



UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOCÊNCIAS

ESTRUTURA E ORGANIZAÇÃO DE DNAs SATÉLITES NO GENOMA DO LAMBARI
Psalidodon paranae (TELEOSTEI, CHARACIFORMES)

Aluno: Rodrigo Zeni dos Santos
Orientador: Ricardo Utsunomia
Co-Orientador: Fábio Porto Foresti

Bauru, SP

2022

RODRIGO ZENI DOS SANTOS

ESTRUTURA E ORGANIZAÇÃO DE DNAs SATÉLITES NO GENOMA DO LAMBARI

Astyanax paranae (TELEOSTEI, CHARACIFORMES)

Dissertação apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Assis, para a obtenção do título de Mestre em Biociências (Área de Conhecimento: Caracterização e Aplicação da Diversidade Biológica)

Orientador: Dr. Ricardo Utsunomia

Co-Orientador: Dr. Fabio Porto Foresti

Bolsista: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

Bauru

2022

S237e

Santos, Rodrigo Zeni dos

Estrutura e organização de dnas satélites no genoma do lambari
Psalidodon paranae (Teleostei, Characiformes) / Rodrigo Zeni dos
Santos. -- Bauru, 2022

76 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp),
Faculdade de Ciências, Bauru

Orientador: Ricardo Utsunomia

Coorientador: Fábio Porto Foresti

1. Genética animal. 2. Citogenética. 3. Bioinformática. I. Título.

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE RODRIGO ZENI DOS SANTOS, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM BIOCÊNCIAS, DA FACULDADE DE CIÊNCIAS - CÂMPUS DE BAURU.

Aos 01 dias do mês de julho do ano de 2022, às 14:00 horas, por meio de Videoconferência, realizou-se a defesa de DISSERTAÇÃO DE MESTRADO de RODRIGO ZENI DOS SANTOS, intitulada **ESTRUTURA E ORGANIZAÇÃO DE DNAs SATÉLITES NO GENOMA DO LAMBARI *Psalidodon paranae* (TELEOSTEI, CHARACIFORMES)**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. RICARDO UTSUNOMIA (Participação Virtual) do(a) UFRRJ/Rio de Janeiro, Prof. Dr. ROBERTO FERREIRA ARTONI (Participação Virtual) do(a) Biologia Estrutural Molecular e Genética / UEPG, Prof. Dr. DUILIO MAZZONI ZERBINATO DE ANDRADE SILVA (Participação Virtual) do(a) Departamento de Morfologia / UNESP/Botucatu. Após a exposição pelo mestrando 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. RICARDO UTSUNOMIA

1

DEDICATÓRIA

2

Dedico este trabalho a minha mãe, Tânia, aos meus tios, Carlos e Ana e a minha prima,

3

Júlia.

4 Agradecimentos

5 A Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP, por toda
6 minha formação ao longo dos quatro anos de graduação e dois de mestrado.

7 Agradeço a família, Tânia, Carlos, Ana e Júlia pelo amor e incentivo em todos os
8 rumos que decidi trilhar em minha vida, sempre me dando todo o apoio necessário em
9 minhas escolhas e sempre estando ao meu lado tanto nos momentos ruins quanto nos
10 bons, amo muito todos vocês.

11 Ao meu orientador Prof. Dr. Ricardo Utsunomia e coorientador Prof. Dr. Fábio
12 Porto Foresti, que me acolheram desde minha IC e me inseriram no mundo da genética
13 e bioinformática, por toda a ajuda e ensinamento que eles constantemente me deram,
14 me tornando no pesquisador que hoje sou e sendo ótimos educadores e pesquisadores,
15 tornando fácil de trilhar nesse caminho, com dois ótimos exemplos ao meu lado.

16 Ao Caio Goes, que desde que entrei no LAGENPE está sempre ao meu lado, me
17 passando todo o aprendizado que conhece, estando sempre pronto pra ajudar sem
18 pensar duas vezes e por ser uma das pessoas mais incríveis que eu conheço.

19 A minha melhor amiga Yasmin Garcia, que me aguenta desde 2016 e partir de
20 2018 passou a me aguentar todos os dias o dia todo e sempre me ajudou e me motivou,
21 muito obrigado por estar sempre ao meu lado. É nítido quando sabemos que alguém
22 será amigo para a vida toda e você é um exemplo perfeito disso.

23 A minha companheira de laboratório, companheira de saída a restaurantes,
24 companheira para conversar o dia todo sem perder o assunto, companheira de séries e
25 filmes, companheira para a elaboração de infinitos projetos futuros, companheira de
26 projetos de academia e por fim, minha companheira, Giovana Funari. Muito obrigado
27 por ser essa pessoa incrível que você é e me dar forças todos os dias para continuar
28 sorrindo, ter te conhecido foi muito melhor do que qualquer outra coisa que eu poderia
29 pedir, estar ao seu lado torna qualquer problema algo tranquilo de se resolver, muito
30 obrigado por me fazer tão bem, te amo.

31 Aos meus dois ex companheiros de trabalho, Pedrão e Milan, muito obrigado
32 por todos os trabalhos realizados juntos, por todos os aprendizados que desenvolvemos
33 durante esses anos trabalhando e por companhias tão incríveis, tornando até os
34 momentos de estresse possíveis de lidar. Aprendi muito com ambos, vocês são pessoas
35 maravilhosas e desejo o melhor na vida de ambos, sei do potencial que vocês têm e sei
36 que vão conseguir trilhar a vida na área de atuação que desejarem, muito obrigado pelo
37 nosso tempo junto.

38 A todos os demais integrantes e ex integrantes do grupo do LAGENPE, todos
39 vocês fizeram os meus dias de trabalho algo muito mais alegre e divertido, todos os

momentos de café, almoço, saídas que demos com o grupo e claro, momentos de trabalho, se tornaram muito melhores com companhia incríveis como as encontradas neste laboratório.

E por último, mas não menos importante, meus grandes amigos, que mantemos contatos diariamente apesar da distância, Ramiro, Lúcio, Tobias, Jota e Barney. Todos vocês têm um espaço importante na minha vida e eu sou muito grato pela amizade de cada um, sei que posso confiar em vocês para qualquer necessidade e também sei que nossas amizades não se quebrarão nem com distância e nem com o tempo, amo todos vocês.

68 **Resumo**

69 Os genomas dos organismos eucariotos estão repletos de sequências de DNA
70 repetitivos. Dentre eles, os DNAs satélites (satDNAs) constituídos em arranjos de
71 sequências repetidas em *tandem*. Historicamente, técnicas distintas foram utilizadas
72 para o isolamento e caracterização destes satDNAs, porém o advento de tecnologias de
73 Sequenciamento de Nova Geração (NGS), possibilitou a utilização de *reads* curtos, que
74 juntos a ferramentas bioinformáticas, foi decisivo para estudos que permitiram a
75 caracterização de coleções de satDNAs (satelitomas) de uma espécie. O primeiro
76 satelitoma caracterizado em peixes foi realizado na espécie *Psalidodon paranae*, com a
77 utilização da *pipeline RepeatExplorer1* e revelou uma quantidade significativa deste tipo
78 de sequência no genoma deste organismo, incluindo várias famílias presentes no
79 cromossomo B desta espécie. Com isso, o objetivo inicial do estudo foi utilizar de uma
80 nova *pipeline*, o TAREAN, para uma reavaliação do satelitoma da espécie *P. paranae* e
81 uma comparação entre diferentes *pipelines* usadas para a caracterização destas
82 sequências, o que resultou em 56 novas famílias caracterizadas para esta espécie. Com
83 o sequenciamento de seis novas bibliotecas de *reads* curtos de *P. paranae*, três
84 indivíduos contendo cromossomo B e três indivíduos não contendo, foram
85 apresentados dados de abundância de todo este satelitoma, com enfoque em satélites
86 presentes no cromossomo B. Os resultados evidenciaram um satélite caracterizado por
87 TAREAN mais abundantes que o ApaSat02, que até o atual trabalho era o segundo mais
88 abundante da espécie, novos quatro DNAs satélites mais abundantes que o ApaSat03,
89 além de novos satélites muito presentes nos cromossomos B da espécie, apresentando
90 essa *pipeline* como um grande auxílio para a expansão dos satelitomas anteriormente
91 caracterizados. Além disso, uma análise mais detalhada em *reads* longos possibilitou a
92 estudo de como cada DNA satélite do catálogo é estruturado ao longo do genoma de
93 *Psalidodon paranae*, apresentando seu tamanho médio e o valor máximo atingindo de
94 cada *arrays* destes DNAs satélites, auxiliando na explicação destes possuírem ou não
95 uma sinais a nível cromossômico e nos apontando futuros candidatos a aplicação da
96 técnica de FISH.

97

98 **Abstract**

99 The genomes of eukaryotic organisms have a large number of repetitive DNA
100 sequences. Among them, the satellite DNAs (satDNAs) constitute arrays of tandem
101 repeated sequences. Historically, different techniques were used for the
102 characterization of satDNAs, the advent of Next Generation Sequencing (NGS)
103 technologies enabled the use of short reads, with bioinformatics tools, was decisive
104 for studies that allowed the characterization collections of satDNAs (satellitomes) in
105 a specie. The first fish satellitome characterized was for the species *Psalidodon*
106 *paranae*, characterized using the pipeline RepeatExplorer1 and revealed a
107 significant amount of this type of sequence in the genome of this organism,
108 including several families present in the B chromosome of this species. Thus, the
109 initial objective of this study was using a new pipeline, TAREAN, to expand the
110 satellite of the *Psalidodon paranae* and a comparison between different pipelines
111 used for the characterization of these sequences, which resulted in 56 new families
112 characterized for this species. We perform the sequencing of six new libraries of
113 short reads of *P. paranae*, three individuals containing B chromosome and three
114 individuals not containing, abundance of all this satellitome were calculated,
115 focusing on satellites present in the B chromosome. Resulting in a satellite
116 characterized by TAREAN more abundant than ApaSat02, until the current work the
117 second most abundant of this species, in addition to four others more abundant
118 than ApaSat03, and new satellites very present in the B chromosomes of the
119 species, contributing for the expansion of previously characterized satellites. In
120 addition, a detailed analysis of long reads was performed and presented the
121 average size and the maximum value reached for each array of these satellite DNAs
122 of the *Psalidodon paranae* species, helping to explain the chromosomal signaling of
123 these satellite DNAs.

124

125

126

127 Sumário

128

129	1	Introdução	10
130	1.1	Apresentação das sequências repetitivas com enfoque em DNAs satélites	10
131	1.2	Montagens de Satelitomas	12
132	1.3	Ordem dos Characiformes, com foco na espécie <i>Psalidodon paranae</i>	15
133	2	Objetivos.....	17
134	3	Material e métodos	17
135	3.1	Estimulação de mitoses	18
136	3.2	Obtenção dos cromossomos metafásicos mitóticos	18
137	3.3	Extração do DNA genômico.....	19
138	3.4	Sequenciamento massivo (Next Generation Sequencing – NGS)	19
139	3.5	Análises bioinformáticas	20
140	3.5.1	Análise estrutural detalhada do satelitoma de <i>Psalidodon paranae</i>	20
141	3.7	Hibridação <i>in situ</i> fluorescente (FISH).....	22
142	4	Capítulo I.....	24
143		Ampliação na caracterização de DNAs satélites em <i>Psalidodon paranae</i> (Teleostei;	
144		Characiformes) e o estudo de suas abundâncias.....	24
145	4.1	Introdução	24
146	4.2	Material e métodos	25
147	4.2.1	Amostras.....	25
148	4.2.2	Extração de DNA e sequenciamento do genoma	26
149	4.2.3	Caracterização do satelitoma por TAREAN	26
150	4.2.4	Estudo do satelitoma em <i>reads</i> longos.....	28
151	4.2.5	Hibridização <i>in situ</i> fluorescente – FISH.....	28
152	5	Resultados	28
153	5.1	Satelitoma de <i>Psalidodon paranae</i>	28
154	5.2	Distribuição de <i>arrays</i> de satDNAs em <i>reads</i> longos de <i>P. paranae</i>	32

155	5.3	Distribuição cromossômica de dez novas famílias de satélite em <i>Psilododon paranae</i>	
156		36	
157	5.4	Estudos comparativos entre bibliotecas OB/1B em <i>Psilododon paranae</i>	39
158	6	Discussão	45
159	6.1	Expansão do satelitoma de <i>Psilododon paranae</i> e um comparativo entre as	
160		ferramentas <i>RepeatExplorer1</i> e TAREAN	45
161	6.2	Análise do satelitoma de <i>Psilododon paranae</i> em <i>reads</i> longos.....	48
162	6.3	Estudo de abundância em bibliotecas de <i>reads</i> curtos OB e 1B	50
163	7	Referências	52
164	8	Material Suplementar	61
165			
166			
167			

1 Introdução

1.1 Apresentação das sequências repetitivas com enfoque em DNAs satélites

Os genomas eucarióticos são compostos por grandes quantidades de sequências repetitivas com características altamente dinâmicas que podem estar dispersas pelo genoma, como os elementos transponíveis (Yuan, 2018), ou organizadas em *tandem*, como as famílias multigênicas e os DNAs satélites (Charlesworth et al. 1994; Garrido-Ramos 2017). Com exceção das famílias multigênicas, estas porções dos genomas não parecem exercer uma função delimitada ou específica. Apesar de poucos dados em relação à expressão, estudos a partir dos anos 2000, revelaram que alguns poucos DNAs satélites podem desempenhar funções específicas, como formação da heterocromatina (Lee et al. 2006; Traynor et al. 2019; Otake et al. 2020), organização do centrômero (Plohl et al. 2014; Escudeiro et al. 2019; Louzada et al. 2020; Vozdova et al. 2020), além de também participaram no processo de evolução dos genomas, principalmente pela sua alta dinâmica de expansão e retração em curtos períodos de tempo (Gregory 2005).

Os DNAs satélites (satDNAs) constituem uma fração não codificante do genoma dos organismos e consistem em longos *arrays* de sequências monoméricas, repetidas em *tandem*, estando localizados preferencialmente em regiões pericentroméricas e subteloméricas, embora também possam ocorrer nas regiões intersticiais dos cromossomos (López-Flores and Garrido-Ramos 2012; Plohl et al. 2012). Estas sequências podem constituir diferentes famílias, que variam em localização, composição, tamanho da unidade de repetição e abundância (Garrido-Ramos 2015). Nesse sentido, por representarem segmentos genômicos altamente dinâmicos e suscetíveis a rápidas mudanças, estas sequências são geralmente espécie- ou gênero-específicas (Vicari et al. 2010; Garrido-Ramos et al. 2015).

A evolução a longo prazo dos satDNAs segue caminhos delimitados pela “hipótese *library*” de evolução (Fry e Salser 1977), que é fundamentada na ideia de que as espécies apresentam uma biblioteca (do inglês, *library*) permanente de DNAs satélites. Esta biblioteca pode ser amplificada de modo independente nos genomas de espécies diferentes, causando uma mudança rápida no perfil de distribuição e abundância do DNA satélite até mesmo em espécies próximas (Plohl et al. 2012). Ruiz-Ruano et al. (2016) complementaram essa proposta, propondo que o nascimento de um satDNA se dá pela duplicação de alguma sequência em um local do genoma,

formando um pequeno *array* de sequências; então, este *array* é disseminado para diversos locais do genoma e podem passar por processos locais de amplificação, resultando no nascimento de um novo satDNA na espécie.

Uma vez que um determinado satDNA esteja disseminado na forma de *arrays*, estas sequências passam por processos de homogeneização intragenômica das mutações, que ocorrem ao longo das repetições, sendo que sua fixação em grupos de organismos reprodutivamente relacionados se dá por um processo estocástico chamado “impulso molecular” (ou *molecular drive*, do inglês) (Dover 1986, 2002). Esse processo é comumente chamado de “evolução em concerto” e é uma combinação de processos intragenômicos (homogeneização) e populacional (fixação) (Dover 2002).

O termo “DNA satélite” teve origem pela descoberta desses elementos em um processo de ultracentrifugação de CsCl, no qual os satDNAs eram apresentados como um pequeno pico neste perfil (Kit 1961). Nas décadas seguintes novas técnicas para a caracterização de DNAs satélites foram criadas, como a cinética de renaturação (Hutton et al. 1973; Botchan, 1974) e a digestão com enzimas de restrição, sendo essa a principal forma de caracterização de satDNAs até a introdução de metodologias genômicas (Horz & Zachau, 1977; Kopecka et al. 1978).

Estes métodos, apesar de precisos, possuíam uma baixa eficiência devido ao caráter aleatório de se caracterizar satélites, pois dependia que um determinado DNA satélite desconhecido possuísse um sítio de restrição da enzima utilizada no experimento. Além disso, mesmo que o resultado por enzimas de restrição fosse positivo, deve-se considerar que os satDNAs são sequências extremamente dinâmicas e com rápidas mutações, tornando-os, na maioria dos casos, espécie- ou gênero-específicos, impossibilitando o estudo deles em espécies adicionais na maioria dos casos. Apesar destas ressalvas, marcadores cromossômicos importantes em peixes Neotropicais foram descobertos por este método, como o caso do DNA satélite As51 (Mestriner et al. 2000), SATH1 e SATH2 (Jesus et al. 2003) e Pph2004 (Vicente et al. 2003).

A partir dos anos de 2013, com a expansão da obtenção de dados de sequenciamento de nova geração, bem como o desenvolvimento de ferramentas bioinformáticas adequadas, os estudos em DNAs satélites foram positivamente impactados. Destaca-se, o desenvolvimento da pipeline *RepeatExplorer1* (Novák et al.

2013), que possibilitou a caracterização de coleções de satDNAs (satelitomas) de qualquer organismo não-modelo, através do sequenciamento de baixa cobertura e análises bioinformáticas. Esse cenário favorável refletiu-se no aumento do número de espécies com satelitomas caracterizados (Ruiz-Ruano et al. 2016; Silva et al. 2017; Utsunomia et al. 2019; Crepaldi & Parise et al. 2020; Serrano-Freitas et al. 2020; Stornioli et al. 2021; Kretschmer et al. 2022) e em um maior entendimento sobre a evolução destas sequências.

No entanto, destaca-se a importância de também avançar no conhecimento sobre a biologia estrutural destes DNAs satélites, sendo necessária a utilização de *reads* longos provenientes de sequenciamento de 3ª geração (Miga 2015; Khost et al. 2017; Phillippy 2017; Harris et al. 2019). Este tipo de biblioteca sequenciada está disponível de modo mais restrito, mas claramente permite uma análise muito mais consistente ao possibilitar a caracterização destes elementos dentro dos seus respectivos *arrays* (Krsticevic et al. 2015; VanBuren et al. 2015; Khost et al. 2017; Cechova et al. 2019).

1.2 Montagens de Satelitomas

O conjunto caracterizado de DNAs satélites em uma espécie é denominado satelitoma (Ruiz-Ruano et al. 2016). Os primeiros satelitomas de peixes Neotropicais (Silva et al. 2017; Utsunomia et al. 2019; Crepaldi & Parise 2020; Serrano-Freitas et al. 2020) foram caracterizados utilizando a *pipeline RepeatExplorer1* (Novák 2010). Este *software* utiliza *shorts reads* para a montagem de *contigs* (determinados pelo consenso entre a sobreposição de *reads* com alto grau de similaridade).

Cada *contig* é representado e associado a um gráfico, onde os vértices representam os *reads* e as arestas representam sobreposição de semelhança entre os *reads* (95% de similaridade e 55% de cobertura), formando um *layout*. Como os satDNAs são sequências *in tandem*, os *layouts* dos seus gráficos tendem a possuir um formato esférico/anelar (Figura 1). A forma do gráfico gerado depende da relação entre tamanho do *read* e do monômero, sendo que o gráfico terá forma de anel quando o *read* for menor do que o monômero do satélite e esférico quando o tamanho do *read* for maior que o monômero do satélite, devido a mais de uma sobreposição deste satélite em um mesmo *read* (Figura 1). Nesse sentido, a determinação de um DNA

satélite é feita com base na análise manual do *layout* destes gráficos no *RepeatExplorer* 1.

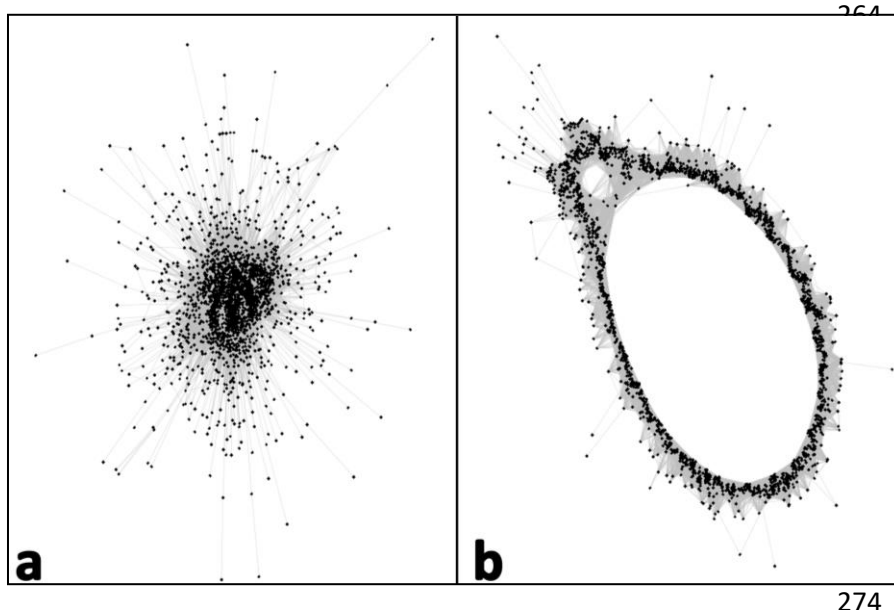


Figura 1. Gráficos resultantes do *RepeatExplorer1*. Em (a), gráfico esférico, resultante da montagem de um satDNA cujos monômeros são menores do que os *reads*, em (b), gráfico anelar, resultante da montagem de um satDNA cujos monômeros são maiores do que os *reads*.

Após a finalização da *pipeline* do *RepeatExplorer1*, é necessária uma varredura em todos os gráficos de formato esférico e anelar obtidos, candidatos a um DNA satélite. Posteriormente, é necessário averiguar individualmente os *contigs* associados a cada gráfico. Essa etapa é realizada manualmente e sua finalização resulta no monômero do possível DNA satélite pertencente ao gráfico e suas variantes. Com os monômeros apresentados, realiza-se uma filtragem da biblioteca original, utilizando o *software* *deconseq* (Schmieder & Edwards, 2011) (https://github.com/fjruizruano/ngs-protocols/blob/master/deconseq_run.py). Esse processo é repetido até nenhum novo satélite ser encontrado na biblioteca genômica da espécie analisada e cada uma destas repetições é conhecida como uma iteração.

