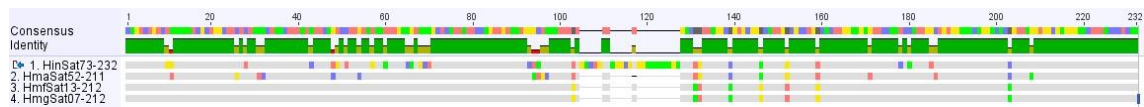
HmaSat42-1756 e HinSat38-1972



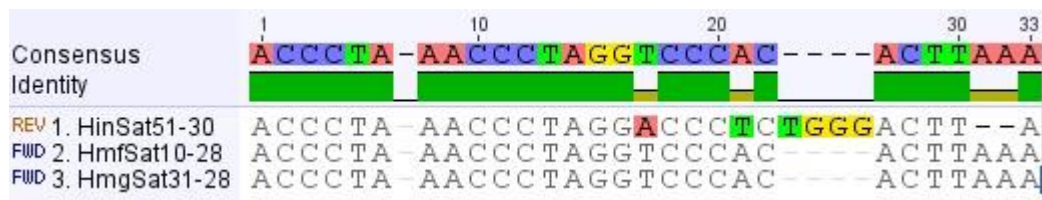
HmaSat47-1420 e HinSat60-1510



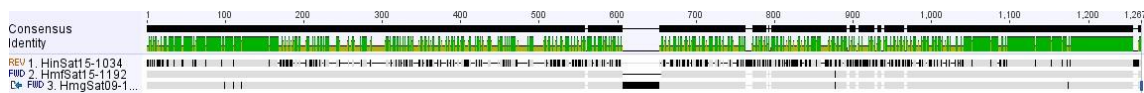
HmaSat52-211, HmfSat13-212, HmgSat07-212 e HinSat73-232



HmfSat10-28, HmgSat31-28 e HinSat51-30



HmfSat15-1192 HmgSat09-1240 e HinSat15-1034



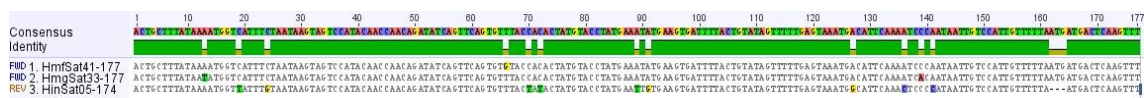
HmfSat17-292, HmgSat12-292, HinSat01-294



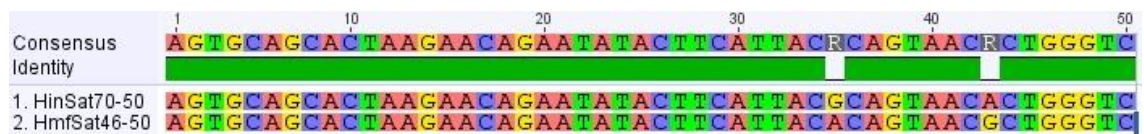
HmfSat26-378 e HinSat58-378



HmfSat41-177, HmgSat33-177 e HinSat05-174



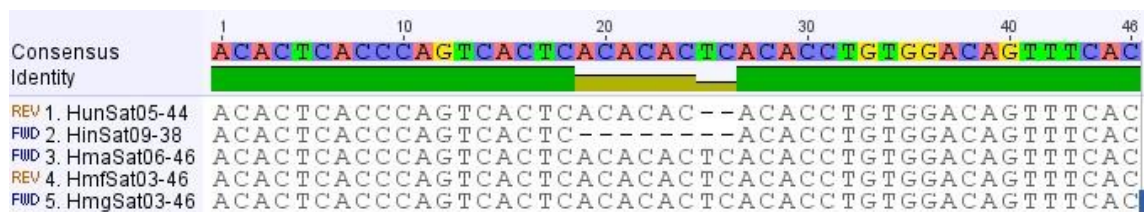
HmfSat46-50 e HinSat70-50



HmgSat37-467 e HinSat31-906



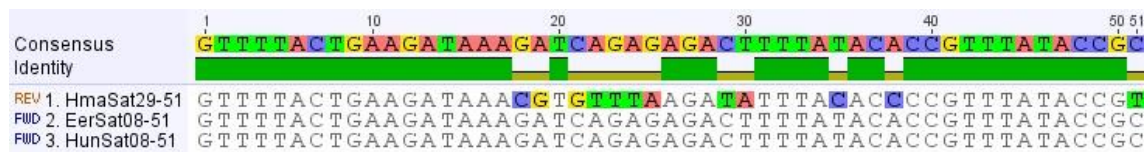
HmaSat06-46, HmfSat03-46, HmgSat03-46, HinSat09-38 e HunSat05-44