Recentemente, Novák et al. (2017), apresentaram uma nova ferramenta para a caracterização de DNAs satélites, também utilizando *shorts reads* e conhecida como TAREAN (*tandem repeat analyser*) (Novák et al. 2017). Uma ferramenta mais automatizada quando comparada ao *RepeatExplorer1*, uma vez que após a montagem

inicial dos *clusters*, uma decomposição em *k-mers* é realizada em todos os gráficos circulares, permitindo a determinação automática da sequência consenso do DNA satélite correspondente. Além disso, uma opção extra denominada “reports” nos apresenta variantes do satélite encontrado. Nesta abordagem, também é necessária a repetição da etapa de filtragem dos satélites encontrados e uma nova realização da *pipeline* TAREAN, até uma completa varredura nos satélites da biblioteca em estudo. As etapas posteriores para a construção e caracterização de todo o satelitoma segue o protocolo do *satMiner* (Ruiz-Ruano et al. 2016) para sua finalização.

Com o desenvolvimento do TAREAN, novos satelitomas foram caracterizados em peixes Neotropicais, como nas espécies *Prochilodus lineatus* (Stornioli et al. 2021) e *Triportheus auritus* (Kretschmer et al. 2022) além de diversos outros grupos, como a ostra *Crassostrea gigas* (Tujic et al. 2021), as moscas *Drosophila virilis* e *D. americana* (Silva et al. 2019), ou o triatomíneo *Rhodnius prolixus*, em que o satelitoma foi caracterizado utilizando TAREAN e RepeatExplorer2 (Montiel et al. 2021).

Além da descrição da biblioteca de satélites de uma espécie, a caracterização de um satelitoma dá suporte para outras análises, por exemplo o desenvolvido por Utsunomia et al. (2019) em *Megaleporinus macrocephalus*. Neste caso, foi possível observar uma maior concentração de DNAs satélites em fêmeas da espécie (1.46% do genoma a mais em relação ao macho), com três satélites exclusivos da biblioteca das fêmeas, além do satélite mais abundante da espécie, o MmaSat01, ser apresentado em uma proporção maior no *landscape* feminino. Em adição a estes dados, 14 dos 30 DNAs satélites analisados por FISH apresentaram sinais somente no cromossomo W da espécie, possibilitando a criação de ferramentas não invasivas para sexagem de indivíduos de *M. macrocephalus*.

Seguindo a abordagem de DNAs satélites em cromossomos sexuais, Kretschmer et al. (2022), realizaram 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. Na técnica de FISH foram revelados 15 DNAs satélites W-específicos em *T. auritus*, evidenciando um acúmulo significativo de DNAs satélites neste cromossomo. O FISH comparativo entre as três espécies revelou a trajetória do cromossomo sexual para este grupo, visto que esta família possui uma origem em comum em seus cromossomos sexuais ZW.

Além do auxílio para a diferenciação cromossômica, o estudo de satelitoma, como apresentado por Stornioli et al. (2021) também é de grande auxílio para estudos em cromossomos supranumerários, ou cromossomos B. Na espécie *Prochilodus lineatus*, foram caracterizadas 51 famílias de DNAs satélites. O mapeamento por FISH revelou 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 foi observado exclusivamente nos cromossomos B e nos dois autossômicos do par 4 (Stornioli et al. 2021).

1.3 Ordem dos Characiformes, com foco na espécie *Psalidodon paranae*

A ictiofauna Neotropical, apresenta uma grande biodiversidade de peixes, com pelo menos 6200 espécies válidas (Reis et al. 2016; Albert et al. 2020). Dentre as mais diversas ordens que pertencem a este grupo, o grupo dos Characiformes se destaca pela riqueza de espécies, com cerca de 2200 espécies classificadas em 24 famílias (Eschmeyer & Fong 2020; Melo et al. 2021), com representantes dessa ordem encontrados também em outras regiões, como África subsaariana e América do Norte (Moreira 2007).

O gênero *Psalidodon* (Characiformes, Characidae) engloba diversas espécies dos popularmente conhecidos “lambaris”, e está inserido na ordem dos Characiformes (família Characidae). Este gênero representa um grupo com relevância para seres humanos, uma vez que é utilizado como recurso alimentício, pesca de subsistência, pesca esportiva e para diversos estudos científicos, devido aos complexos de espécies nelas pertencentes, às variáveis de cromossomos B nelas encontradas e em grandes compartilhamentos de sequências repetitivas entre as espécies (Moreira-Filho, 1991; Artoni et al. 2006; Paiva et al. 2006; Silva et al. 2017, dos Santos et al. 2021).

A espécie *Psalidodon paranae* (Figura 2), antes nomeada como *Astyanax paranae* (Terán et al. 2020), possui número diploide $2n = 50$, podendo conter um macrocromossomo B metacêntrico. Em estudo anterior, Silva et al. (2017) utilizaram o *RepeatExplorer1* e a *pipeline satMiner*, para a caracterização do satelitoma da espécie

Psalidodon paranae, primeiro satelitoma de peixes Neotropicais, contendo um total de 45 famílias de DNAs satélites, com uma predominância de satélites pequenos (menores que 100 pares de bases), sendo 59 pares de bases o valor da mediana destes. Em relação ao cromossomo B, realizaram um estudo entre bibliotecas de *P. paranae* contendo cromossomos B (1B) e não possuintes (OB), com foco em abundância comparativa entre elas, utilizando o *software RepeatMasker* (Smit et al. 2015), seguindo pela realização do log2 do quociente entre os valores de abundância 1B e OB. Como resultante desta análise, caracterizaram 18 satélites que apresentaram uma maior abundância em bibliotecas de indivíduos contendo o cromossomo B, no qual oito foram visualizados pela técnica de hibridização *in situ* fluorescente no cromossomo B da espécie.

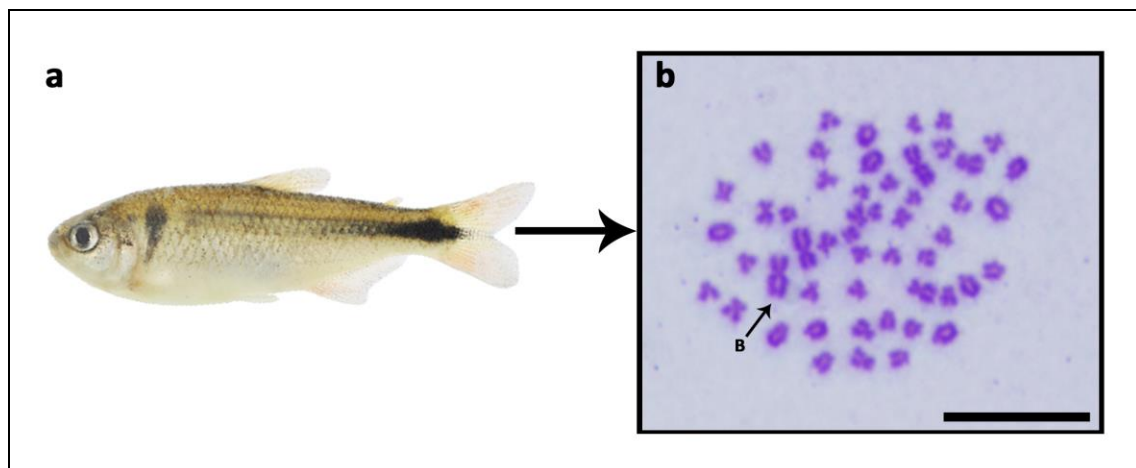


Figura 2. Espécime de *Psalidodon paranae* (a); metáfase da espécie *P. paranae* contendo um cromossomo B. Barra = 10 μ m. (b). Foto retirada de Silva et al. 2014.

Neste contexto, o presente trabalho buscou complementar o satelitoma da espécie *Psalidodon paranae*, no qual 56 novos DNAs satélites foram caracterizados, utilizando da ferramenta TAREAN, apresentando o diferencial desta ferramenta para o *RepeatExplorer1* quanto a caracterização de DNAs satélites. Além de realizamos o mapeamento físico de dez novos DNAs satélites em indivíduos de *P. paranae* contendo o cromossomo B. Analisamos todo o novo catálogo de satélites da espécie, agora contando com 101 sequências, em uma biblioteca de *reads* longos, na busca de uma melhor visualização da estruturação de todo o satelitoma da espécie ao longo de seu genoma. Complementar a estas análises, utilizamos também seis bibliotecas de *P.*

paranae, no qual três são de indivíduos portadores de cromossomos B e três de não portadores, com ênfase em investigar o quanto uma busca por abundância de sequências repetitivas pode ser influenciada ao se utilizar mais de um exemplar para a realização desta análise e o quanto esta influência pode interferir em análises em mapeamentos físicos.

2 Objetivos

O presente estudo teve como objetivo caracterizar o satelitoma de *P. paranae* utilizando o TAREAN, de forma a complementar a informação existente para esta espécie.

De modo específico, pretende-se:

- a) Utilizar a ferramenta TAREAN para a caracterização de novos satélites em *Psolidodon paranae*, trazendo um comparativo em relação ao RepeatExplorer1;
- b) Acessar a estrutura genômica detalhada dos DNAs satélites já descritos em *P. paranae* através de *reads* longos;
- c) Utilizar diferentes bibliotecas de *reads* curtos para comparar a abundância de cada satélite de *Psolidodon paranae* e sua diferenciação quanto a espécimes com e sem cromossomos Bs.

3 Material e métodos

Foram coletados 25 indivíduos de *P. paranae*, com a utilização de redes manuais e varas de pesca, no ribeirão Cascatinha (22°53'22.5" S, 48°29'22.4" W), pertencente a bacia do rio Tietê, região de Botucatu, São Paulo. Para a seguinte fase de amostragem, manutenção e análise dos animais utilizamos o acordo de regras internacionais em experimentação animal, seguidas pelo Comitê de Ética no Uso de Animais da Universidade Estadual Paulista (CEEAA/IBB/UNESP), protocolo 1227.

O primeiro passo foi anestésiar e dissecar cada animal, a fim da utilização de fragmentos de seus tecidos renais, para os estudos citogenéticos e de fragmentos de seu tecido muscular, para os estudos moleculares, fixando estes em etanol 95% para futura extração de DNA.

Todos os indivíduos utilizados neste trabalho foram fixados em formol 10%, conservados em álcool 70% e foram depositados na coleção de peixes do Laboratório de Genética de Peixes, do Departamento de Ciências Biológicas, UNESP- Bauru-SP. Sob os *vouchers* LGP14001-LGP14025.

3.1 Estimulação de mitoses

Para a obtenção das preparações cromossômicas, realizamos a estimulação de mitoses, utilizando fermento biológico com aplicação deste na musculatura dorsolateral do animal, seguindo o protocolo estabelecido por Oliveira et al. (1988) para peixes. A técnica aplicada consiste em:

1. preparação de solução de fermento biológico (Fleischmann) com uma proporção de 0,5g de fermento, 0,5g de açúcar e 7mL de água destilada;
2. incubação da solução em uma estufa (37°C) por cerca de 20 min, ou até a fase em que a solução comece a espumar;
3. injetar a solução na região dorsolateral do peixe, na proporção de 1mL para cada 100g de peso do animal. Para peixes com um baixo peso, como *Psalidodon*, esta etapa pode ser dividida em duas doses, divididas e aplicadas em um período de 48h; e
4. manter o animal em aquário bem aerado a 28°C.

3.2 Obtenção dos cromossomos metafásicos mitóticos

A obtenção de cromossomos metafásicos foi realizada seguindo o protocolo estabelecido por Foresti et al. (1981), nos indivíduos os quais realizamos a estimulação da mitose. Utilizamos de colchicina para a inibição da polimerização dos microtúbulos e seguida da hipotonização das células em suspensão, finalizando com a utilização de um fixador formado por metanol e ácido acético, na proporção 3:1.

1. injetar, na região antero-medial, abaixo da nadadeira peitoral, solução aquosa de colchicina (0,025%) na proporção de aproximadamente 1mL para cada 100g de peso do animal;
2. manter o animal em um aquário areado por um período aproximado de 50 min.;

3. Induzir os animais a morte através da imersão em solução de benzocaína e retirar fragmentos de rim anterior. Transferir o material para uma placa de petri, contendo 7mL de uma solução hipotônica de KCl (0,075M);
4. dissociar o material com o auxílio de pinças de dissecação e seringa de 3mL, resultando em uma solução homogênea;
5. transferir a solução obtida para um tubo de centrífuga e incubar esta solução em estufa a 37°C por 21 min;
6. retirar o tubo da estufa, inserir 7 gotas de fixador gelado (metanol e ácido acético na proporção de 3:1, respectivamente); agitar levemente a mistura com uma pipeta Pasteur e deixar repousar por 5 min à temperatura ambiente;
7. adicionar 6 mL de fixador e novamente agitar a mistura, sem a formação de bolhas e levar à centrífuga (900 ± 100 rpm) por 10 min;
8. retirar o sobrenadante, e inserir 7mL de fixador junto ao material precipitado, centrifugando 7 min a 900 ± 100 rpm;
9. repetir a etapa 8 entre duas ou três vezes para uma completa fixação das células em suspensão;
10. pingar o material em lâminas previamente tratadas em solução de HCL 10%;
11. deixar secar ao ar.

3.3 Extração do DNA genômico

A extração de DNA genômico total dos fragmentos de tecido obtidos foi realizada utilizando o Kit *NucleoSpin® Tissue Columns Collection Tubes* (2ml), seguindo as orientações dos fabricantes. Além disso, um posterior tratamento para remover o RNA residual destas amostras foi realizado, com a adição da RNase A (20 mg/ml - Invitrogen). A qualidade das extrações de DNA obtidas foi verificada com a utilização do equipamento 2100 *Bioanalyzer*, em gel de agarose a 1%.

3.4 Sequenciamento massivo (Next Generation Sequencing – NGS)

O principal desafio deste trabalho foi dissecar a estrutura genômica detalhada dos DNAs satélites em *P. paranae* e o tamanho médio dos *arrays*, tarefa impossível de se cumprir utilizando *reads* curtos. Assim, o sequenciamento PacBio foi realizado para um

indivíduo portador de cromossomo B. Este tipo de sequenciamento produz *reads* com um tamanho médio de 10kb e tamanho máximo de 60kb, facilitando a investigação da estrutura de sequências longas, como é o caso dos DNAs repetitivos. Visando uma cobertura de aproximadamente 3x e considerando uma faixa média de 4–5Gb por célula SMRT, utilizamos uma célula SMRT de sequenciamento para o indivíduo.

Para complementar o estudo de abundância e divergência de cada DNA satélite, além de aprofundar em um estudo comparativo de DNAs satélites presentes e ausentes no cromossomos B da espécie *Psalidodon paranae* foi realizado também o sequenciamento de *shorts reads* para seis indivíduos de *P. paranae*, três portadores de cromossomo B e três não portadores, utilizando da plataforma BGISEQ-500 da BGI (BGI Shenzhen Corporation, Shenzhen, China), maiores detalhes dos sequenciamentos de *short reads* estão apresentados no tópico 4.2.2 do Capítulo I da dissertação.

3.5 Análises bioinformáticas

As análises bioinformáticas foram realizadas de forma remota, utilizando um servidor localizado no Laboratório de Biologia e Genética de Peixes, Departamento de Morfologia, Universidade Estadual Paulista, Brasil.

3.5.1 Análise estrutural detalhada do satelitoma de *Psalidodon paranae*

Utilizando o pacote *NoiseCancellingRepeatFinder* (NCRF) (<https://github.com/makovalab-psu/NoiseCancellingRepeatFinder>) (Harris et al. 2019), foi possível identificar em quais *reads*, na biblioteca de *reads* longos, estavam presentes os satélites contendo monômeros de até 250 pares de bases (pb). Com os resultados gerados pelo NCRF, e utilizando linguagem de programação Python, rodamos simultaneamente, *Raincloud plots* (Allen et al. 2019) (<https://github.com/RainCloudPlots/RainCloudPlots>), apresentados nas figuras 4, 5 e nas figuras 1 e 2 do material suplementar, nos retornando gráficos que permitem uma interpretação visual do tamanho dos *arrays* de cada DNA satélite analisado.

3.6 Desenho dos *primers* e preparação das sondas

A escolha dos satélites para a criação dos *primers* foi baseada em sua abundância na espécie e sua diferenciação entre bibliotecas OB e 1B. Para os satélites cujo monômero possui um tamanho de até 600 pares de base, desenhamos *primers* divergentes, já satélites com seu tamanho superior a 600 pb, desenhamos *primers* convergentes, como apresentado na Figura 3, pretendendo gerar um *amplicon* próximo aos 600 pb, auxiliando na técnica de hibridação *in situ* fluorescente (FISH).

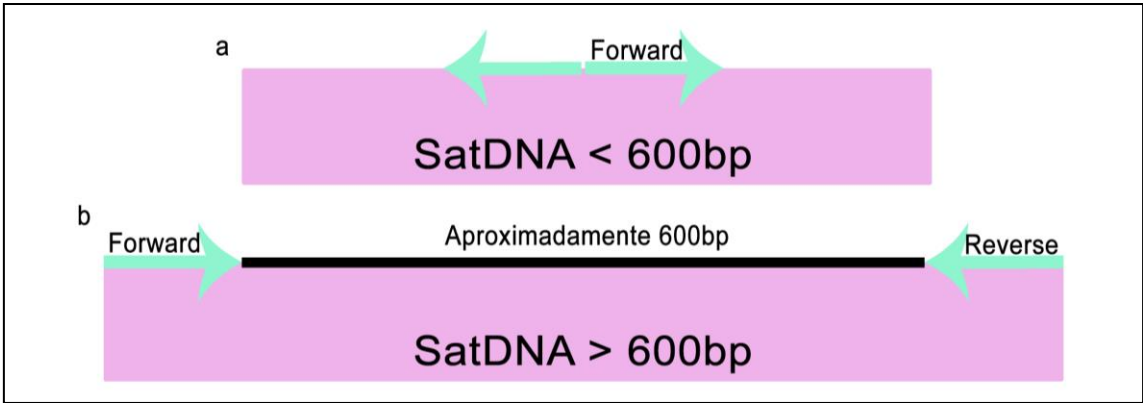


Figura 3. *Primers* divergente (a); e convergentes (b).

Utilizando o DNA genômico de *P. paranae*, realizamos reações de Amplificação em Cadeira da Polimerase (PCR) para cada um dos dez satélites em estudo, utilizando as concentrações: 1x Buffer, 2 µM de dNTPs, 1,5 mM de MgCl₂, 0,2 µM de cada *primer*, 0,25 U de *Taq polymerase* e 1-100 ng de DNA. Para a posterior confecção das sondas, o mesmo padrão foi mantido, com pequenas alterações, no qual houve a troca dos dNTPs regulares por dNTP + ½ dTTP (2 µM), além da incorporação de digoxigenina-11dUTP (Roche Applied Science).

Devido à alta diversidade quanto à abundância e tamanho do monômeros dos DNAs satélites-alvo, as reações no termociclador tiveram variações de acordo com as seguintes condições:

95°C	6 minutos	} 30X
95°C	20 segundos	
54~62°C	10~40 segundos	
72°C	10~40 segundos	
72°C	14 minutos	

3.7 Hibridação *in situ* fluorescente (FISH)

A técnica de FISH foi utilizada de acordo com Pinkel et al. (1986), com algumas modificações implementadas no Laboratório de Genética de Peixes (LAGENPE). Todas as sondas de DNA que utilizamos nestas hibridações foram produzidas a partir de 10 pares de *primers* específicos desenhados para regiões de interesse dos satélites estudados no trabalho.

A técnica de FISH foi realizada sempre em dois dias e seus procedimentos são:

Primeiro dia:

Primeira lavagem:

1. Lavar as lâminas em solução PBS 1x, por 3 min.

Tratamento com Pepsina:

1. Incubar em 0,005% pepsina/10mM HCl por 10 min, a 37°C;
2. Lavar em PBS 1x por 3 min, em temperatura ambiente.

Fixação com formaldeído

1. Fixar em 1% em PBS 1x/50 mM MgCl₂ por 10 min, em temperatura ambiente;
2. Lavar em PBS 1x por 3 min;
3. Desidratar em série alcoólica (70%,85% e 100%) gelada por 3 min. em cada e secar as lâminas.

Mix de Hibridação

1. Preparar o mix de hibridação (15 µl de formamida 70%, 6 µl de sulfato dextrano 10%, 3 µl 2xSSC, entre 1 µl de sonda por lâmina e 5 µl de água);

Pré-Hibridação

1. Desnaturar as lâminas em formamida 70% em 2xSSC, por 1:30min. a 70°C;
2. Desnaturar a sonda em 98°C por 10 min;
3. Desidratar as lâminas em série alcoólica, 3 min cada e secar as lâminas.

Hibridação

1. Montar as lâminas com mix de hibridação, cobrir com lamínula e incubar em câmara úmida a 37°C overnight.

Segundo dia:

Lavagens:

1. Retirar as lamínulas e lavar as lâminas em 1x SSC em 42°C por 5min, realizando a agitação do recipiente;

2. Lavar as lâminas em 4xSSC/0,5% Triton por 3min, em temperatura ambiente realizando a agitação do recipiente.

Bloqueio

1. Incubar as lâminas em Tampão 5% NFDM/4xSSC (2,5g de leite em pó desnatado + 10 mL de 20xSSC + 40mL H₂O) por 15 min a temperatura ambiente;

2. Lavar 2x em 4xSSC/0,5% Triton por 3 min cada, a temperatura ambiente;

Detecção

1. Incubar as lâminas com 100 µL de NFDM/0,5 µL de anti-digoxigenina rodamina por 1h em câmara úmida a 37°C;

2. Lavar 3x em 4xSSC/0,5% Triton por 3 min cada, em temperatura ambiente;

3. Desidratar em série alcoólica, 3 min cada e secar as lâminas;

4. Montar as lâminas inserindo 15 µl de DAPI e cobrir com lamínula, mantendo-as em um local escuro até sua visualização.

4 Capítulo I

Ampliação na caracterização de DNAs satélites em *Psalidodon paranae* (Teleostei; Characiformes) e o estudo de suas abundâncias

4.1 Introdução

O genoma dos seres eucariotos é bastante heterogêneo em sua composição, podendo ser encontrados genes de cópia única, pseudogenes, RNAs não-codificantes, elementos repetitivos, dentre outras sequências. Em geral, as sequências repetitivas podem ser subclassificadas em dois principais grupos: i) elementos dispersos; e ii) repetidos em *tandem* (YUAN et al. 2018). Os elementos dispersos compreendem os elementos transponíveis, como os transposons e retrotransposons, e são sequências móveis, muitas vezes com capacidade autônoma. Por outro lado, as repetições em *tandem* são sequências formadas por uma menor unidade, conhecida como monômero, o qual tendem a se repetir em sentido *head-to-tail*, ou seja, um monômero localizado sequencialmente a outro, formando longos *arrays* desta mesma repetição, variando de dezenas até milhares de vezes (Charlesworth et al. 1994; Garrido-Ramos 2017). Entre as repetições em *tandem* conhecidas, pode-se citar as famílias multigênicas codificantes, como os genes ribossômicos 5S e 18S, e os DNAs satélites (satDNAs), que não são codificantes (Garrido-Ramos 2017). Os satDNAs podem ser divididos em satélites simples, quando os monômeros são menores ou iguais a 10 pares de bases (pb), satélites, quando maior que 10 pb e menor que 100 pb e satélites complexos, possuindo 100 pb ou mais (Wei et al. 2018). Por definição, o conjunto de todas as famílias caracterizadas de DNAs satélites de uma espécie é denominado satelitoma (Ruiz-Ruano et al. 2016).