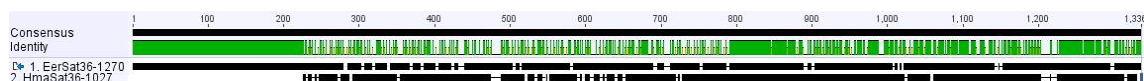
HmaSat19-43, HunSat09-43, EerSat10-43



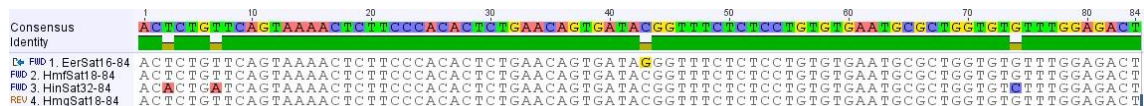
HmaSat29-51, HunSat08-51 e EerSat08-51



HmaSat36-1270 e EerSat36-1027



HmfSat18-84, HmgSat18-84, HinSat32-84, EerSat16-84



HmfSat21-42, HmgSat20-42, HinSat40-42, HunSat12-41 e EerSat09-41

	1	10	20	30	40	42
Consensus	GATAAAGTGGCCAGTGAGTGTTCAGTAGAAGGGGGGTGTTTCT					
Identity						
REV 1. EerSat09-41	GATAAAGTGTCCAGTGAGTGTTCAGTAGA-GGGGGGTGTTTCT					
FWD 2. HunSat12-41	GATAAAGTGGCCAGTGAGTGTTCAGTAGA-GGGGGGTGTTTCT					
FWD 3. HinSat40-42	GATAAAGTGGCCAGTGAGTGTTCAGTAGAAGGGGGGTGTTTCT					
REV 4. HmfSat21-42	GATAAAGTGGCCAGTGAGTGTTCAGTAGAATAAGGGGGGTGTTTCT					
REV 5. HmgSat20-42	AATAAAGTGGCCAGTGAGTGTTCAGTAGATAAAGGGGGGTGTTTCT					

HmfSat39-57, HunSat38-63 e EerSat40-59

	1	10	20	30	40	50	60	63
Consensus	CACACTCTGTGGAGGAGAGACACACACCTCACTAACACACACCTC-----CACACACAATAA							
Identity								
REV 1. HmfSat39-57	CACACTCTGTGGAGGAGAGACACACACCTCACTAACACACACCTC-----CACACACAATAA							
FWD 2. HunSat38-63	CACACTCTGTGGAGGAGAGACACACACCTCACTAACACACACCTCACAACACACACAATAA							
FWD 3. EerSat40-59	CACACTCTGTGGAGGAGAGACACACACCTCACTAACACACACCTC-----CACACACAATAA							

HmfSat40-457 e EerSat37-905

	1	50	100	150	200	250	300	350	400	450	500	550	600	650	700	750	800	850	900	950	970
Consensus																					
Identity																					
FWD 1. HmfSat40-457																					
REV 2. EerSat37-905																					

HinSat23-923, HunSat19-958 e EerSat19-958

	1	50	100	150	173	200	250	300	350	400	450	500	550	600	650	700	750	800	850	900	950	970	
Consensus																							
Identity																							
FWD 1. HunSat19-958																							
REV 2. EerSat19-958																							
FWD 3. HinSat23-923																							

HinSat56-42, HunSat20-42 e EerSat24-42

	1	10	20	30	40	42
Consensus	CTACAGCTGAATCCACTGCAAGAACAACACCAGCACTGACTA					
Identity						
1. HinSat56-42	CTACAGCTGAATCCACTGCAAGGAACAACACCAGCACTGACTA					
2. EerSat24-42	CTACAGCTGAATCCACTGCAAGAACAACACCAGCACTGACTA					
3. HunSat20-42	CTACAGCTGAATCCACTGCAAGAACAACACCAGCACTGACTA					

HinSat36-21, HunSat28-23 e EerSat30-21

	1	10	20	23
Consensus	GC--AAGTGTTCAGTGTAGTGT			
Identity				
FWD 1. EerSat30-21	GC--AAGTGTTCAGTGTAGTGT			
FWD 2. HunSat28-23	GC--AAGTGTTCAGTGTAGTGT			
REV 3. HinSat36-21	GC--AAGTGTTCAGTGTAGTGC			

HinSat62-51 e EerSat46-50

	1	10	20	30	40	50	51
Consensus	KTTNATGGACTAATATCTCCTGTTTTAGTGTAKAKGAGTGTTTATTGGGGTTT						
Identity							
1. EerSat46-50	GT-ATGGACTAATATCTCCTGTTTTAGTGTATAGGAGTGTTTATTGGGGTTT						
2. HinSat62-51	TTAATGGACTAATATCTCCTGTTTTAGTGTAGATGAGTGTTTATTGGGGTTT						

HinSat53-106 e HunSat34-98

	1	10	20	30	40	50	60	70	80	90	100	106	
Consensus	AGTAAATACAGACTCCACTCCGAACTGTAACAGTTATAATCAGTTATAAACAGTGGTCTTCTCAGACTCAGACTGATCTACAGCACAGCGCCCTCTATCTGCTGTAAG												
Identity													
1. HinSat53-106	AGTAAATACAGACTCCACTCCGAACTGTAACAGTTATAATCAGTTATAAACAGTGGTCTTCTCAGACTCAGACTGATCTACAGCACAGCGCCCTCTATCTGCTGTAAG												
2. HunSat34-98	AGTAAATACAGACTCCACTCCGAACTGTAACAGTTATAATCAGTTATAAACAGTGGTCTTCTCAGACTCAGACTGATCTACAGCACAGCGCCCTCTATCTGCTGTAAG												

HunSat01-130 e EerSat02-130

	1	10	20	30	40	50	60	70	80	90	100	110	120	130
Consensus	CACACTAACCTCTGTCGATCAACATTATGAAAAAATAAACAACACCTTTTAGAGCATTCCTGACTTCATAAAACACAGTTCAGTTCCTGCTTATTCCTGATTCAGTTTCTTACTGAAA													
Identity														
1. EerSat02-130	CACACTAACCTCTGTCGATCAACATTATGAAAAAATAAACAACACCTTTTAGAGCATTCCTGACTTCATAAAACACAGTTCAGTTCCTGCTTATTCCTGATTCAGTTTCTTACTGAAA													
2. HunSat01-130	CACACTAACCTCTGTCGATCAACATTATGAAAAAATAAACAACACCTTTTAGAGCATTCCTGACTTCATAAAACACAGTTCAGTTCCTGCTTATTCCTGATTCAGTTTCTTACTGAAA													

HunSat02-141 e EerSat03-141

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	14
Consensus Identity	SAAAGACACTCTCTGTCATTCCTGACTTCATTAAACACATTCCTTTCCTGTTGAAAGCTCCCTGTCCTACGCTTTAATCCGTAATCCTTATCCAGTCAAGACCTCTAGCTTAAGTTAAT														
FWD 1. EerSat03-141	SAAAGACACTCTCTGTCATTCCTGACTTCATTAAACACATTCCTTTCCTGTTGAAAGCTCCCTGTCCTACGCTTTAATCCGTAATCCTTATCCAGTCAAGACCTCTAGCTTAAGTTAAT														
REV 2. HunSat02-141	SAAAGACACTCTCTGTCATTCCTGACTTCATTAAACACATTCCTTTCCTGTTGAAAGCTCCCTGTCCTACGCTTTAATCCGTAATCCTTATCCAGTCAAGACCTCTAGCTTAAGTTAAT														