Na década passada, a caracterização de DNAs repetitivos, de forma massiva, tornou-se muito mais válida, devido a uma ampliação e barateamento das tecnologias de sequenciamento de nova geração, além de uma intensidade maior nas pesquisas destas áreas, sendo criadas então diversas *pipelines* com foco no estudo destas sequências, como *RepeatExplorer1*, *satMiner* e TAREAN (Novák et al. 2013; Ruiz-Ruano et al. 2016; Novák et al. 2017). O primeiro satelitoma desenvolvido para peixes Neotropicais é pertencente a espécie *Psalidodon paranae* e corresponde a 45 famílias

de DNAs satélites, desenvolvido utilizando do *software RepeatExplorer1* (Silva et al.2017). Utilizando de sequencias repetitivas para mapeamento cromossômicos, como 18S, 5S e histonas, foi apresentado o cromossomo B de *Psalidodon paranae* como um isocromossomo devido a apresentar uma grande homologia destas pinturas entre seus braços, além de sua possível origem ser dada pela divisão centromérica incorreta de um cromossomo proto-B acrocêntrico produzindo um iso-B metacêntrico (Silva et al. 2014; Silva et al. 2016). Estes trabalhos, tanto de estudo do cromossomo B quanto o pioneirismo de satelitoma em peixes, além de ser uma espécie de fácil captura, manuseio e realização de técnicas laboratoriais, contribuíram para o lambari *Psalidodon paranae* se tornar um peixe modelo para o estudo de cromossomo B e nas análises de sequencias repetitivas.

No presente trabalho, realizamos o sequenciamento de *reads* curtos em seis novas espécimes de *Psalidodon paranae*, três indivíduos possuintes de cromossomo B e três indivíduos sem estes cromossomos supranumerários, com o objetivo de complementar o satelitoma desta espécie, utilizando da *pipeline* TAREAN, buscando seu diferencial em relação a *RepeatExplorer1*, além de um estudo diferencial destas sequências repetitivas em indivíduos 0B e 1B. Para um estudo quanto a estruturação e melhor entendimento quanto a sinais em cromossomos ao utilizar a técnica de FISH, realizamos também um sequenciamento de *reads* longos de um indivíduo de *P. paranae* possuidor de um cromossomo B, utilizando de ferramentas computacionais para estas análises.

4.2 Material e métodos

4.2.1 Amostras

Indivíduos de *P. paranae* foram coletados com o uso de redes de puçá e vara de pesca no córrego Cascatinha (22°53'22.5" S, 48°29'22.4" W), bacia do rio Tietê, Botucatu, São Paulo, Brasil. Os procedimentos de amostragem, manutenção e análise dos animais foram realizados de acordo com as regras internacionais em experimentação animal, seguidas pelo Comitê de Ética no Uso de Animais da Universidade Estadual Paulista (IBB/UNESP), protocolo CEEAA/IBB/UNESP, número de protocolo 1227. Todos os animais coletados e analisados foram fixados em formol 10% e conservados em álcool 70%, após isso, foram depositados na coleção de peixes do

Laboratório de Genética de Peixes, Departamento de Ciências Biológicas, UNESP-Bauru-
SP.

4.2.2 Extração de DNA e sequenciamento do genoma

Seis indivíduos de *P. paranae* foram selecionados para sequenciamento de nova geração (NGS), sendo três portadores de 1 cromossomo B (1B) e três não-portadores de B (0B). Para estes sequenciamentos utilizamos a plataforma BGISEQ-500 da BGI (BGI Shenzhen Corporation, Shenzhen, China) (*paired-end* 2x100bp) e (*paired-end* 2x150bp), com coberturas de sequenciamento de aproximadamente 4,7x; 4,7x e 0,5x para as bibliotecas contendo cromossomo B e 4,7x; 3,9x e 0,6x para as bibliotecas 0B, considerando que o genoma de *P. paranae* tem tamanho estimado em 1.87 Gb.

Neste caso, extraímos DNA genômico dos músculos preservando-os em etanol 100%, utilizando o *Wizard Genomic DNA Purification Kit* (Promega), seguindo as instruções fornecidas pelo fabricante. Uma etapa adicional de remoção de RNA foi realizada com a utilização de RNase A (Invitrogen). Os resultados gerados pela etapa anterior foram conferidos em gel de agarose a 1%, para a verificação da integridade do DNA genômico extraído.

Para a comprovação de que cada indivíduo sequenciado realmente pertencesse a espécie *P. paranae*, realizamos a montagem do genoma mitocondrial para os seis espécimes de estudo, utilizando o *software* NOVOplasty (Dierckxsens et al. 2016). Após a caracterização do *contig* do mitogenoma (de aproximadamente 16.6 kb), a anotação do genoma mitocondrial foi realizada utilizando o *software* MitoFish (Iwasaki et al. 2013) , pelo site (<http://mitofish.aori.u-tokyo.ac.jp/>) utilizando a aba “fish mitogenome annotation”, nos resultando em um output com o genoma mitocondrial circular de cada indivíduo montado.

4.2.3 Caracterização do satelitoma por TAREAN

Utilizando o satelitoma de *P. paranae* caracterizado por Silva et al. (2017) como *database* referencial, contando com as 45 diferentes famílias de DNAs satélites, caracterizamos o satelitoma da mesma espécie pelo TAREAN (Novák et al. 2017), seguido pela etapa do *RepeatExplorer2* (Novák et al. 2013), utilizando o protocolo *satMiner*, criado por Ruiz-Ruano et al. (2016) para um comparativo entre o TAREAN e o

RepeatExplorer1 (Novák et al. 2013). Todas as bibliotecas passaram por uma etapa de filtragem de qualidade e correção nos tamanhos dos *reads*, utilizando o *software* Trimmomatic (Bolger et al. 2014), utilizando os filtros LEADING:3; TRAILING:3; SLIDINGWINDOW:4:15; MINLEN:100 e CROP:101. Posteriormente realizamos uma seleção aleatória de 2x 500.000 *reads* para a clusterização no TAREAN. Após a prospecção dos DNAs satélites pelo TAREAN, utilizamos o programa DeconSeq (Schmieder & Edwards, 2011) para a remoção de *reads* com similaridade aos satDNAs nas bibliotecas de estudo, escolhendo novamente 2x 500.000 *reads* para a clusterização no TAREAN e repetindo este processo até que mais nenhum satélite fosse relatado como *output* da *pipeline*.

Após finalizado o TAREAN, realizamos o mesmo processo pelo *RepeatExplorer2*, utilizando como base as bibliotecas utilizadas na última iteração do TAREAN, à procura de *clusters* com um gráfico circular ou anelar, seguida pela análise dos *contigs* associados utilizando o *software* Geneious 8.1. Utilizamos o *dotplot* destes *clusters* para a caracterização de novos monômeros e realizamos o *workflow* similar ao TAREAN, até mais nenhum satélite ser detectado.

Com o catálogo completo, seguimos para uma etapa de filtragem e remoção de famílias multigênicas. Finalmente, realizamos uma busca de similaridade entre as sequências do catálogo, utilizando o padrão de estabelecido por Ruiz-Ruano et al. (2016) para classificar os satDNAs quanto a uma mesma superfamília, família ou uma variante.

Para a determinação de parâmetros de abundância e divergência de cada família de satDNA, utilizamos o *software RepeatMasker* (Smit et al. 2017), utilizando a opção “*cross_match*” nas seis diferentes bibliotecas de *reads* curtos, utilizando dez milhões de *reads* para cada amostra para a análise. A estimativa da abundância foi obtida com a proporção entre a quantidade de *reads* mapeados em cada satélite dividido pelo total de nucleotídeos analisados. A partir deste resultado, nomeamos cada família do satelitoma com base em ordem decrescente de abundância, como sugerido por Ruiz-Ruano et al. (2016), inserindo-os aos já pré-existentes. Também como resultante do *RepeatMasker*, averiguamos a divergência média entre as famílias de satDNAs, utilizando a distância entre as sequências, para isso utilizamos o Kimura-2- parâmetro, tendo como base o *script* calcDivergenceFromAlign.pl pertencente ao *RepeatMasker* (Smit et al. 2017).

4.2.4 Estudo do satelitoma em *reads* longos

Utilizando o pacote NoiseCancellingRepeatFinder (NCRF) (<https://github.com/makovalab-psu/NoiseCancellingRepeatFinder>) (Harris et al. 2019), analisamos *reads* longos da biblioteca PacBio no qual cada um dos satélites do estudo, com até 250 pares de bases, estavam presentes. Para esta análise utilizamos os parâmetros: maxnoise = 20%; scoring = pacbio; stats = events e minlength = 15 vezes o tamanho do monômero. Após a conclusão do NCRF, realizamos uma filtragem de seu *output* total, obtendo os tamanhos e quantidades de *arrays* de cada satélite do estudo.

Para a visualização gráfica destes resultados, utilizamos o programa *Raincloud plot* (Allen et al. 2019), com algumas pequenas modificações para um resultado simultâneo de todos os satélites do estudo, como apresentados nas Figuras 1 e 2 do material suplementar, auxiliando para um entendimento visual do tamanho e quantidade de *arrays* de cada um destes DNAs satélites.

4.2.5 Hibridização *in situ* fluorescente – FISH

A técnica de FISH foi procedida como descrito em Pinkel et al. (1986), com algumas modificações estabelecidas pelo nosso laboratório. As sondas utilizadas para todos os DNAs satélites deste trabalho foram marcadas utilizando digoxigenina-11-dUTP, tendo como base a amplificação das sequências satélites no genoma de *P. paranae* e anti-digoxigenina-rodamina para a detecção dos sinais, enquanto os cromossomos foram contra-corados por DAPI. Para cada satélite em estudo, um mínimo de 20 metáfases foi analisado para a confirmação dos resultados.

5 Resultados

5.1 Satelitoma de *Psilododon paranae*

Após nove iterações com TAREAN e uma iteração de *RepeatExplorer2*, 56 novos satélites foram caracterizados. Nas duas primeiras iterações do TAREAN foram caracterizados 29 dos 45 satélites antes caracterizados pelo *RepeatExplorer1*, nas seguintes iterações foram encontrados mais oito dos satélites apresentado por Silva et al. (2017), sendo que os oito restantes (ApaSat22, ApaSat25, ApaSat28, ApaSat30, ApaSat32, ApaSat37, ApaSat41 e ApaSat42) não foram encontrados em nenhuma das

749 iterações. Destaca-se que vários dos novos satélites apresentaram abundância maior em
750 relação aos que haviam sido caracterizados por Silva et al. (2017) (Tabela 1). Neste
751 contexto, para não alterar a denominação de sequências que foram publicadas
752 previamente, todos os novos satélites encontrados aqui foram inseridos a partir do
753 ApaSat46.

754 Em relação aos tamanhos de cada monômero (RUL, de “repeat unit lengths”) dos
755 101 satélites, a variação foi de 6 a 4841 pb, apresentando uma mediana de 70 pb, nos
756 45 primeiros satélites, realizado por *RepeatExplorer1* a mediana era de 59 pb e nos 93
757 satélites encontrados por TAREAN a mediana possui 77 pb. Quanto ao conteúdo A+T, o
758 valor mínimo foi de 30.3% e o valor máximo de 80.5%, possuindo uma média de 57.4%,
759 ocorrendo então uma tendência a satélites ricos em conteúdo A+T, como apresentado
760 por Silva et al. (2017) (Tabela 1).

761 Além da superfamília apresentada em Silva et al. (2017), aqui foram encontradas
762 sete novas (SF2- ApaSat01 e ApaSat46; SF3- ApaSat30, ApaSat69 e ApaSat73; SF4-
763 ApaSat49, ApaSat50 e ApaSat59; SF5- ApaSat61 e ApaSat93; SF6- ApaSat87 e ApaSat88;
764 SF7- ApaSat90 e ApaSat100; SF8- ApaSat95 e ApaSat98) (Tabela 1).

765
766 **Tabela 1:** Características dos 101 satélites de *P. paranae* obtidas por SatMiner e TAREAN com
767 base nas seis bibliotecas em estudo (RUL = “repeat unit length”; A+T (%) = porcentagem de
768 conteúdo AT do monômero).

SF	satDNA	RUL	A+T (%)	Abundância 1B média	Abundância OB média	Divergência Kimura Média (%)
2	ApaSat01	51	54.9	0,052435062	0,034046689	7,185
1	ApaSat02	236	64.8	0,004804598	0,004565518	15,665
	ApaSat03	91	51.6	0,002438174	0,002325241	9,458333
1	ApaSat04	233	63.1	0,001623148	0,001537732	17,83667
	ApaSat05	23	52.2	0,000593099	0,000760427	11,57
	ApaSat06	86	46.4	0,000813441	0,000456893	5,278333
	ApaSat07	6	50.0	0,000186907	0,000145688	16,85667
	ApaSat08	35	62.9	0,000655514	0,000548761	9,433333

	ApaSat09	21	75.0	0,000545433	0,000615273	15,54167
	ApaSat10	179	63.1	0,000560899	0,000492921	4,225
	ApaSat11	22	50.0	0,000306046	0,000339504	10,52833
	ApaSat12	69	60.9	0,000340654	0,000385519	3,546667
	ApaSat13	22	56.5	0,000176354	0,000120015	8,511667
	ApaSat14	184	63.0	0,000222579	0,0001772	6,37
	ApaSat15	51	54.9	0,000320842	0,000242189	7,085
	ApaSat16	54	48.1	0,000269628	0,000286717	9,073333
	ApaSat17	365	51.8	0,000086196	0,000106445	1,243333
	ApaSat18	58	46.6	0,00025751	0,000243381	6,973333
	ApaSat19	77	66.2	0,000315784	0,000313174	5,815
	ApaSat20	18	50.0	0,00016362	3,32533E-05	13,21
	ApaSat21	68	48.5	6,24263E-05	7,58523E-05	1,773333
	ApaSat22	63	56.5	0,000172424	0,000153352	17,01833
	ApaSat23	37	43.2	0,000099719	5,67853E-05	3,665
	ApaSat24	78	56.4	6,94253E-05	5,99083E-05	5,753333
	ApaSat25	27	51.9	6,76247E-05	5,33713E-05	14,115
	ApaSat26	195	65.1	0,000170074	0,000151595	7,966667
	ApaSat27	178	39.9	0,00011931	0,000128565	8,571667
	ApaSat28	52	55.8	0,00015219	0,000159857	15,665
	ApaSat29	52	67.3	0,000129623	0,000135664	13,88333
3	ApaSat30	50	64.0	0,000175153	0,000138532	14,84167
	ApaSat31	165	63.0	0,000144518	0,000165725	10,30667
	ApaSat32	85	58.8	3,84683E-05	2,49297E-05	7,66
	ApaSat33	112	64.3	0,000201549	7,40017E-05	4,616667
	ApaSat34	59	37.3	0,000196155	0,00017164	1,43
	ApaSat35	37	48.6	8,06513E-05	0,000075607	4,18
	ApaSat36	21	57.1	0,000123395	0,00014325	6,565
	ApaSat37	38	57.9	3,82107E-05	1,34133E-05	4,741667
	ApaSat38	107	41.1	0,000108975	0,0001349	16,79167
	ApaSat39	32	50.0	9,22477E-05	0,000008949	8,546667
	ApaSat40	189	65.6	0,000341354	0,000103631	8,565
	ApaSat41	33	30.3	0,000037194	4,14193E-05	10,055

	ApaSat42	90	66.7	7,35167E-05	5,75353E-05	4,786667
	ApaSat43	61	68.9	0,000145389	0,000117577	9,078333
	ApaSat44	21	38.1	0,000105435	3,85947E-05	24,69667
	ApaSat45	113	75.1	5,79233E-05	5,51633E-06	4,446667
2	ApaSat46	39	48.8	0,013966168	0,010090606	1,863333
	ApaSat47	97	44.6	0,00245406	0,001789784	15,4
	ApaSat48	4841	42.0	0,002542092	0,00292509	7,983333
4	ApaSat49	631	73.9	0,003588679	0,003518296	13,375
4	ApaSat50	1633	63.0	0,003168859	0,002689567	9,106667
	ApaSat51	1543	73.7	0,002308124	0,002156108	7,133333
	ApaSat52	1813	51.0	0,001296854	0,00154454	3,23
	ApaSat53	595	55.5	0,001642283	0,00101407	7,648333
	ApaSat54	152	64.5	0,001503833	0,001341473	6,803333
	ApaSat55	42	57.1	0,001002504	0,001038557	13,78
	ApaSat56	92	53.4	0,000803213	0,000878313	21,13333
	ApaSat57	40	57.5	0,000591998	0,000503159	2,466667
	ApaSat58	499	66.3	0,000598357	0,000551366	9,438333
4	ApaSat59	42	50.0	0,000374007	0,000208647	6,85
	ApaSat60	2258	61.0	0,000645128	0,000417107	11,98333
5	ApaSat61	187	68.4	0,000522933	0,000437774	10,5
	ApaSat62	106	60.4	0,000368419	0,000345223	4,21
	ApaSat63	1277	51.8	0,000352373	0,000333233	2,42
	ApaSat64	61	70.5	0,000420323	0,000449695	5,566667
	ApaSat65	29	44.8	0,000307267	0,000267524	16,17167
	ApaSat66	80	71.2	0,000341097	0,000360566	6,748333
	ApaSat67	27	63.0	0,000292668	0,000307149	6,391667
	ApaSat68	43	60.5	0,000140675	0,00006692	4,008333
3	ApaSat69	20	50.0	0,000112334	7,72237E-05	8,766667
	ApaSat70	32	56.2	0,00022819	0,000232747	14,03333
	ApaSat71	24	66.7	0,000195716	0,000157123	26,915
	ApaSat72	75	56.0	0,000239986	0,000289671	9,001667
3	ApaSat73	354	69.5	0,000255035	0,000262451	8,641667
	ApaSat74	58	65.5	0,000175205	0,000179025	5,883333

	ApaSat75	1251	53.5	0,000158948	0,000179529	1,85
	ApaSat76	185	68.6	0,000207487	0,000177954	6,23
	ApaSat77	166	71.7	0,000206439	0,000263488	7,346667
	ApaSat78	393	55.2	0,000155062	0,0001856	8,086667
	ApaSat79	657	53.4	0,000152106	0,000168053	1,036667
	ApaSat80	191	62.3	0,000188692	0,000211636	8,553333
	ApaSat81	40	67.5	0,000132085	0,000126259	14,765
	ApaSat82	1279	53.6	0,000134732	0,000170621	6,188333
	ApaSat83	33	48.5	0,000105129	7,55573E-05	14,29833
	ApaSat84	52	44.2	9,72613E-05	0,000127061	8,76
	ApaSat85	68	58.8	8,36723E-05	8,63883E-05	3,528333
	ApaSat86	648	67.3	0,000162574	0,000168921	8,07
6	ApaSat87	584	58.4	0,000145436	0,000124538	2,36
6	ApaSat88	20	60.0	7,28863E-05	0,000056821	9,04
	ApaSat89	24	55.8	9,25673E-05	5,48467E-05	17,80333
7	ApaSat90	80	58.7	9,57123E-05	0,000125476	7,316667
	ApaSat91	43	58.1	9,04323E-05	0,000111439	7,173333
	ApaSat92	52	47.7	0,000113577	6,21963E-05	13,76667
5	ApaSat93	142	72.5	0,000131237	0,000133467	9,29
	ApaSat94	205	51.7	0,000071144	9,10503E-05	4,955
8	ApaSat95	34	55.9	6,09023E-05	4,07737E-05	5,163333
	ApaSat96	419	49.9	7,81223E-05	0,000097486	1,506667
	ApaSat97	59	54.2	0,000082698	9,48597E-05	3,235
8	ApaSat98	548	56.4	0,000077397	0,000094291	10,47333
	ApaSat99	580	57.1	8,40413E-05	9,33687E-05	5,678333
7	ApaSat100	41	80.5	4,92593E-05	3,29663E-05	7,613333
	ApaSat101	70	57.1	0,000056755	4,38627E-05	6,463333

770

771 5.2 Distribuição de *arrays* de satDNAs em *reads* longos de *P. paranae*

772 *Raincloudplots* para todos os DNAs satélites do catálogo cujo monômero fosse
773 menor do que 250 pares de bases (limitante do *software NoiseCancellingRepeatFinder*;
774 Harris et al. 2019) foram gerados. Nestas figuras é possível compreender a mediana dos
775 *arrays*, que foi apresentada para cada satélite. Tendo em vista que DNAs satélites com

um tamanho de *array* superior a dez mil pares de bases tendem a ser visíveis com a técnica de FISH (dos Santos et al. 2021), o ApaSat54 se tornou um ótimo candidato para a aplicação deste estudo, visto que existem 4 *arrays* superiores a 10 kb de extensão, sendo um destes de 22,5kb (Figura 4), além de apresentar valor de mediana elevada, chegando a quase 4 mil pares de bases, como apresentado na tabela 2. A FISH corroborou com estes dados, pois apresentou sinal em quatro pares de cromossomos autossômicos (Figura 5).

Os dados gerados pelo NCRF e sua interpretação foram importantes para compreender a ocorrência de sinais de FISH para o ApaSat54 e a ausência em ApaSat57 e ApaSat68 (Figura 4). Além de uma possível explicação da falta de sinais visíveis em FISH do ApaSat20, ApaSat36, ApaSat38 e ApaSat44 nos cromossomos autossômicos (Figura 5).

O estudo em reads longos auxiliou em analisar o ApaSat20, que contem *arrays* próximo a 1.5 kb, o ApaSat36, contendo seus *arrays* próximos a 1.3kb, o ApaSat38, contendo *arrays* próximos a 3.5kb e o ApaSat44, contendo *arrays* próximos a 1.4kb, sendo que os *outliers* de todos estes DNAs satélites não apresentaram tamanhos superiores a 10kb e com exceção do ApaSat44, que apresenta 53 *reads*, todo apresentam poucos *reads*, auxiliando na compreensão destes não sinalizarem nos cromossomos autossômicos nas técnicas de FISH (Figura 5).