HunSat04-45 e EerSat04-45

	1	10	20	30	40	45
Consensus	GACACCTGACCTGGTGGTGGTGGTGTGTAGTGTGTGCTCGTTCCTGCGAGTT					
Identity						
1. EerSat04-45	GACACCTGACCTGGTGGTGGTGGTGTGTAGTGTGTGCTCGTTCCTGCGAGTT					
2. HunSat04-45	GACACCTGACCTGGTGGTGGTGGTGTGTAGTGTGTGCTCGTTCCTGCGAGTT					

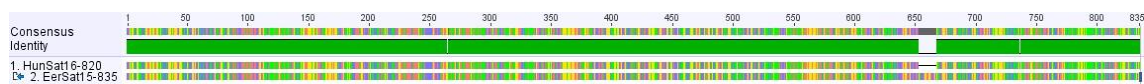
HunSat07-33 e EerSat05-473

	1	25	50	75	100	125	150	175	200	225	250	275	300	325	350	375	400	425	450	473	
Consensus																					
Identity																					
REV 1. EerSat05-473																					
FWD 2. HunSat07-33																					

HunSat14-31 e EerSat07-31

	1	10	20	30	31
Consensus	CGCTTTTCCCTCCCTGTCGTGTCAGAGTCCAAGCC				
Identity					
REV 1. HunSat14-31	CGCTTTTCCCTCCCTGTCGTGTCAGAGTCCAAGCC				
FWD 2. EerSat07-31	CGCTTTTCCCTCCCTGTCGTGTCAGAGTCCAAGCC				

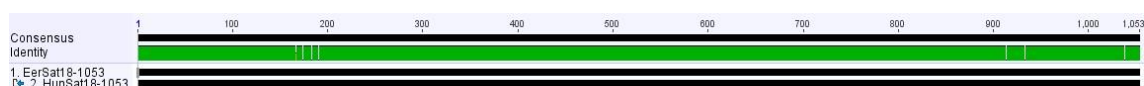
HunSat16-820 e EerSat15-835



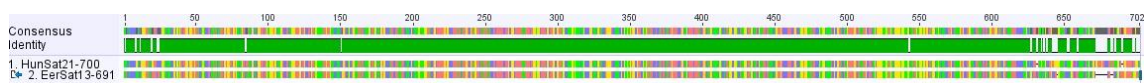
HunSat17-923 e EerSat11-871



HunSat18-1053 e EerSat18-1053



HunSat21-700 e EerSat13-691



HunSat22-62 e EerSat23-62



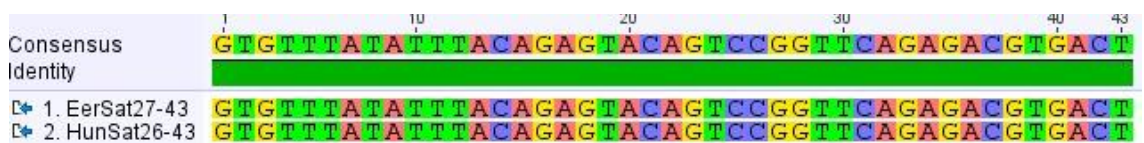
HunSat23-152 e EerSat22-152



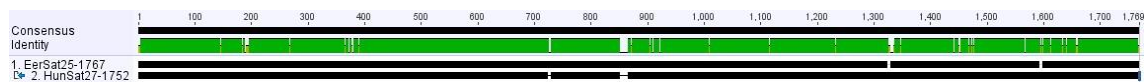
HunSat24-92 e EerSat21-94



HunSat26-43 e EerSat27-43



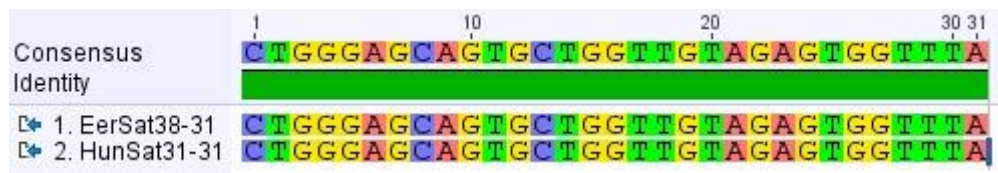
HunSat27-1752 e EerSat25-1767



HunSat29-764 e EerSat29-772



HunSat31-31 e EerSat38-31



HunSat32-427 e EerSat43-427



HunSat37-32 e EerSat47-32

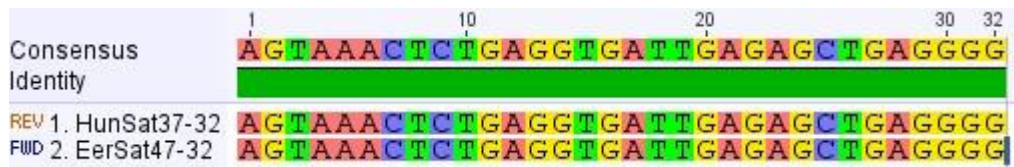


Tabela Suplementar 1 – Satélites compartilhados entre os grupos analisados.

HmaSat	HmfSat	HmgSat	HinSat	HunSat	EerSat
HmaSat20	HmfSat14	HmgSat11	HinSat25	HunSat11	EerSat12
HmaSat23	HmfSat20	HmgSat24	HinSat52	HunSat15	EerSat20
HmaSat41	HmfSat33	HmgSat36	HinSat50	HunSat10	EerSat14
HmaSat48	HmfSat52	HmgSat40	HinSat68	HunSat33	EerSat39
HmaSat51	HmfSat24	HmgSat21	HinSat59	HunSat35*	EerSat33*
<u>HmaSat53</u>	<u>HmfSat45</u>	<u>HmgSat44</u>	<u>HinSat66</u>	<u>HunSat13</u>	<u>EerSat17</u>
HmaSat01	HmfSat04	HmgSat02	-	-	-
HmaSat02	-	HmgSat01	-	-	-
HmaSat05	HmfSat11*	-	-	-	-
HmaSat07	HmfSat42	HmgSat35	-	-	-
HmaSat09	HmfSat47	-	-	-	-
HmaSat13	HmfSat25	HmgSat22	-	-	-
HmaSat18	HmfSat09	HmgSat14	-	-	-
HmaSat24	HmfSat01	HmgSat04	-	-	-
HmaSat32	HmfSat16	HmgSat15*	-	-	-
HmaSat34	HmfSat34	HmgSat26	-	-	-
HmaSat40	HmfSat43	-	-	-	-
HmaSat45	HmfSat50	-	-	-	-
HmaSat46	-	HmgSat32*	-	-	-
HmaSat50	-	HmgSat41	-	-	-
HmaSat54	HmfSat31	-	-	-	-
-	HmfSat07	HmgSat06*	-	-	-
-	HmfSat22	HmgSat30/42*	-	-	-
-	HmfSat28	HmgSat38	-	-	-
-	HmfSat29	HmgSat19	-	-	-
-	HmfSat35	HmgSat28	-	-	-
HmaSat08*	HmfSat02	HmgSat10*	HinSat21	-	-
HmaSat11	HmfSat23	HmgSat16	HinSat29*	-	-
HmaSat15	HmfSat12	HmgSat13	HinSat19	-	-
HmaSat16	HmfSat36	HmgSat23	HinSat42	-	-
HmaSat17	HmfSat06*	HmgSat08	HinSat04/HinSat08*	-	-
HmaSat21	HmfSat49	HmgSat25	HinSat71	-	-
HmaSat25	-	-	HinSat20	-	-
HmaSat30	HmfSat32	HmgSat34	HinSat39	-	-
HmaSat33	HmfSat19	HmgSat17	HinSat43*	-	-
HmaSat35	HmfSat38	HmgSat27	HinSat47	-	-
HmaSat37	HmfSat30	HmgSat29	HinSat44	-	-
HmaSat38	-	-	HinSat35	-	-
HmaSat39	HmfSat44*	-	HinSat34	-	-