Tabela 2. Abundância, mediana e tamanho médio dos *arrays* gerados pelo *Raincloud plot*.

Satélites	Abundância (pares de bases)	Mediana	Média
ApaSat01	8483818	1556,5	1972,063691
ApaSat02	450641	7203	8666,173077
ApaSat06	2995	1497,5	1497,5
ApaSat08	49709	2766,5	3550,642857
ApaSat09	46203	1511	2887,6875
ApaSat11	202058	1940	2377,152941
ApaSat12	198449	2086	2390,951807
ApaSat15	20183	1147,5	1441,642857

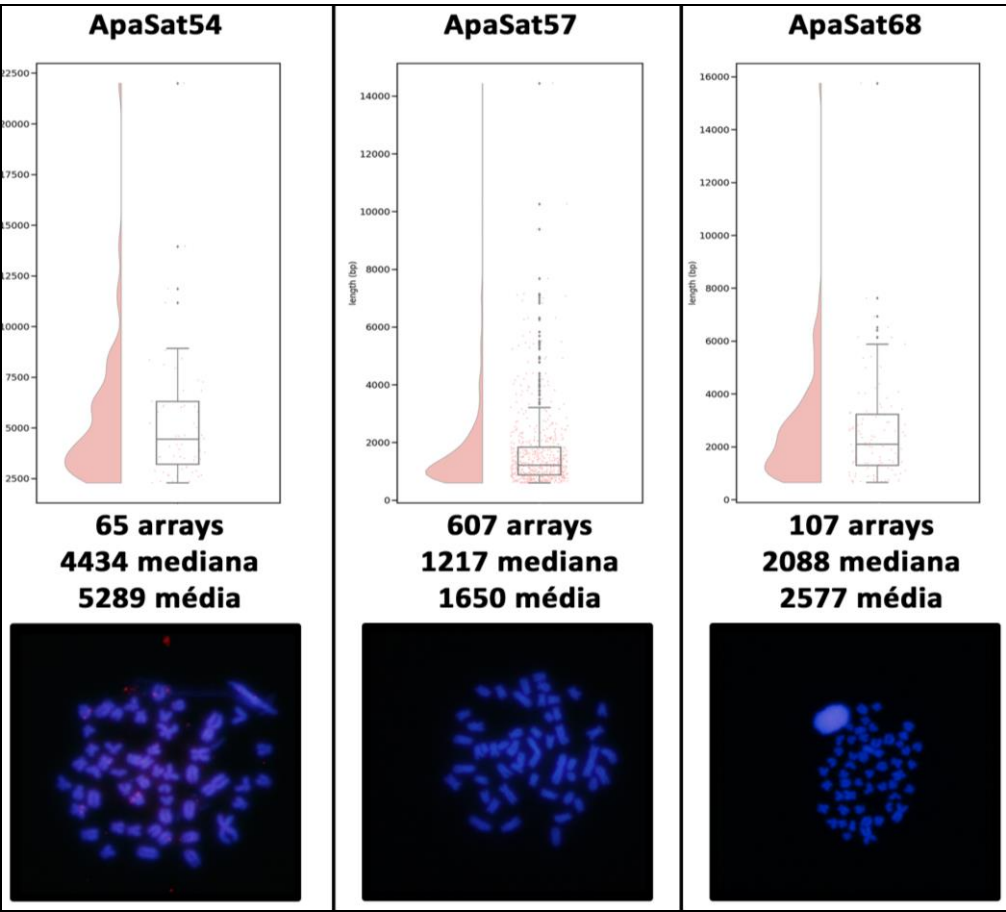
ApaSat18	2033	2033	2033
ApaSat20	36264	1517	1726,857143
ApaSat21	16558	2929,5	2759,666667
ApaSat22	99424	3059	4322,782609
ApaSat23	46232	7941	11351,375
ApaSat24	15954	10923,5	10725,83333
ApaSat25	66126	1445	1889,314286
ApaSat28	2617	870	872,3333333
ApaSat29	9808	1380,5	18087
ApaSat31	2478	2478	2478
ApaSat34	328306	2116	2218,283784
ApaSat35	24918	6609,5	6229,5
ApaSat36	17447	1333,5	1453,916667
ApaSat38	7142	3571	3571
ApaSat39	3094	3094	3094
ApaSat42	14149	4283	4716,333333
ApaSat43	29495	3302	4213,571429
ApaSat44	95135	1422	1795
ApaSat45	2725	2725	2725
ApaSat46	4937866	760	856,8221412
ApaSat54	343849	4434	5289,984615
ApaSat57	1001926	1217	1650,61944
ApaSat59	169024	1335	2086,716049
ApaSat62	184912	2961	4202,545455
ApaSat64	190795	1451	1946,887755
ApaSat65	5674	786	810,5714286
ApaSat66	48720	1887,5	2214,545455
ApaSat67	389879	1526	2499,224359
ApaSat68	275789	2088	2577,46729
ApaSat69	40308	559	660,7868852
ApaSat70	5231	680	871,8333333

ApaSat71	17428	562	670,3076923
ApaSat74	23882	2370	3411,714286
ApaSat76	5999	5999	5999
ApaSat85	116675	2204	2482,446809
ApaSat88	44050	1314	2202,5
ApaSat91	5243	2621,5	2621,5
ApaSat92	12691	1807	2538,2
ApaSat95	44100	2585,5	2450
ApaSat97	18081	1514	1390,846154
ApaSat100	17973	952	1123,3125

797

798

799



800

801

802

803

Figura 4. FISHs e gráficos *Raincloud plot* de três novos satélites do estudo, junto da quantidade de *arrays*, mediana e média de seus *arrays*.

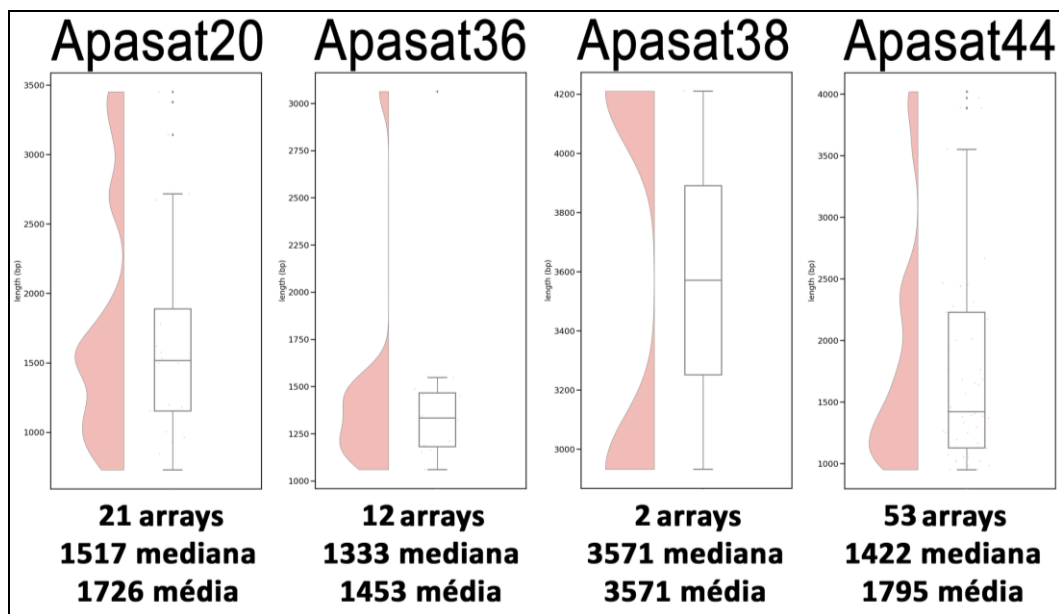


Figura 5. Satélites que não apresentaram cluterização nos cromossomos autossômicos em Silva et al. 2017.

5.3 Distribuição cromossômica de dez novas famílias de satélite em *Psalidodon paranae*

Foram desenhados 10 novos pares de primers contemplando 10 novas famílias de satDNAs, apresentadas na tabela 3, escolhidas a partir da abundância destas famílias (ApaSat47 até ApaSat51) ou do seu diferencial entre indivíduos 1B e 0B (ApaSat53, ApaSat54, ApaSat57, ApaSat60 e ApaSat68). Todas as reações de PCR foram bem sucedidas, sendo possível a confecção de sondas para os mesmos. O mapeamento por FISH detectou *clusters* de três destes novos satélites, ApaSat49, ApaSat53 e ApaSat54. Enquanto o ApaSat49 e o ApaSat53 foram mapeados dispersos pelos cromossomos, o ApaSat54 foi mapeado em 4 pares de cromossomos autossômicos. Apenas o ApaSat53 apresentou sinais conspícuos no cromossomo B da espécie em análise. Os demais satDNAs analisados foram considerados não *cluterizados* a nível cromossômico, devido à ausência de sinal no FISH (Tabela 3; Figura 6).

825 **Tabela 3.** Presença ou ausência de *clusterização* das famílias de DNAs satélites mapeados na
826 espécie *Psalidodon paranae* (C = *clusterizado*; NC = não *clusterizado*).

827

Satélites	Presença de <i>cluster</i>
ApaSat47	NC
ApaSat48	NC
ApaSat49	C
ApaSat50	NC
ApaSat51	NC
ApaSat53	C
ApaSat54	C
ApaSat57	NC
ApaSat60	NC
ApaSat68	NC

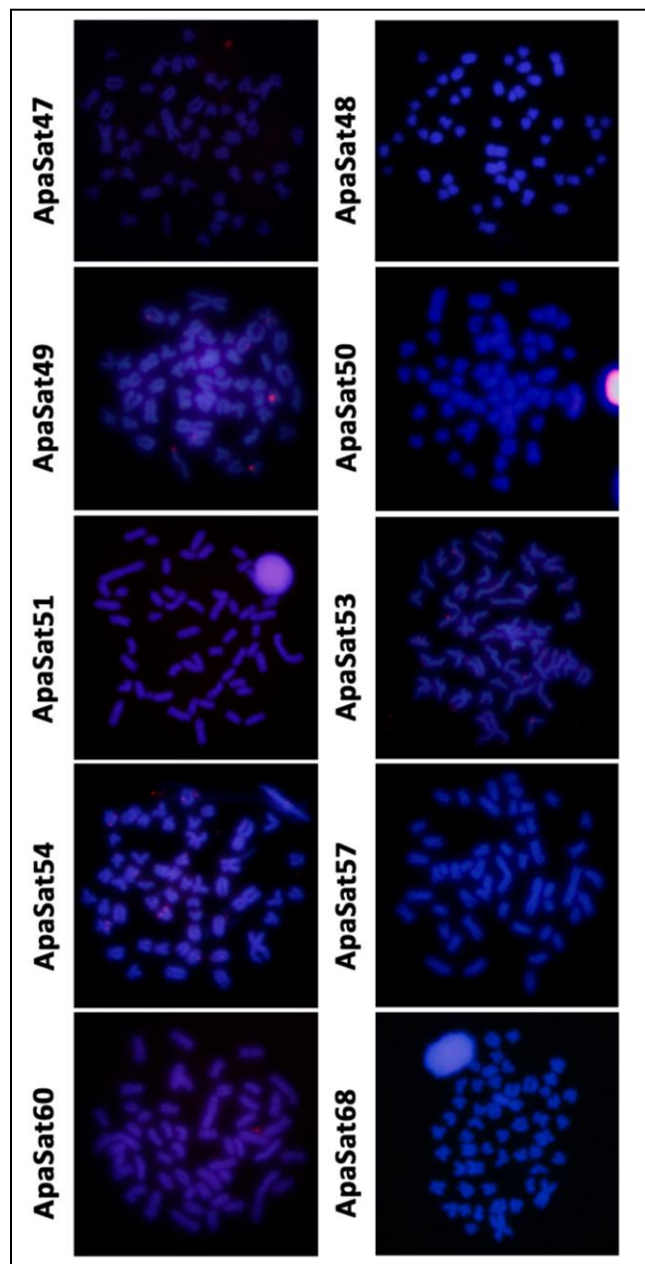


Figura 6. Hibridizações das dez novas famílias de satélites em cromossomos metafásicos de *Psalidodon paranae*.

5.4 Estudos comparativos entre bibliotecas OB/1B em *Psalidodon paranae*

Com o sequenciamento de seis indivíduos de *P. paranae*, três destes 1B e três OB, realizamos uma busca pela abundância e divergência de cada um dos 101 satélites do estudo em cada biblioteca. Para esta análise, fizemos subamostragens de 2x5.000.000 de *reads* de cada biblioteca genômica e depois utilizamos o *software RepeatMasker* (Smit et al. 2017), utilizando a opção “*cross_match*”, para obter valores de abundância e diversidade do catálogo total de 101 satDNAs. Após o estudo de abundância individual, realizamos um estudo de abundância entre as bibliotecas de indivíduos 1B e OB, como apresentado na Tabela 3 do material suplementar, para uma melhor compreensão de marcação que ocorreram nestes cromossomos B.

Para complementar o estudo de abundância entre bibliotecas 1B e OB, realizamos a construção de gráficos de estimativas (*Gardner-Altman estimation plot*) (Ho et al. 2019), com foco nos primeiros 45 satélites caracterizados por Silva et al. (2017). Esta abordagem gráfica, utiliza a média de um grupo controle, no nosso caso as bibliotecas 1B e a normaliza para o valor zero do gráfico, que ao comparada com o valor da média do grupo teste (bibliotecas OB) gera o tamanho da diferença, conhecido como tamanho de efeito (*effect size*), além de apresentar o intervalo de confiança e um gráfico realizando uma reamostragem de cinco mil vezes dos dados utilizados (Ho et al. 2019) (Figura 3 do material suplementar).

Ao realizar os estudos de abundância por razão entre médias, junto ao de estimativa, nota-se que um de suas bibliotecas de indivíduo possuinte do cromossomo B possui valores de abundância muito diferencial em relação a outras duas bibliotecas, como apresentado nos DNAs satélites ApaSat01 e ApaSat44 (Figura 3 do material suplementar), além de uma total ausência do ApaSat31 (Tabela 1 do material suplementar). Realizamos um estudo de cobertura de genes presentes no cromossomo B em *Psalidodon paranae*, notando uma alteração no valor dos picos ao se comparar as outras duas bibliotecas portadores de cromossomo B (Figura 8), dando indícios deste ser uma possível variante do cromossomo B. Para realizar um estudo comparativo, retiramos esta biblioteca na nossa análise, utilizando então duas bibliotecas de indivíduos com cromossomo B e três bibliotecas de indivíduos não portadores deste cromossomo, montando uma tabela com todas as razões entre a média 1B sobre OB, gráficos de estimativa pra todos os DNAs satélites que apresentam sinais de FISH no

cromossomo B e uma tabela contendo todos os valores de estimativa para os 101 satélites de *P. paranae* (Tabela 4; Figura 7; Tabela 4 do material suplementar).

Tabela 4. Razão de abundância entre duas bibliotecas 1B e três bibliotecas 0B de *Psalidodon paranae*, em ordem decrescente.

Satélite	Proporção abundância 1B (duas bibliotecas) / 0B (três bibliotecas)
ApaSat45	14,1468970935
ApaSat39	13,1506872276
ApaSat20	7,1787089014
ApaSat37	4,1751739563
ApaSat44	3,3507738548
ApaSat40	3,1935135239
ApaSat33	3,0771942073
ApaSat68	3,0115660490
ApaSat92	2,5742031953
ApaSat59	2,5437454069
ApaSat06	2,3279373420
ApaSat89	2,1519903975
ApaSat23	2,0645295734
ApaSat100	2,0347273481
ApaSat95	2,0321326673
ApaSat01	1,9671525181
ApaSat69	1,9602928291
ApaSat13	1,9124204264
ApaSat53	1,8291068568
ApaSat25	1,8106536593
ApaSat83	1,7603784323
ApaSat15	1,7327895883
ApaSat46	1,7315103275
ApaSat88	1,7100103835
ApaSat47	1,7068095541
ApaSat07	1,6667879304

ApaSat34	1,5642448065
ApaSat24	1,5319321185
ApaSat32	1,5275174157
ApaSat08	1,4693409149
ApaSat60	1,4555341915
ApaSat26	1,4443944818
ApaSat71	1,4438050525
ApaSat101	1,4217785816
ApaSat57	1,3607041640
ApaSat65	1,3501855906
ApaSat76	1,3298433487
ApaSat10	1,3220403283
ApaSat54	1,3021816808
ApaSat14	1,2799609105
ApaSat81	1,2509524074
ApaSat30	1,2253457075
ApaSat85	1,2171030039
ApaSat22	1,2097983724
ApaSat43	1,1987378484
ApaSat62	1,1798110012
ApaSat58	1,1737881757
ApaSat03	1,1698802404
ApaSat02	1,1636372915
ApaSat42	1,1539778455
ApaSat50	1,1507463408
ApaSat87	1,1460990383
ApaSat63	1,1453851155
ApaSat61	1,1451836296
ApaSat04	1,1278092515
ApaSat27	1,0816774438
ApaSat35	1,0617270888
ApaSat18	1,0469504562
ApaSat51	1,0454617765

ApaSat55	1,0345591577
ApaSat41	1,0035410195
ApaSat67	1,0017597322
ApaSat21	0,9789349482
ApaSat49	0,9649657107
ApaSat16	0,9633314383
ApaSat66	0,9627146739
ApaSat52	0,9606574257
ApaSat29	0,9534866864
ApaSat79	0,9470791953
ApaSat99	0,9470682527
ApaSat82	0,9469643244
ApaSat05	0,9457309055
ApaSat86	0,9414027042
ApaSat93	0,9311163836
ApaSat75	0,9290309644
ApaSat74	0,9282921380
ApaSat28	0,9229549388
ApaSat91	0,9216477176
ApaSat11	0,9189617609
ApaSat78	0,9159310631
ApaSat96	0,9152955296
ApaSat94	0,8925118341
ApaSat73	0,8888230565
ApaSat70	0,8764817592
ApaSat12	0,8653731423
ApaSat17	0,8514517715
ApaSat80	0,8454054046
ApaSat19	0,8437746469
ApaSat36	0,8318689194
ApaSat56	0,8305208088
ApaSat48	0,8211582298
ApaSat38	0,8139975636

ApaSat98	0,8078395605
ApaSat97	0,8006195116
ApaSat09	0,7944335279
ApaSat64	0,7912751976
ApaSat84	0,7573626247
ApaSat72	0,7512246652
ApaSat90	0,6997253119
ApaSat77	0,6841857241
ApaSat31	0,3616027789

897

898

899

900

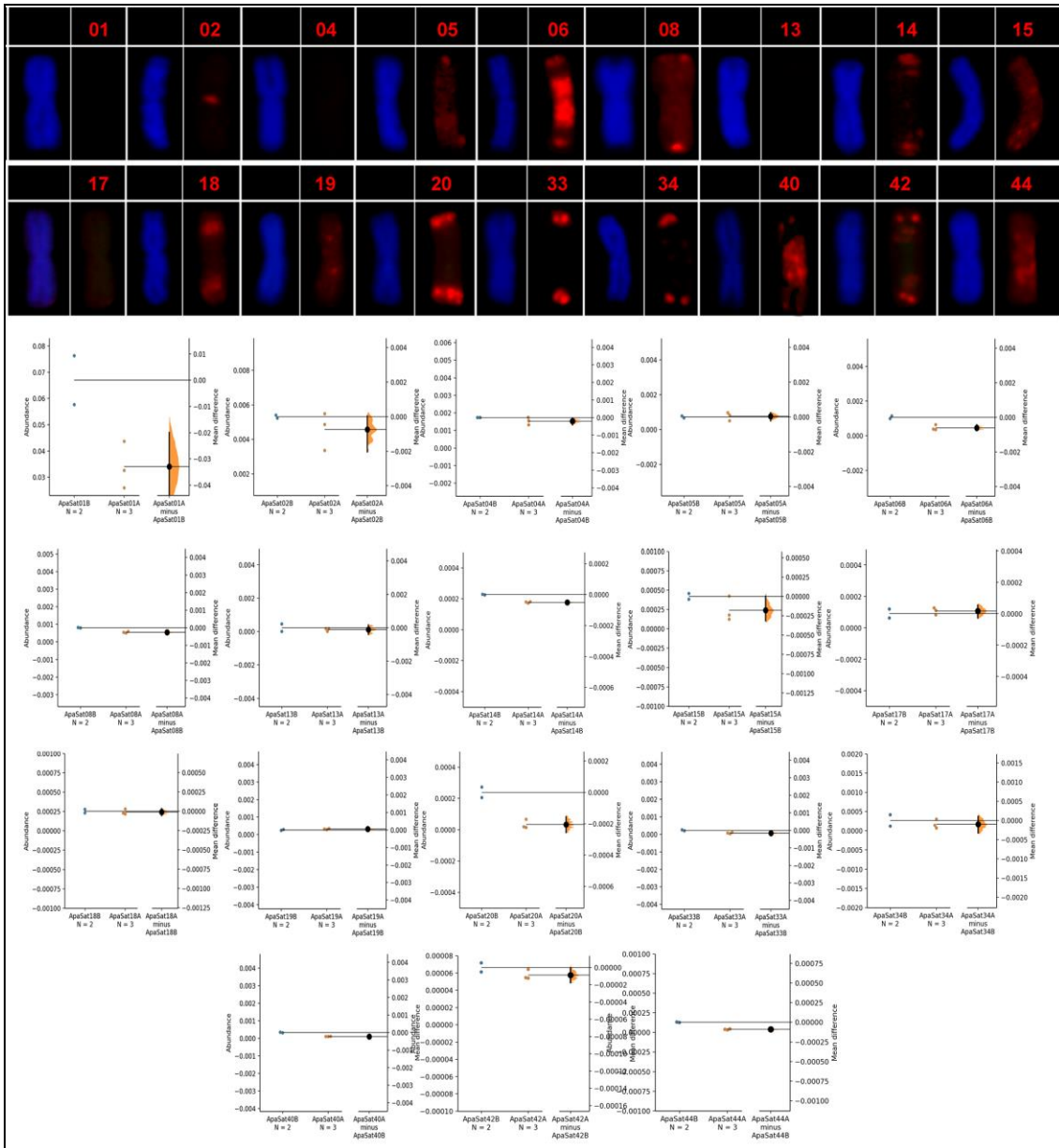


Figura 7. Cromossomos B marcados retirados do trabalho de Silva et al. 2017, no qual cada número representa seu satélite (01 = ApaSat01...44 = ApaSat44) e gráficos de confiança para todos os satélites que apresentaram sinais de FISH no cromossomo B. No gráfico de confiança ambos os grupos são plotados nos eixos esquerdos; a diferença média é plotada em eixos flutuantes à direita como uma distribuição de amostragem *bootstrap*. A diferença média é representada como um ponto; o intervalo de confiança de 95% é indicado pelas extremidades da barra de erro vertical.