HmaSat42	-	-	HinSat38	-	-
HmaSat47	-	-	HinSat60	-	-
HmaSat52	HmfSat13	HmgSat07	HinSat73*	-	-
-	HmfSat10	HmgSat31	HinSat51/72*	-	-
-	HmfSat15	HmgSat09	HinSat15*	-	-
-	HmfSat17	HmgSat12	HinSat01	-	-
-	HmfSat26	-	HinSat58	-	-
-	HmfSat41	HmgSat33	HinSat05	-	-
-	HmfSat46	-	HinSat70	-	-
-	-	HmgSat37	HinSat31*	-	-
HmaSat06	HmfSat03	HmgSat03	HinSat09	HunSat05	-
HmaSat19	-	-	-	HunSat09	EerSat10
HmaSat29*	-	-	-	HunSat08	EerSat08
HmaSat36	-	-	-	-	EerSat36
-	HmfSat18	HmgSat18	HinSat32	-	EerSat16
-	HmfSat21	HmgSat20	HinSat40	HunSat12	EerSat09
-	HmfSat39	-	-	HunSat38	EerSat40
-	HmfSat40	-	-	-	EerSat37
-	-	-	HinSat23*	HunSat19	EerSat19
-	-	-	HinSat56	HunSat20	EerSat24
-	-	-	HinSat36	HunSat28	EerSat30
-	-	-	HinSat62	-	EerSat46
-	-	-	HinSat53	HunSat34	-
-	-	-	-	HunSat01	EerSat02
-	-	-	-	HunSat02	EerSat03
-	-	-	-	HunSat04	EerSat04
-	-	-	-	HunSat07	EerSat05*
-	-	-	-	HunSat14	EerSat07
-	-	-	-	HunSat16	EerSat15
-	-	-	-	HunSat17	EerSat11
-	-	-	-	HunSat18	EerSat18
-	-	-	-	HunSat21	EerSat13
-	-	-	-	HunSat22	EerSat23
-	-	-	-	HunSat23	EerSat22
-	-	-	-	HunSat24	EerSat21
-	-	-	-	HunSat26	EerSat27
-	-	-	-	HunSat27	EerSat25
-	-	-	-	HunSat29	EerSat29
-	-	-	-	HunSat31	EerSat38
-	-	-	-	HunSat32	EerSat43
-	-	-	-	HunSat37	EerSat47

Tabela Suplementar 2 – Primers dos satélites mapeados por FISH.

satDNA family	F	R
Hmf01	CAATGAAAGGCAGAAACAGT	GATTGTAAAAGCAAAGATGC
Hmf02	TGGGTCCAGACAAATCAAAA	TCATAATGATTGGACCTTGT
Hmf04	GTCTCAGATTTCTTCCAAAC	GGTCCTGACGAAGACAGT
Hmf06	CTTCACTCTAGTTGTGTTTC	GATCAGAAAGAGACGCAC
Hmf12	TACACACACAGCTCCTCG	ACTGATGTAGTTCACTGTAG
Hmf13	CAGGCAACACATTTTATTCA	TGTCTGTCAGAAGTGGGATT
Hmf14	ACACCAGAGGAACCCCAT	AGTGGAAGTGTGTTCTCT
Hmf24	GACTCTTCAATTTACACCTG	TCAGTGTTCTTCTGTCATAC
Hmf33	TTATGGCCTTCAGCATTCTG	AAACTGTTACCAGACCACTC
Hmf42	ACCAACATTAAGCTGTAAGG	ATTGATGTTAGAATGGCAGG
Hmf45	AGTTACAGCACCGAAATACT	GTAAACGGCTGTAGCTGG
Hmf52	AGAACACACACCGCTGAA	ACCCGGTTCTATCAGAGC
Hmg08	AATGGACAGAGAATCAGTCT	TCTTCAGTGTGTTGAAAGCA
Hun15	TCGGAGGAGAACTAAC	GTGTGGAGTCTGATGAACAG
Hun21	AGCCAGGTTCAATTCACACT	TTGCCGCTTGTTTTTGTGGT
Eer02	TTTAGAGCATTCTGTAGTT	AGGAAAAGTGAATCAAGCAA
Eer03	AACACATTTGCTTTTGCCTG	AATCAAAGTGAAGGAATGCA