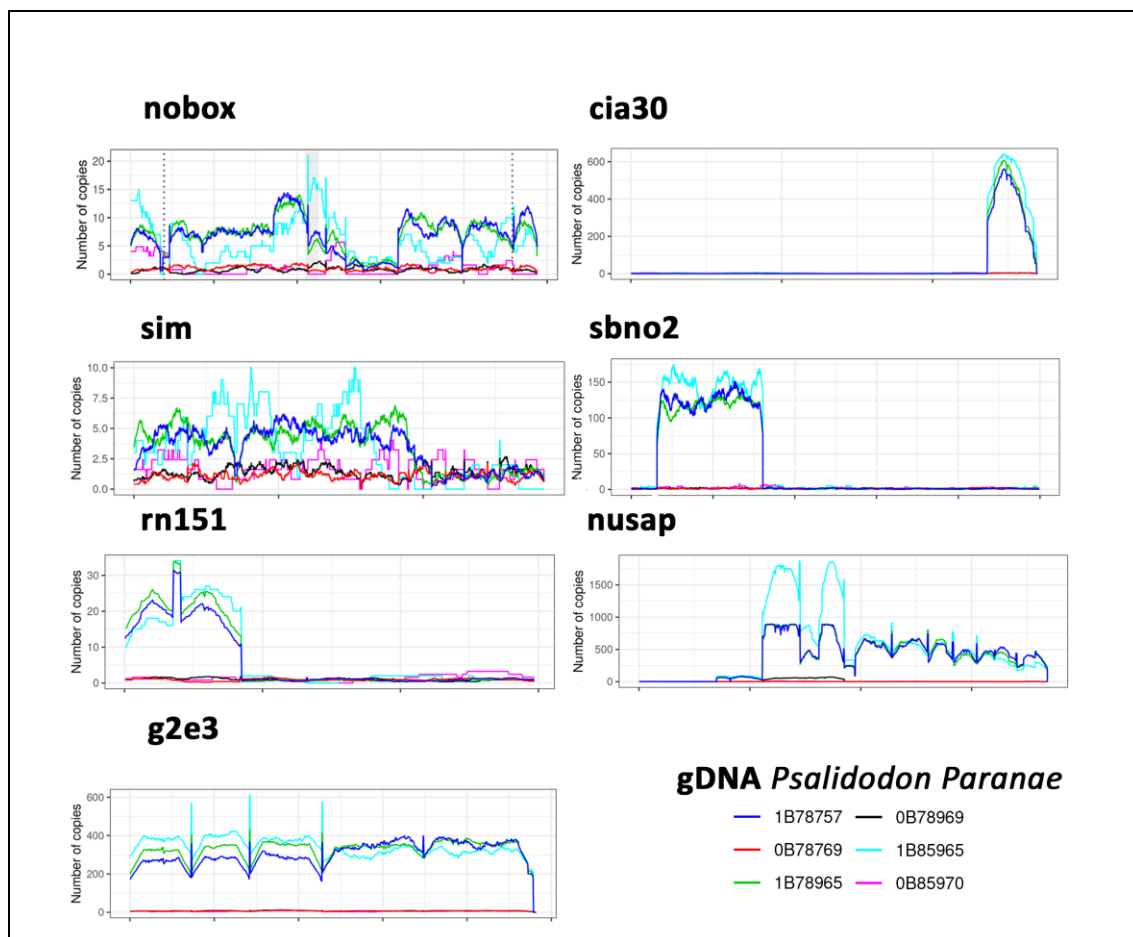


Figura 8. Gráficos de cobertura de genes presentes no cromossomo B de *Psalidodon paranae* para as seis bibliotecas de estudo. Notar a ocorrência de picos diferenciais da biblioteca 1B (85965) em relação às outras duas bibliotecas 1B (78775 e 78965).

6 Discussão

6.1 Expansão do satelitoma de *Psalidodon paranae* e um comparativo entre as ferramentas *RepeatExplorer1* e TAREAN

Utilizando a ferramenta *RepeatExplorer1*, Silva et al. (2017) caracterizou de pela primeira vez o satelitoma de um peixe Neotropical, *Psalidodon paranae*. Em seu trabalho foram apresentadas 45 famílias de DNA satélites para a espécie, nas quais o ApaSat02 e ApaSat04 foram considerados os únicos satélites a pertencerem a uma mesma superfamília, além da apresentação de 16 satélites pertencentes ao cromossomo B (Figura 7), sendo dois deles exclusivos deste cromossomo, ApaSat20 e ApaSat44. Neste contexto, o presente trabalho utilizou mais uma ferramenta de análise, o TAREAN (Novák et al. 2017), caracterizando 56 novos satélites para a espécie *P. paranae* (ApaSat46 até ApaSat101), além de 37 famílias anteriormente já caracterizadas

(totalizando 93 satélites caracterizados pelo TAREAN). Assim, 56 satélites foram caracterizados exclusivamente pelo TAREAN, enquanto o ApaSat22, ApaSat25, ApaSat28, ApaSat30, ApaSat32, ApaSat37, ApaSat41 e ApaSat42 foram DNAs satélites exclusivos da análise de *RepeatExplorer1*. Com a expansão da coleção de DNA satélite em *P. paranae*, setes novas superfamílias de DNA satélites foram caracterizadas, totalizando oito superfamílias para a espécie, dos quais quatro foram encontradas exclusivamente pelo TAREAN (SF2- ApaSat01 e ApaSat46; SF3- ApaSat30, ApaSat69 e ApaSat73; SF4- ApaSat49, ApaSat50 e ApaSat59; SF5- ApaSat61 e ApaSat93; SF6- ApaSat87 e ApaSat88; SF7- ApaSat90 e ApaSat100; SF8- ApaSat95 e ApaSat98).

A Superfamília 2 é de extrema importância, pois o ApaSat01, também conhecido como As51, possui uma variante de 39 pares de bases (ApaSat46), sendo estes os dois satélites mais abundantes da espécie, como notado nas Tabela 1 e 2 do material suplementar, além de dois satélites extremamente conservados no gênero *Psalidodon* (Goes et al. 2022). Utilizando de *reads* longos é possível notar a abundância destes dois satélites, como apresentado nas figuras 1 e 2 do material suplementar, ambos possuindo a maior quantidade de *reads* no atual estudo, além do ApaSat01 contendo *arrays* enormes, chegando a valores de 14 mil pares de bases.

A composição A+T manteve-se maior que 50%, como nos 45 primeiros satélites, tendo uma porcentagem de 57,4% e corroborando a tendência de que DNAs satélites são mais ricos em conteúdo A+T (Silva et al. 2017; Utsunomia et al. 2019; Serrano-Freitas et al. 2020; Stornioli et al. 2021). A composição A+T nos DNAs satélites encontrados exclusivamente por TAREAN (ApaSat46 até ApaSat101) teve um acréscimo de 3.2% em relação aos 45 primeiros satélites da espécie e uma diferenciação de 3.5% de comparado com os satélites exclusivos do *RepeatExplorer1* e uma diferença de 1.97% ao comparar os 45 satélites caracterizados por *RepeatExplorer1* e os 93 caracterizados por TAREAN. Dos sete DNAs satélites que possuem seu conteúdo A+T superior a 70%, quatro deles são considerados satDNAs complexos, porém seis deles foram encontrados utilizando a ferramenta TAREAN, resultando em indícios que não há uma relação direta entre tamanho de satélite e seu conteúdo A+T, mas sim no modo de que estes DNAs satélites foram caracterizados, visto que todos os satélites superiores a 2 mil pares de bases e 85,7% dos satélites mais ricos em AT foram caracterizados pela ferramenta TAREAN.

A utilização do TAREAN para caracterizar um satelitoma também tende à caracterização de satélites complexos, com monômeros superiores a dois mil pares de bases e como apresentado por Novák et al. (2017), podendo chegar em até 5 kb. De fato, no satelitoma da espécie de peixe *Prochilodus lineatus*, caracterizado com TAREAN, um DNA satélite com monômero de 3008 bp (PliSat7) foi apresentado (Stornioli e Goes et al. 2021), enquanto em *Triportheus auritus*, satélites que alcançam 2229 bp (TauSat04) também foram descritos (Kretschmer e Goes et al. 2022). Em contrapartida, análises de satelitômica comparativa que utilizam apenas a *pipeline RepeatExplorer1*, tendem a encontrar monômeros com tamanho reduzido, como visto em Silva et al. (2017), que disponibilizaram um catálogo com poucos satélites complexos, sendo o maior com 365 pares de base (ApaSat17). Em Utsunomia et al. (2019), aproximadamente um terço do satelitoma de *Megaleporinus macrocephalus* é considerado complexo (57 satélites), no entanto não há a ocorrência de nenhum monômero superior a dois mil pares de bases. Também apresentado em Serrano-Freitas et al. (2020), ao realizar a caracterização do satelitoma de *Characidium gomesi*, foi apresentado o CgoSat35 como o maior do satelitoma da espécie, contendo 677 pares de base. Nos novos 56 satélites caracterizados no atual trabalho, foram encontrados pelo TAREAN dois novos satélites com seu tamanho superior a dois mil pares de bases, sendo eles o ApaSat48, contendo 4841 pb e o ApaSat60, contendo 2258 pb, além de que aproximadamente 50% dos satélites caracterizados por TAREAN de *P. paranae* possuem seu monômero superior a 100 bp, sendo considerados complexos.

A complementariedade das duas ferramentas utilizadas (*RepeatExplorer1* e TAREAN) indica que a aplicação de ambos os métodos permite a caracterização de um catálogo mais completo de DNAs satélites em uma espécie pois, dentre os oito satélites exclusivos em *RepeatExplorer1*, alguns, como o ApaSat28, está organizado em pequenos *arrays* (800 pares de base por NCRF), ou, ate mesmo, longos *arrays* (ApaSat22, ApaSat25 e ApaSat42).

Dos 56 novos satDNAs caracterizados, foram produzidas 10 sondas para satélites abundantes ou com diferenciação em bibliotecas 1B/0B (Tabela 4, Figura 6), sendo três destes *clusterizados* em nível cromossômico, o ApaSat49, escolhido devido à sua abundância, ApaSat53 e ApaSat54, escolhidos devido a possível presença no cromossomo B. Os experimentos de FISH do ApaSat49 e ApaSat53 revelaram marcações

dispersas nos cromossomos, enquanto a FISH com ApaSat54 mostrou sinais cromossômicos em 4 pares de cromossomos autossômicos. Considerando os três casos, apenas ApaSat53 foi mapeado visivelmente no cromossomo B. Os *raincloud plots* nos auxiliam na compreensão da *clusterização* do ApaSat54, devido a este possuir *arrays* que chegam em até 22.5 mil pares de base (Figura 2 suplementar), já que *arrays* com um tamanho maior de 10 mil pares de bases tendem a apresentar sinais clusterizados nas técnicas de FISH (dos Santos et al. 2021).

O presente trabalho demonstra que a caracterização de DNA satélites utilizando a ferramenta TAREAN pode ser de grande auxílio para a complementação de satelitomas antes caracterizados por *RepeatExplorer1*, mais que dobrando o catálogo de satélites da espécie *P. paranae*, caracterizando satélites complexos com monômeros de tamanho raramente apresentados anteriormente, além de um aumento do número de superfamílias na espécie, como o caso do As51 (ApaSat01) e sua variante de 39 pb (ApaSat46), sendo estes os DNAs satélites mais abundantes da espécie. A caracterização por TAREAN e o estudo de abundância e divergência também auxiliou na caracterização de novos satDNAs extremamente abundantes, como o caso do ApaSat46, ApaSat47, ApaSat48, ApaSat49 e ApaSat50, todos mais abundantes que o ApaSat03 e DNAs satélites com potencial de *clusterização* a nível cromossômico, como apresentado pelo ApaSat49, ApaSat53 e ApaSat54, além de um decréscimo na média de divergência Kiruma dos satélites caracterizados por TAREAN se comparado aos do *RepeatExplorer1*.

6.2 Análise do satelitoma de *Psalidodon paranae* em *reads* longos

Os *Raincloud plots* foram de grande auxílio para a explicação de dúvidas apresentadas ao longo do trabalho, como por exemplo a ausência de *clusterização* no cromossomo B do ApaSat39 e no ApaSat45 em Silva et al. (2017), sendo que estes dois satélites foram os que apresentaram os maiores valores na proporção 1B/0B nas 6 bibliotecas de estudo (10,5 e 10,3 respectivamente) e gráficos de estimativa com médias diferenciais (figura 7). No entanto, é possível notar nos *boxplots* destes dois satélites (Figura 1 do material suplementar) e em suas análises em *reads* longos, tabela 2, que ambos possuem somente um *read* contendo seu *array* além, de seu pequeno tamanho (3100 bp e 2725 bp, respectivamente), podendo explicar que mesmo

ocorrendo sinais em cromossomos autossômicos há sua ausência no cromossomo B, apesar de sua alta abundância nas bibliotecas 1B.

Estas análises nos ajudam a melhor compreender, por exemplo, a *clusterização* do ApaSat54, pois em seu gráfico *boxplot* existem valores discrepantes (*outliers*) contendo quatro *reads* nos quais os *arrays* são maiores do que 10kb, chegando até 22,5kb (mediana = 4434 pb) para um total de 65 *reads* analisados com 343849 bp totais, contando a soma total dos *arrays*, possibilitando então sua *clusterização* a nível cromossômico (Figura 4). O estudo do ApaSat54 auxilia no entendimento da ausência de sinais a níveis cromossômicos do apaSat57 e ApaSat68, em que ambos possuem uma grande quantidade de *reads* (607 e 107, respectivamente), no entanto, possuem uma mediana destes *reads* muito baixa, sendo em ApaSat57 um valor quatro vezes menor ao de ApaSat54, explicando a ausência de clusterização em nível cromossômico destes DNAs satélites, mesmo que ambos abundantes na espécie, além de corroborar com a ideia de que não somente a quantidade de *arrays*, mas também os respectivos tamanhos são fundamentais para a observação de sinais visíveis de FISH.

Um satélite de grande importância para a citogenética de peixes, o As51 (Mestriner et al. 2000), também conhecido por ApaSat01 (Silva et al. 2017) é um caso atípico nesta análise, pois como apresentado em trabalhos anteriores (Abel et al. 2006; Kantek et al. 2009; Vicari et al. 2010; Vicari et al. 2011), este satélite possui marcações em diversos cromossomos de espécies de *Psalidodon*, como em *P. paranae*. Nossas análises revelam uma baixa mediana e média do tamanho de seus *arrays* (1556 e 1972, respectivamente), porém com sete *arrays* superiores a 10 mil pares de bases e este satélite e sua variante (ApaSat46) possuem as maiores abundâncias no estudo com as bibliotecas de *reads* longos, apresentando 8483818 bp ao longo dos *arrays* deste DNA satélite, sendo estes *arrays* superiores a 10 kb e sua enorme abundância os principais motivos de sinais estarem presente em diversos cromossomos da espécie *P. paranae*.

Os estudos em *reads* longos nos apresentaram um possível foco para futuras análises, o ApaSat62, que mesmo possuindo uma mediana próxima a 3kb, possuem cinco *reads* com *arrays* superiores a 10kb e quase 200000 bp presentes em seus *arrays*, sendo então este DNA satélite um bom candidato para futuros experimentos de FISH. O atual trabalho apresentou o estudo com sequências de *reads* longos uma importante ferramenta para a compreensão de como ocorre a estruturação genômica dos DNAs

satélites encontrados tanto por *RepeatExplorer1* quanto por TAREAN, nos dando indícios de possíveis candidatos a apresentarem marcações cromossômicas, além de ser mais um método de auxílio para o estudo de abundância do catálogo de DNAs satélites de uma espécie, como apresentado uma extrema abundância da Superfamília 2 (ApaSat01 e ApaSat46) no qual, em pares de bases, representam quase 74% do estudo em *reads* longos.

6.3 Estudo de abundância em bibliotecas de *reads* curtos OB e 1B

Com o estudo de abundância em cinco bibliotecas de *reads* curtos, três delas de espécimes OB e duas de espécimes 1B, foi possível caracterizar quais dos satélites pertencentes ao catálogo de *Psilidodon paranae* estão presentes em grande escala nos cromossomos B da espécie, mesmo quando não há marcação cromossômica, como o caso do ApaSat39 e ApaSat45, que possuem valores de abundância muito maiores nas bibliotecas contendo 1B, como apresentado nas tabelas 1 e 2 do material suplementar, além de altos valores na proporção de abundância entre as bibliotecas 1B e OB (13,1 e 14,1, receptivamente) (Tabela 4).

Dois casos interessantes de DNAs satélites exclusivos do cromossomo B em nível cromossômico em *P. paranae* são ApaSat20 e ApaSat44 (Silva et al. 2017), presentes em uma baixa abundância em biblioteca OB. Nossas análises, utilizando agora cinco bibliotecas para a comparação da razão entre a abundância das três bibliotecas 1B pelas três bibliotecas OB, revelou em ApaSat20 uma razão de 7,1 enquanto do ApaSat44 resulta em 3,3, apresentando também uma concentração destes nas bibliotecas de indivíduos com cromossomo B da espécie. Utilizando gráficos de estimativa também é possível notar grande diferença entre as duas bibliotecas 1B, se comparadas as três bibliotecas OB, auxiliando na compreensão de sinais exclusivos aos cromossomos B da espécie (Figura 7).

Referente ao As51, também conhecido por ApaSat01, este possui um alto valor entre a proporção de abundância 1B/OB (1.9) e não é *clusterizado* no cromossomo B. Ao realizar o estudo de estatísticas de estimativa é possível notar que as duas bibliotecas 1B também são dadas como diferentes em relação às três OB (Figura 7). O provável motivo das duas análises apresentarem um resultado contrário aos sinais na FISH pode ser apresentado na análise em *longs reads*, visto que o ApaSat01 e ApaSat46 juntos

representam mais de 70% de toda a análise dos *arrays*, sendo extremamente presentes na espécie *P. paranae*, muito provavelmente possuindo uma quantidade de seus monômeros no cromossomo B, o que leva a essa diferenciação das bibliotecas, mas não a ponto de formar *arrays* de tamanho possíveis a acarretar sinais a níveis cromossômicos, pois a maior porcentagem de seus *arrays* possuem tamanhos próximos a 1500bp (Tabela 2).

Utilizando somente o método de razão entre bibliotecas 1B/0B, com exceção ApaSat05 e ApaSat19, que possuem uma razão inferior a 1, todos os resultados corroboram com os sinais em FISH dos 16 satélites (Tabela 4). Ao utilizar o estudo gráfico Gardner-Altman, nota-se que o ApaSat05, ApaSat18 e ApaSat19 apresentaram um padrão comum entre os dois tipos de bibliotecas, com resultados muito similares ao método de estudo entre a razão, o caso do ApaSat18 pode ser explicado por possuir uma biblioteca 0B (78769) com uma abundância de satélites superior às outras duas (Figura 7; Tabela 2 do material suplementar), podendo ser observados em alguns casos como ApaSat01, ApaSat02, ApaSat15, ApaSat20 e ApaSat34. Por outro lado, ApaSat23, ApaSat37, ApaSat39 e ApaSat45, apresentaram valores diferenciais nos dois métodos de análises a não são clusterizados em nível cromossômico no cromossomo B. Uma alternativa para explicar a não ocorrência de sinais destes 4 satélites no B é que o ApaSat39 e ApaSat45 possuem apenas 1 *array* em nossas análises em *reads* longos, enquanto o ApaSat23 possui somente 16 *arrays* com uma baixa abundância (Tabela 2).

O estudo utilizando mais de uma biblioteca de *reads* curtos quando se deseja um enfoque em algum diferencial, no caso do presente estudo, abundância e cromossomo B, auxilia para dados estatísticos mais concisos, o que retorna uma maior precisão de quais DNAs satélites podem ser candidatos para futuras análises citogenéticas. A utilização de mais de uma biblioteca de *reads* curtos também garante uma maior confiabilidade nos dados gerados em ordem de abundância de cada satélite que forma o satelitoma da espécie, se comparado separadamente os satélites caracterizados por *RepeatExplorer1*, ApaSat01 até o ApaSat45, e TAREAN, ApaSat46 até o ApaSat101. Estas análises retornam uma melhor confiabilidade de quanto um satélite realmente é abundante em uma espécie, visto que a criação do satelitoma é realizada, em grande parte dos casos, utilizando a biblioteca de somente um indivíduo e com isso, satélite que possuem uma baixa abundância pode ser melhor estudado tendo como

base mais de um indivíduo a espécie, como o caso do ApaSat37, que no atual trabalho se tornou o satélite menos abundante da espécie.

7 Referências

- ABEL L. D. D. S.; MANTOVANI, M.; MOREIRA-FILHO, Orlando. Chromosomal distribution of the As51 satellite DNA in two species complexes of the genus *Astyanax* (Pisces, Characidae). **Genetics and Molecular Biology**, v. 29, n. 3, p. 448-452, 2006.
- ALBERT, James S.; TAGLIACOLLO, Victor A.; DAGOSTA, Fernando. Diversification of Neotropical freshwater fishes. **Annual Review of Ecology, Evolution, and Systematics**, v. 51, n. 1, p. 27-53, 2020.
- ALLEN, Micah et al. Raincloud plots: a multi-platform tool for robust data visualization. **Wellcome open research**, v. 4, 2019.
- ARTONI, Roberto Ferreira et al. Banding pattern of A and B chromosomes of *Prochilodus lineatus* (Characiformes, Prochilodontidae), with comments on B chromosomes evolution. **Genetica**, v. 127, n. 1, p. 277-284, 2006.
- ARTONI, Roberto Ferreira et al. *Astyanax* aff. *fasciatus* Cuvier, 1819 (Teleostei; Characidae): evidences of a species complex in the upper rio Tibagi basin (Paraná, Brazil). **Neotropical Ichthyology**, v. 4, n. 2, p. 197-202, 2006.
- BENSON, Gary. Tandem repeats finder: a program to analyze DNA sequences. **Nucleic Acids Research**, v. 27, n. 2, p. 573-580, 1999.
- BERLIN, Konstantin et al. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. **Nature biotechnology**, v. 33, n. 6, p. 623, 2015.
- BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, 2014.
- BOTCHAN, Michael R. Bovine satellite I DNA consists of repetitive units 1,400 base pairs in length. **Nature**, v. 251, n. 5473, p. 288-292, 1974.
- CAMACHO, Juan Pedro M.; SHARBEL, Timothy F.; BEUKEBOOM, Leo W. B-chromosome evolution. **Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences**, v. 355, n. 1394, p. 163-178, 2000.
- CECHOVA, Monika et al. High satellite repeat turnover in great apes studied with short- and long-read technologies. **Molecular biology and evolution**, v. 36, n. 11, p. 2415-2431, 2019.

1151 CHAISSON, Mark J.; TESLER, Glenn. Mapping single molecule sequencing reads using
 1152 basic local alignment with successive refinement (BLASR): application and theory.
 1153 **BMC bioinformatics**, v. 13, n. 1, p. 238, 2012.

1154 CHIN, Chen-Shan et al. Phased diploid genome assembly with single-molecule real-time
 1155 sequencing. **Nature methods**, v. 13, n. 12, p. 1050, 2016.

1156 CIOFFI, M. B. et al. Chromosomal variability among allopatric populations of
 1157 Erythrinidae fish *Hoplias malabaricus*: mapping of three classes of repetitive DNAs.
 1158 **Cytogenetic and Genome Research**, v. 125, n. 2, p. 132-141, 2009.

1159 CIOFFI, M. B.; BERTOLLO, L. A. C. Initial steps in XY chromosome differentiation in
 1160 *Hoplias malabaricus* and the origin of an X 1 X 2 Y sex chromosome system in this
 1161 fish group. **Heredity**, v. 105, n. 6, p. 554-561, 2010.

1162 CREPALDI, Carolina; PARISE-MALTEMPI, Patricia P. Heteromorphic sex chromosomes
 1163 and their DNA content in fish: an insight through satellite DNA accumulation in
 1164 *Megaleporinus elongatus*. **Cytogenetic and Genome Research**, v. 160, n. 1, p. 38-
 1165 46, 2020.

1166 DE JESUS, Célia Maria et al. Molecular characterization and chromosomal localization of
 1167 two families of satellite DNA in *Prochilodus lineatus* (Pisces, Prochilodontidae), a
 1168 species with B chromosomes. **Genetica**, v. 118, n. 1, p. 25-32, 2003.

1169 DIERCKXSENS, Nicolas; MARDULYN, Patrick; SMITS, Guillaume. NOVOPlasty: de novo
 1170 assembly of organelle genomes from whole genome data. **Nucleic acids research**,
 1171 v. 45, n. 4, p. e18-e18, 2016.

1172 DOS SANTOS, Rodrigo Zeni et al. A long-term conserved satellite DNA that remains
 1173 unexpanded in several genomes of Characiformes fish is actively transcribed.
 1174 **Genome biology and evolution**, v. 13, n. 2, p. evab002, 2021.

1175 DOVER, Gabriel A. Molecular drive: a cohesive mode of species evolution. **Nature**, v.
 1176 299, n. 5879, p. 111-117, 1982.

1177 DOVER, Gabriel A. Molecular drive in multigene families: how biological novelties arise,
 1178 spread and are assimilated. **Trends in genetics**, v. 2, p. 159-165, 1986.

1179 DOVER, Gabriel. Molecular drive. **Trends in Genetics**, v. 18, n. 11, p. 587-589, 2002.

1180 FRICKE, R.; ESCHMEYER, W. N.; FONG, J. D. **Eschmeyer's catalog of fishes: species by**
 1181 **family/subfamily**, 2020.

1182 ESCUDEIRO, Ana et al. Conservation, divergence, and functions of centromeric satellite

1183 DNA families in the Bovidae. **Genome biology and evolution**, v. 11, n. 4, p. 1152-
 1184 1165, 2019.

1185 FERREE, Patrick M. Sex differences: satellite DNA directs male-specific gene expression.
 1186 **Current Biology**, v. 27, n. 10, p. R378-R380, 2017.

1187 FERREIRA, I. A.; BERTOLLO, L. A. C.; MARTINS, C. Comparative chromosome mapping of
 1188 5S rDNA and 5SHindIII repetitive sequences in Erythrinidae fishes (Characiformes)
 1189 with emphasis on the *Hoplias malabaricus* 'species complex'. **Cytogenetic and**
 1190 **Genome Research**, v. 118, n. 1, p. 78-83, 2007.

1191 FORESTI, F. et al. Polymorphic nature of nucleolus organizer regions in fishes.
 1192 **Cytogenetic and Genome Research**, v. 31, n. 3, p. 137-144, 1981.

1193 FRY, Kirk; SALSER, Winston. Nucleotide sequences of HS- α satellite DNA from kangaroo
 1194 rat *Dipodomys ordii* and characterization of similar sequences in other rodents.
 1195 **Cell**, v. 12, n. 4, p. 1069-1084, 1977.

1196 GARRIDO-RAMOS, Manuel A. Satellite DNA in plants: more than just rubbish.
 1197 **Cytogenetic and genome research**, v. 146, n. 2, p. 153-170, 2015.

1198 GARRIDO-RAMOS, Manuel A. Satellite DNA: An evolving topic. **Genes**, v. 8, n. 9, p. 230,
 1199 2017.

1200 GREGORY, Richard I. et al. Human RISC couples microRNA biogenesis and
 1201 posttranscriptional gene silencing. **Cell**, v. 123, n. 4, p. 631-640, 2005.

1202 GOES, Caio Augusto Gomes et al. Revealing the satellite DNA history in *Psalidodon* and
 1203 *Astyanax* characid fish by comparative satellitomics. **Frontiers in Genetics**, p. 1567.

1204 HARRIS, Robert S.; CECHOVA, Monika; MAKOVA, Kateryna D. Noise-cancelling repeat
 1205 finder: uncovering tandem repeats in error-prone long-read sequencing data.
 1206 **Bioinformatics**, v. 35, n. 22, p. 4809-4811, 2019.

1207 HO, Jose et al. Moving beyond P values: data analysis with estimation graphics. **Nature**
 1208 **methods**, v. 16, n. 7, p. 565-566, 2019.

1209 HÖRZ, Wolfram; ZACHAU, Hans G. Characterization of distinct segments in mouse
 1210 satellite DNA by restriction nucleases. **European Journal of Biochemistry**, v. 73, n.
 1211 2, p. 383-392, 1977.

1212 HUTTON, James R.; WETMUR, James G. Effect of chemical modification on the rate of
 1213 renaturation of deoxyribonucleic acid. Deaminated and glyoxalated
 1214 deoxyribonucleic acid. **Biochemistry**, v. 12, n. 3, p. 558-563, 1973.

1215 IWASAKI, Wataru et al. MitoFish and MitoAnnotator: a mitochondrial genome database
 1216 of fish with an accurate and automatic annotation pipeline. **Molecular biology and**
 1217 **evolution**, v. 30, n. 11, p. 2531-2540, 2013.

1218 JESUS, C.M. et al. Molecular characterization and chromosomal localization of two
 1219 families of satellite DNA in *Prochilodus lineatus* (Pisces, Prochilodontidae), a
 1220 species with B chromosomes. **Genetica**, v. 118, n. 1, p. 25-32, 2003.

1221 KANTEK, D. L. Z. et al. Chromosomal location and distribution of As51 satellite DNA in
 1222 five species of the genus *Astyanax* (Teleostei, Characidae, Incertae sedis). **Journal**
 1223 **of Fish Biology**, v. 75, n. 2, p. 408-421, 2009.

1224 KAVALCO, Karine Frehner et al. Biogeographic patterns in the chromosomal distribution
 1225 of a satellite DNA in the banded tetra *Astyanax fasciatus* (Teleostei:
 1226 Characiformes). **Organisms Diversity & Evolution**, v. 13, n. 1, p. 67-76, 2013.

1227 KRETSCHMER, Rafael et al. Satellitome analysis illuminates the evolution of ZW sex
 1228 chromosomes of Triportheidae fishes (Teleostei: Characiformes). **Chromosoma**, p.
 1229 1-17, 2022.

1230 KHOST, Daniel E.; EICKBUSH, Danna G.; LARRACUENTE, Amanda M. Single-molecule
 1231 sequencing resolves the detailed structure of complex satellite DNA loci in
 1232 *Drosophila melanogaster*. **Genome research**, v. 27, n. 5, p. 709-721, 2017.

1233 KIM, Young Bun et al. Divergence of *Drosophila melanogaster* repeatomes in response
 1234 to a sharp microclimate contrast in Evolution Canyon, Israel. **Proceedings of the**
 1235 **National Academy of Sciences**, v. 111, n. 29, p. 10630-10635, 2014.

1236 KIT, Saul; HSU, T. C. Relative stability to thermal denaturation of deoxyribonucleic acid
 1237 (DNA) preparations containing bromodeoxyuridine. **Biochemical and Biophysical**
 1238 **Research Communications**, v. 5, n. 2, p. 120-124, 1961.

1239 KOEHLER, M. R. et al. Cytogenetics of the genus *Leporinus* (Pisces, Anostomidae). 1.
 1240 Karyotype analysis, heterochromatin distribution and sex chromosomes.
 1241 **Chromosome Research**, v. 5, n. 1, p. 12-22, 1997.

1242 KOPECKA, Helena et al. Restriction enzyme analysis of satellite DNA components from
 1243 the bovine genome. **European journal of biochemistry**, v. 84, n. 1, p. 189-195,
 1244 1978.

1245 KRSTICEVIC, Flavia J.; SCHRAGO, Carlos G.; CARVALHO, A. Bernardo. Long-read single
 1246 molecule sequencing to resolve tandem gene copies: The Mst77Y region on the

1247 *Drosophila melanogaster* Y chromosome. *G3: Genes, Genomes, Genetics*, v. 5, n. 6,
1248 p. 1145-1150, 2015.

1249 KUHN, Gustavo CS et al. The 1.688 repetitive DNA of *Drosophila*: concerted evolution at
1250 different genomic scales and association with genes. ***Molecular biology and***
1251 ***evolution***, v. 29, n. 1, p. 7-11, 2012.

1252 LEE, Hye-Ran et al. Transcription and evolutionary dynamics of the centromeric satellite
1253 repeat CentO in rice. ***Molecular biology and evolution***, v. 23, n. 12, p. 2505-2520,
1254 2006.

1255 LÓPEZ-FLORES, I.; GARRIDO-RAMOS, M. A. The repetitive DNA content of eukaryotic
1256 genomes. In: Repetitive DNA. **Karger Publishers**, 2012. p. 1-28.

1257 LOUZADA, Sandra et al. Decoding the role of satellite DNA in genome architecture and
1258 plasticity—An evolutionary and clinical affair. ***Genes***, v. 11, n. 1, p. 72, 2020.

1259 MESTRINER, Carlos A. et al. Structural and functional evidence that a B chromosome in
1260 the characid fish *Astyanax scabripinnis* is an isochromosome. ***Heredity***, v. 85, n. 1,
1261 p. 1-9, 2000.

1262 MIGA, Karen H. Completing the human genome: the progress and challenge of satellite
1263 DNA assembly. ***Chromosome research***, v. 23, n. 3, p. 421-426, 2015.

1264 MONTIEL, Eugenia E. et al. Satellitome analysis of *Rhodnius prolixus*, one of the main
1265 Chagas disease vector species. ***International journal of molecular sciences***, v. 22, n.
1266 11, p. 6052, 2021.

1267 MOREIRA-FILHO, Orlando. *Astyanax scabripinnis* (Pisces, Characidae): a species
1268 complex. Brazil J. ***Genet.***, v. 14, p. 331-357, 1991. NOVÁK, Petr; NEUMANN, Pavel;
1269 MACAS, Jiří. Graph-based clustering and characterization of repetitive sequences in
1270 next-generation sequencing data. ***BMC bioinformatics***, v. 11, n. 1, p. 378, 2010.

1271 MELO, Bruno F. et al. Biogeography of curimatid fishes reveals multiple lowland–upland
1272 river transitions and differential diversification in the Neotropics (Teleostei,
1273 Curimatidae). ***Ecology and evolution***, v. 11, n. 22, p. 15815-15832, 2021.

1274 MOREIRA, Cristiano Rangel et al. Relações filogenéticas na ordem Characiformes
1275 (Teleostei: Ostariophysi). **Unpublished Ph. D. Dissertation, Universidade de São**
1276 **Paulo**, v. 468, 2007.

1277 NOVÁK, Petr et al. RepeatExplorer: a Galaxy-based web server for genome-wide
1278 characterization of eukaryotic repetitive elements from next-generation sequence

reads. **Bioinformatics**, v. 29, n. 6, p. 792-793, 2013.

NOVÁK, Petr et al. TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. **Nucleic acids research**, v. 45, n. 12, p. e111-e111, 2017.

OLIVEIRA, Claudio et al. Chromosome formulae of Neotropical freshwater fishes. **Revista Brasileira de Genética**, v. 11, n. 3, p. 577-624, 1988.

OTAKE, Koichiro et al. CENP-B creates alternative epigenetic chromatin states permissive for CENP-A or heterochromatin assembly. **Journal of cell science**, v. 133, n. 15, p. jcs243303, 2020.

PAIVA, Samuel Rezende; DERGAM, Jorge Abdala; MACHADO, Flavia. Determining management units in southeastern Brazil: the case of *Astyanax bimaculatus* (Linnaeus, 1758)(Teleostei: Ostariophysi: Characidae). **Hydrobiologia**, v. 560, n. 1, p. 393-404, 2006.

PALACIOS-GIMENEZ, Octavio Manuel et al. High-throughput analysis of the satellitome revealed enormous diversity of satellite DNAs in the neo-Y chromosome of the cricket *Eneoptera surinamensis*. **Scientific reports**, v. 7, n. 1, p. 1-11, 2017.

PAZZA, R.; KAVALCO, K. Frehner; BERTOLLO, L. A. C. Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 2–Chromosomal location of a satellite DNA. **Cytogenetic and Genome Research**, v. 122, n. 1, p. 61-66, 2008.

PHILLIPPY, Adam M. New advances in sequence assembly. 2017.

PINKEL, Dr; STRAUME, T.; GRAY, J. W. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. **Proceedings of the National Academy of Sciences**, v. 83, n. 9, p. 2934-2938, 1986.

PITA, Sebastian et al. Comparative repeatome analysis on *Triatoma infestans* Andean and Non-Andean lineages, main vector of Chagas disease. **PLoS One**, v. 12, n. 7, 2017.

PLOHL, Miroslav; MEŠTROVIĆ, Nevenka; MRAVINAC, Brankica. Satellite DNA evolution. In: **Repetitive DNA**. Karger Publishers, 2012. p. 126-152.

PLOHL, Miroslav; MEŠTROVIĆ, Nevenka; MRAVINAC, Brankica. Centromere identity from the DNA point of view. **Chromosoma**, v. 123, n. 4, p. 313-325, 2014.

REIS, R. E. et al. Fish biodiversity and conservation in South America. **Journal of fish biology**, v. 89, n. 1, p. 12-47, 2016.

1311 RUIZ-RUANO, Francisco J. et al. High-throughput analysis of the satellitome illuminates
 1312 satellite DNA evolution. **Scientific reports**, v. 6, p. 28333, 2016.

1313 SERRANO-FREITAS, Érica A. et al. Satellite DNA content of B chromosomes in the
 1314 characid fish *Characidium gomesi* supports their origin from sex chromosomes.
 1315 **Molecular Genetics and Genomics**, v. 295, n. 1, p. 195-207, 2020.

1316 SILVA, Bráulio SML et al. De novo identification of satellite DNAs in the sequenced
 1317 genomes of *Drosophila virilis* and *D. americana* using the RepeatExplorer and
 1318 TAREAN pipelines. **PLoS one**, v. 14, n. 12, p. e0223466, 2019.

1319 SILVA, D. M. Z. A. et al. Delimiting the origin of a B chromosome by FISH mapping,
 1320 chromosome painting and DNA sequence analysis in *Astyanax paranae* (Teleostei,
 1321 Characiformes). **PLoS One**, v. 9, n. 4, p. e94896, 2014.

1322 SILVA, D. M. Z. A. et al. Chromosomal mapping of repetitive DNA sequences in five
 1323 species of *Astyanax* (Characiformes, Characidae) reveals independent location of
 1324 U1 and U2 snRNA sites and association of U1 snRNA and 5S rDNA. **Cytogenetic and**
 1325 **genome research**, v. 146, n. 2, p. 144-152, 2015.

1326 SILVA, D.M.Z.A. et al. Origin of B chromosomes in the genus *Astyanax* (Characiformes,
 1327 Characidae) and the limits of chromosome painting. **Molecular Genetics and**
 1328 **Genomics**, v. 291, n. 3, p. 1407-1418, 2016.

1329 SILVA, D. M. Z. A. et al. High-throughput analysis unveils a highly shared satellite DNA
 1330 library among three species of fish genus *Astyanax*. **Scientific reports**, v. 7, n. 1, p.
 1331 1-12, 2017.

1332 SMIT, A. F. A.; HUBLEY, R.; GREEN, P. RepeatMasker Open-4.0. 2013–2015. 2015.

1333 SMIT, A. F. A.; HUBLEY, R.; GREEN, P. 1996–2010. RepeatMasker Open-3.0. 2017.

1334 SMITH, George P. Evolution of repeated DNA sequences by unequal crossover. **Science**,
 1335 v. 191, n. 4227, p. 528-535, 1976.

1336 STORNIOLI, José Henrique Forte et al. The B chromosomes of *Prochilodus lineatus*
 1337 (Teleostei, Characiformes) are highly enriched in satellite DNAs. **Cells**, v. 10, n. 6, p.
 1338 1527, 2021.

1339 TERÁN, Guillermo E.; BENITEZ, Mauricio F.; MIRANDE, J. Marcos. Opening the Trojan
 1340 horse: phylogeny of *Astyanax*, two new genera and resurrection of *Psalidodon*
 1341 (Teleostei: Characidae). **Zoological Journal of the Linnean Society**, v. 190, n. 4, p.
 1342 1217-1234, 2020.

1343 TRAYNOR, Sofie et al. Remodeling and destabilization of chromosome 1 pericentromeric
 1344 heterochromatin by SSX proteins. **Nucleic acids research**, v. 47, n. 13, p. 6668-
 1345 6684, 2019.

1346 TUNJIĆ-CVITANIĆ, Monika et al. Satellitome Analysis of the Pacific Oyster *Crassostrea*
 1347 *gigas* Reveals New Pattern of Satellite DNA Organization, Highly Scattered across
 1348 the Genome. **International journal of molecular sciences**, v. 22, n. 13, p. 6798,
 1349 2021.

1350 UTSUNOMIA, Ricardo et al. A glimpse into the satellite DNA library in characidae fish
 1351 (Teleostei, Characiformes). **Frontiers in genetics**, v. 8, p. 103, 2017.

1352 UTSUNOMIA, Ricardo et al. Uncovering the ancestry of B chromosomes in *Moenkhausia*
 1353 *sanctaeofilomenae* (Teleostei, Characidae). **PLoS One**, v. 11, n. 3, 2016.

1354 UTSUNOMIA, Ricardo et al. Satellitome landscape analysis of *Megaleporinus*
 1355 *macrocephalus* (Teleostei, Anostomidae) reveals intense accumulation of satellite
 1356 sequences on the heteromorphic sex chromosome. **Scientific reports**, v. 9, n. 1, p.
 1357 1-10, 2019.

1358 VANBUREN, Robert et al. Single-molecule sequencing of the desiccation-tolerant grass
 1359 *Oropetium thomaeum*. **Nature**, v. 527, n. 7579, p. 508-511, 2015.

1360 VICARI, M. R. et al. Colocalization of repetitive DNAs and silencing of major rRNA genes.
 1361 A case report of the fish *Astyanax janaeirensis*. **Cytogenetic and genome research**,
 1362 v. 122, n. 1, p. 67-72, 2008.

1363 VICARI, M. R. et al. Satellite DNA and chromosomes in Neotropical fishes: methods,
 1364 applications and perspectives. **Journal of Fish Biology**, v. 76, n. 5, p. 1094-1116,
 1365 2010.

1366 VICARI, Marcelo Ricardo et al. New insights on the origin of B chromosomes in *Astyanax*
 1367 *scabripinnis* obtained by chromosome painting and FISH. **Genetica**, v. 139, n. 8, p.
 1368 1073-1081, 2011.

1369 VICENTE, Vera Elisa et al. Origin and differentiation of a sex chromosome system in
 1370 *Parodon hilarii* (Pisces, Parodontidae). Satellite DNA, G-and C-banding. **Genetica**, v.
 1371 119, n. 2, p. 115-120, 2003.

1372 VITORINO, C. A. et al. Molecular cytogenetics and its contribution to the understanding
 1373 of the chromosomal diversification in *Hoplias malabaricus* (Characiformes). **Journal**
 1374 **of fish biology**, v. 78, n. 4, p. 1239-1248, 2011.

VOZDOVA, Miluse et al. Sequence Analysis and FISH Mapping of Four Satellite DNA Families among Cervidae. **Genes**, v. 11, n. 5, p. 584, 2020.

WEI, Kevin H.-C. et al. Correlated variation and population differentiation in satellite DNA abundance among lines of *Drosophila melanogaster*. **Proceedings of the National Academy of Sciences**, v. 111, n. 52, p. 18793-18798, 2014.

YUAN, Zihao et al. Comparative genome analysis of 52 fish species suggests differential associations of repetitive elements with their living aquatic environments. **BMC genomics**, v. 19, n. 1, p. 1-10, 2018.

Tabela 1. Abundância, média e desvio padrão das três bibliotecas 1B de *P. paranae*.

Satélites	Biblioteca 78757_B	Biblioteca 78965_B	Biblioteca 85965_B	Média da Abundância	Desvio Padrão 1B
ApaSat01	0,076289398	0,057660662	0,023355125	0,052435062	0,0268512474
ApaSat02	0,00539699	0,005228224	0,003788581	0,004804598	0,0008839338
ApaSat03	0,002854778	0,002585729	0,001874014	0,002438174	0,0005067583
ApaSat04	0,001733889	0,001734647	0,001400907	0,001623148	0,0001924664
ApaSat05	0,000772962	0,000665356	0,000340979	0,000593099	0,0002248736
ApaSat06	0,001135622	0,000991613	0,000313088	0,000813441	0,0004392602
ApaSat07	0,000244313	0,000241349	0,000075058	0,000186907	0,0000968751
ApaSat08	0,000795429	0,000817204	0,00035391	0,000655514	0,0002614238
ApaSat09	0,000415547	0,00056204	0,000658713	0,000545433	0,0001224306
ApaSat10	0,000625501	0,000677821	0,000379376	0,000560899	0,0001593656
ApaSat11	0,000312832	0,000311151	0,000294154	0,000306046	0,0000103327
ApaSat12	0,000309438	0,000357797	0,000354728	0,000340654	0,0000270777
ApaSat13	0,000452479	0,000006558	0,000070025	0,000176354	0,0002412276
ApaSat14	0,000224886	0,000228733	0,000214117	0,000222579	0,0000075763
ApaSat15	0,000456844	0,00038248	0,000123203	0,000320842	0,0001751527
ApaSat16	0,00027705	0,000275357	0,000256477	0,000269628	0,0000114205
ApaSat17	0,00011912	0,000062145	0,000077323	0,000086196	0,0000295057
ApaSat18	0,000231817	0,000277798	0,000262914	0,00025751	0,0000234621
ApaSat19	0,000237916	0,00029058	0,000418856	0,000315784	0,0000930658
ApaSat20	0,000272253	0,000205179	0,000013429	0,00016362	0,0001343235
ApaSat21	0,000060126	0,000088383	0,00003877	6,24263E-05	0,0000248864
ApaSat22	0,000172541	0,000198509	0,000146222	0,000172424	0,0000261437
ApaSat23	0,00014471	0,00008976	0,000064687	0,000099719	0,0000409305
ApaSat24	0,000096539	0,000087012	0,000024725	6,94253E-05	0,0000390036
ApaSat25	0,000117834	0,00007544	0,00000096	6,76247E-05	0,0000545386
ApaSat26	0,000172336	0,000265589	0,000072296	0,000170074	0,0000966664
ApaSat27	0,000116091	0,00016204	0,000079798	0,00011931	0,0000412154
ApaSat28	0,000116272	0,000178809	0,000161488	0,00015219	0,0000322887

ApaSat29	0,000120507	0,0001382	0,000130163	0,000129623	0,0000088588
ApaSat30	0,000153246	0,000186254	0,00018596	0,000175153	0,0000189729
ApaSat31	0,000119853	0,000162536	0	0,000144518	0,0000221042
ApaSat32	0,000041246	0,000034915	0,000039244	3,84683E-05	0,0000032360
ApaSat33	0,000202229	0,000253206	0,000149212	0,000201549	0,0000520003
ApaSat34	0,000120416	0,000416559	0,000051489	0,000196155	0,0001939621
ApaSat35	0,000076872	0,000083676	0,000081406	8,06513E-05	0,0000034642
ApaSat36	0,000107491	0,00013084	0,000131853	0,000123395	0,0000137823
ApaSat37	0,000046628	0,000065378	0,000002626	3,82107E-05	0,0000322117
ApaSat38	0,00008977	0,000129846	0,000107308	0,000108975	0,0000200899
ApaSat39	0,000075561	0,00015981	0,000041372	9,22477E-05	0,0000609567
ApaSat40	0,000315293	0,000346601	0,000362168	0,000341354	0,0000238739
ApaSat41	0,00004146	0,000041672	0,00002845	0,000037194	0,0000075733
ApaSat42	0,000061151	0,000071638	0,000087761	7,35167E-05	0,0000134041
ApaSat43	0,000103339	0,000178549	0,000154279	0,000145389	0,0000383850
ApaSat44	0,000132261	0,000126383	0,00005766	0,000105435	0,0000414783
ApaSat45	0,000010373	0,000145705	0,000017692	5,79233E-05	0,0000761092
ApaSat46	0,019614674	0,015329303	0,006954526	0,013966168	0,0064392112
ApaSat47	0,003388467	0,002721175	0,001252538	0,00245406	0,0010927310
ApaSat48	0,002378881	0,002425043	0,002822352	0,002542092	0,0002438073
ApaSat49	0,003373745	0,003416325	0,003975968	0,003588679	0,0003360768
ApaSat50	0,003011697	0,003178321	0,003316558	0,003168859	0,0001526506
ApaSat51	0,002141573	0,002366684	0,002416116	0,002308124	0,0001463400
ApaSat52	0,001925733	0,001041814	0,000923014	0,001296854	0,0005478552
ApaSat53	0,0019323	0,001777386	0,001217162	0,001642283	0,0003762250
ApaSat54	0,001602993	0,001890691	0,001017814	0,001503833	0,0004448069
ApaSat55	0,001126315	0,001022583	0,000858613	0,001002504	0,0001349758
ApaSat56	0,000658111	0,000800804	0,000950724	0,000803213	0,0001463214
ApaSat57	0,000666081	0,000703221	0,000406692	0,000591998	0,0001615505
ApaSat58	0,000680836	0,000613537	0,000500699	0,000598357	0,0000910228
ApaSat59	0,000621811	0,000439677	0,000060534	0,000374007	0,0002863431
ApaSat60	0,000664993	0,000549234	0,000721157	0,000645128	0,0000876661
ApaSat61	0,000500939	0,000501725	0,000566136	0,000522933	0,0000374167

ApaSat62	0,000416163	0,000398432	0,000290663	0,000368419	0,0000679200
ApaSat63	0,00038187	0,000381491	0,000293757	0,000352373	0,0000507630
ApaSat64	0,000321168	0,000390497	0,000549304	0,000420323	0,0001169560
ApaSat65	0,000379628	0,000342787	0,000199385	0,000307267	0,0000952269
ApaSat66	0,000346644	0,000347601	0,000329045	0,000341097	0,0000104480
ApaSat67	0,000297247	0,000318132	0,000262624	0,000292668	0,0000280359
ApaSat68	0,000234329	0,000168739	0,000018957	0,000140675	0,0001103946
ApaSat69	0,00019086	0,000111902	0,000034241	0,000112334	0,0000783104
ApaSat70	0,000182697	0,0002253	0,000276574	0,00022819	0,0000470052
ApaSat71	0,000227871	0,000225838	0,00013344	0,000195716	0,0000539425
ApaSat72	0,000210438	0,000224778	0,000284743	0,000239986	0,0000394180
ApaSat73	0,000238177	0,000228368	0,000298561	0,000255035	0,0000380121
ApaSat74	0,000150517	0,000181858	0,000193239	0,000175205	0,0000221245
ApaSat75	0,000167395	0,000166181	0,000143269	0,000158948	0,0000135923
ApaSat76	0,000197006	0,000276295	0,000149159	0,000207487	0,0000642127
ApaSat77	0,000164067	0,000196482	0,000258768	0,000206439	0,0000481293
ApaSat78	0,000178426	0,000161567	0,000125192	0,000155062	0,0000272067
ApaSat79	0,000161005	0,000157314	0,000137998	0,000152106	0,0000123562
ApaSat80	0,000190742	0,000167095	0,000208239	0,000188692	0,0000206485
ApaSat81	0,000164464	0,000151424	0,000080366	0,000132085	0,0000452617
ApaSat82	0,000139511	0,000183633	0,000081051	0,000134732	0,0000514577
ApaSat83	0,000139205	0,000126814	0,000049368	0,000105129	0,0000486863
ApaSat84	0,000098727	0,000093735	0,000099322	9,72613E-05	0,0000030684
ApaSat85	0,000082754	0,000127533	0,00004073	8,36723E-05	0,0000434088
ApaSat86	0,000154299	0,000163747	0,000169677	0,000162574	0,0000077558
ApaSat87	0,000153741	0,000131724	0,000150844	0,000145436	0,0000119632
ApaSat88	0,000081646	0,000112683	0,00002433	7,28863E-05	0,0000448231
ApaSat89	0,000133989	0,00010207	0,000041643	9,25673E-05	0,0000469007
ApaSat90	0,000070985	0,000104612	0,00011154	9,57123E-05	0,0000216929
ApaSat91	0,000103971	0,000101444	0,000065882	9,04323E-05	0,0000212987
ApaSat92	0,000154663	0,000165549	0,00002052	0,000113577	0,0000807736
ApaSat93	0,000112852	0,000135694	0,000145164	0,000131237	0,0000166108
ApaSat94	0,000078968	0,000083559	0,000050905	0,000071144	0,0000176772

ApaSat95	0,000086535	0,00007918	0,000016992	6,09023E-05	0,0000382049
ApaSat96	0,000076583	0,000101874	0,00005591	7,81223E-05	0,0000230206
ApaSat97	0,000074394	0,000077499	0,000096201	0,000082698	0,0000117965
ApaSat98	0,000065292	0,000087052	0,000079847	0,000077397	0,0000110850
ApaSat99	0,000082402	0,000094451	0,000075271	8,40413E-05	0,0000096945
ApaSat100	0,000070997	0,000063158	0,000013623	4,92593E-05	0,0000311099
ApaSat101	0,00006176	0,000062966	0,000045539	0,000056755	0,0000097320

1410

1411

1412 **Tabela 2.** Abundância, média e desvio padrão das três bibliotecas OB de *P. paranae*.

Satélites	Biblioteca 78769	Biblioteca 78969	Biblioteca 85970	Média da Abundância	Desvio Padrão OB
ApaSat01	0,043656434	0,025911086	0,032572547	0,034046689	0,0073191149
ApaSat02	0,005480538	0,003355543	0,004860473	0,004565518	0,0008922443
ApaSat03	0,001918363	0,001789742	0,003267618	0,002325241	0,0006684268
ApaSat04	0,001745319	0,00132027	0,001547606	0,001537732	0,0001736659
ApaSat05	0,000820675	0,00049822	0,000962385	0,000760427	0,0001942244
ApaSat06	0,000638331	0,000352726	0,000379621	0,000456893	0,0001287653
ApaSat07	0,000241439	0,000102777	0,000092848	0,000145688	0,0000678274
ApaSat08	0,000596079	0,000507715	0,000542488	0,000548761	0,0000363461
ApaSat09	0,00056403	0,000657071	0,000624718	0,000615273	0,0000385665
ApaSat10	0,000667553	0,000378804	0,000432405	0,000492921	0,0001254076
ApaSat11	0,000369628	0,000327027	0,000321858	0,000339504	0,0000214049
ApaSat12	0,000396934	0,000381743	0,000377879	0,000385519	0,0000082246
ApaSat13	0,000011147	0,000177288	0,000171609	0,000120015	0,0000770160
ApaSat14	0,000180881	0,000179034	0,000171686	0,0001772	0,0000039715
ApaSat15	0,000424049	0,000178462	0,000124055	0,000242189	0,0001304988
ApaSat16	0,000305688	0,000265257	0,000289206	0,000286717	0,0000165995
ApaSat17	0,000126465	0,000082316	0,000110553	0,000106445	0,0000182564
ApaSat18	0,000230152	0,000217755	0,000282235	0,000243381	0,0000279364
ApaSat19	0,000314985	0,000281225	0,000343311	0,000313174	0,0000253788
ApaSat20	0,000066616	0,000018627	0,000014517	3,32533E-05	0,0000236506
ApaSat21	0,000096436	0,000053131	0,00007799	7,58523E-05	0,0000177437

ApaSat22	0,000205826	0,000119603	0,000134627	0,000153352	0,0000376082
ApaSat23	0,000019418	0,000044732	0,000106206	5,67853E-05	0,0000364417
ApaSat24	0,000089756	0,000055271	0,000034698	5,99083E-05	0,0000227153
ApaSat25	0,000104691	0,000032869	0,000022554	5,33713E-05	0,0000365320
ApaSat26	0,000157805	0,000167637	0,000129342	0,000151595	0,0000162389
ApaSat27	0,000105922	0,000141025	0,000138747	0,000128565	0,0000160378
ApaSat28	0,00017996	0,000157118	0,000142492	0,000159857	0,0000154183
ApaSat29	0,000144899	0,000144293	0,000117799	0,000135664	0,0000126346
ApaSat30	0,00013818	0,000146901	0,000130516	0,000138532	0,0000066938
ApaSat31	0,000135969	0,000182161	0,000179044	0,000165725	0,0000210789
ApaSat32	0,000068962	0,000005672	0,000000155	2,49297E-05	0,0000312169
ApaSat33	0,000112676	0,000075532	0,000033797	7,40017E-05	0,0000322204
ApaSat34	0,000302007	0,000071511	0,000141403	0,00017164	0,0000964981
ApaSat35	0,000061685	0,000082715	0,000082421	0,000075607	0,0000098451
ApaSat36	0,000142815	0,000152617	0,000134319	0,00014325	0,0000074765
ApaSat37	0,000037711	0,000002092	0,000000437	1,34133E-05	0,0000171943
ApaSat38	0,000174696	0,000081704	0,000148299	0,0001349	0,0000391283
ApaSat39	0,000013448	0,0000018	0,000011599	0,000008949	0,0000051112
ApaSat40	0,000099808	0,000100535	0,00011055	0,000103631	0,0000049015
ApaSat41	0,000034474	0,000029098	0,000060686	4,14193E-05	0,0000137992
ApaSat42	0,000053957	0,000064165	0,000054484	5,75353E-05	0,0000046928
ApaSat43	0,000135839	0,000128735	0,000088157	0,000117577	0,0000210043
ApaSat44	0,000039106	0,000032298	0,00004438	3,85947E-05	0,0000049457
ApaSat45	0,000002208	0,000010269	0,000004072	5,51633E-06	0,0000034457
ApaSat46	0,013919456	0,006891176	0,009461186	0,010090606	0,0029035963
ApaSat47	0,002479199	0,001518547	0,001371607	0,001789784	0,0004911668
ApaSat48	0,002519209	0,002381845	0,003874217	0,00292509	0,0006734727
ApaSat49	0,003261246	0,003445212	0,00384843	0,003518296	0,0002452240
ApaSat50	0,002546825	0,002826398	0,002695477	0,002689567	0,0001142117
ApaSat51	0,001963134	0,00230068	0,00220451	0,002156108	0,0001419892
ApaSat52	0,002053265	0,001054137	0,001526217	0,00154454	0,0004080980
ApaSat53	0,001090026	0,000874406	0,001077779	0,00101407	0,0000988841
ApaSat54	0,001462514	0,001124408	0,001437498	0,001341473	0,0001538278

ApaSat55	0,001157212	0,000924461	0,001033999	0,001038557	0,0000950749
ApaSat56	0,000856901	0,000883524	0,000894515	0,000878313	0,0000157917
ApaSat57	0,000615658	0,000472202	0,000421618	0,000503159	0,0000821854
ApaSat58	0,000612293	0,000491577	0,000550227	0,000551366	0,0000492887
ApaSat59	0,000387397	0,000148883	0,00008966	0,000208647	0,0001286872
ApaSat60	0,000481949	0,000402506	0,000366866	0,000417107	0,0000481035
ApaSat61	0,000477244	0,000459181	0,000376898	0,000437774	0,0000436731
ApaSat62	0,000397041	0,000321349	0,000317278	0,000345223	0,0000366788
ApaSat63	0,000368198	0,00023873	0,000392772	0,000333233	0,0000675728
ApaSat64	0,000410106	0,000494872	0,000444107	0,000449695	0,0000348304
ApaSat65	0,000351067	0,000243292	0,000208214	0,000267524	0,0000607846
ApaSat66	0,000362484	0,000370125	0,00034909	0,000360566	0,0000086939
ApaSat67	0,000291075	0,000335649	0,000294723	0,000307149	0,0000202075
ApaSat68	0,00014465	0,000030398	0,000025712	0,00006692	0,0000549967
ApaSat69	0,000133529	0,000062645	0,000035497	7,72237E-05	0,0000413277
ApaSat70	0,000230509	0,000225301	0,000242431	0,000232747	0,0000071701
ApaSat71	0,000211374	0,000144566	0,000115428	0,000157123	0,0000401635
ApaSat72	0,000256553	0,000300658	0,000311802	0,000289671	0,0000238558
ApaSat73	0,000231107	0,000294419	0,000261827	0,000262451	0,0000258508
ApaSat74	0,00019614	0,000176307	0,000164628	0,000179025	0,0000130075
ApaSat75	0,00018278	0,000137861	0,000217946	0,000179529	0,0000327753
ApaSat76	0,000229804	0,000189245	0,000114812	0,000177954	0,0000476194
ApaSat77	0,00020915	0,000301324	0,000279989	0,000263488	0,0000393974
ApaSat78	0,000176988	0,000165104	0,000214707	0,0001856	0,0000211461
ApaSat79	0,000176744	0,000126445	0,00020097	0,000168053	0,0000310392
ApaSat80	0,000207607	0,00019626	0,000231042	0,000211636	0,0000144827
ApaSat81	0,000151391	0,000097698	0,000129688	0,000126259	0,0000220538
ApaSat82	0,00015472	0,000151824	0,000205319	0,000170621	0,0000245637
ApaSat83	0,000123878	0,000054801	0,000047993	7,55573E-05	0,0000342807
ApaSat84	0,000118474	0,000118872	0,000143836	0,000127061	0,0000118631
ApaSat85	0,000136649	0,000067342	0,000055174	8,63883E-05	0,0000358852
ApaSat86	0,000181108	0,000175458	0,000150198	0,000168921	0,0000134388
ApaSat87	0,00015718	0,000109714	0,000106719	0,000124538	0,0000231140

ApaSat88	0,000089299	0,000049038	0,000032126	0,000056821	0,0000239808
ApaSat89	0,000058587	0,000067115	0,000038838	5,48467E-05	0,0000118431
ApaSat90	0,000130625	0,000128051	0,000117751	0,000125476	0,0000055623
ApaSat91	0,000138109	0,000089702	0,000106506	0,000111439	0,0000200676
ApaSat92	0,000107613	0,000029426	0,00004955	6,21963E-05	0,0000331486
ApaSat93	0,00011169	0,000162515	0,000126195	0,000133467	0,0000213768
ApaSat94	0,000154075	0,000047434	0,000071642	9,10503E-05	0,0000456478
ApaSat95	0,000083095	0,000020218	0,000019008	4,07737E-05	0,0000299298
ApaSat96	0,000115814	0,000067484	0,00010916	0,000097486	0,0000213878
ApaSat97	0,000089209	0,000091038	0,000104332	9,48597E-05	0,0000067394
ApaSat98	0,000102875	0,000075473	0,000104525	0,000094291	0,0000133234
ApaSat99	0,00011145	0,000076778	0,000091878	9,33687E-05	0,0000141940
ApaSat100	0,000072978	0,000017297	0,000008624	3,29663E-05	0,0000285132
ApaSat101	0,000054389	0,000042077	0,000035122	4,38627E-05	0,0000079664

1413

1414

1415 **Tabela 3.** Razão de abundância entre três bibliotecas 1B, contendo uma possível variação de
1416 cromossomo B e três bibliotecas 0B de *Psilidodon paranae*, em ordem decrescente.

1417

Satélite	Proporção abundância 1B/0B
ApaSat45	10,5003323464
ApaSat39	10,3081536112
ApaSat20	4,9204190056
ApaSat40	3,2939371424
ApaSat37	2,8487077535
ApaSat44	2,7318455054
ApaSat33	2,7235737934
ApaSat68	2,1021368799
ApaSat92	1,8261097921
ApaSat59	1,7925392210
ApaSat06	1,7803765728
ApaSat23	1,7560696424
ApaSat89	1,6877476601

ApaSat53	1,6194958206
ApaSat60	1,5466726763
ApaSat32	1,5430745163
ApaSat01	1,5400928315
ApaSat100	1,4942314887
ApaSat95	1,4936682990
ApaSat13	1,4694370688
ApaSat69	1,4546619991
ApaSat83	1,3913804969
ApaSat46	1,3840762058
ApaSat47	1,3711484419
ApaSat15	1,3247619624
ApaSat101	1,2939249780
ApaSat07	1,2829242399
ApaSat88	1,2827358430
ApaSat42	1,2777655470
ApaSat25	1,2670597199
ApaSat30	1,2643498389
ApaSat14	1,2560849208
ApaSat71	1,2456276200
ApaSat43	1,2365428613
ApaSat08	1,1945359301
ApaSat61	1,1945271651
ApaSat50	1,1782041717
ApaSat57	1,1765616988
ApaSat87	1,1678100066
ApaSat76	1,1659589294
ApaSat24	1,1588593685
ApaSat65	1,1485559569
ApaSat34	1,1428238506
ApaSat10	1,1379099544
ApaSat22	1,1243674683
ApaSat26	1,1218974282

ApaSat54	1,1210306081
ApaSat58	1,0852277708
ApaSat51	1,0705049716
ApaSat62	1,0671933477
ApaSat35	1,0667178083
ApaSat18	1,0580530910
ApaSat63	1,0574352306
ApaSat04	1,0555467523
ApaSat02	1,0523665296
ApaSat03	1,0485681556
ApaSat81	1,0461406052
ApaSat49	1,0200049494
ApaSat19	1,0083350984
ApaSat93	0,9832917083
ApaSat70	0,9804222324
ApaSat74	0,9786603361
ApaSat73	0,9717445669
ApaSat85	0,9685605695
ApaSat55	0,9652848567
ApaSat86	0,9624262971
ApaSat29	0,9554756739
ApaSat67	0,9528524158
ApaSat28	0,9520382843
ApaSat66	0,9460025386
ApaSat16	0,9403976744
ApaSat64	0,9346846196
ApaSat27	0,9280128807
ApaSat56	0,9144948272
ApaSat79	0,9051053338
ApaSat11	0,9014484842
ApaSat99	0,9001021042
ApaSat41	0,8979864476
ApaSat80	0,8915860383

ApaSat09	0,8864899538
ApaSat75	0,8853629961
ApaSat12	0,8836260415
ApaSat31	0,8720367517
ApaSat97	0,8717930698
ApaSat48	0,8690644426
ApaSat36	0,8613918292
ApaSat52	0,8396376569
ApaSat78	0,8354630666
ApaSat72	0,8284789756
ApaSat21	0,8229981939
ApaSat98	0,8208312564
ApaSat91	0,8114962745
ApaSat17	0,8097728397
ApaSat38	0,8078201330
ApaSat96	0,8013697693
ApaSat82	0,7896546537
ApaSat77	0,7834863871
ApaSat94	0,7813700115
ApaSat05	0,7799555513
ApaSat84	0,7654716120
ApaSat90	0,7627959737

1418

1419

1420

1421

1422

1423

1424

1425

1426

1427

1428

Tabela 4. Gráfico contendo as estatísticas de estimativa, apresentado a média das duas bibliotecas 1B como Grupo Controle, a média das duas bibliotecas 0B como Grupo Teste, o Limite inferior e superior atingido pelo Grupo Teste, e, quando há diferença entre os grupos, demarcado por “S” na quinta coluna e por “N” quando a diferença não ocorre.

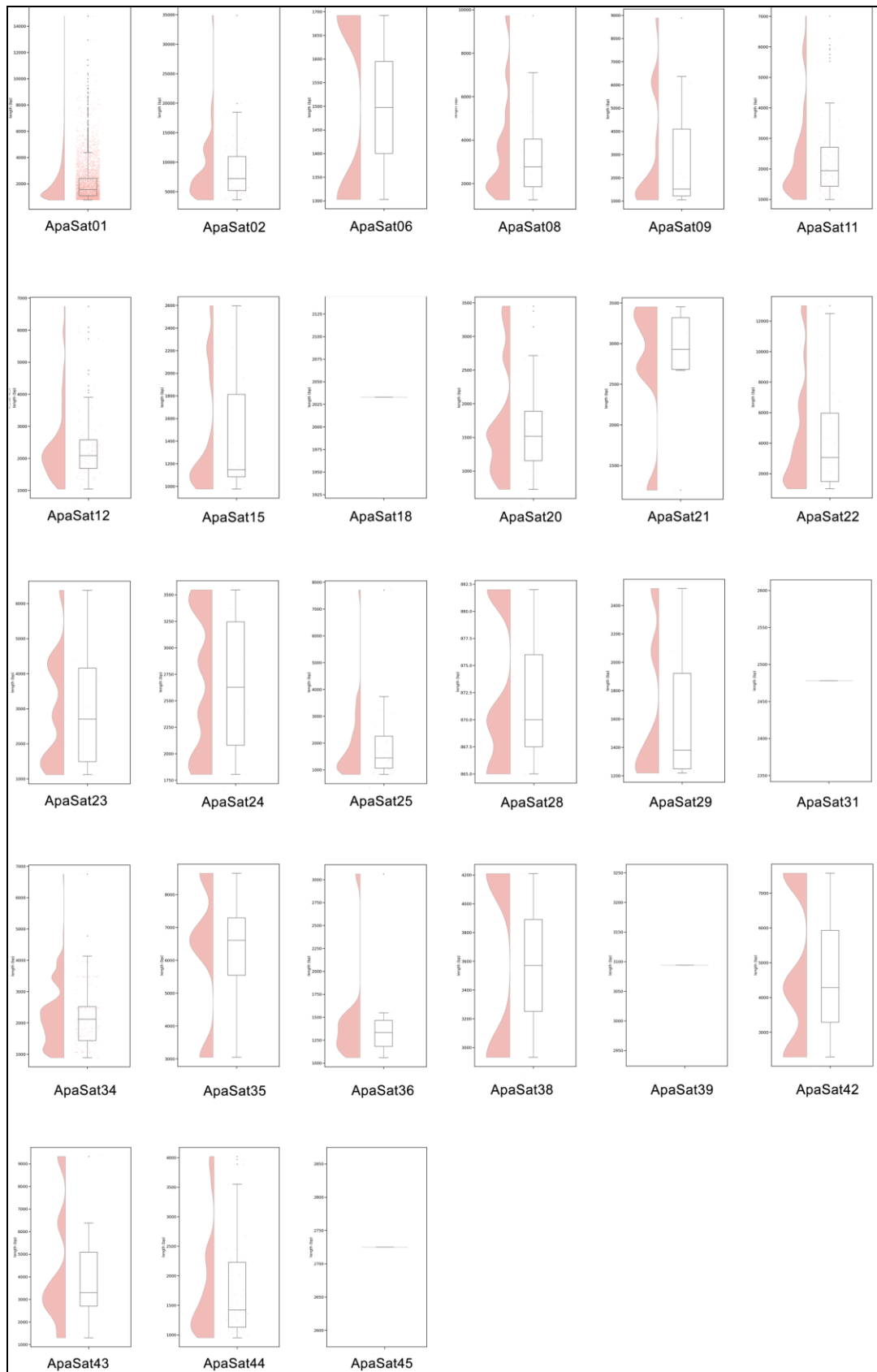
<i>Grupo Controle</i>	<i>Grupo Teste</i>	<i>Limite inferior</i>	<i>Limite superior</i>	<i>Diferença entre bibliotecas</i>	<i>Sinais no cromossomo B em FISH</i>
<i>ApaSat01B</i>	<i>ApaSat01A</i>	-	-	S	
		0,048157825000000	0,019919344000000		
<i>ApaSat02B</i>	<i>ApaSat02A</i>	-	0,000045625666667	N	X
		0,002041447000000			
<i>ApaSat03B</i>	<i>ApaSat03A</i>	-	0,000547364500000	N	
		0,001022162333333			
<i>ApaSat04B</i>	<i>ApaSat04A</i>	-	-	S	
		0,000413998000000	0,000054474333333		
<i>ApaSat05B</i>	<i>ApaSat05A</i>	-	0,000243226000000	N	X
		0,000220939000000			
<i>ApaSat06B</i>	<i>ApaSat06A</i>	-	-	S	X
		0,000773931000000	0,000439518666667		
<i>ApaSat07B</i>	<i>ApaSat07A</i>	-	-	S	
		0,000148501000000	0,000001392000000		
<i>ApaSat08B</i>	<i>ApaSat08A</i>	-	-	S	X
		0,000298601500000	0,000217213666667		
<i>ApaSat09B</i>	<i>ApaSat09A</i>	0,000022219333333	0,000219955333333	S	
<i>ApaSat10B</i>	<i>ApaSat10A</i>	-	0,000015892000000	N	
		0,000281150000000			
<i>ApaSat11B</i>	<i>ApaSat11A</i>	0,000010707000000	0,000057636500000	S	
<i>ApaSat12B</i>	<i>ApaSat12A</i>	0,000021370000000	0,000081144333333	S	
<i>ApaSat13B</i>	<i>ApaSat13A</i>	-	0,000166944000000	N	
		0,000387844666667			
<i>ApaSat14B</i>	<i>ApaSat14A</i>	-	-	S	X
		0,000057047000000	0,000045928500000		
<i>ApaSat15B</i>	<i>ApaSat15A</i>	-	0,000004387000000	N	X
		0,000314653333333			
<i>ApaSat16B</i>	<i>ApaSat16A</i>	-	0,000024837000000	N	
		0,000010946500000			
<i>ApaSat17B</i>	<i>ApaSat17A</i>	-	0,000053712000000	N	
		0,000027391666667			
<i>ApaSat18B</i>	<i>ApaSat18A</i>	-	0,000033057000000	N	X
		0,000055910666667			
<i>ApaSat19B</i>	<i>ApaSat19A</i>	0,000001898333333	0,000095953000000	S	X
<i>ApaSat20B</i>	<i>ApaSat20A</i>	-	-	S	X
		0,000256366000000	0,000155929333333		
<i>ApaSat21B</i>	<i>ApaSat21A</i>	-	0,000024012666667	N	
		0,000026965666667			
<i>ApaSat22B</i>	<i>ApaSat22A</i>	-	0,000020301000000	N	
		0,000073898000000			
<i>ApaSat23B</i>	<i>ApaSat23A</i>	-	-	S	
		0,000116854000000	0,000004045333333		
<i>ApaSat24B</i>	<i>ApaSat24A</i>	-	-	S	
		0,000057077500000	0,000002019500000		
<i>ApaSat25B</i>	<i>ApaSat25A</i>	-	0,000005310333333	S	
		0,000091841666667			
<i>ApaSat26B</i>	<i>ApaSat26A</i>	-	-	S	
		0,000136247000000	0,000011253666667		
<i>ApaSat27B</i>	<i>ApaSat27A</i>	-	0,000023415333333	N	
		0,000045176333333			
<i>ApaSat28B</i>	<i>ApaSat28A</i>	-	0,000051198666667	N	
		0,000031441666667			
<i>ApaSat29B</i>	<i>ApaSat29A</i>	-	0,000023786000000	N	
		0,000020401000000			
<i>ApaSat30B</i>	<i>ApaSat30A</i>	-	-	S	
		0,000053183333333	0,000011806666667		

<i>ApaSat31B</i>	ApaSat31A	-	0,000060230000000	N	
		0,000012208666667			
<i>ApaSat32B</i>	ApaSat32A	-	0,000030881500000	N	
		0,000039252000000			
<i>ApaSat33B</i>	ApaSat33A	-	-	S	X
		0,000219409000000	0,000114315666667		
<i>ApaSat34B</i>	ApaSat34A	-	0,000104759000000	N	X
		0,000321750666667			
<i>ApaSat35B</i>	ApaSat35A	-	0,000005549000000	N	
		0,000021991000000			
<i>ApaSat36B</i>	ApaSat36A	0,000006311000000	0,000039026666667	S	
<i>ApaSat37B</i>	ApaSat37A	-	-	S	
		0,000064389333333	0,000020790000000		
<i>ApaSat38B</i>	ApaSat38A	-	0,000076127000000	N	
		0,000028104000000			
<i>ApaSat39B</i>	ApaSat39A	-	-	S	
		0,000154743666667	0,000063345666667		
<i>ApaSat40B</i>	ApaSat40A	-	-	S	X
		0,000246550666667	0,000208081333333		
<i>ApaSat41B</i>	ApaSat41A	-	0,000010488666667	N	
		0,000012468000000			
<i>ApaSat42B</i>	ApaSat42A	-	-	S	X
		0,000017505333333	0,000038866666600		
<i>ApaSat43B</i>	ApaSat43A	-	0,000027764000000	N	
		0,000090392000000			
<i>ApaSat44B</i>	ApaSat44A	-	-	S	X
		0,000097693666667	0,000084942000000		
<i>ApaSat45B</i>	ApaSat45A	-	-	S	
		0,000142875666667	0,000002791000000		
<i>ApaSat46B</i>	ApaSat46A	-	-	S	
		0,011866828000000	0,003552532500000		
<i>ApaSat47B</i>	ApaSat47A	-	-	S	
		0,001967880000000	0,000575622000000		
<i>ApaSat48B</i>	ApaSat48A	-	0,001472255000000	N	
		0,000020117000000			
<i>ApaSat49B</i>	ApaSat49A	-	0,000340279000000	N	
		0,000133789000000			
<i>ApaSat50B</i>	ApaSat50A	-	-	S	
		0,000581945333333	0,000272579666667		
<i>ApaSat51B</i>	ApaSat51A	-	0,000094993666667	N	
		0,000323091333333			
<i>ApaSat52B</i>	ApaSat52A	-	0,000678408333333	N	
		0,000714236000000			
<i>ApaSat53B</i>	ApaSat53A	-	-	S	X
		0,000990103000000	0,000691442333333		
<i>ApaSat54B</i>	ApaSat54A	-	-	S	
		0,000661919666667	0,000157156333333		
<i>ApaSat55B</i>	ApaSat55A	-	0,000082763000000	N	
		0,000165341333333			
<i>ApaSat56B</i>	ApaSat56A	0,000064971333333	0,000229076666667	S	
<i>ApaSat57B</i>	ApaSat57A	-	-	S	
		0,000264741666667	0,000068993000000		
<i>ApaSat58B</i>	ApaSat58A	-	-	S	
		0,000169709000000	0,000034893500000		
<i>ApaSat59B</i>	ApaSat59A	-	-	S	
		0,000512410000000	0,000143347000000		
<i>ApaSat60B</i>	ApaSat60A	-	-	S	
		0,000286247000000	0,000093766000000		
<i>ApaSat61B</i>	ApaSat61A	-	-	S	
		0,000124041000000	0,000024088000000		
<i>ApaSat62B</i>	ApaSat62A	-	-	S	
		0,000097528000000	0,000010256500000		
<i>ApaSat63B</i>	ApaSat63A	-	0,000003089666667	N	
		0,000143140000000			
<i>ApaSat64B</i>	ApaSat64A	0,000030942666667	0,000145448666667	S	
<i>ApaSat65B</i>	ApaSat65A	-	-	S	
		0,000159721333333	0,000010140500000		
<i>ApaSat66B</i>	ApaSat66A	0,000001967500000	0,000022524000000	S	
<i>ApaSat67B</i>	ApaSat67A	-	0,000024760000000	N	
		0,000025841000000			

<i>ApaSat68B</i>	ApaSat68A	-	-	S
		0,000207055000000	0,000062173000000	
<i>ApaSat69B</i>	ApaSat69A	-	-	S
		0,000146313666667	0,000011050333333	
<i>ApaSat70B</i>	ApaSat70A	0,000001737000000	0,000055760000000	N
<i>ApaSat71B</i>	ApaSat71A	-	-	S
		0,000102730333333	0,000015480500000	
<i>ApaSat72B</i>	ApaSat72A	0,000031775000000	0,000093934666667	S
<i>ApaSat73B</i>	ApaSat73A	-	0,000055187000000	N
		0,000002165500000		
<i>ApaSat74B</i>	ApaSat74A	-	0,000035119000000	N
		0,000013337000000		
<i>ApaSat75B</i>	ApaSat75A	-	0,000039436000000	N
		0,000029534000000		
<i>ApaSat76B</i>	ApaSat76A	-	0,000005758666667	N
		0,000161483000000		
<i>ApaSat77B</i>	ApaSat77A	0,000012668000000	0,000121049500000	S
<i>ApaSat78B</i>	ApaSat78A	-	0,000044710500000	N
		0,000009360666667		
<i>ApaSat79B</i>	ApaSat79A	-	0,000033735166667	N
		0,000034560000000		
<i>ApaSat80B</i>	ApaSat80A	0,000009300333333	0,000056135333333	S
<i>ApaSat81B</i>	ApaSat81A	-	-	S
		0,000060246000000	0,000007267333333	
<i>ApaSat82B</i>	ApaSat82A	-	0,000048941666667	N
		0,000030843666667		
<i>ApaSat83B</i>	ApaSat83A	-	-	S
		0,000088942666667	0,000009131500000	
<i>ApaSat84B</i>	ApaSat84A	0,000019879666667	0,000047605000000	S
<i>ApaSat85B</i>	ApaSat85A	-	0,000030792666667	N
		0,000068303000000		
<i>ApaSat86B</i>	ApaSat86A	-	0,000023042333333	N
		0,000013549000000		
<i>ApaSat87B</i>	ApaSat87A	-	0,000009634000000	N
		0,000046023666667		
<i>ApaSat88B</i>	ApaSat88A	-	-	S
		0,000074919666667	0,000007865500000	
<i>ApaSat89B</i>	ApaSat89A	-	-	S
		0,000095151000000	0,000040640333333	
<i>ApaSat90B</i>	ApaSat90A	0,000016572333333	0,000057924000000	S
<i>ApaSat91B</i>	ApaSat91A	-	0,000035401500000	N
		0,000013005500000		
<i>ApaSat92B</i>	ApaSat92A	-	-	S
		0,000130680000000	0,000052493000000	
<i>ApaSat93B</i>	ApaSat93A	-	0,000037556333333	N
		0,000019169000000		
<i>ApaSat94B</i>	ApaSat94A	-	0,000072811500000	N
		0,000033829500000		
<i>ApaSat95B</i>	ApaSat95A	-	0,000002375000000	N
		0,000067123666667		
<i>ApaSat96B</i>	ApaSat96A	-	0,000034795000000	N
		0,000021744500000		
<i>ApaSat97B</i>	ApaSat97A	0,000012319666667	0,000028385500000	S
<i>ApaSat98B</i>	ApaSat98A	-	0,000038683000000	N
		0,000002445000000		
<i>ApaSat99B</i>	ApaSat99A	-	0,000022524000000	N
		0,000012639666667		
<i>ApaSat100B</i>	ApaSat100A	-	0,000005900500000	N
		0,000059482000000		
<i>ApaSat101B</i>	ApaSat101A	-	-	S
		0,000024922666667	0,000007371000000	

1434
1435
1436
1437
1438
1439

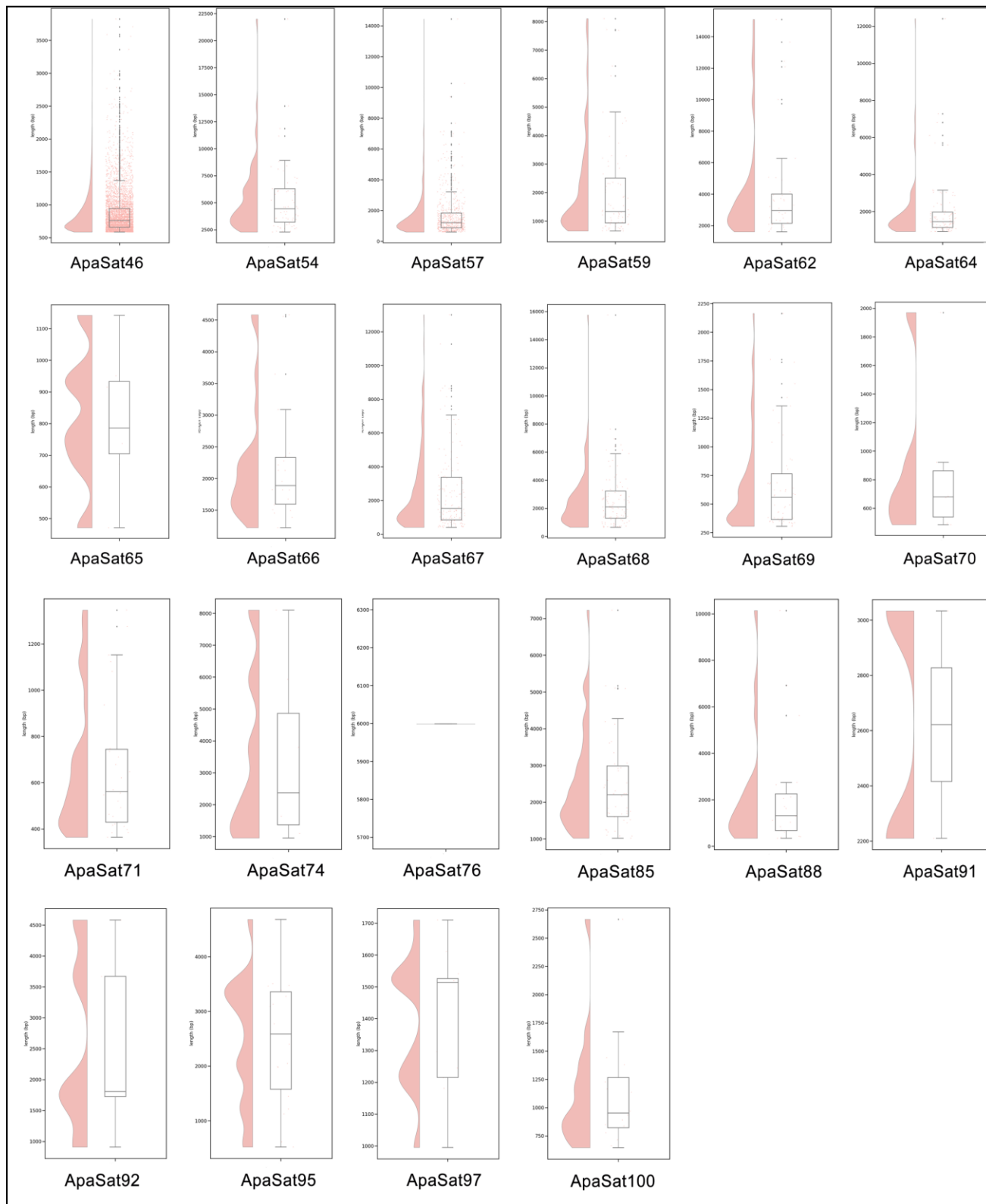
1440



1441

1442

Figura 1. Raincloud plots dos satélites caracterizados por Silva et al. (2017)



1443 **Figura 2.** Raincloud plots dos novos satélites caracterizados no atual trabalho.

1444

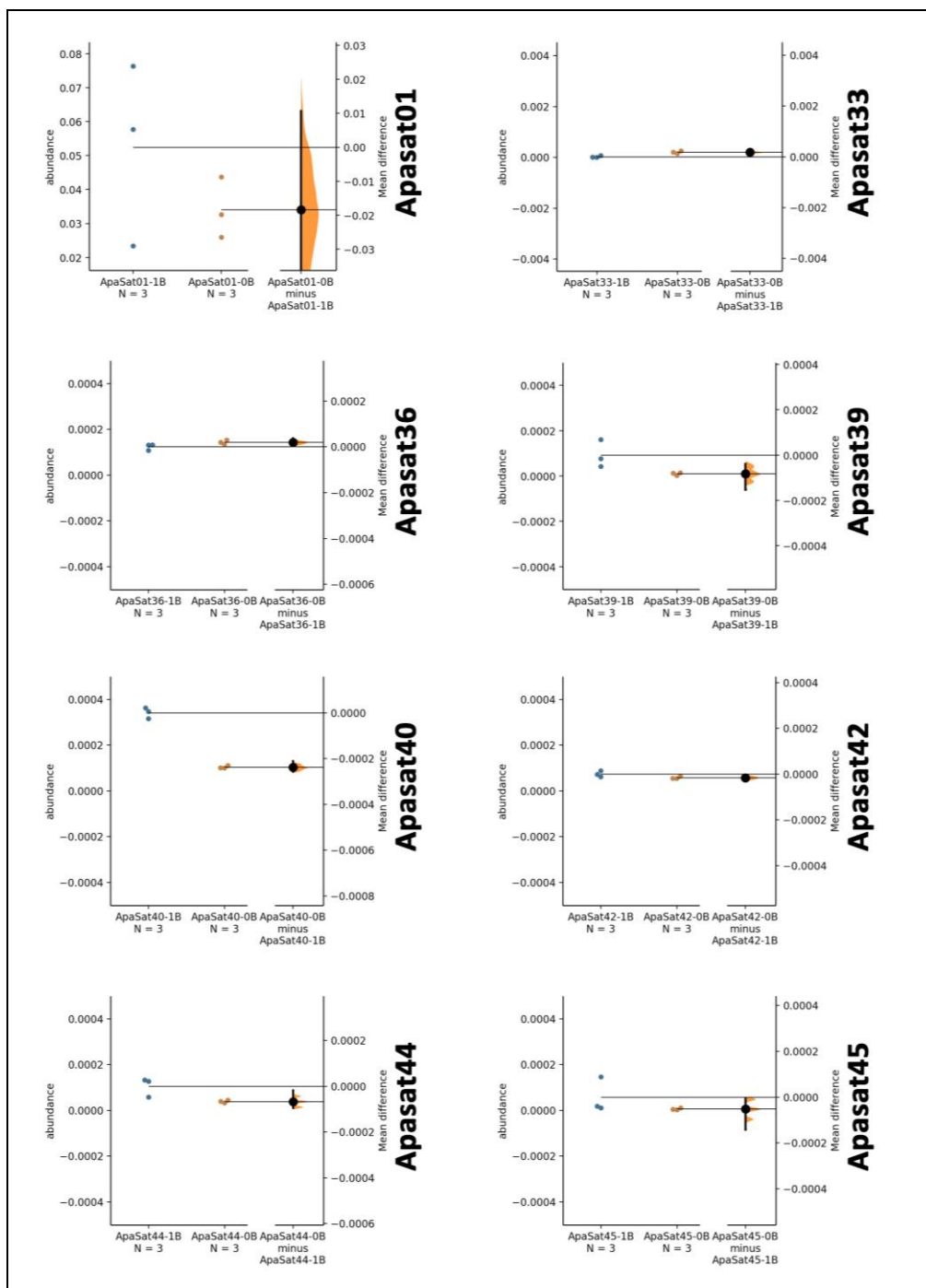


Figura 3. Gráficos de confiança apresentando um gráfico em que não há diferença entre as bibliotecas (ApaSat01) e todos os que apresentaram diferença média entre 1B e 0B. Ambos os grupos são plotados nos eixos esquerdos; a diferença média é plotada em eixos flutuantes à direita como uma distribuição de amostragem *bootstrap*. A diferença média é representada como um ponto; o intervalo de confiança de 95% é indicado pelas extremidades da barra de erro vertical. Nestes gráficos foi inserido a terceira biblioteca 1B, sendo possível notar a diferença dela para as demais em casos como ApaSat01, ApaSat39 e ApaSat44.